TU Delft

Flood event of July 2021: How extreme were the floods in the Ahr, Vesdre and Geul?

Anne Thewissen

Flood event of July 2021

How extreme were the floods in the Ahr, Vesdre and Geul catchments?

By

Anne Thewissen

to obtain the degree of Master of Science
at the faculty of Civil Engineering and Geosciences at Delft University of Technology,
in the tracks of Water Management and Hydraulic Engineering,
to be defended publicly on Monday October 31, 2022 at 13:30 h.

Student number: 4574125

Thesis committee:

Dr. M. Hrachowitz,	TU Delft (Chair)
Dr. Ir. A. Blom	TU Delft
Prof. Dr. Ir. M. Kok	TU Delft
Dr. Ir. L. Bouaziz	Deltares
Prof. Dr. Ir. B. Dewals	ULiège

Cover photo: Insul, Ahr Valley on July 15 2021 by Michael Probst, AP Photo

TU Delft

Preface

Most people dislike staying in tents during heavy rain. So when they announced exceptional rainfall on July 13 and 14 2021, I postponed my camping trip to the Ardennes. Instead, I went to the Belgian-Dutch border to look at the quickly rising water levels of the Meuse. Having lived nearby all my life, I have seen this river in all possible stages – I even walked through it during summer -, but I had never seen such a massive moving mass of water so close to the edge. I would start my thesis in November and finding a thesis topic had been on my mind for some time. When I finally drove through the Vesdre valley on July 16 and saw the damage from the floods, my subject was clear as day. Floods had driven me to pursue the study of Civil Engineering and even in another country, at a university renowned for its water knowledge. There would be no better way to finish this six-year adventure than with the very beginning.

Transforming the initial idea of this thesis topic into this document was not an easy task. It required an openness for exploring research possibilities, but also a resolute determination to preserve my initial, intrinsic motivation of being helpful to the victims of this flood event. The international study area complicated the assimilation and processing of data, as well as the multi-lingual literature studies. Many of these studies had not yet been published at the start of the thesis and the constant influx of new information proved to be a test in resilience.

I could not have completed this transformative journey of both my thesis research and my own skills without my thesis committee, that I hereby would like to thank. Markus, for the expert insight in the overall direction of my research, but also the small details. Astrid, for the many meetings during which the story and structure of this thesis slowly came to the surface. Laurène, for providing the wflow_sbm base models and data input and for all the help to navigate the wflow_sbm model. Matthijs and Benjamin, for their intelligent remarks during the committee meetings that helped me to dive further into certain aspects. I would also like to thank all of my friends for reminding me that fun diversions are a (maybe the most) necessary aspect of completing a thesis. Finally, the greatest gratitude go out to my parents, who have always supported me in everything I do, motivated me to keep pushing and helped me wherever they could, even if were late night spell checks.

Anne Thewissen
Delft, October 2022

Abstract

In mid-July 2021, heavy precipitation caused flooding for multiple tributaries of the Meuse and Rhine rivers which resulted in many deaths and exceptional damage. The Ahr in Germany, the Vesdre in Belgium and the Geul in the Netherlands suffered the greatest damage in each country. The first step in reducing such high damage from international, minor catchments is to build an understanding of the flood event.
The goal of this study was to analyse how extreme several hydrological aspects of the flood event of July 13-16 were for the Ahr, Vesdre and Geul. Analysing the extremity of an event requires a description of that event which helps with building knowledge. Its extremity places that knowledge into context. A flood with a 10-year return period will require a different approach than one with a 1000-year return period to balance the risk reduction and required effort. Three international catchments have been selected as case studies to allow an in-depth analysis of catchments and compliment current research on the flood event with a transboundary perspective.
The hydrological aspects considered for this study concern the forcing, flow mechanisms and response. For the forcing and response, a data analysis has been performed to study the extremity of the origin, magnitude, spatial and temporal variation of the precipitation, as well as the magnitude of the flow. Characteristics of both were used to determine the flood type. The data analysis on flow magnitude and seasonality has been supplemented by historical information to give more insight. The flow mechanisms have been derived from two models with varying structure and spatial variation by analysing parameter values and internal states.

The origin of the precipitation is considered as extreme since the cold core low, a low pressure area with a cold centre, was accompanied by several factors in a rare combination that likely contributed to the high precipitation amounts. The precipitation amount of July 14, the wettest day of the event, was of extreme magnitude according to the return periods of over 1000 years for several gauges. Although there was a large spatial variation in precipitation and thus their return periods, the spatial pattern showed a positive correlation of medium strength with the climatic pattern and is thus not considered as extreme. The temporal variation of the event is also not extreme. So during the event, only the amount of precipitation was extreme, but the precipitation amount in the preceding month was also extreme.
The peak discharges have been estimated to be around 990, 600 and 90 $[m^3/s]$ at the downstream gauges of the Ahr, Vesdre and Geul. The higher peak discharges coincide with sharper hydrographs and a stronger relief and impermeable soils. The return periods of the peak discharges were over 1000 years. Such extreme return periods are not surprising as such discharges had never been measured before. However, historical reconstructions show that similar flood events have occurred at least once for the Ahr. The assumed rare occurrence of summer floods in the regions was also debunked by historical information as 21 to 44% of the historical floods occurred during the summer half-year. Another assumption by reports on the flood event was the flood type identification of a flash flood. The large spatial and temporal scale, as well as lag times for downstream gauges of 13 to 24 hours, do not resemble a flash flood, but rather a long precipitation event with a fast response.
Two models, a lumped, simple model based on the HBV model and a distributed more complex model based on the SBM model, have been calibrated for several years preceding the flood event, resulting in a good performance. Their performance for the flood event was low due to a strong underestimation of the peak discharge. This underestimation was resolved by a re-calibration. The re-calibrated parameters and internal states of the models show an increase in fast responding mechanisms, up to unprecedented magnitudes. The SBM model includes more processes and shows a rare amount of overland flow due to saturation excess. The fast response mechanisms, both in magnitude and relative runoff contribution, were unprecedented and can be considered as extreme.

The magnitudes of the precipitation, flow mechanisms and discharge were extreme for the 2021 flood event. The extreme flows were possible due to fast flow mechanisms, most likely overland flow due to saturation excess, which can be linked to the extreme pre-conditions. The strong relief and impermeable soils also seem to play a role in this extreme response. All catchments are capable of a fast and strong runoff response and although the forcing of the event was extreme, a new flood event in the future can not be excluded and so a higher level of preparedness and protection is required.

This study should be considered as an exploratory first step for action on the flood event. In order to move forward and increase the protection of the region, more detailed research is required. Such research can focus on several areas. Varying the study regions could be helpful for specific purpose. Several examples are an analysis of the entire region for a precipitation analysis, one catchment study for flood measures, or a catchment comparison for detecting hydrological influences. The anthropogenic influences can also provide more insight and opportunities and could be included in these future researches. The use of models was limited in this research but their further development is essential for flood risk management. Uncertainty is another topic that deserves more attention. An important factor in decreasing the uncertainty is to construct robust and extensive measurement infrastructure. Nevertheless, without more international cooperation and data accessibility the additional data will not reach its full potential. International cooperation also provides more value to the previous recommendations. This research shows that the transboundary catchments of the case study can be compared and give more insights to the overall behaviour.

Contents

Chapter 1

Introduction

1.1 Background

Hochwasser, overstroming, inondation, no matter the language, the danger of a flood is universally understood. Defined as the "temporary covering by water of land not normally covered by water" (Directive 2007/60/EC, 2007), floods are researched for flood risk management by combining the scientific fields of hydrology and hydraulic engineering. Hydrology analyses the rainfall-runoff processes, while hydraulic engineering focuses on the flow behaviour and potential structural measures. The focal point of flood risk management has evolved throughout the years. Large coastal floods have been an integral part of history (Brázdil et al., 2012; De Kraker, 2013) and preventing such disasters has changed the landscape in which we live, especially in the Netherlands. In recent decades, major river floods have been given increased attention with projects such as Room for the River along the Meuse (van Stokkom et al., 2005; Rijke et al., 2012). However, on July 13-16 2021 a seemingly different type of flood event hit Western Europe.

On July 10, the European Flood Awareness System (EFAS) sent the first warnings for high risk of flooding to several regions in the Meuse and Rhine catchments (EMS, 2021). In the first hours of July 13, a precipitation event started that would last until the morning of July 16 due to a stationary cold core low over Western Europe. The first cases of flooding throughout Western Europe occurred in the night of July 13-14 (Ziemetz et al., 2021; Waterschap Limburg, 2021). In the evening of July 14, the water of several tributaries of the Meuse and Rhine rose with heights up to the ground floor and locally even the first floor (TFFF, 2021) and with a quick rising velocity that reminded some witnesses of a tsunami (Noulet, 2021; Ziemetz et al., 2021). During July 15, emergency services were the first to witness the devastating damage that had occurred during the night. The flood event resulted in at least 240 fatalities and insured damage of around 38 billion euros across several countries (Kramer & Ware, 2021).

The Ahr valley in Germany was the most fatally hit region with 134 deaths (Jordans & Casert, 2022). The majority of the 39 Belgian victims (Merckx, 2022) lost their lives in the Vesdre catchment (Le Vif, 2022). No fatalities were reported in the Netherlands. As for Belgium and Germany, the damage was mostly situated around small rivers, especially the Geul (TFFF, 2021). The density of flooded towns in fig. 1.1, which matches a similar heatmap of Verkade (2021), confirms that these three catchments were critical catchments in their respective country during the flood event. Since selecting multiple case studies allows for an in-depth analysis, but still preserves the possibility of generalisation, the critical catchments of the Ahr (Germany), Vesdre (Belgium) and Geul (mainly the Netherlands) have been selected as case studies for this research.

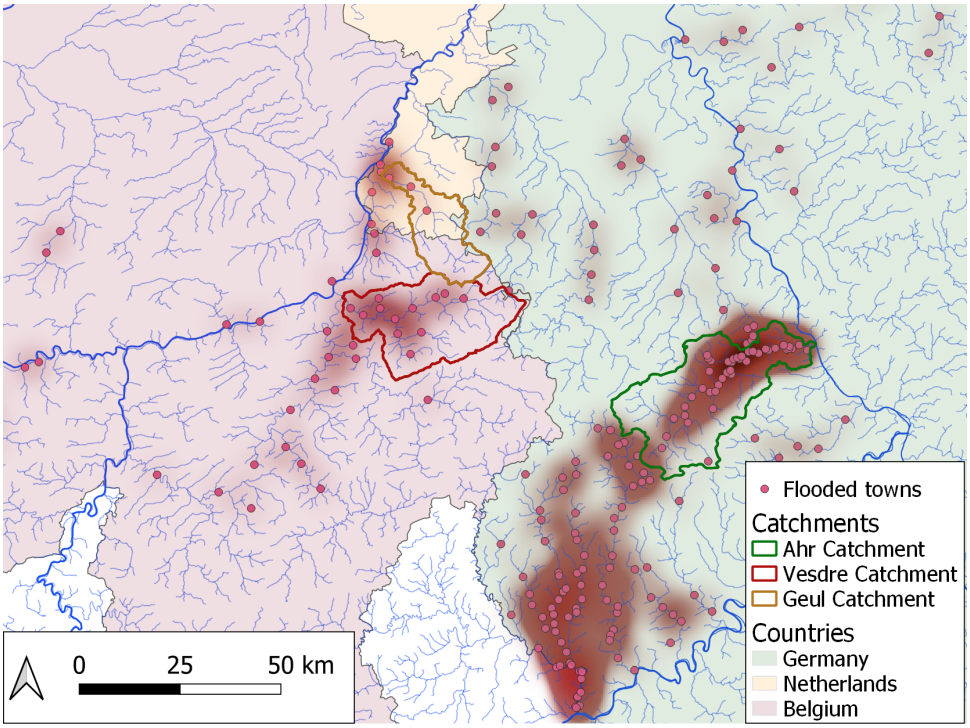


Figure 1.1: Map of the central area of the cold core low (Germany, Belgium, and the Netherlands), with the flooded towns indicated by a red dot, which together form the heat map that show the critical areas. The criteria for a flooded town was multiple flooded houses. The major rivers are indicated in dark blue and the minor ones in light blue lines. The catchment boundaries of the Ahr, Vesdre and Geul are respectively indicated by green, red and orange.

Not only are the case studies transboundary, all three catchments are tributaries of major rivers such as the Meuse and the Rhine. This differs from the aforementioned flood types along coasts and major rivers and could therefore be considered as a different flood type, one that has not yet been given our full attention (Cornwall, 2021). Flood categorisation is often source-based by differentiating in coastal, fluvial, pluvial floods and more (Camus et al., 2020) and are applied in European policies (Directive 2007/60/EC, 2007) and research (Paprotny et al., 2018). The 2021 flood event can be considered as a fluvial flood. The potential compounding effect of pluvial flooding and flooding due to water release from reservoirs is not considered in-depth. Fluvial floods can be categorised in more detail, for example as a flash flood. The categorization into flood types allows further research to derive general properties, which lead to an increased and wider understanding of a specific flood event. The fast rise in water levels suggest a flash flood, as to which the flood event has been referred by several sources (Ziemetz et al., 2021; TFFF, 2021).

Marchi et al. (2010) have acknowledged the knowledge gap on the characterization of flash floods and argued that this is partly due to the systematic lack of observational data, both for precipitation and discharge data, as well as their temporal and spatial detail. A decade later, Eeckman et al. (2021) reported on poor performance for modelling flash floods and highlighted the importance of complicated runoff patterns that are dependent on the soil type and its saturation. Additionally, many research papers, see table A.1 in appendix A, focus either on small flood events in one country with an in-depth analysis or performed a general analysis of flash floods of many international catchments. Researching floods of an extreme extent such as the 2021 flood event in international catchments provides unique challenges and insights. The knowledge gap on flash floods is not only present for their general characterisation but also for this study region in Western Europe. If the flood event of 2021 differs considerably from flash floods or any other flood type, it could be considered as a new flood type which would have an even greater knowledge gap.

The flood event also has a large societal importance due to the large damage and death toll. The most impacted regions still bear the scars of what happened. Knowing what happened and how such a disaster could take place, may help victims in their emotional processing. With the rebuilding in full speed, not considering new insights in areas and their flood risk would be a missed opportunity of improvement.

Most importantly, understanding this flood behaviour can help to increase our protection against similar floods in the future.

1.2 Research objective

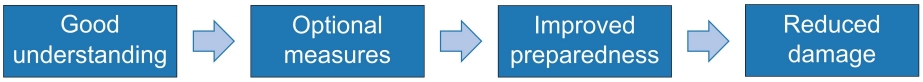


Figure 1.2: Flow chart of how good understanding of a flood event can reduce damage.

With good understanding of a flood event, optional measures can be taken and the preparedness can be improved. Fig. 1.2 shows that through this mechanism the damage can be reduced. Learning lessons from a flood event helps to take measures for future events. The relevance of these lessons depends on the extremity of the event. The goal of flood risk management is not to reduce the risk to zero, but to reduce the risk to a minimal level with reasonable effort. If the flood event or one of its aspects, such as peak discharge, is very extreme, the probability of occurrence of a similar event could be so low that the required effort for total flood protection would not be in balance with the risk reduction. Differentiating the flood event in different aspects increases the potential range of valuable insights. To combine learning lessons from a flood event and placing them into a bigger picture, the research question is the following:

> Which hydrological aspects of the flood event from July 13-16 2021 were extreme in the catchments of the Ahr, Vesdre and Geul?

Since the ultimate goal of flood risk management is to reduce damage, the most valuable study areas for building the understanding of flood events are those with the highest damage. This damage can be quantified by the death toll, financial damage, flooded buildings, et cetera. The catchments of the Ahr, Vesdre and Geul have been selected as case studies based on their critical damage. McPhillips et al. (2018) state that damage should not be included in the definition of an extreme event to assure the assessment of resilience throughout different events. For this reason and to limit the scope of this study, damage is not included as one of the considered aspects for the research question and only hydrological aspects are taken into account.

1.3 Methodology

The extremity of a flood event is not well defined and depends on spatial and temporal scales (Eisenbies et al., 2007; Seneviratne et al., 2012). In the IPCC reports an extreme weather event is defined as a rare event with rarity being quantified by the 10th or 90th percentile of a probability density function estimated from observations (IPCC, 2018). For more extreme cases, Extreme Value Theory can be applied (Seneviratne et al., 2012). Absolute thresholds can also be used, such as the waterlevel exceeding the outer bank level of a river. These various methods of quantification are usually only considered for peak magnitudes such as the discharge, but they can also be applied to other aspects of the extreme weather event. The next paragraphs describe which aspects are analysed for the research question and how their extremity is quantified. Due to the varying nature of each aspect, different quantification methods of extremity had to be applied.

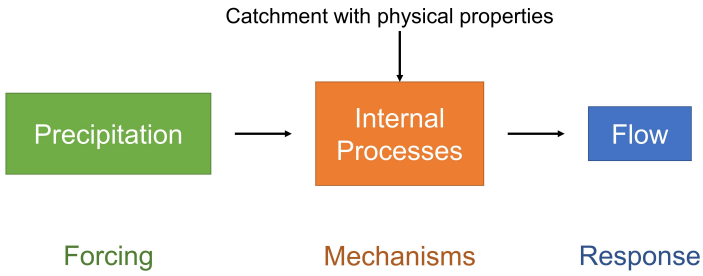


Figure 1.3: Schematization of the forcing, mechanisms and response of a flood event

The potentially extreme aspects can be categorized into 3 groups: the forcing, mechanisms and response. In fig. 1.3, the complex hydrological and hydraulic functioning of a catchment is reduced to a simple in- and output concept. The precipitation is the forcing of the event. The response to that forcing is the resulting flow. In between, many internal processes play a role that are dependent on the physical properties of the catchment. These processes are the mechanisms connecting the forcing and the response.

The forcing and response of a flood event can be observed and thus be subjected to a data analysis. As mentioned earlier, a good understanding of the flood event can improve flood protection. It also serves as a base to analyse its extremity. Therefore, the first sub-question of the research is the following:

1. How can the 2021 flood event be described in terms of forcing and response?

The flood event of July 13-16 2021 is from here on referred to as the 2021 flood event. Its description with the forcing and response is based on observations complemented with estimates since part of the measurement infrastructure was destroyed or rendered inaccurate. The forcing of the event is analysed for its origin, temporal and spatial variation of the precipitation and the soil moisture content. The response is analysed for the temporal and spatial variation of the discharge and the runoff coefficient. The flood type is an overall description of the flood event that includes both the forcing and response and is thus included in the consideration of this question. The description of the forcing and response during the flood event can then be used for answering the second sub-question, which is the following:

2. How extreme were the observed or estimated aspects of the forcing and response of the 2021 flood event?

For the forcing, the extremity of five aspects is considered. Firstly, the origin of the precipitation event, the cold core low, is discussed. Due to the complexities of this atmospheric event, the extremity is not quantified. The return periods of the measured daily precipitation quantify extremity with Extreme Value Theory. The extremity of the spatial variation of the precipitation during the event is quantified with the Pearson correlation coefficient. The extremity of the temporal variation is indicated with the 90th percentile. For the pre-conditions, which is the accumulated rain amount in the month preceding the event, the 90th percentile is applied as well. The response is analysed for the peak magnitude of the discharge with Extreme Value Theory. The extremity of the peak discharge is extended with historical information, as well as the seasonality of the forcing and response. The extremity of these additional aspects is not quantified due to the uncertainties. Nevertheless, they offer interesting insights.

The mechanisms of the flood event have not been considered so far since they cannot be measured directly and therefore require a different approach than a data analysis. With the use of models the mechanisms can be simulated for regular flows and the 2021 flood event. Comparing the mechanisms in both simulations helps us to answer the final sub-question:

3. What extreme mechanisms can be identified based on changes in model parametrization and internal states between a model made for the 2021 flood event and its preceding period?

If the performance of a model for regular flows is not the same for the 2021 flood event, this indicates that the mechanisms of the 2021 flood event differ considerably and can thus be considered as rare or extreme. This is only the case if the input data is correct, which is assumed for now. Which mechanisms

cause this potential low performance for the flood event can be identified through changes in parameters and internal states required to increase the performance of the model for the flood event. This method allows the analysis of the event extremity to be extended beyond a data analysis.

The answers to all three sub-questions form the answer of the main research question by discussing various aspects of the forcing, mechanisms and response and their extremity. The sub-questions form the structure of this thesis, which is discussed in the next section.

1.4 Reader guide

Before answering the sub-questions, the catchment characteristics are described to the reader in chapter 2 with a focus on context for spatial understanding, such as gauge locations, and hydrological properties. The topics of topography, measurement infrastructure, climate and the river system will be discussed in depth. The first sub-question on the description of the forcing and response of the 2021 flood event is answered in chapter 3. This also includes the determination of the flood type. The extremity of the forcing and response is analysed in chapter 4 and answers the second sub-question. The data analysis based on observations and estimates is supplemented with the magnitude and seasonality of historical flood events. The last sub-question is treated in two chapters. Firstly, chapter 5 describes the model set-up by elaborating on the purpose and methodology, which is followed by the structure of the used models. The input data is validated here as well. Secondly, chapter 6 gives the model results. This includes the performance of the models, the derived mechanisms and a discussion on the limitations. The conclusion in chapter 7 summarises all findings as an answer to the main research question and gives recommendations for future research.

Chapter 2

Catchment characteristics

This chapter analyses the characteristics of the three catchments to build spatial understanding of the catchments for the next chapters and to provide context for potential differences in hydrological behaviour. Some of these characteristics influence the hydrological response directly, such as topographic features like the mean elevation, the slope and geological features. The possible consequences of such features on the hydrological and hydraulic behaviour are also discussed. Local flooding influences such as backwater effects from bridges are not included, as this is beyond the scope of this research. The infrastructure of measurement devices does not influence the flow behaviour, but is described in the following section to give an idea of important stations for the rest of this chapter and report. The climate is another relevant topic to provide context of the three catchments. The flow duration curves and anthropogenic influences of the river system are also discussed. In conclusion, the characteristics of the three catchments are explained within the topics of topography, measurement infrastructure, climate and the river system.

2.1 Topography

Topography describes the physical features of an area. The elevation, shape, slope, geology and land cover are features that are analysed here for the three catchments. Since the goal of this chapter is to provide spatial context, the focus will be on physical features that influence the hydrological forcing and response of the catchment. Although this section shortly describes the river system, it is of such importance in this research that it will be discussed in detail in a separate section to allow a more in-depth analysis.

Fig. 2.1 shows the main topographic features of the catchments, such as the elevation, national borders and river systems. All catchments are part of a low mountain range in western Europe, the Rhine Massif, specifically the Ardennes (for the Vesdre and the south of the Geul catchment) and the Eifel (for the Ahr). The northern part of the Geul is on the foothills of the Ardennes. This division for the Geul is due to the national border of Belgium and the Netherlands which runs through the Geul catchment. The catchment of the Geul is mainly situated in the Netherlands (52%), as it is referred to in this thesis. Belgium (42%) and Germany (6%) make up the rest of the catchment area (Klein, 2022). The Vesdre is also a transboundary catchment, but the drainage area in Germany is negligible. The Ahr catchment is the only catchment situated in just one country: Germany. The Ahr river flows into the Rhine at Sinzig. The Vesdre and Geul are tributaries of the Meuse, although the Vesdre first flows into the Ourthe at Chaudfontaine. The Geul flows underneath the Juliana Canal (Julianakanaal) through a culvert several kilometers before entering the Meuse. For this study, the Geul catchment is considered down to this culvert since the flow bifurcated here during the 2021 flood event due to the limited flow capacity of the culvert (TFFF, 2021). The shared location of the three catchments within the Rhine Massif is likely to result similar topographic features, while the varying international locations complicate data accessibility.

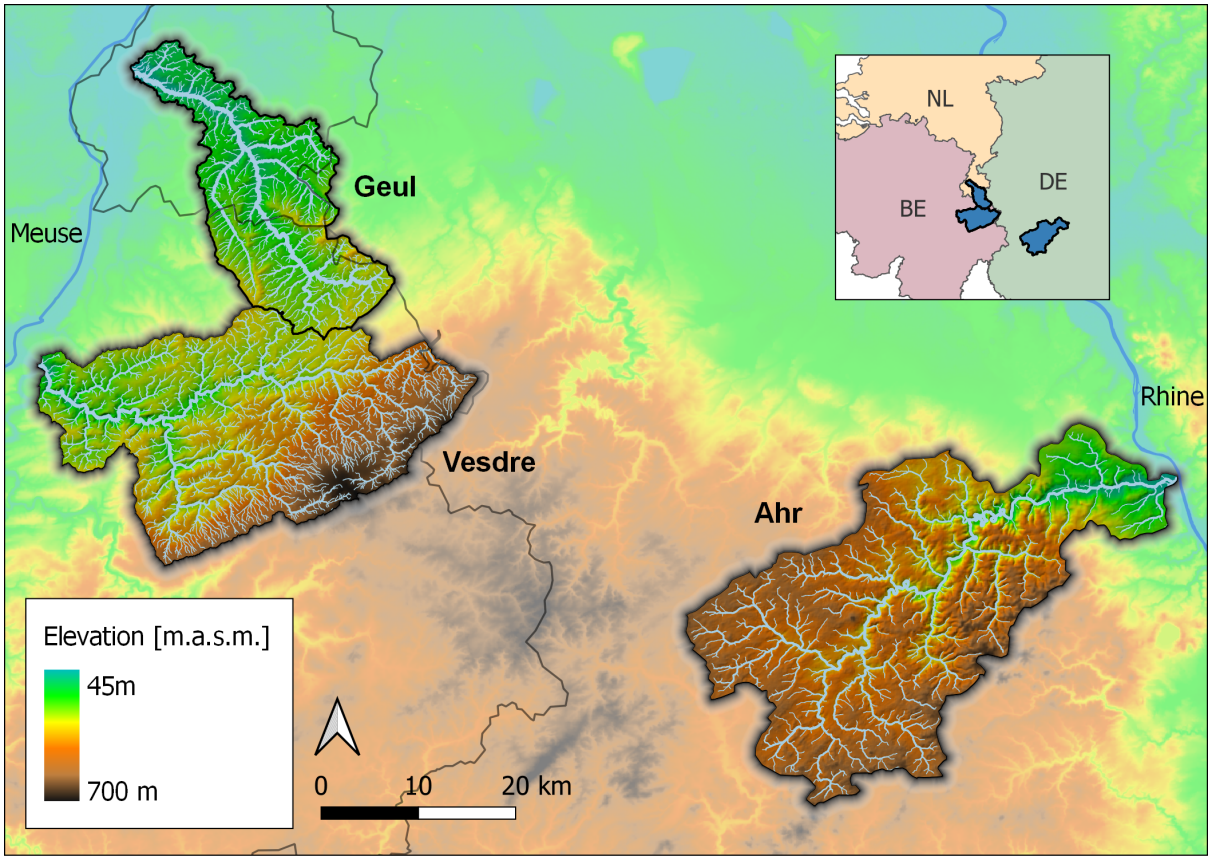


Figure 2.1: Elevation map of the Ahr, Vesdre and Geul catchments, based on the EU DEM v1.1 data, with a subfigure on the top right of their location in Germany (DE), Belgium (BE) and the Netherlands (NL).

From the information in fig. 2.1, several topographic characteristics can be derived. These are given in table 2.1 and have been calculated with the PCRaster plugin in QGIS in combination with the zonal statistics tool. For the length of the rivers, the following sources have been used:

1. Landesamt für Natur, Umwelt und Verbracherschutz Nordrhein-Westfalen (2019)
2. Contrat de rivière de Vesdre (2022)

Table 2.1: Topographic catchment characteristics, derived from the EU-DEM v1.1 data.

Characteristic	**Ahr**	**Vesdre**	**Geul**
Area [km^2]	898	699	336
Length [km]	89[1]	64[1,2]	60[1]
Mean elevation [m. a.s.m.]	410	358	191
River slope [-]	0.006	0.007	0.004
Mean landscape slope [%]	14.5	8.6	6.6
Elongation ratio	0.65	0.66	0.59
Relief ratio	0.012	0.014	0.009

Table 2.1 indicates that all catchment sizes are of the same order of magnitude, between 100 and 1000 [km^2]. The Ahr catchment has the largest area, followed by the Vesdre, which has double the area of the Geul catchment. As expected, the river lengths follow the same ranking from high to low, although the difference between the Vesdre and the Geul is now smaller. This indicates a slightly different shape of the catchment. The mean elevation and slope also decrease in magnitude from the Ahr to the Vesdre and finally to the Geul. The only parameter that does not have the highest value for the Ahr is the river slope, which is slightly higher for the Vesdre. For all topographic features of catchment area, river

length, mean elevation and slope, the highest values are for the Ahr, followed by the Vesdre and the values are smallest for the Geul.

The elongation ratio and relief ratio are two parameters, proposed by Schumm (1956) to compare basin shapes. The elongation ratio reflects the catchment shape. A value towards 0 reflects a thinly stretched catchment and a value of 1 is a perfect circle. The relief ratio reflects the relief relative to the catchment length. A high value indicates a steeper relief. The exact definitions of the ratios and their derivation can be found in appendix B.1. As the elongation ratios are very similar for all catchments, their forms are similar. The relief ratios are also similar. The highest relief ratio for the Vesdre catchment is in line with the previous statement that the Vesdre has the highest elevation gradient for the river. The similarity of the catchments for both parameters, as well as the other parameters in table 2.1, indicates a similarity in shape and slope. These two characteristics form the base of estimating the time of concentration (Kirpich, 1940), which is widely used to estimate peak discharges (Fang et al., 2008). In this regard, the similar shape and slope of the catchments indicate similar peak discharge behaviour.

Analysing the slopes within the catchments helps to identify the plateaus, hills and valleys and create a better understanding of the relief. Detailed figures of the slopes of the Ahr, Vesdre and Geul can be found in appendix B.2, respectively in fig. B.2, fig. B.3 and fig. B.4. In all catchments, the river network has cut deeply into the landscape. This happened during a phase of tectonic uplift, leading to V-shaped valleys (Kreienkamp et al., 2021). These steep valleys are the most pronounced for the Ahr and Vesdre catchment. Additionally, the valleys are quite narrow. The most downstream section of 20 km of the Geul is characterized by a wider valley, as well as the most downstream section of 10 km of the Ahr. For both the Vesdre and Geul, plateaus are also part of the landscape. Especially, the High Fens (Hoge Venen or Hautes Fanges) form a large plateau in the most upstream part of the Vesdre, where the river springs. Overall, the Ahr has the steepest slopes and most narrow valleys, followed by the Vesdre and the Geul has the least extreme elevation profile.

According to several reports on the 2021 flood event, there are several consequences of the catchment reliefs for the hydrological behaviour. The narrow valleys provide little flow area for the run-off, thus small floodplains, which leads high water levels (Bartholomeus, 2021). In case of a flood event, the narrow and steep valleys act as a funnel (Kreienkamp et al., 2021), leading to possibly dangerous situations. Additionally, the steep slopes lead to quick runoff (Van Heeringen et al., 2022; Roggenkamp & Herget, 2022; Mohr et al., 2022; Bachmann, 2021). TFFF (2021) even mentions this principle specifically for all case study areas of the Ahr, Vesdre and Geul. The conclusion of high water levels in the narrow valleys due to small flow areas is an understandable deduction. However, directly linking steep slopes to quick runoff processes is a crude assumption.

It is well known that a higher hydraulic gradient results in higher flow velocities for shallow water flow. In case of uniform flow, the hydraulic gradient is equal to the elevation gradient and then river flow can be approximated with the simple Chézy and Manning equations. Even when the flow is not uniform, a high elevation gradient such as in the study area will likely lead to faster river flow compared to a flat catchment. For overland flow outside a river bed, there are more factors, such as surface roughness(Meshkat et al., 2019), vegetation (Parsons et al., 1996), rills (Abrahams et al., 1996), that influence the runoff process. The statement that a high elevation gradient leads to faster overland flow can be true, but is not certain due to those other potential influences. A similar reasoning can be used for subsurface flow, which is also driven by the hydraulic gradient. In case of steep slopes, rapid subsurface flows can be the dominant source of storm runoff (Rumynin, 2015). However, this is only possible in case of preferential flow paths. These include macropores or thin, saturated soils above the impermeable bedrock layer. In conclusion, it is possible that steep slopes cause quick runoff, but it is not always the case.

The shape and slope of the catchment are not the only characteristics influencing the hydrological and hydrodynamic behaviour. Geology has a crucial influence on the flow paths that the water follows between precipitation and stream flow. Both the surface and bedrock geology are important, respectively for subsurface flow and groundwater flow, although that distinction is not strict. Surface geology is a crucial control of infiltration capacity (Jorgensen & Gardner, 1987), topsoil water storage and hydraulic conductivity (Bormann & Klaassen, 2008). Bedrock geology plays an important role in water storage, mixing and release (Pfister et al., 2017). When developing a distributed hydrological model, catchment-scale soil properties, such as water storage capacity and hydraulic conductivity, are essential and often based on soil databases (Vannier et al., 2014). An in-depth analysis of multiple soil layers, their age and hydrological properties is beyond the scope of this research The geology description below is limited to a

general soil description, the surface geology and the main hydrological properties, with a short validation using aquifer properties to include surface and bedrock geology.

All catchments are part of the Rhenish Massif, and due to that mutual formation, shown in appendix B.3, the region likely has common geological properties. Kreienkamp et al. (2021) describe this region as geologically characterised by paleozoic siliciclastic sedimentary rocks, regionally also limestone, locally interbedded with paleogene and neogene volcanic rocks. At the surface of these rock formations are relatively shallow soils. Although the catchments share the properties of shallow soils and sedimentary rocks, there are some spatial differences which are explained further.

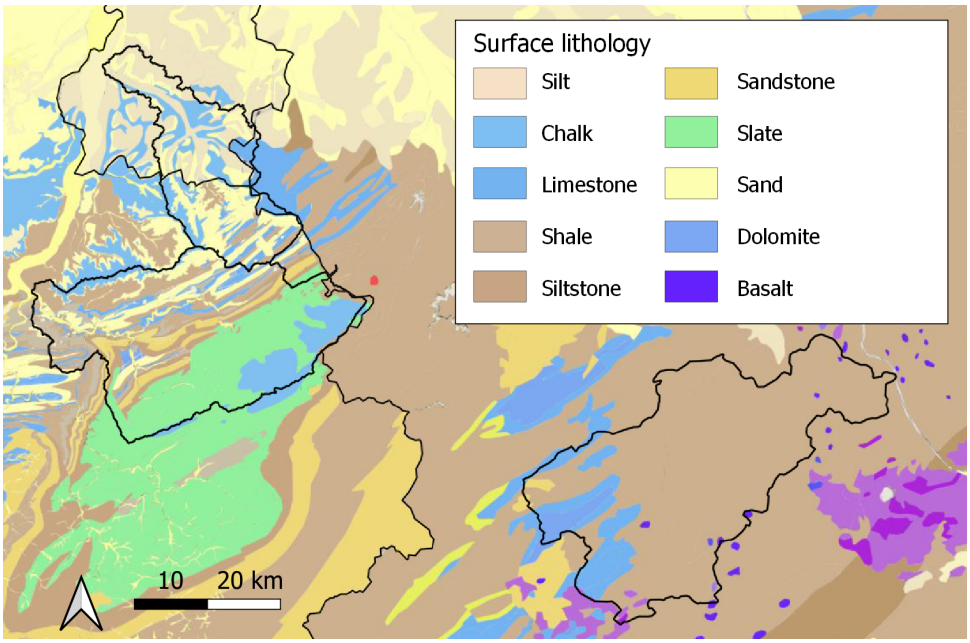


Figure 2.2: EGDI 1:1 Million pan-european Surface Geology, data from EGDI with INSPIRE conformity.

Fig. 2.2 shows the surface geology of the three catchments. The Ahr is mainly dominated by shale soils, with small patches of limestone, dolomite and basalt. Overall, the soil is very homogeneous, especially compared to the other catchments. Considering the soil regions of Belgium (KULeuven, 2020), the region south of the Vesdre river is considered as the Ardennes and is quite homogeneous, dominated by slate and some chalk. The region north of the main river is considered as the Land of Herve, and is heterogeneous, consisting of strips of sand, sandstone, shale and limestone. The southern, Belgian part of the Geul catchment is also located in the Land of Herve and consists of mainly sand, with parts of shale, chalk and limestone. The northern part is mainly chalk, silt and sand.

Several hydrological properties can be derived from this soil description. The main focus is the storage capacity and hydrological conductivity due to their importance in the hydrological subsurface processes. Firstly, the overall siliceous rocks in the Eifel and Ardennes region have a low water capacity due to their high skeletal content (Kreienkamp et al., 2021). For the Ahr, the thin soil in combination with the low hydraulic conductivity property of shale (Sallam, 2006) results in limited subsurface and deep groundwater flows. The south of the Vesdre catchment has a similar description, only with slate instead of shale which also has very low hydraulic conductivity. Describing the geohydrological properties of the northern part of the Vesdre catchment is difficult due to the heterogeneity. The strips of sand allow for water flow into the soil, but due to the alternation with siltstone, sandstone and limestone, this flow is still limited. In general, the flow through the rocks is expected to be small for the Vesdre, although some cracks in the rocks can carry water (Berger, 1992). With the overall impermeable soil types, the average annual precipitation discharge ratio for the Ahr and Vesdre will probably be higher than for permeable soils (Pfister et al., 2017).

In contrast to the Ahr and Vesdre, subsurface flows play an important role for the northern part of the Geul catchment since the soil consists mostly of sand, chalk and silt which all allow flow in the soil. Sand has a high permeability, leading to deep percolation into aquifers with a resulting delay and reduction of the flood response (Gaál et al., 2012). Chalk plays a similar role since it is a porous medium with fissures, resulting in high permeability (Rahman & Rosolem, 2017). Silt is less permeable but is situated

mostly underneath the downstream river and its main tributaries and as such plays a smaller role in runoff generation. The more permeable soils of the Geul, compared to the Ahr and Vesdre, will likely lead to a damped response in incoming precipitation due to the larger storage (Pfister et al., 2017).

Appendix B.3 contains the IHME1500 map, fig. B.6, showing the aquifers of the region. The Ahr and Vesdre are mostly classified as non-aquiferous rocks, which aligns with the impermeable soil properties. Both qualities indicate that subsurface flows will likely be small, especially groundwater flow. The Geul also has a region with non-aquiferous rocks but is dominated by highly productive porous aquifers. The aquifer map matches with the hydrological properties of the soil. In conclusion, subsurface flow is expected to be largest for the Geul and needs to be accounted for in the modelling efforts. For the Vesdre and the Ahr, the subsurface flows may be limited.

The surface hydrology is largely influenced by the land cover. In fig. 2.3, the five main land cover categories from the commonly used CORINE land cover dataset are shown with their coverage fractions for each catchment. Artificial surfaces include urban fabrics (with impermeable features such as roads and buildings), transport units, construction sites and artificial, non-vegetated areas. Agricultural areas include arable land, permanent crops, pastures and heterogeneous agricultural areas. Forests and semi-natural areas include forests, shrub and/or herbaceous vegetation and open land with little to no vegetation. Both wetlands and water bodies include inland and coastal locations. The land cover types of the CORINE dataset aid in identifying different land covers of the catchments and potential consequences on surface hydrology processes.

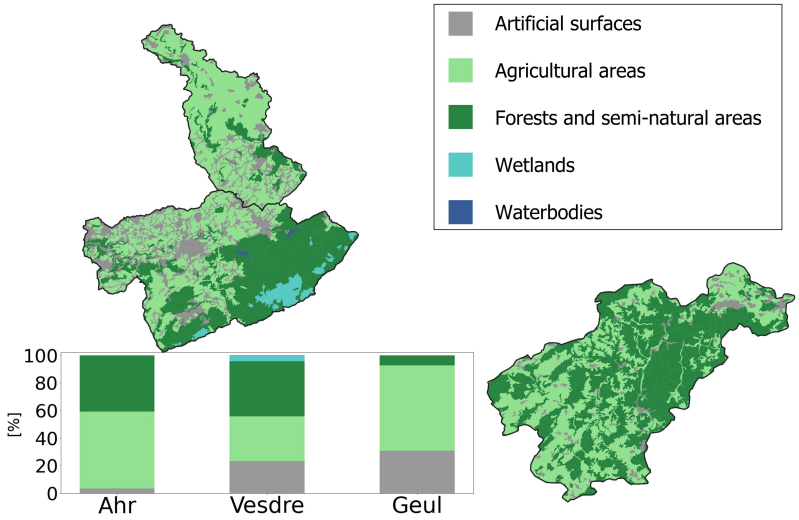


Figure 2.3: CORINE land cover map (CLC2018) from Copernicus, with the coverage fractions of each category for the catchments presented in a bar graph.

Fig. 2.3 shows that the fraction of the artificial surfaces is highest for the Geul (31%), followed closely by the Vesdre (23%). For the Ahr, it is substantially lower (3%). The fraction of agricultural area is relatively high for all catchments, but mostly for the Geul (62%) and Ahr (56%), where it is the dominant land cover. The Vesdre has a lower fraction of agricultural area (32%) and the dominant land cover is forests and semi-natural areas (40%). This fraction is even higher for the Ahr (41%) and very low for the Geul (7%). Additional land cover types for the Vesdre are wetlands (4%) and water bodies (0.3%). The Vesdre has a balanced variation of artificial, agricultural and natural land cover types, while the Ahr has little artificial surfaces and the Geul has little natural surfaces.

The land cover provides significant differences between the catchments which can influence the hydrological behaviour since the land cover is an important influence on the infiltration-runoff diversion. Low permeable surfaces can cause more surface runoff, while highly permeable surfaces such as forests can allow for more infiltration. The response times of these flow paths differ considerably. Although there are many other factors which play a role in the flow path partitioning, the higher fraction of impermeable surfaces in the Geul catchment may lead to a quicker run-off process compared to the Ahr and Vesdre in which natural land covers are more dominant.

The land cover differences are also reflected in the population density of the catchments. As could be expected from the low percentage of artificial surfaces, the Ahr catchment has the lowest pop-

ulation density of 112 [$inhabitants/km^2$]. The Vesdre and Geul catchments are similar in terms of artificial surfaces, which is also reflected in the population density. The Vesdre catchment has a population density of 316 [$inhabitants/km^2$], while the Geul catchment has the highest population density of 320 [$inhabitants/km^2$]. The population density is derived from the GEOSTAT 2018 grid dataset, from EuroGeographics, which is a 1 km by 1 km grid. Due to the strong relief, many of these inhabitants live near the river valleys which are easier for construction. This can be seen in fig. B.10, fig. B.11 and fig. B.12 in appendix B.5. Although the grid size is coarse, visual observations in the catchments confirm that the inhabitants live close to the river itself. With a total number of over 400.000 inhabitants in these three catchments, the societal relevance cannot be understated. Their high exposure increases their flood risk.

2.2 Measurement infrastructure

The rest of this chapter and research uses data from different measurement stations. Their locations are shown in fig. 2.4 to give some spatial context. To assure legibility, fig. B.7, fig. B.8 and fig. B.9 in appendix B.4 include the station names on separate figures for respectively the Ahr, the Vesdre and the Geul catchment. Several characteristics of the measurement stations such as the coordinates, record period, measurement frequency and source are given in the table B.1, table B.2 and table B.3 for the respective catchments.

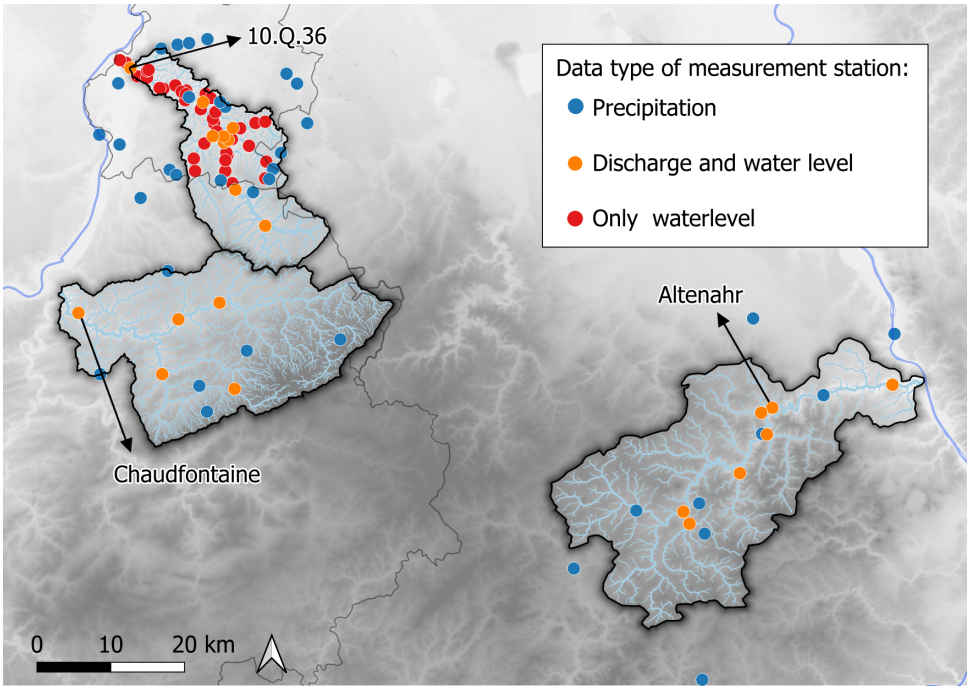


Figure 2.4: Locations of all measurement stations for precipitation, discharge and water level considered for this research. The main gauges of each catchment are labelled.

The main gauges indicated in fig. 2.4 are the most downstream discharge gauges used for most of the analysis in the rest of the report. For the Ahr, this is the "Altenahr" station, named after the village where it is located. The distance along the river between the gauge and the outlet is around 31 km. There is an additional station more downstream of Altenahr at Bad Bodendorf, but its record length is very short since it has only been constructed recently. Additionally, most reports concerning the 2021 flood event in the Ahr utilise the Altenahr station as a main point of reference (Mohr et al., 2022; Roggenkamp & Herget, 2022; Schäfer et al., 2021). For the Vesdre, the nearest gauge to the outlet is "Chaudfontaine", with a distance of 6 km to the confluence of the Vesdre and the Ourthe. For the Geul, the main gauge is "10.Q.36" at Meerssen, with a distance of 3 km to the confluence with the Meuse. The stations of Altenahr, Chaudfontaine and 10.Q.36, Meerssen will be considered as the main gauges for the Ahr, Vesdre and Geul in the rest of this research.

Fig. 2.4 indicates a higher station density for the Dutch part of the Geul compared to the other catchments. This may be due to a better familiarity with local organisations. Many of the gauges represent waterlevels, which are not considered in this study. The argumentation for this choice is given in section 3.2. A part of the precipitation gauges only have available data for the flood event. The high station density gives an overestimation in comparison to the stations used in this research. It contrasts with the Belgian part of the Geul which barely has any measurement stations. The Ahr and Vesdre have a similar station density. Having multiple measurement stations in all catchments gives the option of analysing and including spatial variability and offers secondary measurements in case of failures.

2.3 Climate

Since climate is defined as the average weather (IPCC, 2018) and weather mainly determines the forcing of a catchment, climate has a strong influence on the hydrological behaviour of the catchments. Understanding some main characteristics of the climate for the different catchments is imperative to understand the overall hydrological behaviour. According to the commonly used Köppen-Geiger classification system, the climate of all catchments is classified as Cfb, or a temperate oceanic climate (Peel et al., 2007). Its main characteristics are the lack of a dry summer or winter, yearly moderate temperatures and a warm summer. This climatic zone is quite large in Western Europe, spreading from Northern Spain to Germany and Ireland. Although the Köppen-Geiger classification shows that all catchments are within the same climatic zone and thus have a similar climatic forcing, this climatic zone is of such scale that a more in-depth look is required.

Another relevant climatic classification scheme is that of Budyko (1974), which shows climate characteristics of catchments. The Budyko plot is based on the coupled energy-water balance under the assumption that the long-term water balance is primarily determined by the atmospheric supply and demand of water (Berghuijs et al., 2014). Each catchment can be plotted on the Budyko plot when the precipitation, potential evaporation and actual evaporation or discharge are known. There are multiple uses, but in this section the Budyko plot is used to analyse the aridity of the catchments. The aridity index can be found on the horizontal axis and is quantified by ratio of potential evaporation and precipitation. A value higher than 1 indicates an arid catchment, while a value lower than 1 indicates a wet catchment. A more detailed description of the Budyko plot and how it has been applied for the gauges in this research can be found in appendix B.6.

According to fig. 2.5, the catchments of the Vesdre and Geul have an aridity index lower than 1. In this case, a catchment is wet, or in other terms: energy-limited. The Ahr has an aridity index higher than 1, indicating a dry or water-limited catchment. Since the Köppen-Geiger classification only utilizes monthly averages of precipitation, it is not exceptional that a more detailed analysis, such as the Budyko plot, shows one catchment that varies from the others. The possibility that the Ahr is a dry catchment according to the Budyko plot, in contrast with the wet catchments of the Vesdre and Geul, requires a further analysis.

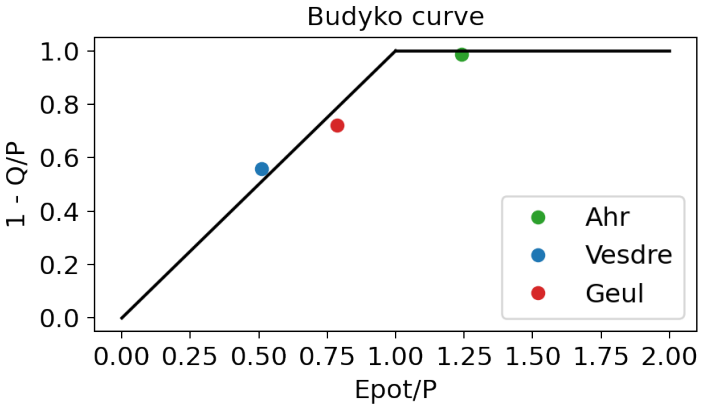


Figure 2.5: Budyko plot of the main gauges of each catchment, showing the climate classification according to Budyko (1974). The plot of all stations is given and discussed in appendix B.6, showing that the main gauges are representative for the entire catchment.

The precipitation is the most important factor in this study that is influenced by the climate. The spatial variation is the main interest since it shows how much variation there is within the catchments and how much precipitation averages vary between catchments. Low precipitation averages for the Ahr would offer an explanation for the Budyko results.

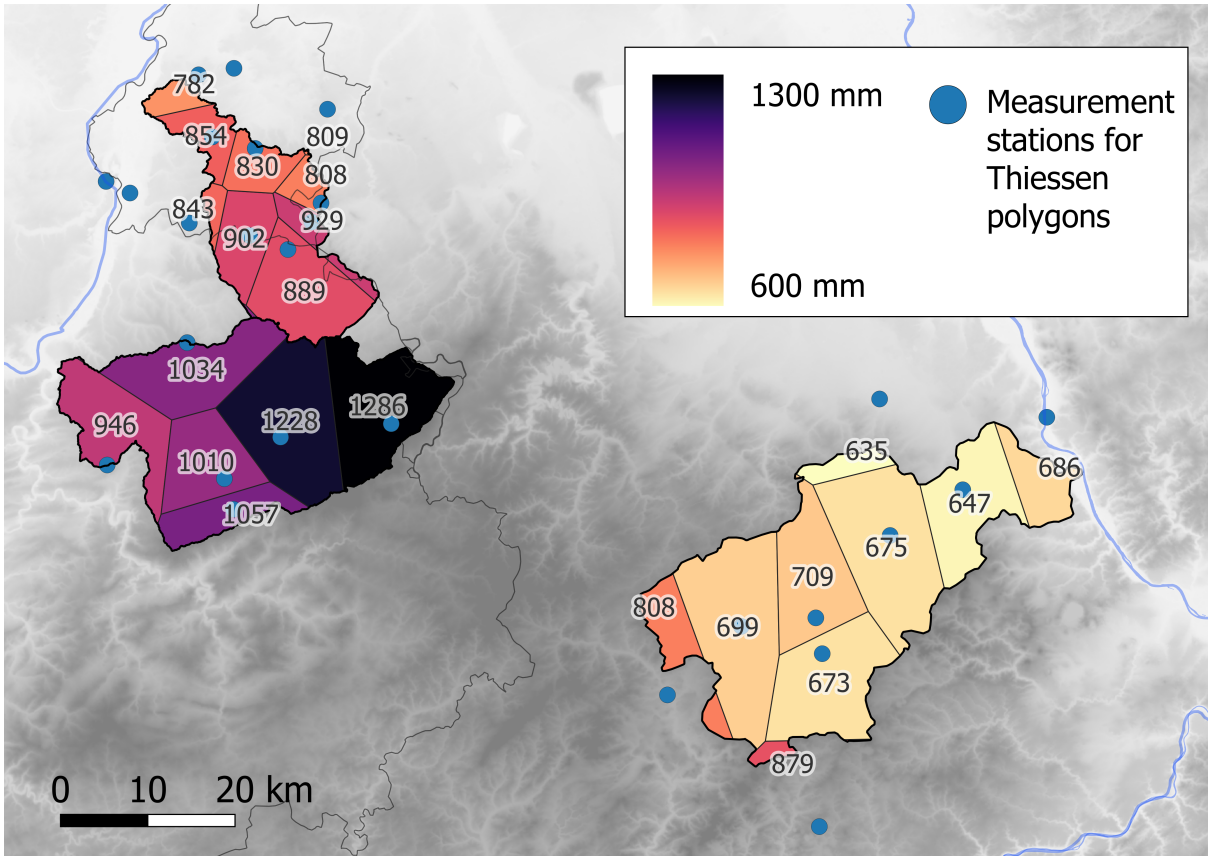


Figure 2.6: The mean annual precipitation amount for different areas of the catchments using areal interpolation with Thiessen polygons based on the measurement stations with precipitation records of over 10 years. The label indicates the mean annual precipitation amount for that area [mm/y]. More information on the stations and the making of this figure can be found in appendix B.7.

Fig 2.6 shows that the spatial differences of the mean annual precipitation are vary large. The values range between 647 and 1286 [mm/y]. The mean annual precipitation for the entire catchment can be calculated using the areal fractions as weights for the annual precipitation averages. The areal fractions are the areas of each Thiessen polygon divided by the total area of the catchment. For the Ahr, this amounts to an average annual precipitation amount of 697 [mm/y]. The Vesdre is noticeable by its dark shades in the figure, indicating high precipitation amounts. Its average value is the highest with 1120 [mm/y]. The Geul has an average value of 870 [mm/y], making it drier than the Vesdre but more wet than the Ahr. When discussing the event precipitation for the catchments, it is important to keep in mind that the average annual precipitation is relatively high for the Vesdre and low for the Ahr.

The low average precipitation of the Ahr can offer an explanation on why it qualifies as a water-limited and dry catchment according to the Budyko plot. In case the potential evaporation remains similar, which would not be surprising for a region in the same climatic zone with a similar elevation, lower precipitation amounts would result in a higher aridity index according to the Budyko plot. The ratios between the average annual precipitation of the catchments are similar to the ratios of their aridity index. To ensure that the location of the Ahr on the Budyko plot is not caused by a data error, a data check is performed in section 5.4.

Considering the relatively small distance between the catchments and their common climatic Köppen-Geiger classification, the large magnitude of climatic differences is not expected. However, the orography of the region may offer an explanation since the catchments are situated near or within the elevated areas of the Ardennes and Eifel. In case of winds from a westerly to north-westerly direction, clouds originating from the North Sea roll up these hills and due to orographic lifting, the humid air is forced to dew point temperature which can then result in increased precipitation. Brisson et al. (2011) suggested a similar pattern to explain the higher precipitation in the Ardennes compared to the rest of Belgium. This pattern could explain the lower precipitation amounts of the Ahr due to orographic shading from the Ardennes. Depending on the clouds, the wind direction and the exact orography, local precipitation amounts can

vary considerably for each precipitation event.

The relief can also influence the precipitation at a smaller scale. Specific locations of a gauge, such as in a valley or on a hilltop, could influence smaller differences but they are not taken into account here. It is important to realise that even the long term signal of precipitation shows considerable differences between stations within a singular catchment. For the Ahr, Vesdre and Geul the maximum range within the catchments is respectively 244, 340 and 147 [mm/year], which are significant values.

2.4 River system

In addition to the river network shown in fig. 2.1 and some characteristics of the river system in table 2.1, this section provides more details by discussing two more aspects of the river systems. So far, this chapter has described natural characteristics of the catchments. The land cover can be considered as natural and anthropogenic, but there is no intent of influencing the river system with the land cover. However, there are also deliberate anthropogenic measures which impact the river system and should be discussed. This chapter concludes with the flow duration curves which show several flow characteristics of the river systems that can be connected to previously described catchment characteristics.

Smaller water management measures such as weirs are present in all catchments. From personal observation, their presence is most numerous in the Geul, since a lot of tributaries are controlled by small weirs in the densely populated regions. The Vesdre has less, but larger weirs. The Ahr seems the least impacted by weirs. The weirs in the catchments regulate the flow regime and thereby dampen the natural flow signal. This could disturb signals in the data analysis or complicate the flow simulation in the models, which do not include weirs. Although there is no public data of the amount of weirs in the catchments, all rivers are affected by their influence, some to a higher degree than others.

Another anthropogenic influence that is noticeable for all rivers is the limited flow area in urban areas. Due to the steeper hills and advantages of waterways, urbanization has developed nearby rivers. The flow area of the rivers has been decreased in favor of urban development. The river has often been reduced to a straight, deep channel surrounded by retaining walls or even buildings. The city centre of Valkenburg provides an example, as shown on fig. 2.7. It is also an example of what happens when the flow exceeds the capacity of the channel as most of the city centre flooded during the 2021 flood event. Flow capacity in the rivers is also limited by bridges transversing the channels. Their flow obstruction becomes stronger in case of high water levels and debris accumulation. The International Meuse Commission (2005) confirms the considerable hydromorphological pressure for the Vesdre catchment, which is mostly caused by artificial banks and transversal obstacles. A limited flow area decreases the flow capacity and can create bottlenecks in the river system, leading to local floods. The straightening and deepening of channels in urban areas decreases the response time of catchments since the flow distance is shorter and flow resistance decreases from the lower roughness. Although there are multiple effects of a limited flow area for the river system, its influence is mostly local and is therefore not included in the rest of this research.

Figure 2.7: The main channel of the Geul in the centre of Valkenburg, limited in its course by retainment walls and buildings. In the background, a temporary bridge has been constructed since the original bridge of the Emmalaan was destroyed. Although the picture was taken on 11-02-2022, 7 months after the event, the water level of the flood event is still vaguely visible on the marl walls, as can be seen on the top right corner picture. Both pictures have been taken by the author of this research.

One of the most potent and relevant anthropogenic influences are controlled reservoirs. In the Vesdre catchment, two large reservoirs have been constructed with the purpose to provide drinking water, electricity, emergency storage for floods and conserve the biological and chemical river quality, (Ziemetz et al., 2021). The reservoirs of Gilleppe and Eupen have been inaugurated in 1878 and 1950 (Delforge, 2021), which can be considered as the starting year of their effect on the river. As shown in fig. 2.8, the reservoir of Eupen is situated upstream of the city Eupen and it has a volume of 25 $[10^6 m^3]$ and drainage area of 106 $[km^2]$, which is 15% of the total drainage area of the Vesdre catchment. The reservoir of Gileppe is situated upstream of Limbourg and has a volume of 26 $[10^6 m^3]$ and a drainage area of 54 $[km^2]$, which is 8% of the total drainage area. The two reservoirs block the natural runoff of 23% of the drainage area. This area has the highest precipitation (Berger, 1992), which can also be seen in fig. 2.6. The large storage capacity of the reservoirs and the human factor in their control can strongly affect the flow regime. Since there are no public gauges upstream of the reservoirs there is no unaffected flow information available. Due to the complexity, the influence of the reservoirs is not considered further in the data analysis.

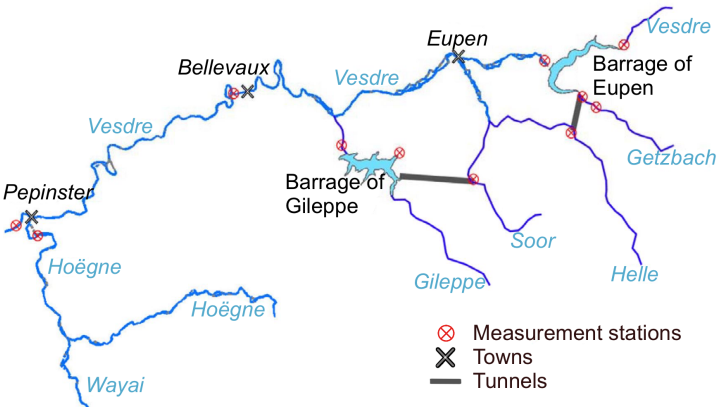


Figure 2.8: Situation of the Eupen and Gileppe Reservoir in the Vesdre catchment, from (Ziemetz et al., 2021).

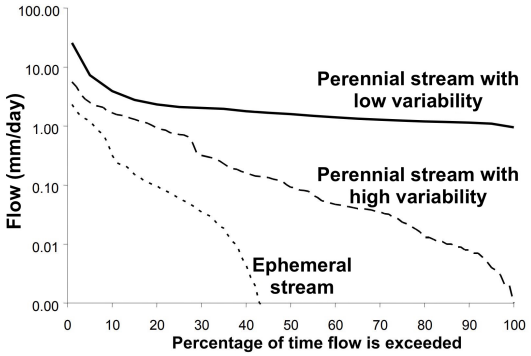


Figure 2.9: Typical flow duration curves for perennial and ephemeral streams, from Brown et al. (2006).

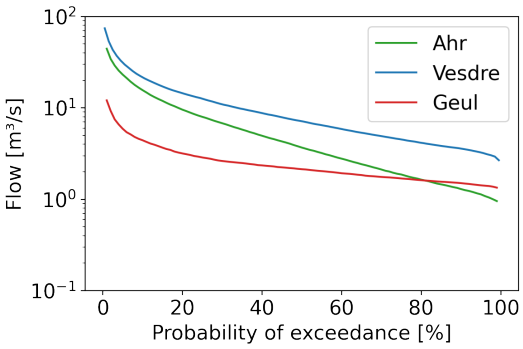


Figure 2.10: The average annual flow duration curves for the main gauges of the Ahr, Vesdre and Geul.

Many of the discussed catchment characteristics influence the flow regime. Flow duration curves can help to identify several characteristics of the flow regime by visualising how much of the time a certain discharge is exceeded. Fig. 2.9 shows three typical flow duration curves as an example for those characteristics. When the exceedance percentage for the lowest flow is less than 100, the river has periods of no flow. This defines an ephemeral river. The opposite is a perennial river, which has flow all year round. As the rivers in this research are mostly rain-fed and situated in an oceanic, non-arid climate, they are expected to be perennial. In fig. 2.10, the flow duration curves of the main gauges are given, as determined in appendix B.8. This appendix also includes the flow duration curves of each available year which show no climatic pattern. All curves of fig. 2.10 reach an exceedance probability of 100 which indicates that all main rivers are perennial.

The curvature of the flow duration curve indicates the variability of the flows. Fig. 2.9 shows that a more flat curve is a characteristic of a stream with low variability, whereas a steep curve is typical for a stream with high variability. The variability of streamflow can have multiple causes such as the precipitation variability or the storage capacity of the catchment. Anthropogenic influences, such as structural measures, around the stream can also affect the flow duration curve as most measures decrease the natural variability and thus create a more flat curve. The flow duration curves of fig. 2.10 have a comparable flow range, but the curve is steeper for the Ahr than for the Vesdre and Geul, which has the most flat curve. This indicates that the Ahr has the most variable flow, likely due to low storage capacity in the soil and limited anthropogenic measures, whereas the opposite is true for the Geul.

The flow regimes can partially be explained by the catchment characteristics. Steep slopes lead to high elevation gradients. The combination with thin, impermeable soils creates a high hydraulic gradient, which can lead to faster runoff processes. Due to the strong relief, this effect is mostly expected in the Ahr. The low storage capacity due to the impermeable soils and this fast runoff could be the reason of the high variability of the Ahr. The limited amount of anthropogenic influences is also in line with the high variability. The Vesdre forms the middle ground in most characteristics and also does so for the flow duration curves. The Geul has the weakest relief and most permeable soils, potentially leading to slower runoff processes and more infiltration. This creates a damped response which is reflected by its least steep flow duration curve for all catchments. The land cover would have an opposite effect, creating less infiltration in the Ahr and less in the Geul, but this effect seems counteracted by the other characteristics. So even though the Ahr receives less precipitation than the Vesdre and Geul, its high variability could still lead to very high discharges. The opposite is true for the Geul, where the more damped response of the catchment could lead to lower peak discharges. Overall, the catchments do not differ in such a strong way that they cannot be compared.

Chapter 3

Flood event characteristics

The 2021 flood event was characterised by an intense precipitation event, followed by high flows. This chapter answers the first sub-question of the research by describing the forcing and response of the event. Their extremity is analysed in the next chapter. The forcing and response also form the foundation for the chapters on the models.

3.1 Forcing

The flood event of July 13-16 2021 started with a precipitation event of multiple days with varying intensity. The precipitation did not only vary in time, but also in space. The highest precipitation amount of 293 $[mm/4d]$ was recorded in the Vesdre, while the lowest precipitation amount was recorded in the Ahr with a value of 100 $[mm/4d]$. The forcing of the flood event is analysed for the origin, magnitude, temporal and spatial pattern.

Origin: cold core low

The extreme precipitation on July 13-16 originated from the cold core low "Bernd". A cold core low consists of a low pressure area accompanied by cold air in the upper reaches of the atmosphere. This low pressure area does not coincide with the usual frontal systems and attracts humid air from a large region. During summer, the temperature differences between the surface and that upper cold air provide energy for strong convection and rainfall formation (TFFF, 2021). Moreover, the low pressure area was blocked by high pressure areas in Eastern Europe and the Atlantic Ocean, as can be seen in fig. 3.1, which made the low pressure system stationary. The warm surface temperature in the Mediterranean region seemed to have been a source of warm and unusually humid air (Kreienkamp et al., 2021; Schäfer et al., 2021). The North Sea and Baltic Sea are also considered as major moisture sources (Mohr et al., 2022). The large inflow of humid air led to high moisture masses in the cold core low. Due to its stationarity, it caused a long period of consistent rainfall in the same region.

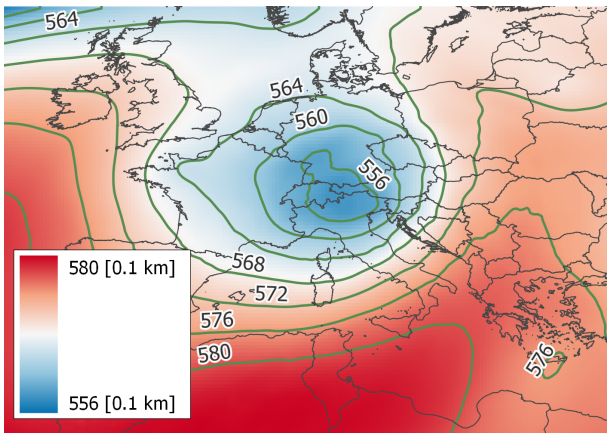


Figure 3.1: 500-hPa Geopotential height [0.1 km] on July 14 2021 14:00 h. Data from ERA5 Hourly Reanalysis of pressure levels

Temporal analysis

The precipitation pattern varied considerably in space and time. The temporal analysis of the precipitation is based on gauge observations as they provide the most accurate time series for one location. The hourly precipitation measurements for each catchment are plotted both directly (known as the hyetograph) and cumulatively (known as the mass curve). On the mass curve, the centroid and its moment of occurrence have been plotted for the events that have no data gaps. The centroid indicates the centre

of mass, in this case 50% of the total precipitation volume. Its moment of occurrence can be used to determine the lag time, which is helpful for comparing runoff velocities. The time range for the 2021 flood event is taken from July 12 to July 16, as little to no rainfall occurred before and after these days. The direct and cumulative precipitation plots build an understanding of the temporal variation, intensity and magnitude of the precipitation event.

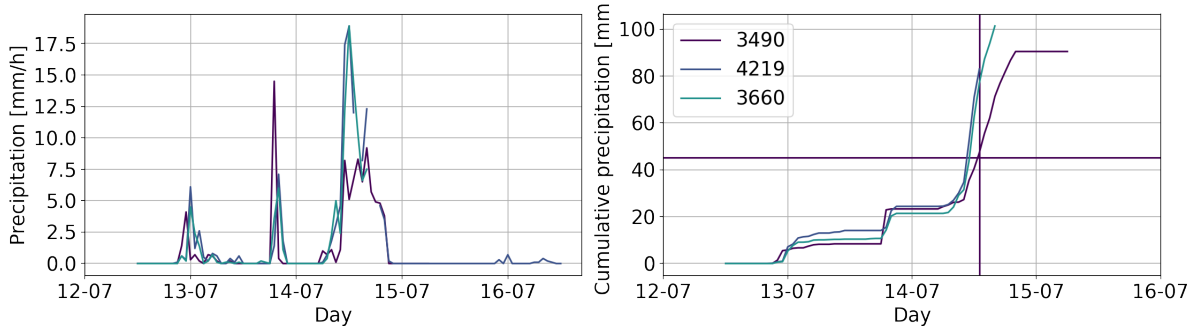


Figure 3.2: Left: Hourly observed precipitation from July 12-16 2021 for gauges in the Ahr catchment. Right: cumulative plot of the same data.

The hyetograph for the Ahr catchment in fig. 3.2 shows three distinguished peaks, with the first occurring at the start of July 13 around midnight. The next peak is situated between 19:00 h and 20:00 h on July 13. The majority of the event precipitation, around 80%, fell on July 14, when it rained from the early morning until the evening. This pattern can be seen in the mass curve on the right as there are flat parts with sharp rises which indicate low precipitation periods followed by intense peaks. The centroid of the rain event can only be plotted for 1 gauge due to many data gaps. Its moment of occurrence is on July 14 at 13:00 h. The maximal precipitation amount for the event is around 100 $[mm/4d]$.

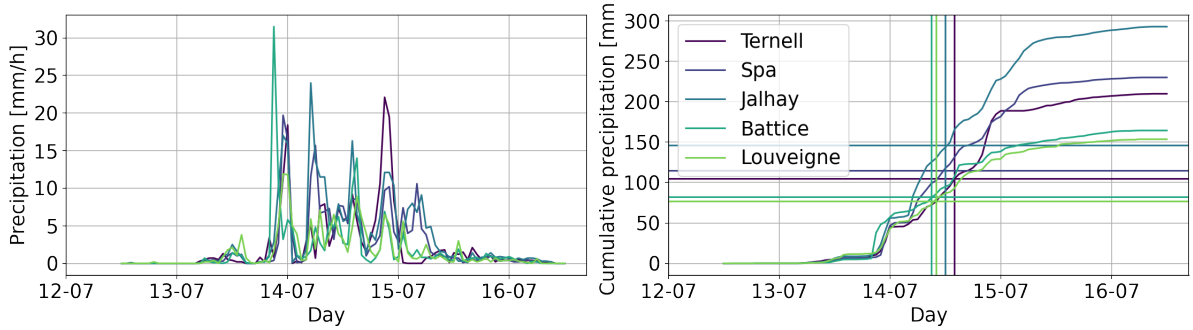


Figure 3.3: Left: Hourly observed precipitation from July 12-16 2021 for gauges in the Vesdre catchment. Right: cumulative plot of the same data.

The hyetograph of the Vesdre in fig. 3.3 shows three major and two minor peaks with constant precipitation throughout, which is in contrast with the three separate peaks of the Ahr. The precipitation of the Vesdre started later and lasted longer as well. This difference between the precipitation patterns of the Vesdre and Ahr is also visible in the mass curves as there is a more constant rise. The moment of occurrence of the centroids is in the morning of July 14 which is several hours before the Ahr. The most noticeable difference is the amount of cumulative precipitation. The highest recorded amount in the Vesdre catchment is 293 $[mm/4d]$, which has been recorded for the station of Jalhay and is almost triple the maximum recorded amount of the Ahr. Each station of the Vesdre catchment has recorded a higher total precipitation amount than any station of the Ahr catchment. The total precipitation volumes per catchment will be discussed later on. The event precipitation of the Vesdre is characterised by high total sums of precipitation due to a prolonged, uninterrupted period of rain.

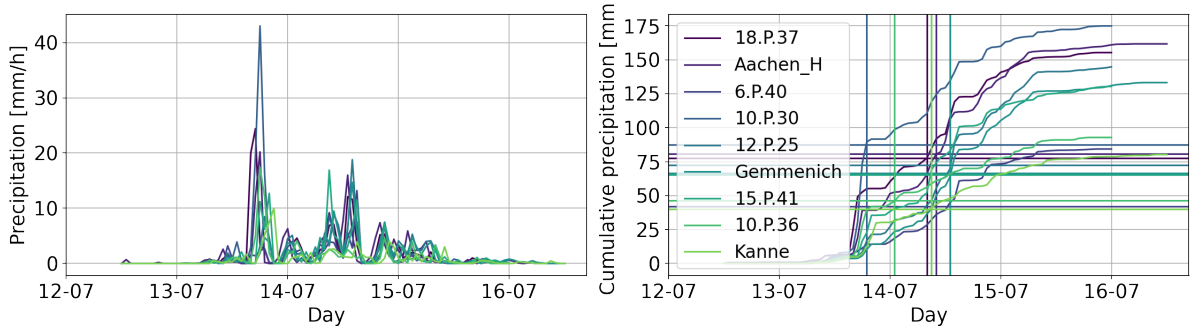


Figure 3.4: Left: Hourly observed precipitation from July 12-16 2021 for gauges in and surrounding the Geul catchment. Right: cumulative plot of the same data.

The temporal precipitation pattern of the Geul of fig. 3.4 consists of multiple peaks which are rather high. The main precipitation peak occurs in the afternoon of July 13, followed by low intensity rainfall until a new, but lower peak occurs in the afternoon of July 14. The duration and end of the event is comparable to the Vesdre, as well as the timing of most centroids. However, the magnitude of the cumulative precipitation is less extreme than for the Vesdre, but it is still higher than the Ahr in more than half of the stations.

Overall, the precipitation event was more significant in its total precipitation amounts than the hourly intensities. Although there were some noticeable hourly peaks, such as a maximum intensity of 43 $[mm/h]$ in the Geul, the total precipitation amounts are more outstanding. Fig. 3.5 gives an idea of how large the precipitation amounts of the event were by comparing them to the average yearly precipitation amount. The percentages range between 15 and 25%, with the station of Jalhay again as an outlier with 27%. There is little difference between the catchments, except for the Geul which has noticeably more internal variation than the Vesdre and Ahr. With such a large percentage of the yearly precipitation falling in 3 to 4 days, the total precipitation amount of the event can be seen as a significant characteristic of the flood event.

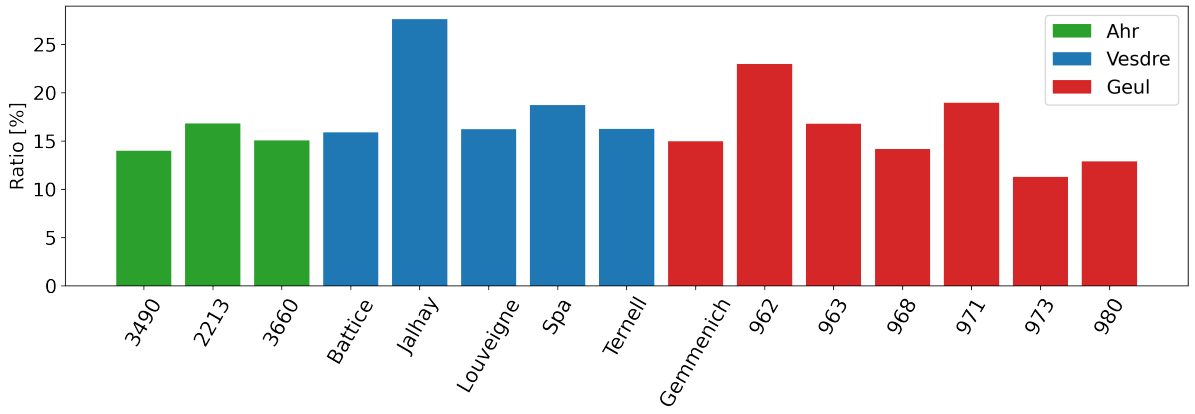


Figure 3.5: Percentage of the precipitation sum of the flood event compared to the average annual total precipitation amount for each relevant station.

Spatial analysis

The mass curves of the different catchments have shown that the precipitation amounts varied considerably within the catchment and between the catchments. Radar data has been used for the spatial analysis of the precipitation since radar data has a higher spatial resolution than rain gauge networks (Scofield & Kuligowski, 2003).

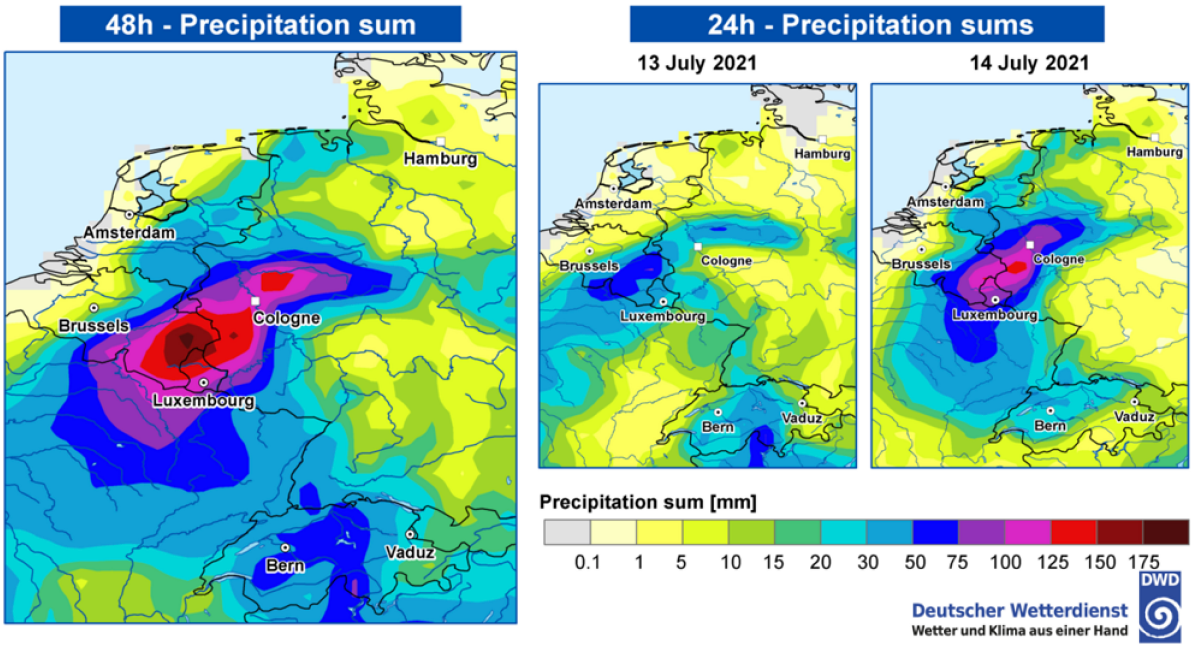


Figure 3.6: Left: 48 hour precipitation sum (13 July 00:00 UTC - 15 July 2021 00:00 UTC or 12 July 22:00 h - 14 July 22:00 h in local time). Right: 24 hour precipitation sum (13 July and 14 July 2021). DWD (Deutscher Wetterdienst: German Meteorological Service) extended the E-OBS dataset by numerous station measurements. (Kreienkamp et al., 2021).

Fig. 3.6 visualises the large-scale spatial pattern for the 48 hour sum of July 13 and 14, as well as for the 24 hour sum of each day. The precipitation amount of the 48 hour period has a centre around the Vesdre, and extends towards the Ahr and the Belgian part of the Geul. The location of all catchments near the region of the highest precipitation confirms their choice as critical catchments for the case studies. The rain mainly fell in the south of Belgium on July 13, affecting the Vesdre and the Geul. The majority of the total precipitation amount fell on July 14 with the highest amount close to the Ahr. This is in line with the hyetographs which showed that especially the Ahr had a high precipitation peak on July 14.

Fig. C.1 in appendix C.1 provides the daily and total precipitation amount of each catchment in the units of precipitation height $[mm/d]$ and volume $[m^3/d]$. The values are derived from the radar data discussed in section 5.4 and confirm the spatial pattern of fig. 3.6. The Vesdre received the highest volume of rain during the entire event, 127 $[10^6\, m^3]$, followed by the Ahr with 90 $[10^6\, m^3]$ and the lowest rain volume fell in the Geul, 47 $[10^6\, m^3]$. The average precipitation amount, expressed in $[mm/d]$, does not account for the area and results in a higher precipitation amount for the Geul than the Ahr. Even though all catchments were closely located near the centre of the precipitation event, the Vesdre received considerably more rain than the Ahr and the smallest volume of rain fell in the Geul which is partly due to its smaller catchment area.

Soil moisture content

The soil moisture content is part of the forcing since it is a dynamic state that is strongly and directly influenced by the precipitation. Increasing amounts of precipitation before and during the event slowly saturate the soil. A fully saturated soil can lead to saturation excess overland flow. Antecedent soil moisture content is considered one of the main drivers behind the runoff coefficient's spatial and temporal variability, according to research in Austrian catchments by Merz & Blöschl (2009). Since soil moisture content plays an important role in flood generation (Stein et al., 2020), a short description of the soil moisture content during the 2021 flood is provided here.

The three weeks before the event were exceptionally wet (Cornwall, 2021; Kreienkamp et al., 2021). Mohr et al. (2022) quantify this statement for the Ahr with the Antecedent Precipitation Index, API, confirming the wet period. API is discussed in more detail in section 4.1. The same study points to overland flow from infiltration and saturation excess as a driver of the rapid rainfall-runoff transformation due to these wet soils and the steep slopes. Junghänel et al. (2021) use a map of DWD indicating the free soil water storage between 0 and 60 $[cm]$ soil depth, as can be seen in fig. 3.7. The Ahr has an average

available water storage of 30 to 75 [mm], which is indicates a low storage availability, but not the lowest of the region. Some locations in the south and west of the catchment indicate an available water storage of less than 30 [mm]. Roggenkamp & Herget (2022) state that the infiltration capacity of the soil was exceeded for a large area of the Ahr. Considering the amount of available storage and the amount of precipitation, this is a reasonable hypothesis. Although there is little quantified proof, especially for the Vesdre and Geul, there is consensus in the current research that the soil was under wet conditions before the 2021 flood event.

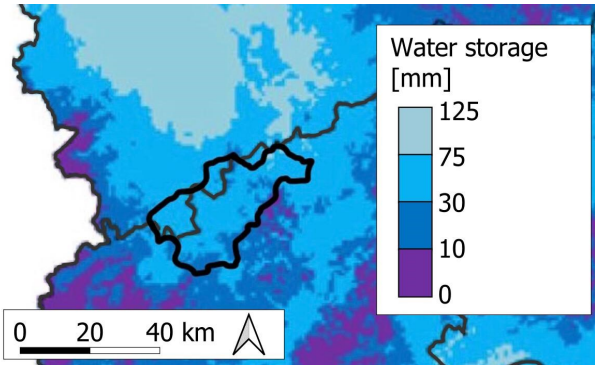


Figure 3.7: Free water storage in soil under grass up to a depth of 60 [cm]. Figure from Junghänel et al. (2021), based on DWD data.

3.2 Response

The most important aspect of this extremity analysis is the flow of the 2021 flood event. Whereas the precipitation was previously considered as the forcing of the event, this section analyses the response to the forcing by quantifying the flow and identifying key properties. This analysis not only offers valuable insight in the river's response but also provides the reference baseline for the extremity analysis and performance of the models.

The response of a flood event can be quantified by several parameters, such as the discharge, water level and the flood area. Here, the discharge is used as the quantified parameter because it allows a comparison in space and time. The water level and flood area both depend on the flow area which varies at each location. This variability makes it difficult to compare their values for different locations. The water level is the only parameter that can be measured directly leading to less uncertainty. The flood area can be derived from pictures, satellite imagery or even flood marks, but the large scale makes a precise estimate difficult. Discharges are mostly derived from water levels and stage-discharge relationships which can lead to increased uncertainty (Coxon et al., 2015), especially for high flows during flood events (Di Baldassarre & Claps, 2011). Kreienkamp et al. (2021) appoint discharge as the most suitable flow parameter for event description and the comparison to long term measurements. Despite the uncertainty potential of the discharge, the possibility of spatial and temporal comparison is considered to be of higher importance, hence the selection of discharge as the metric of the response.

Firstly, this section analyses the response in time and space for the different catchments. With the known precipitation and discharge, the runoff coefficient can be computed indicating how much of the precipitation became runoff after a certain period. This gives a first insight in the mechanisms during the event, as well as the differences between the catchments and their flow behaviours.

Discharge observations and estimates

The most common way to present flow data is a hydrograph, which shows the time series of a flow parameter. Here, only the hydrographs of the discharge are discussed. The shape of the hydrograph, the peak discharge and the timing are important properties that can be read directly from this graph.

Discharge observations for the Ahr are publicly available at the Wasserportal of the Ministerium für Klimaschutz, Umwelt, Energie und Mobilität (Ministry of Climate Protection, Environment, Energy and Mobility) of Rhineland-Palatinate. Daily mean discharge observations are verified, but not for the three most recent years. The exact start and end time of the measurements of each station can be

found in table B.1. Hourly discharge measurements of the flood event have been supplied by the LFU (Landesamt für Umwelt: State Agency for Environment) in Rhineland-Palatinate. This authority is responsible for all measurements of the river system and have filled the data gaps for the flood event with preliminary estimates based on their local knowledge.

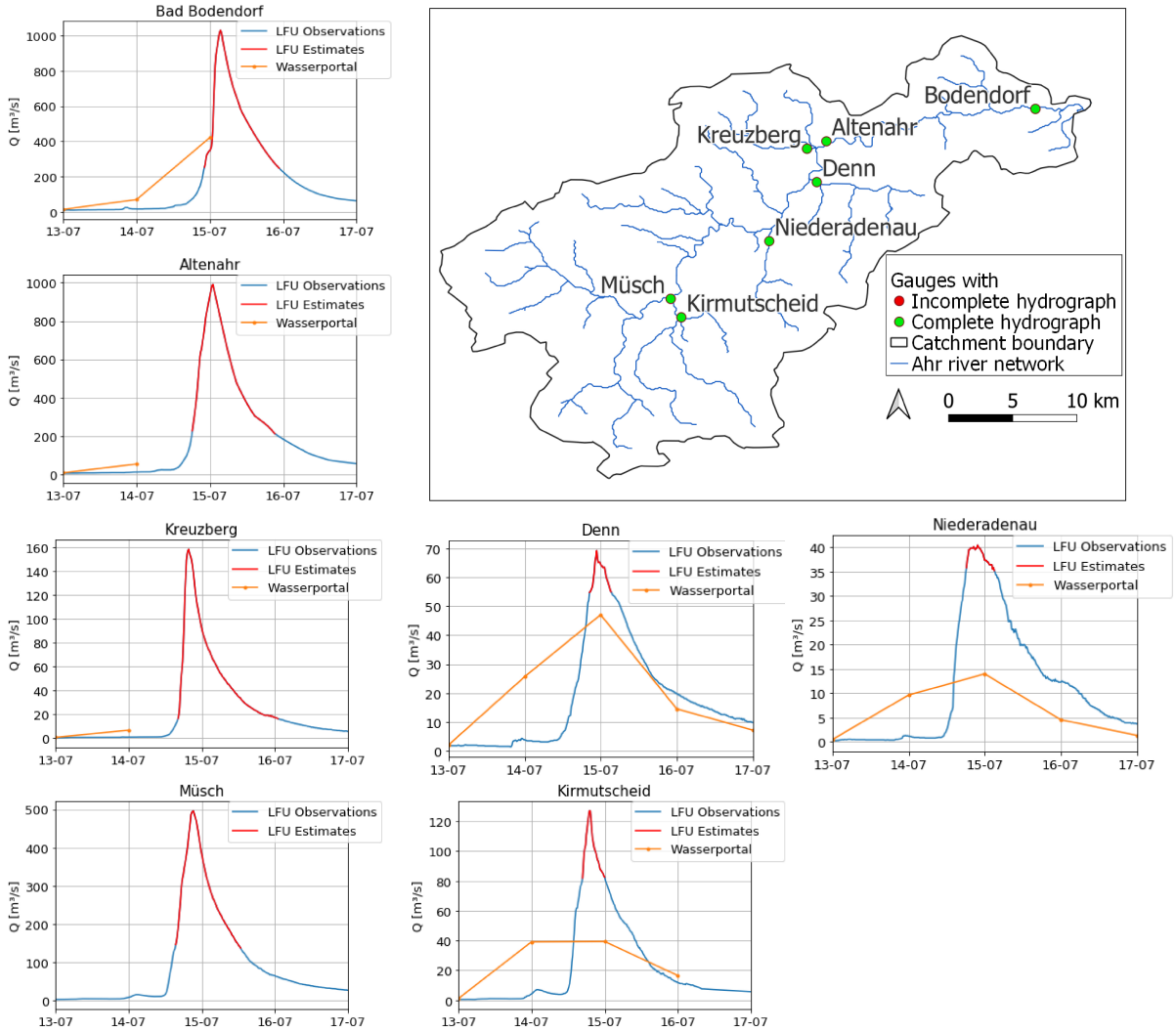


Figure 3.8: Hydrographs for the Ahr on July 13-17 2021, from public records (Wasserportal) and preliminary datasets (LFU). The estimated sections of the LFU data are shown in red to highlight their uncertainty. The map shows the locations of the gauges. As (estimated) hydrographs are complete for the event, their locations are indicated by a green marker.

The hydrographs of fig. 3.8 are complete for all gauges due to the LFU estimates. This extensive insight shows very high peak discharges for the gauges of the main river (from upstream to downstream: Müsch, Altenahr, Bodendorf). The main gauge at Altenahr has a peak discharge of 991 $[m^3]$. The peak discharges of the tributaries are lower due to the smaller drainage area, but are still high. The gauges in the tributaries have also recorded a larger period of the flood event, probably due to the lower and thus less destructive flows. The hydrographs are not only remarkable for their peak discharge, also the shape of the hydrographs is significant due to their sharp peaks, mostly for the gauges of the main river.

The hydrographs of fig. 3.8 are only estimates, even though they are made by the responsible authority with expertise. To assess their uncertainty, a comparison is made to peak discharge estimates of several reports on the flood event. Roggenkamp & Herget (2022) have estimated peak discharges in various locations throughout the catchments based on flood marks and the empirical Manning equation. More information on the methodology and estimate per location can be found in appendix E.1. Their estimates vary considerably from the LFU estimates for the smaller tributaries (relative error of 58%), but their estimates are a close approximation for the gauges of the main river (relative error of 11%). Their

estimated peak discharge just downstream of Altenahr was 1120 $[m^3/s]$. Kreienkamp et al. (2021) have collected multiple peak discharge estimates of which many are still unpublished. The overall estimate was around 1200[±200] $[m^3/s]$ around Altenahr. The methods varied between flood mark derivations and uncalibrated models. Another method mentioned in the study was to use discharge measurements of the Rhine upstream and downstream of the outlet and their timeshift. This resulted in a peak discharge of 1130 $[m^3/s]$ for the outlet. As the mean estimate of 1000 $[m^3/s]$ in the report of Kreienkamp et al. (2021) is in line with the peak discharge of 991 $[m^3/s]$ at Altenahr station according to the LFU time series and considering the expertise and inside knowledge of the LFU, their detailed data sets are considered for further analysis.

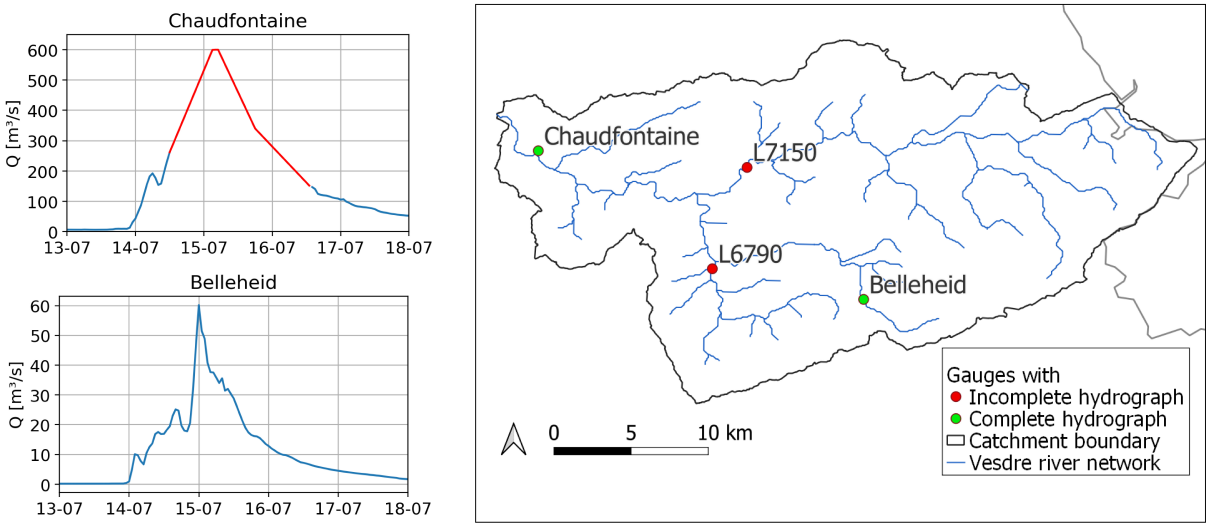


Figure 3.9: Hydrographs for the Vesdre on July 13-18 2021, from Aqualim. The estimated sections are shown in red to highlight their uncertainty. The map shows the locations of the gauges. Gauges with a complete hydrograph are indicated by a green marker. If no observation nor estimate is available, the gauge is indicated by a red marker.

Fig. 3.9 shows the hydrographs for the Vesdre catchments for which the data has been retrieved from Aqualim, a service of SPW Direction des Cours d'Eau non navigables (Department of Non-navigable Waterways of the Wallonian government). Only the station of Belleheid captured the entire event. The hydrograph of Chaudfontaine has been completed manually due to its importance as the main gauge. It has been based on the following principles. Ziemetz et al. (2021) have performed a large study on the flood event in the Vesdre. For the return periods, they used a peak discharge of 600 $[m^3/s]$ which lies within the range from other estimates (TFFF, 2021)(Archambeau et al., 2021). According to witnesses, the peak occurred during the early night of July 15 (Ziemetz et al., 2021). The report also mentions a potential double peak due to a large water release from the dam at Eupen, but due to uncertainty this is not taken into account for deriving the hydrograph shape here. The hydrograph shape is then estimated as a linear interpolation between the observation and the peak discharge and its moment of occurrence. The recession of the hydrograph includes an inflexion point to be more similar to standard recession curves.

Since there are only two complete hydrographs available for the Vesdre, its analysis is less detailed than for the Ahr. The peak discharge of 600 $[m^3/s]$ for Chaudfontaine is high. Since the gauge of Belleheid only has a small drainage area, its peak discharge is much lower. This lower discharge is probably the reason why the event has been fully recorded, which is a similar principle as for the tributary gauges of the Ahr. The shape of the hydrograph in Chaudfontaine in the main river has a narrow shape to reach that high peak discharge of 600 $[m^3/s]$, while the hydrograph of Belleheid has a more irregular shape resulting in a peak discharge of only 60 $[m^3/s]$.

Peak discharge estimates have been made for multiple locations in the Vesdre. These estimates have been made using multiple methods and have been collected by Ziemetz et al. (2021). The results can be found in fig. E.2 and table E.3 in appendix E.3, along with their methods and more details on the dam of Eupen, which may have caused the second peak. The estimated peak discharge for Belleheid,

as an average of two different methods, is around 115 $[m^3/s]$, which is almost twice the measured peak discharge. This difference in peak discharge and the unusually sharp hydrograph rise decrease the reliability of the Belleheid station. The peak discharge of Chaudfontaine has been derived from flood marks and is estimated to be around 650[±30%] $[m^3/s]$. Archambeau et al. (2021) have made an estimate of 580-660 $[m^3/s]$. These estimates create a considerable uncertainty envelope around the peak discharge of 600 $[m^3/s]$ for Chaudfontaine. Due to the lack of alternatives, the hydrographs of both stations will be used in the rest of this research. However, the uncertainty of the hydrographs, especially for the main gauge with its already uncertain peak discharge, is quite high and thus the resulting findings should also be considered as a first rough estimate and not precise results.

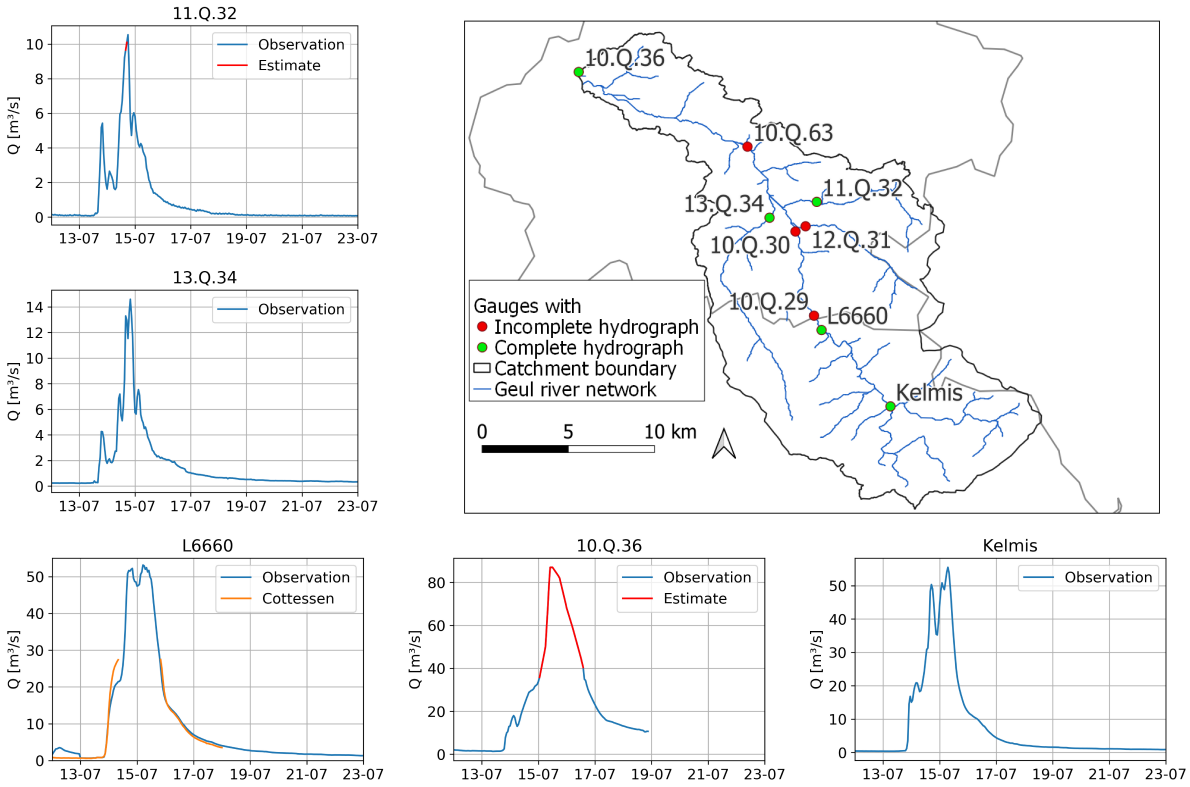


Figure 3.10: Hydrographs for the Geul on July 13-23 2021, from WL and Aqualim. The estimated sections are shown in red to highlight their uncertainty. The map shows the locations of the gauges. Gauges with a complete hydrograph are indicated by a green marker. If no observation nor estimate is available, the gauge is indicated by a red marker.

The discharge data sets for the Dutch area of the Geul have been provided by WL (Waterschap Limburg: Water Board of Limburg). The data sets for the Belgian region (stations L6660 and Kelmis) have been retrieved from Aqualim. Only four gauges have captured a (nearly) complete hydrograph, as shown in fig. 3.10. The 2 hour gap of station 11.Q.32 has been filled with a linear interpolation. Van der Veen (2021) provides a valuable fifth hydrograph for the main gauge, 10.Q.36 at Meerssen, based on the peak discharge estimate by WL of 85 to 90 $[m^3/s]$, using the average of 88 $[m^3/s]$, which occurred on July 15 at 10:00 h. The plot of station L6660 at Sippenaeken contains a second hydrograph, from station 10.Q.29 at Cottessen which is situated 1 km downstream. This downstream gauge is a flume with a maximum measurement capacity of 25 $[m^3/s]$ which did not record the peak of the flood event as a result. Since the station of Sippenaeken is situated just upstream of Cottessen with no tributaries between both gauges and is considered as very stable (Direction des Cours d'eau non navigables, 2022), the hydrograph of station L6660 is used in the analysis. This results in four measured hydrographs and one hydrograph estimated with expert knowledge for the Geul analysis.

The five hydrographs of fig. 3.10 show a variety of hydrograph shapes. The upstream gauges of L6660 at Sippenaeken and Kelmis have a double peak. It is also noticeable how the peak discharge of 53.1 $[m^3/s]$ at L6660 is lower than the peak discharge of 55.5 $[m^3/s]$ at Kelmis, although the drainage area of station L6660 is around 1.6 times larger. This difference could be due to flooding between both

gauges, flood wave attenuation, a measurement or conversion error at Kelmis or Sippenaeken or due to an unidentified weir in between. A detailed analysis of this difference is not within the scope of this study. The shape of the hydrographs of the tributaries, 11.Q.32 and 13.Q.34, does not have this double peak but instead, they show a very sharp, singular peak, albeit with a lower peak discharge. Although the drainage area of Kelmis is only 55% larger than of 13.Q.34, the peak discharge is 280% larger showing a spatial heterogeneity in forcing, flow mechanisms or flood wave attenuation. The main gauge, 10.Q.36, has a wider shape of the hydrograph than the tributaries and lacks the double peak of the upstream gauges. With an estimated peak discharge of 88 $[m^3/s]$ it has the highest peak discharge of all gauges. The variety in peak discharges and hydrograph shapes throughout the Geul catchment is a significant characteristic in comparison to the other catchments.

Peak discharge estimates have been made for more locations than those in fig. 3.10. Asselman et al. (2022) have simulated discharges and flood areas throughout the Geul catchment, with a focus on Valkenburg, which is located 3 km downstream of station 10.Q.63. The estimated peak discharge of 134 $[m^3/s]$ is higher than the early expert judgement estimate of WL of 100-120 $[m^3/s]$. Furthermore, the peak discharge of the simulation occurs on July 14 around 16:00 h, while the actual gauge only failed at 23:25 h and TFFF (2021) estimates that the peak discharge occurred on July 15 at 03:00 h. That same model also calculated a peak discharge of 100 $[m^3/s]$ at Cottessen, which is almost double the observation of L6660 at Sippenaeken. The simulated peak discharge of 102 $[m^3/s]$ for 10.Q.36 at Meerssen is the closest estimate of all simulated locations. TFFF (2021) has estimated a peak discharge of 100 $[m^3/s]$ with an unknown method at an unknown location, most likely at Meerssen. Although all sources use uncertain or even unknown methods of estimation, the peak discharge of Meerssen of 88 $[m^3/s]$ could be an underestimation to a range of 30%.

Although only two gauges are located in Belgium, they are valuable due to their complete hydrographs. Several reports have missed this information by focusing on the Dutch part of the Geul (TFFF, 2021; Asselman et al., 2022). This highlights the importance of using transboundary catchment information.

With the exception of the Ahr, most hydrographs for the main river are incomplete whereas the lower flows at tributaries are often recorded. An exception is the most downstream gauge which is a focus of estimation studies due to its relevance for the entire catchment. The difference between observations and estimates highlights the uncertainty of these hydrographs. As mentioned before, the discharge time series is the only flood metric that allows for temporal and spatial comparison. Therefore, it will serve as the base for the results of this research despite its uncertainty. Quantifying the uncertainty and its propagating through the results is beyond the scope of this research.

Another consequence of the uncertainty of these hydrographs is the limited possibility to derive properties from these hydrographs. The peak discharge and timing of the peak can be derived by looking at the location of the maximum discharge. In case of the estimated hydrographs, they form the basis of the shape approximation. The studies mentioned above also focused on these two key properties. Other properties such as the time base and time to recession (Bhunya et al., 2011) are more difficult to confidently derive from these estimated and uncertain hydrographs and are thus not included in this study.

Fig. 3.11 combines the hydrographs of the main gauge of each catchment to compare the magnitude, timing and shape of the hydrograph. The timing of the flood peak is similar, although there is a small delay of the Geul. The hydrograph shape has more contrast between the catchments. The hydrograph of the main gauge of the Ahr has a much more narrow shape than the Vesdre and the Geul. Both the rise and recession are steeper for the Ahr, which points towards a more flashy response. The contrast in shape can also be seen in the peak discharge which is much higher for the Ahr (around 990 $[m^3/s]$), followed by the Vesdre (600 $[m^3/s]$ and the Geul (around 90 $[m^3/s]$). The peak discharge for the Ahr was more than a factor 10 larger than for the Geul. The peak discharge and the hydrograph shape differ significantly for the main gauges of all three catchments.

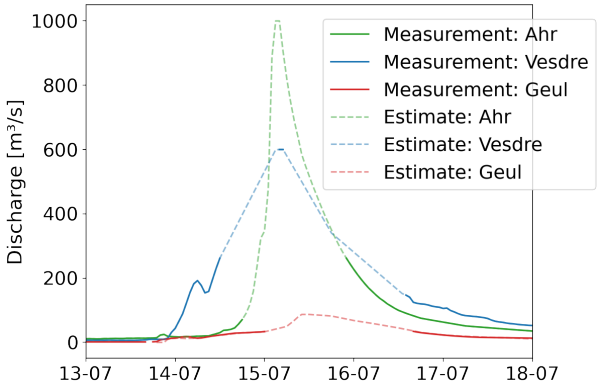


Figure 3.11: Hydrographs of the main of the Ahr (Altenahr), Vesdre (Chaudfontaine) and Geul (10.Q.36, Meerssen) for the 2021 flood event.

Runoff coefficient

The runoff coefficient is the ratio of the cumulative precipitation and discharge volume. In other words, it is the amount of water that has flowed through a certain point relative to the amount of water fallen in its drainage area. A low runoff coefficient means little precipitation has become runoff which could point to high infiltration and storage rates. Although the runoff coefficient is often given as one value, it actually varies in time and is thus plotted as time-dependent in fig. 3.12. Due to the limited data range, the runoff coefficients are only calculated until July 19. The runoff coefficients vary considerably. With a runoff coefficient of 0.3 after 6 days for the Geul, only 30% of the precipitation has become runoff indicating a large storage capacity for the Geul. The runoff coefficients of the Vesdre and Ahr are considerably higher. The Ahr reaches a 90% runoff ratio after less than 5 days.

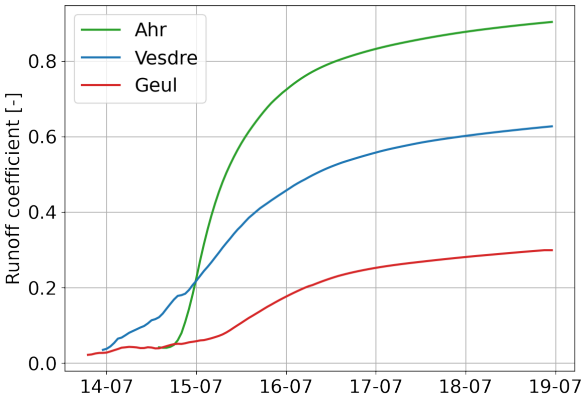


Figure 3.12: Runoff coefficients of the Ahr (Altenahr), Vesdre (Chaudfontaine) and Geul (10.Q.36) for the 2021 flood event.

The steepness of the curves shows that the transformation from precipitation to runoff and actual flow through the river system is much faster for the Ahr than the Vesdre and Geul. Especially the flat slope of the Geul indicates a slow runoff process. Both the maximum value of the runoff coefficient and the steepness of its curve fit the catchment characteristics and the flow characteristics of the flow duration curves. The steeper slopes and rocky soils of the Ahr and Vesdre decrease the storage capacity and could have led to the quick-runoff process, while the less steep slopes and more permeable soil likely led to a damped response for the Geul.

3.3 Flood typology

Differentiating flood events into types can offer valuable insights for both the practical and scientific fields of flood risk management. Common flood typology is based on causative processes and differentiates in coastal, fluvial and pluvial floods. The 2021 flood event was a fluvial flood event as it originated from rivers, with locally small pluvial floods. A more detailed typology can teach us more. Merz & Blöschl (2003) define five types of fluvial floods: long-rain floods, short-rain floods, flash floods, rain-on-snow floods and snow melt floods. This classification is based on a catchment perspective which focuses on the catchment state and the atmospheric inputs, instead of causative processes or general atmospheric circulation patterns.

Classifying flood types on a more detailed level is important for several reasons. Typology can be considered as a similarity analysis tool based on spatial – statistical relationships (Syvitski, 2003). A flood typology based on process understanding can help to detect flood events with similar processes. As a result, the contribution of the different mechanisms of flow response for each flood type can be identified (Viglione et al., 2010). The main flood type of a catchment can be seen as a fingerprint of that catchment since the flood type incorporates both precipitation characteristics (forcing), runoff generation (mechanisms) and runoff properties (response) (Gaál et al., 2012). Coupling such catchment characteristics to a certain flood type can help to make estimated guesses on a catchment when its type is known, or on the flow characteristics when the data is scarce but the catchment characteristics are known (Blöschl et al., 2007). The characteristics of a flood type can also be used as a performance indicator of new framework (Viglione et al., 2010). From these examples, it shows that flood typology is useful in the scientific field to improve the understanding of the catchment and hydrological characteristics.

Typology can also be used to detect trends, such as seasonality (Gaume et al., 2009), or it can be applied in frequency analysis. Flood frequency analysis is currently done by fitting a statistical distribution to a sample of events (Merz & Blöschl, 2003), but when this sample of events is caused by different processes, this may lead to uncertain predictions (Potter, 1958). Nied et al. (2014) highlight the importance to shift from a pure flood frequency analysis to a frequency analysis based on process understanding.

Including this process knowledge by implementing typology classification, would lead to an improved flood probability quantification.

Quantifying the flood probability for different flood types could help with setting up effective forecasting and monitoring tools for precipitation, as well as flood forecasting models (Borga et al., 2008). Flood hazard prediction models also benefit from understanding the different event characteristics for flood generation (Norbiato et al., 2008; Eeckman et al., 2021). Flood typology can also play a role in engineering practice, by using them for design hydrographs (Brunner et al., 2017) or by linking the flood type to flood duration and thus discharge volume which is important for many engineering purposes (Gaál et al., 2012). Determining detailed flood types is thus not only useful for the hydrological understanding of catchments and frequency analysis, but also for practical purposes.

A flash flood is a flood type that is mentioned specifically in relation to the 2021 flood event (Ziemetz et al., 2021; Kreienkamp et al., 2021) and confirming if the 2021 flood event was a flash flood is an important purpose of this section. The danger of flash floods is considerable. In Europe, almost 50% of the severe floods (differentiating in riverine floods, flash floods and storm surge floods) between 1950-2005 have been identified as flash floods (Barredo, 2007). Severe floods were defined as floods with exceptional damage or at least 10 fatalities. Flash floods are often characterized by quickly rising and unexpected water levels in combination with high flow velocities (FEMA, 1981; Castro, 1996), creating dangerous situations (Jonkman et al., 2008). Flash floods are not just relevant on a societal level. Scientifically, the response of a catchment during flash flood events exposes the rate-limiting processes of the response to extreme floods and how they depend on catchment properties and flood severity (Marchi et al., 2010). The relation to the 2021 flood event, societal and scientific relevance create a special interest for flash floods.

Process Type	Long-Rain Floods	Short-Rain Floods	Flash Floods	Rain-on-Snow Floods	Snowmelt Floods
Timing of floods	no pronounced seasonality	no pronounced seasonality	floods and extreme rainfall mainly in summer or late summer	mainly occur at the change between cold and warm periods	floods in spring to summer
Storm duration	long duration (>1-day)	duration of several hours to 1-day	short duration (<90 min), high intensities	moderate rainfall events can cause large floods	rainfall unimportant
Rainfall depths, snowmelt	substantial rainfall depths	moderate to substantial rainfall	small to moderate rainfall depths	snowmelt and rainfall	snowmelt, no or minor rainfall
Catchment state (SWE, soil moisture)	wet due to persistent rainfall	wet for large flood events	dry or wet	wet, snow covered	wet, snow covered
Runoff response dynamics	slow response	fast response	flashy response	fast or slow response	medium or slow response
Spatial coherence	large spatial extent of storms and floods (>10^4 km^2)	local or regional extent	limited spatial extent of storms and floods (<30 km^2)	limited to areas of snow cover	medium spatial extent of floods

Figure 3.13: Indicators for identifying flood process types at the regional scale according to Merz & Blöschl (2003).

The indicators for each flood type according to Merz & Blöschl (2003) can be found in fig. 3.13. As this research concerns a flood event during summer, in which snow played no role, the flood types of rain-on-snow and snowmelt floods are not discussed here. The distinction in the different flood types is based on the timing, storm duration, rainfall depth, catchment state, runoff response and spatial scale.

According to Merz & Blöschl (2003), flash floods mostly occur during summer, the precipitation is only of a short duration (<90 min.) with high intensities and small to moderate rainfall depths which result in a flashy response. The spatial extent of the storm and flood is small (<30 [km^2]). Other studies use a wider definition of flash floods. In general, the main characteristic of flash floods is a quick runoff process leading to sudden floods (usually within several hours after the precipitation event) (Marchi et al., 2010). Georgakakos (1986) indicates that there is no sharp distinction between floods and flash floods. As such, response times under 6 hours are a main indicator, but even longer events of up to 16 hours can be denoted as flash floods (Marchi et al., 2010). The catchment size is limited to around 1000 (Braud et al., 2014) or 2000 [km^2] (Gaume et al., 2009). A catchment of a larger scale is less geomorphically effective, meaning that it allows for fast runoff processes, since the orography is less complex (Marchi et al., 2016). The orography influences the forcing by augmenting the precipitation and anchors the convection, as well as the response since the streamflow is quickly concentrated by the steep relief (Kobiyama & Goerl, 2007). Besides complex orography, a convective and local origin of the precipitation event is a characteristic of flash floods, as well as infiltration excess mechanisms triggering

the flood (Tarasova et al., 2019). The exact definition of a flash flood is not well defined, but its main description is a short and intense precipitation event in a small catchment followed by a quick response of suddenly high flows.

Following the indicators of fig. 3.13, several flood types could apply to the 2021 flood event. The event can be considered as a long-rain flood as the storm duration was longer than 1 day. Secondly, the spatial extent was large for both the storm and the flood if the entire region in which floods occurred is considered. A third indicator that fits with a long-rain flood type are the substantial rainfall depth, since multiple locations received over 150-200 mm in 48 hours. The runoff response dynamic is defined as the ratio between the peak discharge and the average discharge on the day that the peak discharge occurred. When this ratio approaches 1, the response is slow. A higher value indicates a fast response. Although the hydrograph shapes are only estimates, fig. 3.11 shows that the peak discharge is higher than the average discharge and so, the responses of the Ahr and Vesdre are not slow. Only the Geul might satisfy this condition of slow response as a long-rain event. Due to the lack of quantification of the indicators by Merz & Blöschl (2003), the 2021 flood event could also qualify as a short-rain flood event. Although the duration of the storm was slightly longer than the maximum of 1 day, the moderate to substantial rainfall, fast response and regional extent all fit the requirements of a short-rain flood. When applying the definition of Merz & Blöschl (2003) for flash floods, only one characteristic of the 2021 flood event, the timing in summer, would fit the classification of a flash flood. The duration was certainly longer than the 90 minutes of fig. 3.13, the rainfall depths were not small or moderate, the response is not exactly flashy based on the runoff response dynamic and the spatial extent is much larger than 30 $[km^2]$. According to the conditions of fig. 3.13, the 2021 flood event would fall between the characteristics of a long-rain and short-rain flood. However, when the wider definition of flash floods is used, another image arises. Reports of extremely sudden water level rises on July 14 (Schmid-Johanssen et al., 2021; Noulet, 2021) suggest flash flood behaviour (Georgakakos, 1986). All catchments are smaller than 1000 $[km^2]$, which was one of the requirements. The complex orography which could lead to the typical quickly concentrated runoff is present. The precipitation was from a convective origin (Kreienkamp et al., 2021). One of the most dominant characteristics of flash floods is the fast response, which can be quantified by the lag time. The lag time is defined as the time difference between the centroid of the precipitation volume and the peak discharge. From fig. 3.14, the time difference for the most downstream gauge of the Ahr, Vesdre and Geul are respectively 13, 17 and 24 hours. The Ahr and possibly the Vesdre could be considered as flash floods if the lag time is defined less strictly, as was done by Marchi et al. (2010). The Geul has a lag time that extends well past the rough limit of 16 hours of (Marchi et al., 2010). The 2021 flood events fits several general characteristics of a flash flood, but the lag time, which is one of the main flash flood characteristics, is quite long for the 2021 flood event to qualify as a flash flood.

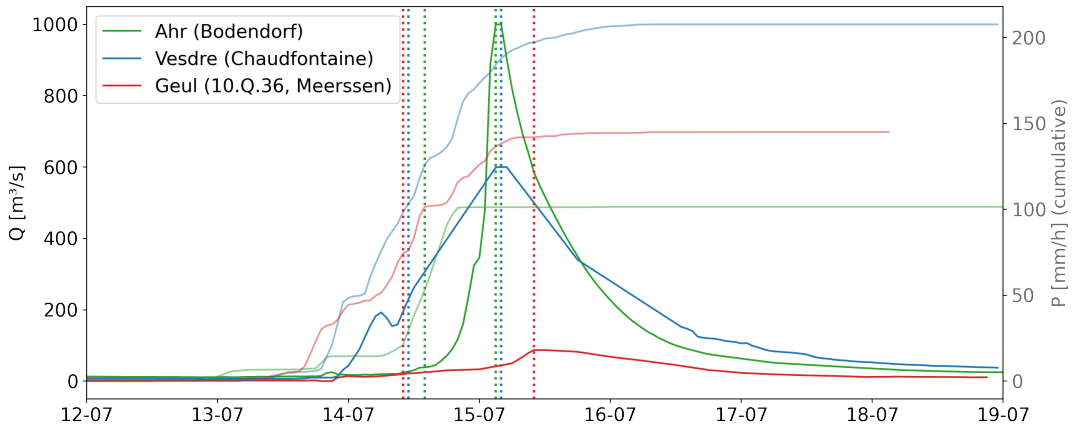


Figure 3.14: The lag times of the three catchments can be derived from the vertical, dashed lines which show the centroid of the precipitation event and the moment of peak discharge. The hydrograph (dark) and cumulative hyetograph (light) are shown as well.

According to the classification of Merz & Blöschl (2003), the 2021 flood event qualifies as a short-rain event, or a long-rain event with a quick response. If a more general definition of flash floods is applied, the 2021 flood event could also be considered as a flash flood, albeit of with a slightly slower response than usual. In conclusion, there is no flood type that clearly fits this event, which hinders the derivation of certain characteristics. Nevertheless, Viglione et al. (2010) state that temporal, spatial and movement components are all important in runoff generation and routing for short-rain flood events. The importance of spatial scale for flash floods is also highlighted by Kreienkamp et al. (2021). Whether it is a flash flood or not, such quickly responding floods are unusual, as no flash flood has been documented in this region by previous research (Marchi et al., 2010; Barredo, 2007). The lack of a generally accepted and quantified definition of flood types makes it difficult to classify the 2021 flood event, but the fast response is still a noteworthy characteristic.

Chapter 4

Extremity of the forcing and response

The event description of the previous chapter can be used to analyse which aspects of the 2021 flood event were extreme. This chapter focuses on the forcing and response of the 2021 flood event and how they compare to flow conditions before July 2021 and as such, answers the second sub-question of the research. For the forcing, the aspects of the origin, daily precipitation, temporal and spatial variation and pre-conditions are considered as potential extreme aspects. The peak discharge is the only aspect of the response that is analysed for its extremity. The extremity of each aspect is quantified, but due to the nature of some aspects this could not be done with one quantification method of extreme. Some aspects are analysed with the use of return periods, others with the 90th percentile and one aspect required the use of correlation coefficient. To consider a period that is longer than the measurement records, historical information is included to give more insights in the magnitude of the peak discharge and the seasonality of floods in the catchments.

4.1 Forcing

Origin

The high precipitation amounts of the 2021 flood event originated from a low pressure system that was stationary due to blockage of high pressure systems. Similar precipitation events that are associated with quasi-stationary low pressure systems, their fronts, or convective systems which are located on the western flank of persistent blocking systems are common in Europe during summertime (Mohr et al., 2022). The low pressure system "Bernd" is therefore not rare or extreme. However, aggravating factors had an influence on the origin of the 2021 flood event. The stationarity of the cold core low for several days was uncommon (TFFF, 2021). Attracting humid air of a large region is a characteristic of such cold core lows, but unusual sea surface temperature anomalies (of up to 8K) over the Baltic Sea contributed to high and rare values of total precipitable water (Mohr et al., 2022). Due to the stationarity, a lot of this humid air became precipitation for one region, leading to high local precipitation amounts. Additionally, the rising air leading to precipitation was re-enforced by several troughs within the system (KMI, 2021). A similar effect was caused by the low mountain ranges (Ardennes, Eifel) which forced orographic uplift and damming of the precipitation system which led to persistent heavy rain for several areas (Kreienkamp et al., 2021). Due to the complex atmospheric processes and little research on such precipitation events from stationary cold core lows in Europe, the extremity of the origin can not be quantified. Nevertheless, the main conditions of the origin of the rain event in combination with multiple adverse local circumstances indicate that the origin of this storm event was rare and can thus be considered as extreme.

Return periods

A widely used method of quantifying the extremity of a precipitation event is to compute the return period of the precipitation intensity, which is the precipitation amount that has been measured in a certain amount of time. A return period is the average time between events of the same magnitude. A rain event with a 100-year return period has a probability of occurrence of 1% each year.

One method of computing the return period is the Generalized Extreme Value distribution, also known as GEV, developed by Jenkinson (1955). It combines three extreme value distributions by adding a shape parameter to the Gumbel distribution to influence the shape of the tail of the distribution. This distribution has been widely used for hydro-meteorological extremes (Villarini et al., 2011).

It can be applied to radar data or point data from gauge measurements. The latter was chosen for this research since it provided longer record periods. Reliable estimates from the GEV distribution require long records (Papalexiou & Koutsoyiannis, 2013). Only stations with over 15 years of data have been used for calculating the return periods here. Such record lengths are still quite short, but due to the limited record length of all stations such a low limit is deemed sufficient. Although there are no strict rules in scientific literature, Ziemetz et al. (2021) state that return periods that are three times higher than the record period become unreliable.

Furthermore, the return periods have been calculated for the precipitation intensity of 1 day due to the limited data availability for long term records. Since July 14 was the day with the highest precipitation fraction of the entire event (80% for the Ahr, 64% for the Vesdre, and 50% for the Geul, see fig. C.1), the return periods of July 14 have been computed. For this calculation, the precipitation of the event has not been included in deriving the GEV distribution since the goal is to compare the 2021 flood event to the conditions preceding the event. The return periods of the precipitation of July 14 are not precise estimates due to the limited record lengths, but they give a first insight in how extreme the precipitation amount of the 2021 flood event was.

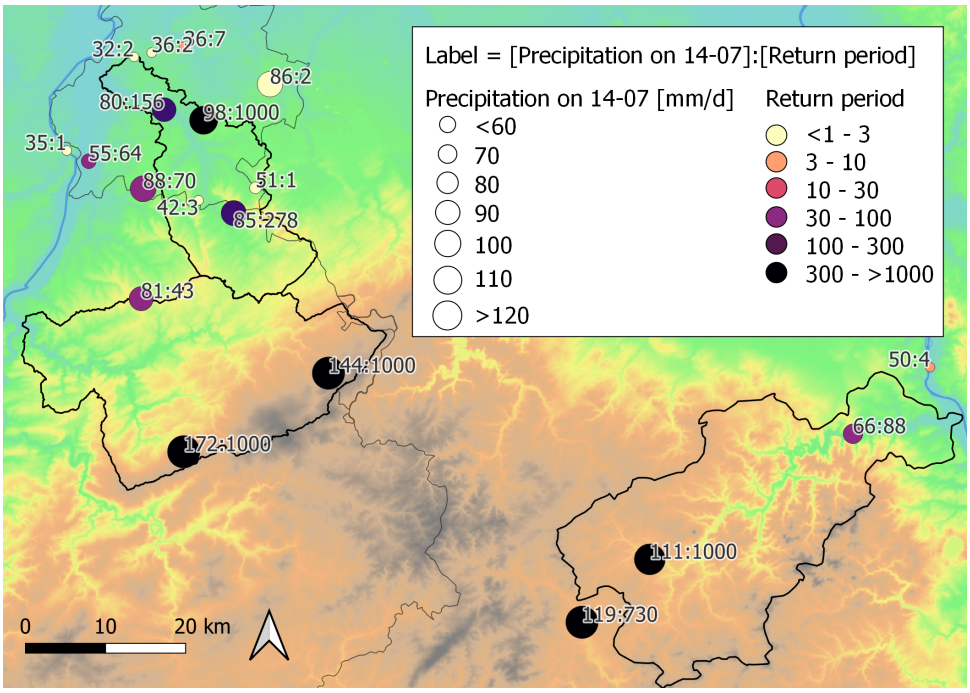


Figure 4.1: For each gauge with sufficient data the precipitation amount on July 14 is given, as well as its return period from the GEV distribution, based on the preceding records. The precipitation amount is indicated by the size of the marker and the return period is indicated by the color scale.

Fig. 4.1 shows how multiple stations have extreme return periods of more than 1000 years. At the same time, the precipitation amounts and thus return periods vary spatially. The estimated return periods vary between 1 and over 1000 years. The extremely low return periods have been recorded especially in and around the Geul catchment, but also the Vesdre and Ahr have some stations with low return periods. Extreme magnitudes of more than 1000 years appear in each catchment. With the rule of the thumb to only have accurate return period estimates up to three times the record length, the limited record periods of 15 to 90 years lead to estimates that are only accurate up to 45 or 270 years, depending on the station. Values of above 100 years should thus be considered as rough estimates. For return periods higher than 1000 years, no exact value is given due to the high uncertainty for such extreme return periods. The extreme magnitude of some return periods indicates that the precipitation amount on July 14 was extreme for multiple gauges.

In appendix D.2, the lower and upper bounds are given for the 95% confidence interval of the return periods. Except for some stations that show no sign of extremity for the entire confidence interval, almost all stations show an extreme return period of over 1000 years for the upper bound. Even for the lower bound, multiple stations have an extreme return period, although most estimates decrease by an order of magnitude compared to the mean estimate shown in fig. 4.1. The confidence interval confirms that some stations received precipitations amounts of extreme return periods, while some received unexceptional amounts.

Appendix D.2 also includes the return periods for 2-day and 3-day precipitation amounts. Most return periods increase when a longer event duration is considered. Only for the Ahr is there a slight decrease. It seems that the extremity of the precipitation event in the Ahr was mainly caused by the precipitation of July 14 and not the other days. The previous conclusion that the precipitation amount was extreme for many stations still holds when considering a longer precipitation event.

Other studies on the 2021 flood event have also made estimates for the return periods by using various methods. For the Ahr, the following studies give return period estimates based on radar data. Schäfer et al. (2021) gives an exceedance of 800 years for the 24-h precipitation in the Ahr, based on RADOLAN event data and HYRAS long-term data. TFFF (2021) estimates that the 48-h precipitation return period is 3.800 years, based on E-OBS data and a Gumbel distribution. This method has been applied to return periods for all catchment areas. The highest estimate is from Kreienkamp et al. (2021), with a value of at least 15.000 years, which is based on REGNIE data while using the GEV distribution excluding the event. For the Vesdre, estimates have been made from radar data and gauge measurements. TFFF (2021) gives a minimum return period of 10.000 years for the 48-h sum. Kreienkamp et al. (2021) have looked at the 48-h precipitation sum for the entire Belgian Meuse system which resulted in a 1000-year return period. This is in contrast with the results of the official weather authority, KMI (2021), who used gauge measurements and found low return periods of 30 years. The stations of Spa and Jalhay were an exception with 100-year return periods. Ziemetz et al. (2021) have done an in-depth analysis for the station of Jalhay. Using the same method as for fig. 4.1, the minimum return period was estimated to be around 1000 years for the 72-h sum. The return periods for 24-h and 48-h sum resulted in lower values. For the Geul, TFFF (2021) made an estimate for the return period of 1000 years. Van Heeringen et al. (2022) have used the same radar dataset as in this thesis, the KNMI re-analysis product. Comparing the 48-h sum for the entire Geul catchment to the average Dutch precipitation statistics resulted in a return period of 900 years. This estimate is acknowledged as a possible overestimation by a factor 2 since the region receives more rain than the rest of the Netherlands. Although all of the mentioned studies in this paragraph use varying methods and data sources which result in different estimates of the return periods, the high values of those estimates confirm the high return periods found in fig. 4.1 and the overall extremity of the precipitation amount.

Spatial variation

The return periods of the previous section indicate how extreme the precipitation was for 1 specific location (or 1 entire region for several studies). Spatial variations within the catchment are not part of the return period consideration. This section analyses if the spatial variation of the precipitation during the 2021 flood event was extreme by comparing the precipitation pattern of the event to the pattern of the average annual precipitation.

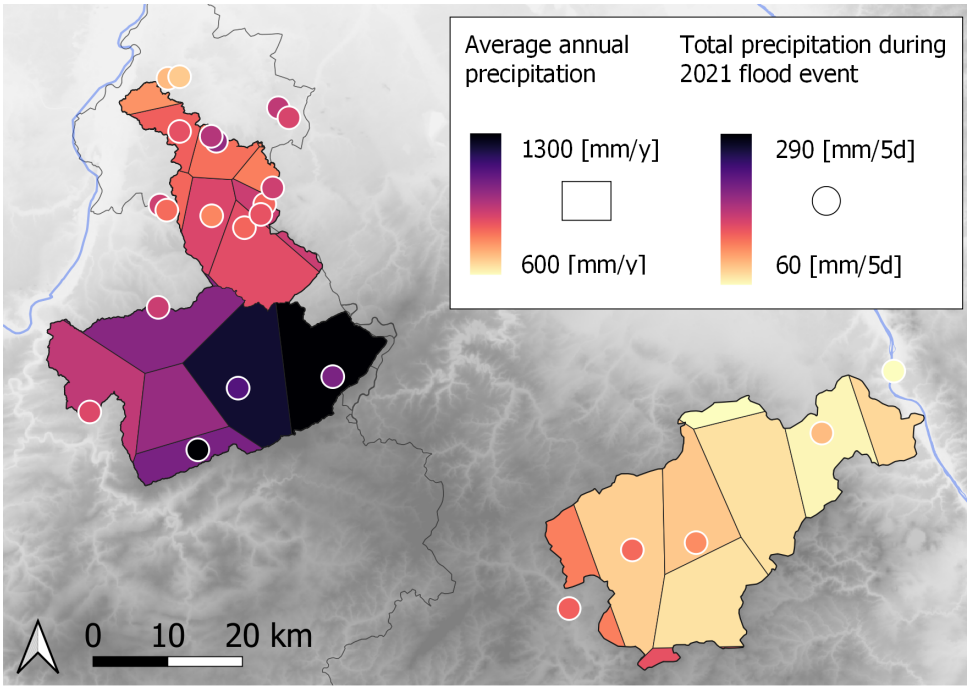


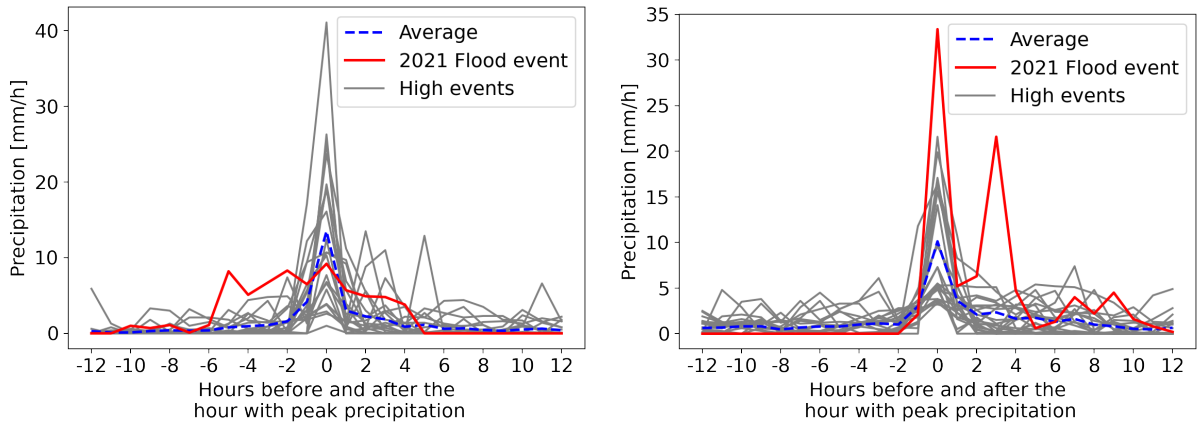
Figure 4.2: Comparison of the spatial variation of the long term precipitation pattern and the event precipitation. The polygons are taken from fig. 2.6 and indicate the average annual precipitation. The round markers represent the total precipitation sum of the entire 2021 flood event (5 days to cover the event for all gauges and limit influences of temporal variations)

Fig. 4.2 shows that the spatial pattern of the precipitation of the 2021 flood event is similar to the average annual precipitation pattern, which can be considered as the climatic pattern. The precipitation sums of the event are higher in the Vesdre, lower in the Geul and they are the lowest in the Ahr. Even within the catchments, the overall pattern seems to follow the climatic pattern. There are still several outliers, which is possible due to local orographic effects or faulty measurement devices. Some spatial variation from a precipitation event is always expected due to the cloud pattern, stationarity and orographic effects, but in general the event pattern seems to be similar to the climatic pattern. The similarity was expected because the patterns have mutual influencing forces such as the orography.

The similarity between both spatial patterns can be quantified with the Pearson correlation coefficient. Appendix C.3 gives more detail on this correlation coefficient and its calculation. With a value of 0.356, the correlation coefficient indicates a positive correlation of medium strength. As there is no negative or very poor correlation, the spatial variation of the precipitation during the 2021 flood event was similar to the climatic pattern and is therefore considered as not extreme.

Temporal variation

Precipitation events can have temporal variations, ranging from multiple days with low intensity to very short and high intense events. The extremity of this aspect can be analysed by comparing the temporal variation of the flood event with that of other high precipitation events. Since the temporal resolution is the more accurate for gauges than for radar data, gauge measurements are used for this analysis.



(a) Ahr (3490) *(b) Vesdre (Battice)*

Figure 4.3: Hyetographs for the labelled gauges, the hyetographs of five other gauges can be found in fig. C.2 in appendix C.2. The hyetograph is given in grey for the 30 precipitation events with the highest daily average. The hourly maximum of that day is located in the middle with 12 hours before and 12 hours after that hour of maximum precipitation. The average of the hyetographs of these high precipitation events is given in a dashed blue line and the hyetograph of the 2021 flood event is indicated in red.

Fig. 4.3 shows that there are many events with a higher hourly peak, thus a more intense precipitation pattern. There are also events with a lower and more constant intensity. Although the hyetograph of the flood event usually looks more flat than the average hyetograph, which is shaped more as a sharp peak with low precipitation in the surrounding hours, the temporal variation does not seem extreme compared to previous events.

The extremity of the temporal variation can be quantified by calculating the time during which 90% of the event precipitation is measured. The event is defined as 12 hours before and 12 hours after the peak precipitation, as shown in fig. 4.3. The duration for 90% of that event precipitation is calculated for the 100 days with the highest daily precipitation. If the 2021 flood event lies within the 90th percentile of all events the temporal variation can be considered as extreme. The number of events for this analysis is increased from 30 to 100 compared to fig. 4.3 to have a sufficiently large sample for detecting the 90th percentile. The analysis periods are limited to the record lengths of the hourly data sets and the analysis is only done for the gauges of fig. 4.3 and fig. C.2 in appendix C.2. In case more than 30% of the hours of one event contain no value due to data gaps, the event is not considered for the analysis. The limit of 30% was chosen as a benchmark after trial-and-error and it did not influence the results. Applying this method resulted in an extreme temporal variation for the Gemmenich station in the Geul, but none of the stations in the Ahr or the Vesdre. The method was also applied for a 48 hour period, so 24 hours before and 24 hours after the hour with the maximum precipitation. The results were the same as for the 24 hour period. In general, most of the stations did not exceed the 90th percentile and thus the temporal variation can be considered as not extreme.

Pre-conditions

Including hydrological pre-conditions can strengthen the link between weather events and flood occurrence, but they are commonly disregarded when forming that link (Nied et al., 2014). The same study has quantified the hydrological pre-conditions by modelling the soil moisture content. Since a specific model for soil moisture content would be outside the scope of this research and measurements of soil moisture content during the event and preceding period are sparse, this study uses the preceding rain amount as an indicator of the hydrological pre-conditions. Mohr et al. (2022) and Schröter et al. (2015) have used a similar method, respectively for the flood event of July 2021 and June 2013 in Germany, by computing the Antecedent Precipitation Indicator (API). Using the API as a parameter to approximate catchment wetness is a well-established method (Teng et al., 1993; Ahmed, 1995; Blanchard et al., 1981). Defined by Kohler & Linsley (1951), the API sums the precipitation amount on a number of days before the event with a weight for each day while including a decay constant. The recommended values for the decay constant vary (Viessman & Lewis, 1996) and depend on the catchment characteristics and season (Heggen, 2001). This study simplifies the API by excluding the weight and its decay factor and just adding up the precipitation amounts of the preceding period. This period consists of 30 days, as

done by Mohr et al. (2022) and Schröter et al. (2015). The pre-conditions are represented as the total precipitation amount in the 30 days preceding the high precipitation event.

To quantify extreme pre-conditions and form a reference of normal conditions, the preceding precipitation amounts of other events with precipitation amounts over 30 $[mm/d]$ have been calculated as well. In case of consecutive days with precipitation exceeding this limit, only the first day is used as this is the start of the precipitation event. For the 2021 flood event, July 13 or July 14 is the indicative day, depending on which day exceeded the event rain limit of 30 $[mm/d]$. The limit for extreme pre-conditions was set at the 90th percentile of preceding precipitation for all events. How these choices in the method influence the results is discussed later.

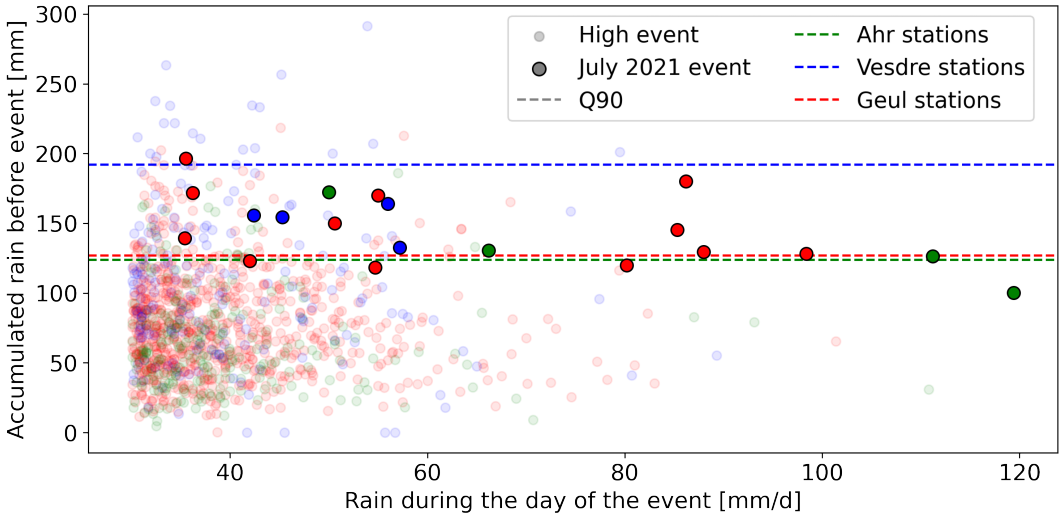


Figure 4.4: Pre-conditions for the relevant stations of the Ahr (green), Vesdre (blue) and Geul (red), see appendix C.4. The horizontal axis shows the precipitation amount on the day of the event that exceeded the 30 $[mm/d]$ limit. The vertical axis shows the precipitation amount in the 30 days preceding that event. The full marker shows the pre-conditions for each station for the July 2021 event, while the transparent markers indicate the pre-conditions for other events. The dashed line indicates the 90^{th} percentile of the preceding precipitation amount for each catchment.

In fig. 4.4, the precipitation amounts preceding the 2021 flood event exceed the 90th percentile for most gauges. Some stations received almost 300 [mm] in 31 days, which is a considerable amount compared to the yearly precipitation amounts of 600 to 1100 $[mm/y]$. For the Ahr and Geul catchment, 75% of the stations exceed the 90th percentile for the 2021 flood event which indicates extreme pre-conditions. As this is the majority of all stations, both catchments are considered as extreme in terms of pre-conditions.

Appendix C.4 provides more information on how fig. 4.4 was made, for example which relevant stations were selected. It also tests various choices for the method, such as choice of the first day of consecutive events, the 90th percentile limit for the definition of extreme, the duration of 30 days preceding the events and the event rain limit of 30 $[mm/d]$. The choices for the 30 days and 30 $[mm/d]$ values are robust and representative for the results. Selecting the second day for consecutive events, leads to extreme pre-conditions for half of the Vesdre stations. Lowering the extremity limit to the 80th percentile results in even higher percentages of extreme stations. Overall, appendix C.4 validates the method and shows that all catchments had unusually high pre-conditions.

4.2 Response

Return periods

The most common method to assess if the response, in this case the discharge, was extreme, is to calculate the return period. More specifically, the return period is calculated for the peak discharge, using the same method as for the precipitation: the GEV distribution. The number of locations is limited since the number of known peak discharges of the 2021 flood event is limited. Additionally, a sufficiently

long record period of the gauges was required. Due to shorter data sets for the discharge, the minimum record length is decreased to 10 years, instead of the 15 years for the return periods of the precipitation. Similarly to the return periods for the precipitation, the discharge of the 2021 flood event has been excluded for the derivation of the GEV distribution to keep consistency in comparing the 2021 flood event to the preceding period. When the current event is included for the GEV distribution, the return periods are lower, 30% of the gauges have a decreased return period by one order of magnitude (see table D.4 in appendix D.3).

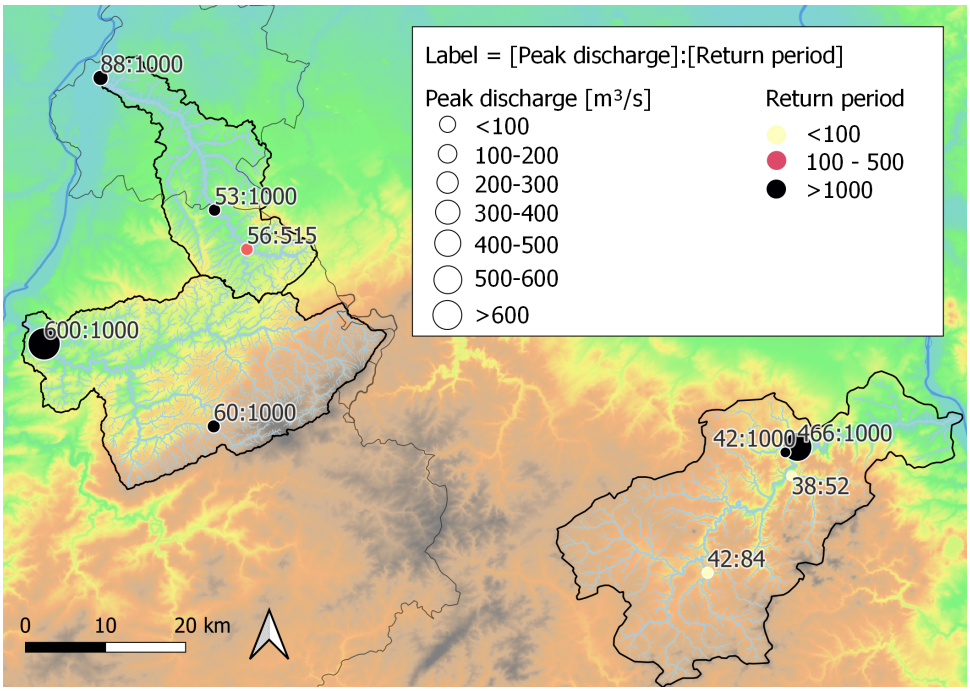


Figure 4.5: For each station with sufficient data the peak discharge is shown, as well as its return period from the GEV distribution, based on the preceding records. The peak discharge value is indicated by the size of the marker and the return period is indicated by the color scale.

Almost all return periods in fig. 4.5 exceed 1000 years, which is the highest distinction. This shows that the peak discharge was extreme. The data sets show that the peak discharges were considerable higher than the highest recorded peak discharge. It is possible that the discharge of other high flow events were also missed by the gauges in the data records, but since the data sets are verified by the authorities this is an unlikely explanation. More likely is the fact that the peak discharge of this event was unprecedented. Only two small tributaries in the Ahr catchment have less extreme return periods. This could be due to less precipitation in their drainage areas. Due to the sparsity of rain gauges in this area, this hypothesis is difficult to substantiate, and is therefore not analysed in more detail. With the exception of these two small tributaries, all gauges have a very high return period and therefore, the peak discharge can be considered as extreme for the 2021 flood event.

In fig. 4.5, the discharge values in the Ahr catchment are considerably lower than the peak discharges in fig. 3.8. The long term records were only available as daily averages. Therefore, the daily averages of the discharge were computed for the event and the maximal values of those averages (for either July 14 or 15) were used for deriving the return periods. Since the peak discharge occurred during the night, the maximum discharge averaged over 24 hours is split up when computing the daily averages. The peak daily discharges and the associated return periods for the absolute 24 hour maximum can be found in appendix D.3. The return periods then increase with an order of magnitude, compared to return periods based on the daily maximum. Appendix D.3 also provides the upper and lower bounds of the 95% interval for the interval period. All stations have a return period of over 1000 years as an upper bound, with one exception of 725 years. The lower bounds of the return periods show varying results, with one half of the stations having return periods of over 1000 years and the other half having return periods that are lower than 200 years. It can be concluded that even for a wide confidence interval, most stations show an extreme peak discharge magnitude from the high return periods.

The return periods of the peak discharge have been analysed by several reports. Instead of calculating the return periods, Schäfer et al. (2021) and Mohr et al. (2022) have compared the peak discharge of Müsch and Altenahr, estimated by the LFU, to the peak discharge with a 100-year return period, HQ100, that has been computed from earlier measurements. The peak discharges of the 2021 flood event were larger than the HQ100 discharge by a factor of 3.3 to 4, which already shows the extreme magnitude of the peak discharges. Johst & Demuth (2021) of the LFU have estimated a minimum return period of 1000 years for the flood event based on 80 years of measurements. Since this estimate excluded the historical floods of Roggenkamp & Herget (2014), Thieken et al. (2021) have recalculated the return periods with the inclusion of these historical floods which still resulted in a minimum return period of 1000 years. These historical floods are discussed in more detail in the next section. Ziemetz et al. (2021) have estimated several return periods for the Vesdre. For Chaudfontaine and Belleheid the return periods exceed 10.000 years, although the peak discharge of Belleheid used in this study was higher than in fig. 4.5 by a factor of 1.8. The high value of the return periods entails a high uncertainty. For Verviers, a city halfway along the Vesdre, the return period is estimated to be 200 years. The Geul has no known water levels for a 1000-year return period so the estimate of TFFF (2021) remains between 100 and 1000 years for the waterlevels throughout the river system. The literature offers no explanation or even occurrence of low return periods for small tributaries. The headwater gauges seem to have been the main focus of all studies mentioned in this paragraph. The high return periods of fig. 4.5 are thus confirmed by these studies and although the exact values of the return periods remain uncertain, they indicate that the peak discharge of the 2021 flood event were extreme.

4.3 Historical context

Measurement devices have not always been around which makes their data sets limited in length. Paleohydrology fills this gap by researching earlier hydrologic systems. Such research is work-intensive and thus the amount of research is limited. Information of the hydrologic system that cannot be derived from measurements is hereby referred to as historical information. For the Ahr, an important piece of such historical information is the study by Roggenkamp & Herget (2014) who have reconstructed historical peak discharges. This is the first topic of this section: the historical flood magnitude. There have been no similar studies for the Vesdre and the Geul. Nevertheless, the timing of flood events is often recorded in studies or even newspapers. The second topic of this section discusses the seasonality of historical floods from literature. The seasonality of the measured precipitation and discharge is also considered as comparison. As the uncertainty from historical information is relatively high, the findings of this section are not quantified as an answer to the second sub-question considering the extreme aspects of the data analysis.

Historical flood magnitude

Roggenkamp & Herget (2014) have reconstructed five pre-instrumental floods in the Ahr catchment for which there was sufficient information. The discharge of one location can be determined by multiplying the flow velocity and the flow area. Using flood marks (on buildings or from pictures) and topographic maps, the flow area was derived for each location. The mean velocity was derived with the empirical Manning formula. The exact method is described in appendix E.1. The Manning equation assumes uniform flow which means that the flow is constant in space and time. Since backwater effects disturb uniform flow, the flood mark locations have been carefully selected to minimise the occurrence of backwater effects. The assumption of uniform flow is still quite simplistic for complex flow conditions during a flood. The computation also includes other uncertain factors due to the estimation of the flow area, water level and the roughness coefficients. Since the uniform flow assumption and parameter estimates can cause uncertainty in the peak discharge estimations, the method was validated for two flood events when gauges have been installed, in 1984 and 1988. This resulted in an underestimation of the mean peak discharge of 5 to 7%. Although there are some simplistic assumptions in the method of Roggenkamp & Herget (2014), their study gives a valuable insight in the range of possible historical peak discharges.

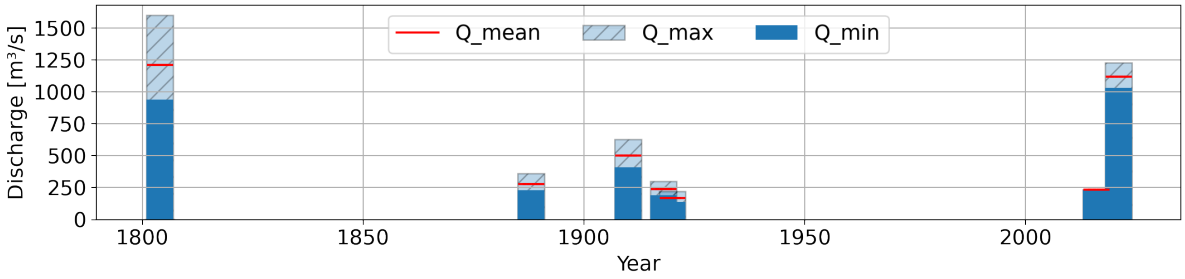


Figure 4.6: Bar chart of the historical discharge estimates of Roggenkamp & Herget (2014) for the Ahr, the maximum measurement, which occurred in 2016, and the estimated peak discharge of the 2021 flood event from Roggenkamp & Herget (2022). The light blue bar indicates the range between the minimum and maximum estimate and the red line indicates the mean estimate.

The reconstructed discharges of fig. 4.6 show that at least one flood event with a peak discharge of a similar order of magnitude as the 2021 flood event has occurred before and that the flood event of 2021 was not unique in the Ahr.

Historical seasonality

The seasonality of the flood event was said to be unique since it occurred during summer. The seasonality of the yearly maximum precipitation and discharge events are described in appendix F.1 and show that for all catchments the maximum discharge events mostly occur during the winter half-year. Only 22, 30 and 21% of the high discharge events occurred during the summer half-year (April-September) in the Ahr, Vesdre and Geul. On the contrary, the precipitation has the highest density of maximum events during the summer half-year, between 74 and 79%. This seasonal difference between high flow and high precipitation events is noteworthy, but not analysed further in this research. Kreienkamp et al. (2021) and International Meuse Commission (2005) confirm that winter and spring are the usual seasons of high flows in the catchments. Nevertheless, other catchments in Western Europe have experienced large summer floods (Verkade, 2021). According to the measurements and literature studies, high discharge events do not usually occur during the summer half-year. As a consequence, it can be expected that flood events in the summer half-year are unusual but not rare.

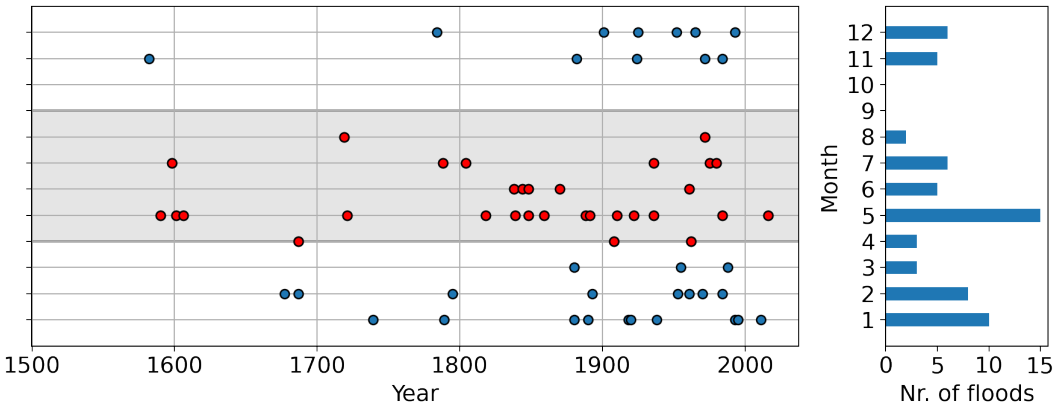


Figure 4.7: Seasonality of historical floods of the Ahr river, with years on the horizontal axis and the month of occurrence on the vertical axis. The amount of flood events in a single month are visualised on the distribution on the right. The red dots indicate the events that take place in the grey zone, which indicates the summer half-year.

To take a closer look at the seasonality of flood events in the catchments, historical flood events have been analysed. Fig. 4.7 shows that historical flood events in the Ahr catchment have occurred during the summer half-year. The data of the main source for this table (Seel, 1983) is complemented with more flood events from recent sources (Raths, 1985; Kess, 1986, 1989; Landesamt für Umwelt, 2022). Between 1410 and 2021, 75 floods have been documented of varying severity, with consequences ranging from flooded agricultural lands to multiple deaths. The month during which the flood occurred is known

for 63 events. During the summer half-year, 28 of those flood events occurred, which is 49%. The floods during the winter half-year are mostly related to ice and snowmelt, while the summer flood events are usually caused by thunderstorms. The latter are characterised by rapid water level increases and high flow rates. In conclusion, summer flood events in the Ahr catchment are not exceptional, instead they are quite common.

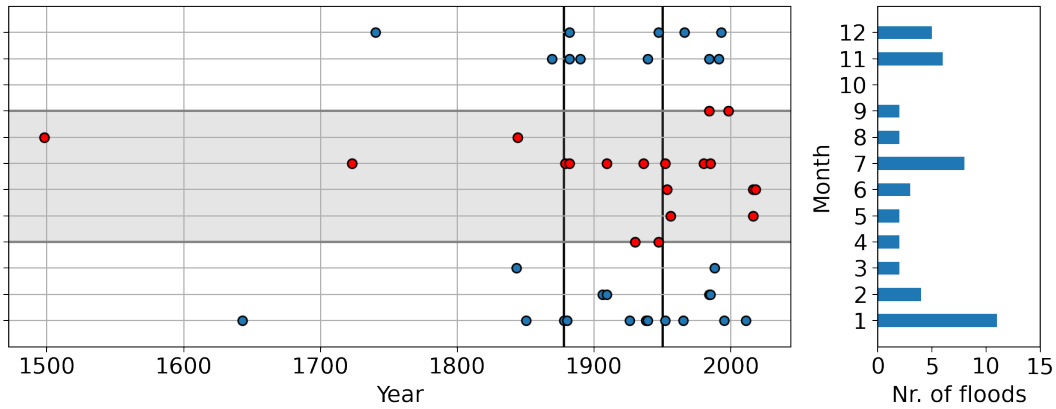


Figure 4.8: Timing of historical floods of the Vesdre river, with years on the horizontal axis and the month of occurrence on the vertical axis. The amount of flood events in a single month are visualised on the distribution on the right. The red dots indicate the events that take place in the grey zone, which indicates the summer half-year. The vertical, black lines indicate the inauguration of the dams of Gileppe and Eupen, respectively in 1878 and 1950.

Delforge (2021) has compiled information on many historical floods of the Vesdre river. He reports 51 floods between 1498 and 2018, of which 47 floods have a known month of occurrence. Of these 47 floods, 17 occurred during the summer half-year. July was the month with the most summer floods. The total percentage of floods during the summer half-year is 40%. Despite being less than the percentage of the Ahr, this is still a considerable percentage. Summer floods can thus not be considered as rare in the Vesdre catchment.

The available information on floods for the Geul is very limited. Historical records usually focus on the Meuse because flood events in the Meuse overshadow the smaller events of the Geul due to the difference in impact. No research has been done on the flooding of the Geul exclusively, as has been done for the Ahr and for the Vesdre. However, a short scan of newspapers provides a first insight in the historical flooding behaviour of the Geul. The resulting flood list, table F.1, can be found in appendix F.2, as well as how the list was composed. An event was considered as a flood event if the river exceeded the natural river bank. Therefore, the list consists of flood events of varying severity. The goal was to identify the timing of the floods, therefore this varying severity is not of major importance.

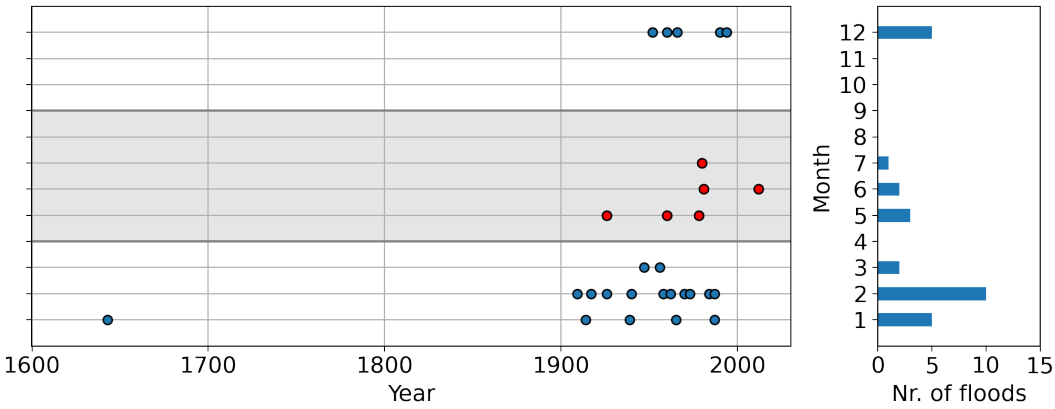


Figure 4.9: Timing of historical floods of the Geul river, with years on the horizontal axis and the month of occurrence on the vertical axis. The amount of flood events in a single month are visualised on the distribution on the right. The red dots indicate the events that take place in the grey zone, which indicates the summer half-year.

The list is visualised in fig. 4.9, in a similar figure as for the Ahr and Vesdre. Only 21% of the historical floods in the Geul occurred during the summer half-year, which is substantially less than for the Ahr and Vesdre. The danger of floods during winter is mostly caused by snow melt or ice blockages as the sources of many winter floods emphasize. The floods that occurred during summer were mostly caused by convective thunderstorms.

The historical analysis of flood seasonality in the Ahr and Vesdre shows a percentage of flood events in the summer half-year which is double than that of the measurements for high flow events. For the Geul there is no change. Although the historical analyses are not an in-depth study of flood events with separate severity classes, they confirm that flood events during the summer half-year are unusual but not rare, confirming the expectation from the seasonality of the measurements.

Chapter 5

Model set-up

Identifying extreme flow mechanisms, also known as flow paths in hydrology, is the goal of the third sub-question of this research. An extreme flow path or a difference in dominant flow paths during the 2021 flood event is not only interesting for further research, but it can have important consequences for flood risk management in the region. Measures that influence one specific flow path can be less useful if another path is dominant during extreme flows. Additionally, insight in the dominant flow paths for each catchment can also help to understand the observed differences between the catchments. Identifying extreme flow paths for the 2021 flood event helps to build our overall hydrological understanding of the catchments which is useful for many practices.

Table 5.1: The 5 main flow paths and their temporal scale, accompanying figure 5.1.

Flow path	**Temporal scale**
1. Interception-evaporation	1 min. - days
2. Hortonian overland flow	1 min. - 15 min.
3. Fast subsurface flow	10 min. - 1 day
4. Saturation overland flow	1 min. - 15 min.
5. Groundwater flow	1 day - years

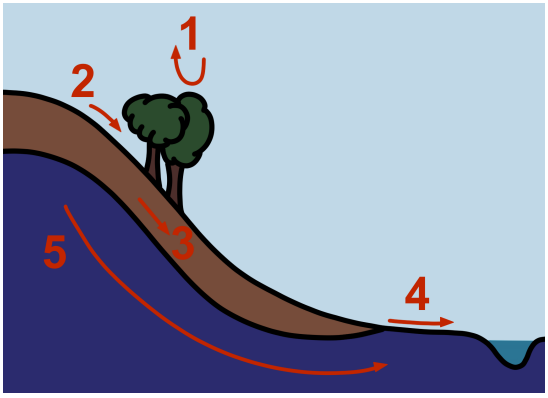


Figure 5.1: Conceptual visualisation of the main flow paths in a hilly landscape.

Fig. 5.1 and table 5.1 give an oversight of the main flow paths. Interception is obstruction and storage of precipitation by vegetation. After the precipitation is intercepted, it can move to the ground surface through stemflow and throughfall. Alternatively, the intercepted water can evaporate. Only the evaporation by interception is indicated in this figure for simplicity. Hortonian overland flow, also known as infiltration excess flow, occurs when the precipitation rate is larger than the infiltration capacity. The excess water runs off over land. Another quick runoff process is the fast subsurface flow. Water infiltrates into the soil but still flows quickly to the river through permeable soil layers. Saturation overland flow has a similar temporal scale as infiltration excess flow. At the bottom of hill slopes the soil can get saturated from a high groundwater level to such an extent that the water level rises above the soil surface. The soil can also become fully saturated in other locations due to prolonged precipitation, halting the infiltration process. Finally, water can flow to the deeper soil layers as groundwater, which is generally the slowest flow process. The hydrological behaviour of the catchments is determined by these five flow paths that vary in spatial and temporal scales and depend on the catchment characteristics.

Flow paths cannot be derived directly from observations, unlike the forcing and response. Models can be used to identify the dominant mechanisms since the mechanisms can be conceptually included in the model structure. A model that is calibrated to optimize the performance for a measurable metric such

as the discharge then likely represents the behaviour of the flow mechanisms through its structure. The selection of the models is the topic of the next section. This also includes the reasons to use two models. The structure of those models is explained in the following sections. Next, the input data is validated with a data check. Finally, the method of deriving the flow mechanisms with the models is discussed.

5.1 Purpose and methodology

Modelling is not the only method to derive flow paths. Other methods exist, such as regionalization and hydrologic signatures (McMillan, 2021). In regionalization, knowledge on flow paths is used from nearby, similar catchments. In modelling, dominant processes can be found through the sensitive parameters and internal model states. Finally, hydrological signatures combine features of the flow time series. Through extensive research, some signatures have been linked to hydrologic processes. However, both regionalization and hydrologic signatures are methods for long-term hydrologic behaviour as they use longer time series for their derivation. To assess the flow paths for the flood event, only modelling offers the opportunity to derive properties of such a short time series.

Model results have various sources of uncertainty such as input data, calibration data, parameters and the structure of the model (Moges et al., 2020). For example, the specific structure of a model can miss or incorrectly reflect important hydrological processes. In climatology, the model structure error is quantified by ensemble modelling (Huang et al., 2017), a method that combines the results of multiple models and is starting to be applied more often in hydrology. Ensemble modelling does not only enable the quantification of structural uncertainty, it also reduces the uncertainty of the results and therefore performs better than a single model (Moges et al., 2016). This uncertainty reduction is the result of error compensation, increased consistency and reliability (Hagedorn et al., 2005). In contrast with the previously mentioned studies, the goal of this research is not to produce the most accurate model results such as discharge, but to use the model structure and calibration to detect the dominant hydrological processes. Only two models are applied for this goal since the detection of dominant mechanisms is only part of the research objective and so, ensemble modelling with more than two models would be outside the scope of the research. Using two models with different structures allows for a comparison between findings and therefore adds robustness to these findings.

The robustness of the findings can be increased by maximally diversifying the two models. The selected models are a simple, lumped model, based on the HBV model, and a more complex and spatially distributed model, based on the SBM model. For simplicity, the rest of this research refers to them as the HBV model and SBM model. Both models are conceptual because they mimic the physical mechanisms within the hydrologic cycle through simplified concepts.

The main difference between the two models is that a lumped model represents the entire catchment as one homogeneous grid cell, while distributed models can consist of many grid cells and thus reflect the heterogeneity of the catchment. Nevertheless, all distributed models in fact remain lumped models, but only to a smaller spatial scale. Despite the apparent advantage of including spatial variation, distributed models do not always outperform more simple, lumped models (Beven, 1989; Grayson et al., 1992). Since the distributed model requires a parameter set for each grid cell, the determination of parameter values is difficult (Rosso et al., 1994) and using automatic calibration for the parametrization can result in equifinality. Equifinality occurs when there are so many parameters that one parameter value can compensate for another, resulting in many varying parameter sets with similar performance (Beven, 1993). To avoid equifinality as a disadvantage of distributed models, the parameter values can be estimated from physical characteristics such as the geology or slope. However, this can lead to inaccurate parametrization. The possibility of equifinality is not only higher for the SBM model due to its spatial distribution, but also due to the inclusion of more hydrological processes. The higher process complexity leads to a higher number of parameters. Comparing a lumped model with a smaller amount of processes and a distributed model with a large variety of processes can provide an additional insight to see which model performs better. The widely varying models mainly contribute to the main goal to optimize the robustness by being opposite in spatial scale (lumped and distributed), model structure (simple and complex) and parametrization (calibration and physically-based).

There are many models within these previously described model types. The selected models are appropriate for the flow characteristics of the flood event and for the catchment characteristics. An important factor for the selection of these models was the familiarity and access to guidance for the author.

5.2 Lumped HBV model

The lumped model used in this research is adopted from the course CIE4431 Hydrological Modelling. It represents a simplified version of the HBV model. The HBV model or Hydrologiska Byråns Vattenbalansavdelning model, has been developed by Bergström (1976) for Scandinavian catchments. Many different versions have been established since then. The original model is a semi-distributed, conceptual model. A lumped version was used here for simplicity. Fig. 5.2 visualises the included processes as a scheme of stores and fluxes. A detailed overview of all accounted processes and equations can be found in appendix G.1. A short but comprehensive description of the HBV model is given below with a focus on the model properties that are relevant for interpreting the model results in the next chapter.

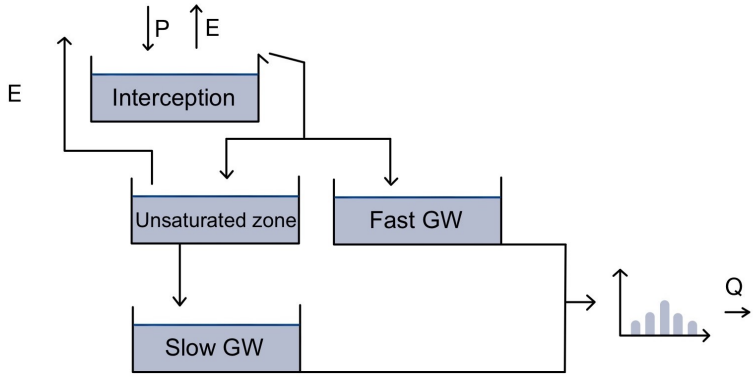


Figure 5.2: Schematization of the lumped HBV model with the stores and external fluxes (P: precipitation, E: evapotranspiration, Q: flow).

For each time step, the input of precipitation moves through the stores and fluxes in the following order. First, the precipitation enters the interception store, where it can evaporate back to the atmosphere or flow to one of two subsurface store. One represents the fast subsurface runoff as the fast responding groundwater store ("Fast GW"). The other subsurface store ("Unsaturated zone") represents the unsaturated zone from which more water can evaporate or flow to the slow groundwater storage. The water from both the slow and fast groundwater stores are added up and delayed with a lagging function to simulate the final discharge. With only two fluxes contributing to the total flow, a slow and fast one, the HBV model is relatively easy to interpret.

The following parameters are considered in the lumped model. More information on their meaning and their value ranges can be found in appendix G.1.

- **Imax**: Maximal interception storage: reflects the storage capacity of the vegetation
- **Sumax**: Maximal storage capacity of the unsaturated zone: reflects the storage capacity of the top soil, which depends on the thickness and permeability of the soil. The ratio of the amount of water in the store to its capacity is the filling degree of the unsaturated zone which plays a role in all processes connected to this store
- **β**: Exponent for unsaturated/ fast groundwater flow: influences the amount of water going to the fast groundwater flow store. A value above 1 leads to less flow to this fast store. The filling degree of the unsaturated store has a larger influence on this flow than β
- **Ce**: Transpiration parameter: reflects the potential for soil evaporation and transpiration which is related to the soil properties and land cover. A high value indicates little evapotranspiration
- **Pmax**: Percolation factor: reflects the vertical hydraulic conductivity between the unsaturated zone and slow groundwater zone. A high value leads to a higher flux between these stores
- **Kf**: Horizontal conductivity of the fast responding groundwater store to river discharge. A higher value indicates a higher flux
- **Ks**: Horizontal conductivity of the slow responding groundwater store to river discharge. A higher value indicates a higher flux

- **Tlag**: Time delay of river discharge. The discharge is delayed over Tlag days in the shape of a symmetrical triangle

There are several limitations of this model. Firstly, there is no explicit flux for overland flow, although this flux could be indirectly accounted for in the fast responding groundwater flux. Secondly, the water of the unsaturated zone can only run off through the slow groundwater zone. Since the filling degree of this zone determines the partitioning of slow and fast runoff and many other fluxes, little water in the unsaturated zone could thus never lead to fast runoff which excludes the process of Hortonian overland flow. Lastly, one delay function is applied to the sum of the fast and slow groundwater flux, although both could have different delays. Their varying temporal scales are reflected by their varying horizontal conductivity, but not the delay function. This lumped model is quite limited and over-simplified, but this makes the calibration quicker and its use easier.

5.3 Distributed SBM model

Similar to the HBV model, many variations of the SBM model have been developed since its introduction. They are all based on the Simple Bucket Model (SBM) principle, which works similar to the HBV model since the structure of both models is based on different stores that can be filled and drained with water. For this study, the wflow_sbm model from Deltares has been used. Wflow is a modelling platform that includes multiple distributed models. The wflow_sbm model proved the best option, since it did not require extensive calibration (in contrast with the wflow_hbv model), it was completely finished (in contrast with the wflow_topoflex model) and it fitted the characteristics of the catchment. The processes that are included in the wflow_sbm are given in fig. 5.3.

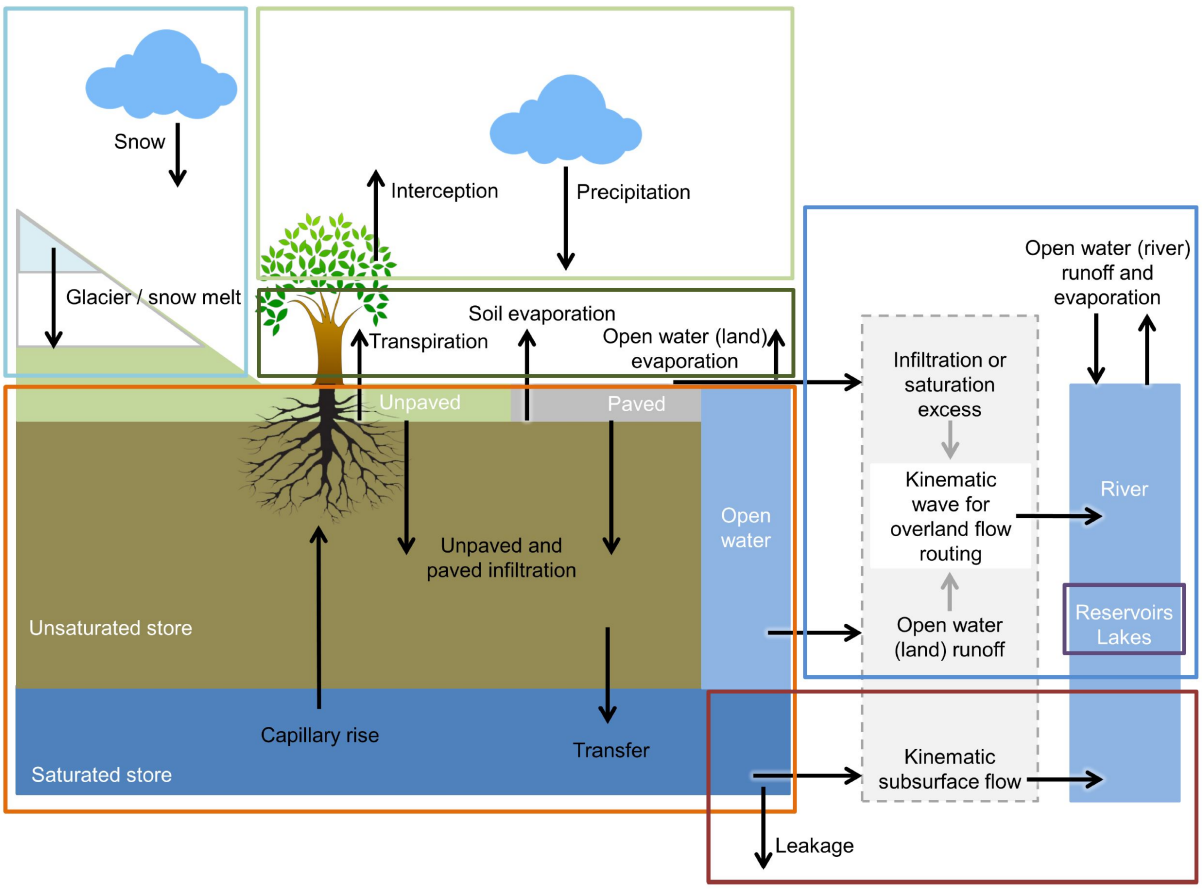


Figure 5.3: Schematization of the wflow_sbm model, from Van Verseveld et al. (2022).

The precipitation and potential evaporation run through a set order of processes as shown in fig. 5.3. At each process, represented by a store, a certain amount of water is subtracted from the flux of water

that can move to the next store. The SBM model splits the cells of its raster in land cells and river cells, each with different fluxes. A land cell can still contain an area of open water, for example when a lake is present. In this configuration, cells are identified as river cells when the drainage area surpasses 20 [km^2]. The precipitation and potential evaporation of the land cells go through the following vertical processes in this order: interception, open water evaporation, infiltration, flow between soil layers (percolation), evaporation from soil layers, transpiration, excess water, capillary rise and leakage from the saturated zone. The water then moves laterally as river flow for the river cells and as overland and groundwater flow (from the saturated zone) for the land cells. The lateral movement of water is based on the topography and the kinematic wave approximation, which will be discussed later. The saturation overland flow as indicated as number 4 in fig. 5.1 is also included and added to the surface runoff after updating the groundwater levels from the lateral subsurface flow. This saturation overland flow only concerns the saturation overland flow at the bottom of hill slopes due to high groundwater levels. The model is also capable of including snow and glaciers, as well as reservoirs. The exact representation in the model for each process and other information on the SBM model can be found in appendix G.2. The SBM model includes many flow paths which gives it a more detailed representation of reality, even though the concept of the SBM model and HBV model is similar.

With so many processes come a lot of variables. Several parameters are more sensitive according to Imhoff et al. (2020) and are shortly discussed here as they are relevant for the model interpretation later on. The most sensitive parameters are c, Kext, KsatHorFrac, M, RootingDepth, Sl, Swood, thetaS and TT. These parameters mostly influence the water storage, fluxes between layers and cells and the interception. The most important parameter in this study is KsatHorFrac, since it has a high degree of conceptualization. It reflects the horizontal hydraulic conductivity of the saturated zone.

The high amount of parameters increases the risk of equifinality during automatic calibration. Therefore, most parameters are derived from look-up tables and (pedo)transfer functions (PTF) which are both linked to actual catchment properties such as land cover and geology. Deltares has provided the base models for the Ahr, Vesdre and Geul catchments as cut-outs from their larger Rhine and Meuse models.The base maps were already of sufficient quality and only some parameters had to be scaled or offset differently for the smaller catchments in this research. The grid cells have a size of 600 by 900 [m], creating between 625 and 1581 grid cells, depending on the catchment. Although the SBM model has many parameters and therefore a risk of equifinality, the standard derivation of parameters from catchment characteristics and the higher sensitivity of several parameters decrease the required effort for a manual calibration.

There are several limitations to the wflow_sbm model (Deltares, 2019). Results for soils deeper than 2 [m] may be unrealistic, but it is assumed that this causes no significant error since the soils have been described as thin in chapter 2 and the permeable soil layer above bedrock is thinner than 2 [m], according to Pelletier et al. (2016). The lateral movement of water could be wrong in non-steep terrain due to the use of the kinematic wave approximation for lateral shallow water flow. The approximation is based on the assumption that lateral water flow is driven by elevation gradients instead of hydraulic gradient. This could lead to incorrect flow in flat reliefs where the hydraulic gradient differs from the elevation gradient. Due to the overall strong relief and narrow valleys in the three catchments, this limitation is a minor issue. None of the limitations of the SBM model reduce the performance of the models for these catchments in such a way that the model cannot be applied.

5.4 Data check

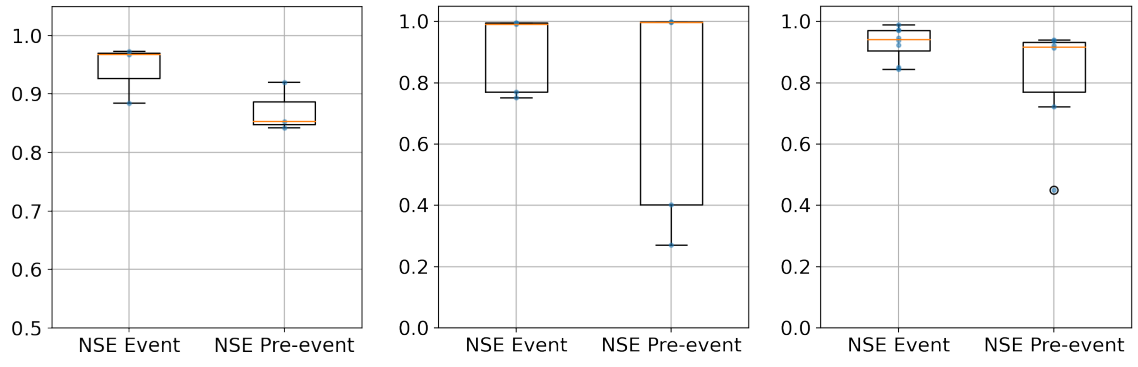
The quality of the model results is limited by the quality of its input. Point measurement errors at gauges can always occur, but most of the measurements used in this research have been verified by the responsible authority. The discharge data was not always verified for recent years and therefore, it has been checked with the Budyko curve appendix B.6 and with a visual check. No clear indications of data errors have been found. Due to the importance of the spatial scale for fast floods (Kreienkamp et al., 2021; Viglione et al., 2010), spatially varying input data is valuable. A simple spatial interpolation such as Thiessen polygons for the gauges could lead to inaccurate results. Radar data offers a higher spatial resolution. Since radar data is not measured directly, validating the radar data with the measurements is required before using it as model input.

There are multiple radar data sets available for Western Europe, and multiple data sets have been analysed for the different catchments. Only the data sets with the best fit are discussed here. Furthermore, temperature and potential evaporation are taken from the ERA5 dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF). The temperature only plays a limited role in the models and potential evaporation is rarely measured in the catchments. Therefore, validating the ERA5 temperature and potential evaporation is not included in this research. The precipitation plays a crucial role and requires validation. The ERA5 precipitation dataset was not the most accurate dataset for any catchment and is not used. Each catchment uses a different radar product. For the Ahr, the RADOLAN dataset of the Deutscher Wetterdienst (DWD) provided the best results. It combines data of 17 radar stations throughout Germany and more from neighbouring countries, which provides sufficient spatial coverage for the Ahr. The temporal coverage is from 2018-01-03 to 2021-12-31. For the Vesdre, the REGNIE dataset is provided by Deltares for the period of 2005-01-01 to 2018-02-21 and 2021-03-27 to 2021-07-19. It is derived from the operational system of Rijkswaterstaat (Directorate-General for Public Works and Water Management in the Netherlands) and is made by interpolating station data with a climatological grid. For most areas, the grid is derived from the REGNIE dataset, hence the name, with the genRE method (Van Osnabrugge et al., 2017). For the Meuse catchment, which includes the Vesdre catchment, the climatological grid is based on the E-OBS data set, which is based on rain gauge interpolation for the precipitation. A spatial re-analysis of the KNMI (Koninklijk Nederlands Meteorologisch Instituut: Royal Dutch Meteorological Institute) for the flood event (Overeem & Leijnse, 2021) has been used for the Geul and was available for the period of 2019-01-01 to 2021-12-31. The re-analysis product is based on international radar data and validated station data.

Although the REGNIE dataset is not a radar product, this research uses radar products as an umbrella term for all three spatially distributed precipitation data sets that have been described above, to avoid confusion with gauge measurements and a simplistic spatial interpolation from Thiessen polygons. These three selected data sets for each catchment have a different source and time span, which hinders an analogue analysis for all catchments. Nevertheless, a high similarity between the radar products and the gauge measurements and a higher spatial resolution are considered as more important aspects than the homogeneity of the data for all catchments. For this reason, these three varying data sets have been selected instead of one data set, such as ERA5.

The analysis periods for the models are split in two parts. The pre-event or non-event period (nev) consists from the start of the available data until 11 July 23:00 h. For the Vesdre, the pre-event period is considered until 21 February 2018 23:00 h to avoid a data gap of 3 years. The Vesdre has the largest coverage for the pre-event period but misses 3 years before 2021, whereas the Ahr and Geul cover less time (respectively 3.5 and 2.5 years) but have no data gaps up to the 2021 flood event.

The 2021 flood event period is considered from 12 July 00:00 h until 18 July 23:00 h, which covers the earliest precipitation and the recession of the hydrograph.



(a) Ahr *(b) Vesdre* *(c) Geul*

Figure 5.4: Boxplots of NSE values for the stations with available data for the Ahr (a), Vesdre (b) and Geul (c). The NSE values are given for both the event (from 12-07-2021 until 17-07-2021) and the preceding period for which radar data was available. The exact meaning of NSE values is described in the next section, as it is the calibration objective function.

Fig. 5.4 shows that the radar data performs slightly worse for the pre-event period, especially for the stations of Jalhay and Spa in the Vesdre catchment and station 380 in Maastricht in the Geul catchment. In appendix H.1, the figures H.1, H.2, H.3 show the comparison of the instantaneous precipitation during the event for the stations used in fig. 5.4. With overall values of more than 0.8, it can be concluded that the precipitation data sets are of sufficient quality to use in this study.

5.5 How to derive flow mechanisms from models

The extremity of the flow mechanisms can be analysed with the two models by deriving and comparing the dominant flow paths for the 2021 flood event and the pre-event period. This goal requires two model runs: one calibrated for the period before July 2021 and one calibrated for the flood event itself. A variety of derivations can be made from both calibrations. A first insight in possibly extreme flow mechanisms is given by the performance of the model that is calibrated for the pre-event period to replicate the flood event discharge. If the model is capable of replicating normal flow conditions (during the pre-event period), the model structure can simulate the important flow mechanisms. If the model cannot simulate the 2021 flood event with the same parameter set, there are two possibilities. In case that a re-calibration still results in a poor performance, the model structure does not represent the flow mechanisms sufficiently to reproduce the flow. In case that a re-calibration results in a good performance for the flood event, the model structure suffices and the parameter changes can be analysed to detect the required changes in flow mechanisms to simulate the flood event. These required changes in flow mechanisms can be validated by analysing the internal states of the models. The stores and fluxes of the models can be compared between both calibrations to detect differences from the re-calibration. These internal states can finally be used to identify extreme flow mechanisms of the 2021 flood event.

Both models require a different calibration method due to their varying model structure. The lumped HBV model has 8 parameters. With the consequent low risk of equifinality, automatic calibration can be applied. The Monte Carlo-calibration has been selected for this purpose. For each iteration, the model is run with a parameter set that is randomly generated from a given range of values. The parameter set with the optimal performance, quantified by a performance indicator, is used as the final parameter set. The initial value range for the parameters can be found in table H.2 in appendix H.3. Only for the Geul some changes were needed after trail-and-error iterations to optimize model performance. The resulting parameter range is given in table H.3. The Monte Carlo-calibration is based on 10.000 iterations and the performance indicator is the Nash Sutcliffe efficiency (NSE), eq. (5.1), since this is a common metric for assessing the performance of hydrological models (Knoben et al., 2019).

$$NSE = 1 - \frac{\sum_{t=1}^{t=T}(Q_{sim}(t) - Q_{obs}(t))^2}{\sum_{t=1}^{t=T}(Q_{obs}(t) - \overline{Q_{obs}})^2} \tag{5.1}$$

An NSE value smaller than 0 indicates that the model has a worse performance than a fictional model consisting of the mean observation for each time step. A value of 1 indicates a perfect correspondence between the model and the observation. NSE values larger than 0.5 show a satisfactory performance (Moriasi et al., 2007). Although there are no strict limits of which values indicate a good model, values over 0.7 or 0.8 can be interpreted as a good performance. Due to the mathematical properties of the NSE value, see eq. (5.1), the NSE value inherently increases for a smaller sample size and is thus no good measure to compare its value for data sets of different lengths. The NSE value is only computed for one calibration variable: the hourly discharges, since the goal is to analyse the flow mechanisms during different flow regimes without focusing on specific aspects such as the water volume balance. The NSE value for the hourly discharges is used as the performance indicator for the calibration process and indicates a satisfactory performance of the model with a value of more than 0.5.

The performance of the SBM model is also analysed with the NSE value. An automatic calibration such as the Monte-Carlo calibration would likely result in equifinality due to the large number of parameters for each grid cell. Furthermore, the SBM model has a long run time of 5 to 40 minutes due to the spatial variability. To optimize the performance of the model, some of the sensitive parameters that were described in section 5.3 have been changed manually. Firstly, this had to be done for the original Ahr model since its performance might have been sufficient for the entire Rhine area, but not for its own catchment during normal flow conditions. The KsatHorFrac value, which reflects the horizontal

hydraulic conductivity, has been decreased to a constant value of 50, which is lower than the usual 250 (Van Verseveld et al., 2022). The low value is in line with the impermeable soils found in section 2.1. Manually changing certain parameter values as a calibration method is more subjective than an automatic calibration. The manual calibration has been continued until a good performance was reached with the sensitive parameters that could not be improved further with reasonable efforts. Further performance improvements would require much more runs for varying parameters to a more detailed level and due to the long run times, this would require a large amount of time. Since the SBM model is only a small part of this research, only limited time and thus calibration quality is available for its manual calibration. The manual calibration was a process of testing the effects of the sensitive parameters described in the SBM model description of section 5.3, combining high performance values of several parameters and by trial-and-error optimising the overall performance.

The data check of section 5.4 mentioned the available data periods of the model input. The periods for the Ahr and Geul of 2 to 3 years are relatively short. These short periods of several years would be insufficient for models with a main goal of replicating discharges for flood prediction or other practical applications. Such purposes require a considerable calibration and validation period. This research uses the models for detecting dominant flow mechanisms and is therefore less dependent on an optimal flow replication. Nevertheless, a shorter data set includes less variations in flow patterns. With the limited length of around 3 years, a particularly dry summer or high flood event could have been excluded from the data set which means that the data set might not include all possible patterns in flow mechanisms. This reduction in insight of flow mechanisms due to the short data sets is the reason that validation with an independent data set of observations is not included here. A wider pattern of flow patterns of an already limited period was deemed more important than the model validation.

An additional downside of the short data periods is the use of short warm-up periods. After some initial model runs, 2 months seemed sufficient for this purpose. Whether this assumption was correct, is explored in the discussion, section 6.4. Since the warming-up period is not relevant for the calibration and its objective function, the calibration periods are even shorter than the data periods. The calibration periods are given in table H.1 in appendix H.2. Since the Vesdre has a longer data period, starting in 2005, a warming-up period of 1 year has been used. Due to the long run times of the model, especially considering the manual calibration for the SBM model, only the period starting from 2016 was used for the calibration of the Vesdre models. The flood event period is equal to that of the previous section, so from 12 July 00:00 until 18 July 23:00. The flood event models have a warming-up period of at least 3 months (Vesdre, Ahr SBM) or start with the outstate of the pre-event model (Geul, Ahr HBV). The model of the Ahr for the pre-event period has been calibrated by averaging the hourly discharge for each day since the observed discharge has a daily time step. For the event period, this was not necessary. In conclusion, calibrating the models for the pre-event period is hindered by the short length of the forcing data and consequently short warm-up periods and a limited variety of flow patterns.

For legibility of the next chapter, the model calibrated for the period preceding the flood event will be called the original model or in short (nev), reflecting the “non-event period”. The model calibrated for the flood event itself will be called the re-calibrated model or in short (ev).

The data analysis for the forcing and response included most stations, but the flow mechanisms will only be analysed for the main stations. These are for the Ahr, Vesdre and Geul respectively Altenahr, Chaudfontaine and Meerssen (10.Q.36). Due to computational time of the models and the interpretation of all the states, only the processes for the entire catchment are considered and not the tributaries.

Chapter 6

Model results

By following the model set-up of the previous chapter, the dominant mechanisms of the flood event can be deducted and compared to the usual flow mechanisms. The relevance of that deduction depends on multiple factors, but the general performance of the model to replicate observed flows is an important indicator. The estimated flows for the 2021 flood event are included in the term of observed flow to maintain a clear distinction between the reference flow and modelled flow. If the model is not able to replicate observed flows, the model is not a reliable representation of reality and thus the emerging flow mechanisms are not reliable as well. Therefore, the model performance is discussed first. The performance is analysed for the pre-event period and the 2021 flood event according to the calibration indicator. Additionally, other indicators are studied to evaluated the model performance for the pre-event period. In the third section, the emerging flow mechanisms are deducted from the parameter changes and the internal states. How extreme these flow mechanisms are, is included in this section. Finally, the discussion summarizes the results of the chapter, but also provides the limitations of these findings.

6.1 Calibration performance

Table 6.1: NSE values for the direct hydrograph for the three main gauges for both models as a result of the calibration for the pre-event period. The values indicate the performance of the original models.

	Direct	
	HBV	**SBM**
Ahr	0.79	0.71
Vesdre	0.79	0.65
Geul	0.67	0.40

The calibration of the models for the pre-event period as described in section 5.5 has resulted in the NSE values of table 6.1. The values range from 0.65 to 0.79 with an outlier of 0.4 for the SBM model of the Geul. They indicate a good performance of both the HBV and the SBM model (except for the Geul) of the simulation of the discharge. The HBV model outperforms the SBM model. The overall good performance of both models show that they represent the flow mechanisms sufficiently to simulate the normal flow conditions. Therefore, the models can be used for the next steps to derive the possibly extreme flow mechanisms of the 2021 flood event.

The direct hydrographs of the models that form the base of table 6.1 can be found in appendix H.5 and show how the observed and simulated discharge differ. The observed hydrograph of the Ahr has a very low baseflow and sharp discharge peaks. The baseflow is well represented by both models. The higher peaks are mostly underestimated by both models, but some peaks are overestimated by the SBM model. The recession after flood peaks is often too short for the SBM model. Such low baseflows and sharp peaks are not present in the flow pattern of the Vesdre. The HBV and SBM model are not good at replicating the varying baseflow. Especially after a high flow period, the HBV model has a recession curve that is too fast and underestimates the baseflow. The SBM model overestimates many peaks. For the Geul, the HBV model has a slight overestimation for low flows and underestimates the higher peaks.

The SBM model underestimates the base flow and varies in its overestimation and underestimation of the peaks. Most models of all catchments are not capable of accurately replicating the baseflow, but the discontinuities between the models and observations are largest for the high flows, especially the peaks.

Re-calibrating the original models for the 2021 flood event shows how much the model performance can be improved. The changes in parameter values can give insight in the changes of the flow mechanisms, which has been pointed out in section 5.5. The optimal dataset of the original and re-calibrated HBV models for the Ahr, Vesdre and Geul can be found in appendix H.4.

Table 6.2: NSE values for the direct hydrograph of the 2021 flood event for the three main gauges for the original (nev) and re-calibrated (ev) models.

	HBV (Nev)	**HBV (Ev)**	**SBM (Nev)**	**SBM (Ev)**
Ahr	0.19	0.91	0.11	0.65
Vesdre	0.29	0.99	0.90	0.93
Geul	0.48	0.85	-0.17	0.96

In table 6.2, the NSE values of the simulated flood event discharge for the original and re-calibrated models are given. The NSE values of the original models range between -0.17 and 0.48 (with an outlier of 0.9 for the SBM model of the Vesdre) which indicates an overall poor performance. The good performance of the models for the pre-event period is not extended to the flood event. The sample size for the flood event is much smaller than for the pre-event period which naturally leads to higher NSE values. The NSE values of table 6.1 and table 6.2 can therefore not be compared in absolute terms. The high values of the re-calibrated models indicate a good performance, regardless of the small sample size. The considerable improvement in performance from re-calibration and the poor performance of the original model for the flood event show that the difference in flow mechanisms during the 2021 flood event could not be captured by the original model configuration and is a first indication towards extreme events.

Ahr

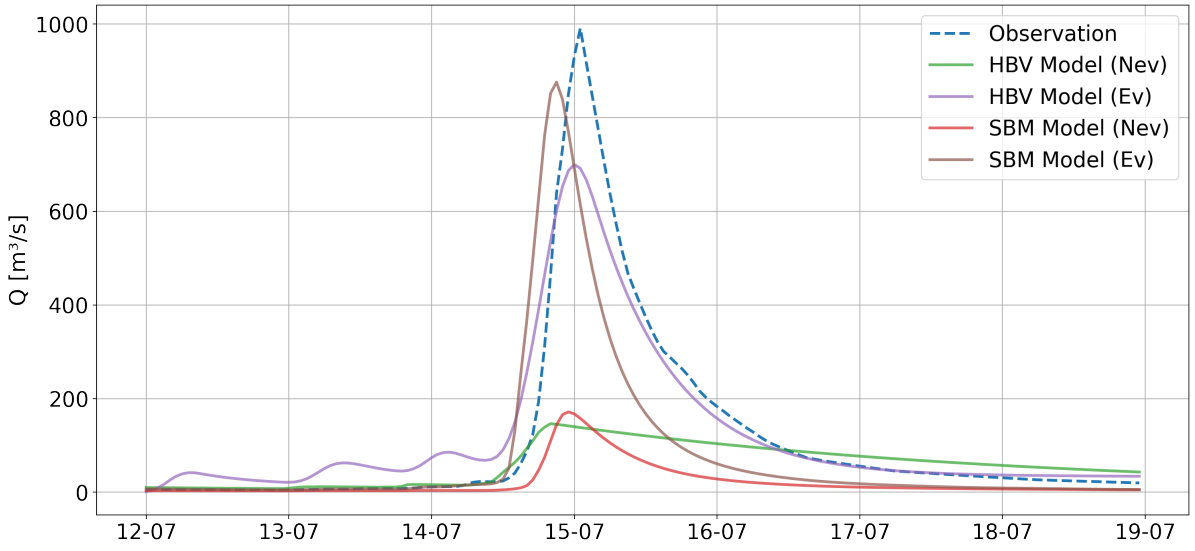


Figure 6.1: Hydrograph of the flood event for the Ahr (Altenahr station), for the HBV and SBM model, both calibrated for the period before the event (Nev) and during the event (Ev). The observed discharge from fig. 3.8 is given as a dashed line.

To provide more context to the NSE values of table 6.2, the flood event hydrographs are given for the observation (in all catchments a partial estimate) and all models. This helps to identify the cause of low or high performance of certain models. For the Ahr, fig. 6.1 shows why the performance of the original models is very poor for the flood event. The observed peak discharge is underestimated with a factor of 5 by the original models. Re-calibrating the model to optimize the NSE value for the event results in a great improvement of the performance. The re-calibrated SBM model achieves the highest peak discharge of 876 $[m^3/s]$, which is still an underestimation of the estimated peak discharge of 11%. The peak discharge is achieved 4 hours before the estimated peak discharge. The re-calibrated HBV

model has a better approach for the timing and recession curve. Both re-calibrated models have a better performance than the original models by having increased the modelled peak discharge which resulted in the highest peak discharge of the SBM model, but the best timing for the HBV model.

Vesdre

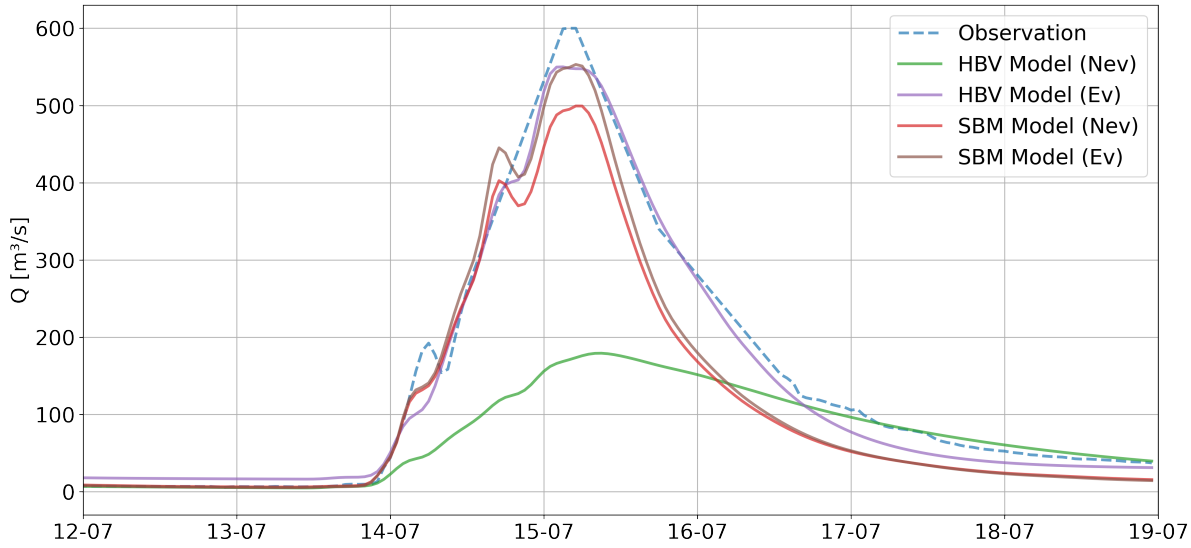


Figure 6.2: Hydrograph of the flood event for the Vesdre (Chaudfontaine station), for the HBV and SBM model, both calibrated for the period before the event (Nev) and during the event (Ev). The observed discharge from fig. 3.9 is given as a dashed line.

Fig. 6.2 shows how most models have a similar shape as the hydrograph for Chaudfontaine, which was a rough estimate based on the peak discharge, its moment of occurrence and linear interpolation. The original HBV model is the only model that strongly underestimates the peak discharge. The re-calibration is able to increase its peak discharge close to the observed peak discharge. Additionally, it approximates the recession of the hydrograph well. Of all original models of all catchments, the original SBM model for the Vesdre has the best performance, both in the timing and the peak discharge magnitude. This explains the exceptionally high performance compared to the other original models for the flood event in table 6.2. The re-calibration model only has a slightly higher peak discharge. Although the original HBV model had a poor performance, the re-calibrated HBV model results in the best hydrograph approximation which is in contrast with the SBM model that has a good performance for both versions.

Geul

As for the Ahr and Vesdre, the original HBV model underestimates the estimated peak discharge, now by a factor of 2. The re-calibration of the HBV model results in a higher peak discharge, but the hydrograph shape is still different. Especially the recession curve is too gradual. The original SBM model is the only model for any catchment that overestimates the peak discharge. Not only the higher peak discharge is noticeable, the time difference between the peak discharges of 16 hours is also a distinct deviation. The poor similarity between the observed and modelled discharge explains the outlier of 0.4 of the SBM model for the Geul in table 6.1. The re-calibrated SBM model only has a slight underestimation of the peak discharge and has the best approximation of the hydrograph shape. The original models for the Geul had a overestimation and underestimation for the peak discharge but the re-calibrated models were an improvement for the estimated peak discharge and hydrograph shape.

While the original models are good at modelling the flow before the 2021 flood event, their poor performance (with the exception of SBM model of the Vesdre) for the flood event confirm that there is a change in behaviour during the flood event which was not captured by the original calibration of the models. This enables the proposed method of analysing the re-calibrated parameters to detect possibly different flow mechanisms.

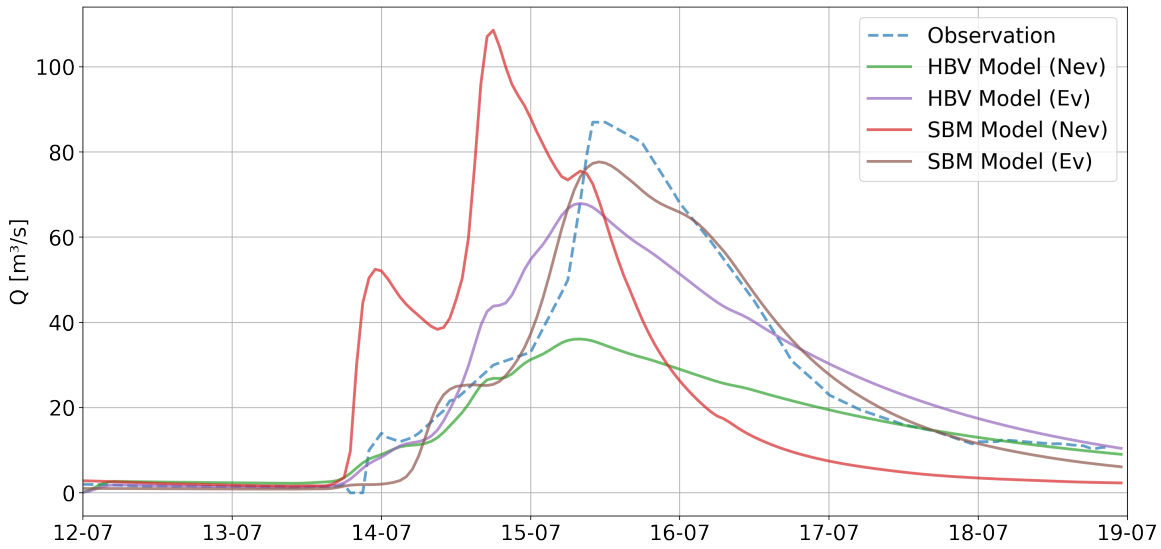


Figure 6.3: Hydrograph of the flood event for the Geul (Meerssen, 10.Q.36, station), for the HBV and SBM model, both calibrated for the period before the event (Nev) and during the event (Ev). The observed discharge from fig. 3.10 is given as a dashed line.

6.2 Evaluation performance

The previous section only analysed the performance of the models for the calibration period with the same indicator as the calibration, the Nash Sutcliffe coefficient for the direct hydrograph. This showed that the current model can perform well for its own calibration period and objective function. To assess if this specific model can make accurate predictions in general, the model results should be validated by testing the performance of the model beyond its calibration period (Refsgaard & Storm, 1990). However, due to the limited length of the data sets and the preference to include as much data as possible for the calibration, there is no data set available for validation in each catchment. Instead of validation, the performance of the model can still be evaluated by testing for different metrics than during the calibration.

Testing for new performance metrics can be done in multiple ways. A different objective function, also known as a performance criterion, can be selected, for example the Kling-Gupta Efficiency (KGE) instead of NSE. Alternatively, the evaluation can be based on a different set of observations or variables, such as the actual evaporation. The currently used variable, the discharge, can also be changed for the frequency domain, for example by fitting flow duration curves. In this section, both the method of different objective functions as well as transforming the observation set with a different frequency domain are chosen as evaluation methods.

The alternative objective function that has been chosen is the Kling-Gupta Efficiency of Gupta et al. (2009) which combines the correlation, bias and ratio of variances in a more balanced way than NSE (Liu, 2020). Due to this increased balance, the KGE is becoming an increasingly used objective function in hydrology instead of NSE (Knoben et al., 2019). Although the NSE already favors high flows due to the squared error it includes, the performance is worse than with KGE calibration for modelling annual peak flows (Mizukami et al., 2019). Another way to focus on high flows is by splitting the NSE computation for the direct hydrograph in days or hours with a flow above or below the mean flow.

Table 6.3: Performance metric results for the direct hydrograph with NSE, KGE and NSE for the days with flows higher (+mean) and lower (-mean) than the mean observed flow. The evaluation periods are the same as the calibration periods.

	NSE		**KGE**		**NSE (+mean)**		**NSE (-mean)**	
	HBV	SBM	HBV	SBM	HBV	SBM	HBV	SBM
Ahr	0.79	0.71	0.81	0.56	0.71	0.43	-2.13	0.41
Vesdre	0.79	0.65	0.83	0.61	0.66	0.26	-1.28	-2.93
Geul	0.67	0.40	0.82	0.65	0.60	0.24	-2.93	-5.61

Table 6.3 provides the results of the evaluation with varying objective functions. The metrics NSE and KGE cannot be compared directly since the benchmark value is not the same. The benchmark value is the value of the objective function if the modelled series only consists of the mean observation. This benchmark value is 0 for NSE and -0.41 for KGE (Knoben et al., 2019). The value for optimal performance is 1 for both metrics. Additionally, there are no clear thresholds for the original KGE metric to determine if the model performance is "good" or "satisfactory". This lack of clear thresholds complicates the interpretation of the KGE and is the main reason why it was not chosen as the main objective function for the calibration. With the KGE as the performance indicator for the direct hydrograph, the HBV model still performs better than the SBM model. In contrast with the NSE value, the Geul models have similar performance as the Ahr and Vesdre models.

From the analysis of the divided hydrograph, it can be seen that the NSE values of the days of observed flows above the mean are higher than for the flows below the mean. This shows that the performance of the model is better for high flows, which is most likely due to the use of the NSE value for calibration and its inclusion of the squared error. This gives more weight to higher flows. The good performance for high flows of the pre-event models is beneficial for this research.

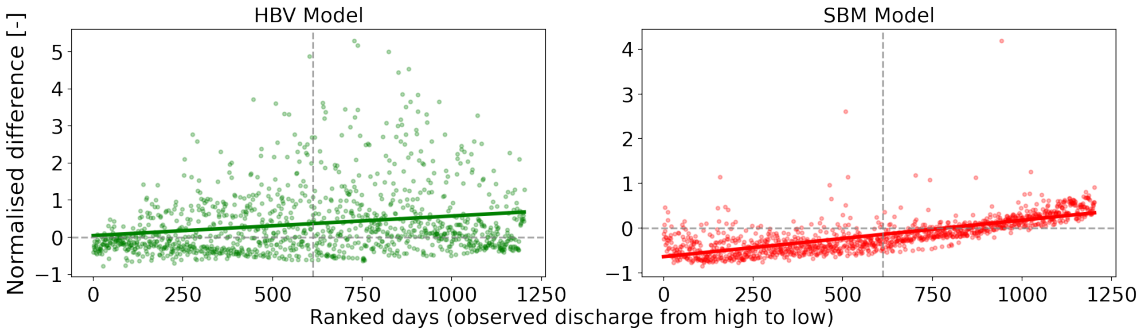


Figure 6.4: Normalised differences between the simulated and observed discharge of the Ahr (Altenahr station) for each day, ranked from high to low discharge values. Both the HBV model (left) and SBM model (right) discharges are shown. The hourly discharges of the simulation have been averaged for each day, the observed discharge are already provided as daily means.

Fig. 6.4 shows the HBV and SBM models of the Ahr and how they overestimate and underestimate different discharge values. The highest discharges are underestimated by both models. As the discharges decrease, the precision of the HBV model decreases and the degree of underestimation decreases for the SBM model. This leads to a trend line which indicates overestimation of the low flows for both models. Both models of the Ahr show an underestimation of the high flows and overestimation of the base flows.

Fig. H.7 and fig. H.8 in appendix H.6 are similar figures for the Vesdre and the Geul. These figures have more scatter, which hinders a similar conclusion as for the Ahr. Neither of the HBV models show any trend. The SBM model of the Vesdre shows an overestimation for all discharge values, but it increases in strength for decreasing discharges. The SBM model of the Geul has an overall underestimation that increases in strength for decreasing discharges. All models have relatively higher factors of overestimation than of underestimation.

Table 6.4: NSE values for the modelled cumulative hydrograph, autocorrelation function and the flow duration curves of appendix H.6.

	Cumulative		**Autocorrelation**		**FDC**	
	HBV	**SBM**	**HBV**	**SBM**	**HBV**	**SBM**
Ahr	0.96	0.43	0.96	0.66	0.95	0.85
Vesdre	0.95	0.51	0.98	0.85	0.95	0.85
Geul	0.98	0.93	0.37	0.97	0.91	0.90

The model performance can also be evaluated by reshaping the discharge data set. Table 6.4 shows the evaluation results for three reshaping methods. The most common method of reshaping discharge data is by considering the cumulative hydrograph. The performance of the HBV model is high, while the performance of the SBM model for the Ahr and Vesdre is on the limit of being good (>0.5). The good performance of the HBV model can be explained by its simplicity and its conservation of the water balance. The manual calibration of the SBM model did not include the cumulative hydrograph which can explain why its performance is lower than for the direct hydrograph. Overall, the performance for simulating the cumulative hydrograph can be considered as sufficient.

The autocorrelation function is a known method to validate output (Barlas, 1990). Here, it is applied for the evaluation. The function calculates the correlation coefficient between the original data set and its copy that shifts with increasing time steps. A standard number of 30 time steps has been applied here. The temporal resolution of the discharge data was decreased to a daily time step, since this is common practice with the autocorrelation function. The resulting autocorrelation functions are visualised in appendix H.6. Table 6.4 indicates that the performance of the HBV model is better than the SBM model for the Ahr and Vesdre. The Geul shows an opposite pattern with a very high performance of the SBM model, and low performance of the HBV model. Overall, the original models are capable of replicating the autocorrelation function to an acceptable level.

The flow duration curve is a similar performance indicator as the autocorrelation function. The flow duration curves based on the long-term observed discharge have already been shown in fig. 2.10. The observed and modelled flow duration curves for the calibration period can be found in appendix H.6. The HBV model has a slightly better performance than the SBM model. In general, the performance of the models to represent the flow duration curve is good.

The various methods of assessing the performance of the discharge simulations show that the original models are capable of reflecting more flow characteristics than only the direct hydrograph. Although this evaluation is not equal to a validation, it can still be concluded that the performance of the original models is sufficient to continue with the proposed method of deriving the flow mechanisms. The evaluation also confirms that the HBV model has an overall better performance than the SBM model.

6.3 Flow mechanisms

Two catchment characteristics of chapter 2, impermeable soils and steep orography, led to the expectation of fast runoff processes. For the 2021 flood event, the fast runoff processes could have been increased by the high amount of precipitation in the preceding period and consequent high soil moisture content. This could even have led to saturation excess overland flow since the water cannot infiltrate in the saturated soils. To derive the flow mechanisms from the models and compare these to our expectations, this section first analyses the HBV model, followed by the SBM model. The catchments are discussed separately, in the usual order of the Ahr, Vesdre and Geul. For each model and catchment, the parameter changes from the re-calibration and their expected effects on the model are discussed first. The internal states are then compared to these effects in order to determine general changes in flow mechanisms. As such, the dominant flow mechanisms for the 2021 flood event can be detected. Finally, their extremity is considered.

HBV Model

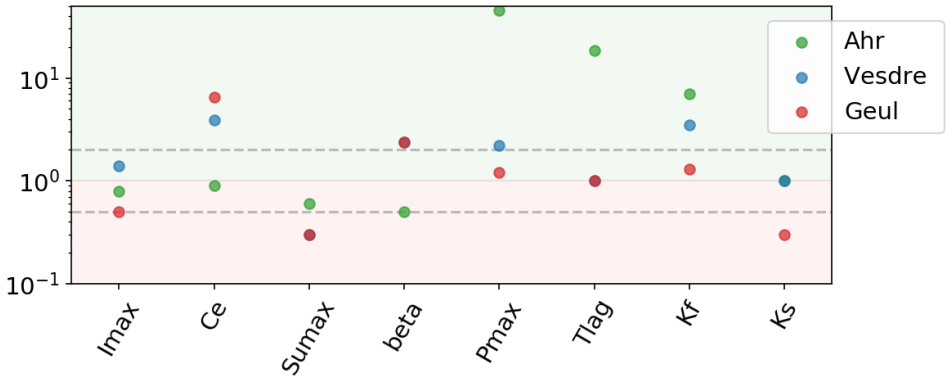
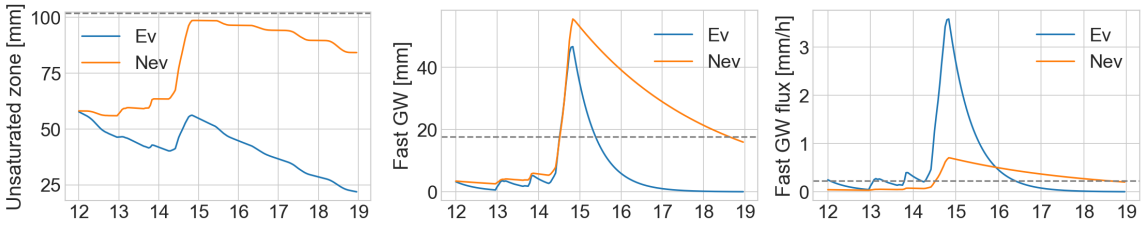


Figure 6.5: Ratios of the re-calibrated to original values of the HBV parameters, so the increasing (green) or decreasing (red) factor after re-calibration. Parameter changes of more than half or double its original value are significant, these limits are indicated by the dashed grey line.

Fig. 6.5 shows how the parameter values have changed after the re-calibration for each catchment. The exact values of the parameters are given in appendix H.4, along with an elaborate explanation of their changes. The likeliness of the parameter values compared to the real world situation are not analysed in this research, only their changes are considered. For the Ahr, there were significant changes in Pmax, Sumax, Tlag and Kf. Fig. 6.5 shows how the change in Pmax is the largest parameter change for any catchment or parameter. This high increase in Pmax, the percolation parameter, increases the drainage from the unsaturated store to the slow responding store. A fast drainage of the unsaturated store will decrease its filling degree. Since this filling degree partitions the infiltration between the unsaturated store and fast responding store, its decrease will also lead to less flow to the fast responding store. A faster recession of the hydrograph can then be expected. The filling degree is also altered by the decrease of Sumax, the maximum capacity of the unsaturated zone and the denominator of the filling degree. The decrease of Sumax leads to an increase of the filling degree. The increase in Tlag will cause a delay of the peak discharge. The large increase of Kf will increase the flux from the fast responding store to the runoff per timestep, thus creating a faster runoff process. Combining all these effects, it is likely that the re-calibration of the Ahr will result in a higher flux from the fast responding store, as well as a delay and faster recession for the hydrograph.



(a) Unsaturated zone *(b) Fast GW* *(c) Fast GW flux*

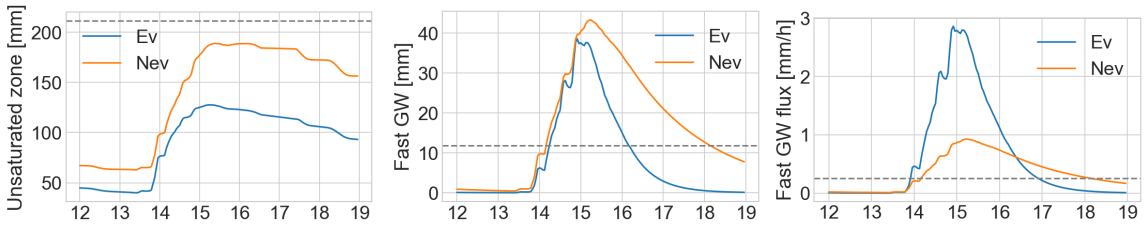
Figure 6.6: Water amounts in the two most relevant stores and the fast responding flux for the original (nev) and re-calibrated model (ev) in the Ahr during the flood event (the day of July is given on the horizontal axis). The dashed grey line shows the maximum value of the pre-event period.

The changes in parameters result in a different response and thus different flow mechanisms. Fig. 6.6 shows the most relevant stores. Other states of the stores and fluxes during the event are shown in fig. I.2 in appendix I.1. The hydrograph modelled with the original parameters strongly underestimated the peak discharge and had a very slow recession. To improve the performance of the model, the automatic calibration led to changes resolving those issues. As predicted from the increase of Pmax, the percolation has increased, leading to more water going to the slow responding store. The consequent faster decrease of the fast responding store is visible in fig. 6.6. Due to the higher Kf, the flux from the fast responding store has strongly increased so it is now higher (with a factor of 5 at the maximum) and with a faster recession. The increase of flow from the slow responding curve has little effect since this

flux contributes only 5% to the peak discharge (due to the smaller horizontal conductivity for the slow responding store). Although Sumax decreased, the maximum filling degree remained equal with a value of 0.7. However, the recession after the maximum peak was much faster. The improved simulation of the hydrograph is due to the higher peak discharge, caused by the higher horizontal conductivity for the fast responding flux, the faster recession, caused by the increased percolation factor, and a better timing of the peak discharge, caused by the delay factor.

By comparing the states during the event to the states in preceding years (2018-2021), which are shown in fig. I.1 in appendix I.1, it can seen that the slow and fast responding stores of groundwater were more full than usual during the 2021 flood event, especially for the summer time. As a consequence, the fluxes to runoff were also much higher, the fast response for the re-calibrated model was almost 16 times higher than the maximum value in the preceding years. The unusually high amount of water and flux from the fast responding store can be seen from the grey lines in fig. 6.6. The magnitudes of several flow mechanisms was unprecedented in those 3 years.

For the HBV model of the Vesdre, fig. 6.5 indicates that the parameters of Pmax and Kf have increased significantly, but to a smaller degree than the Ahr. The increase in percolation and runoff from the fast response store will thus be somewhat lower than for the Ahr. The decrease in Sumax is also significant and will likely lead to more fast response runoff. The parameter β has increased, which dampens the effect of the Sumax decrease on the partitioning to the fast response store. Sumax still has a larger effect here. Another new significant parameter change occurs for Ce, the evapotranspiration parameter. Its increase will likely lead to less evapotranspiration from the unsaturated soil. Without this loss, the filling degree remains higher and so strengthens the effect of the decreased Sumax. Overall, the parameter changes are expected to lead to more runoff from the fast response store.



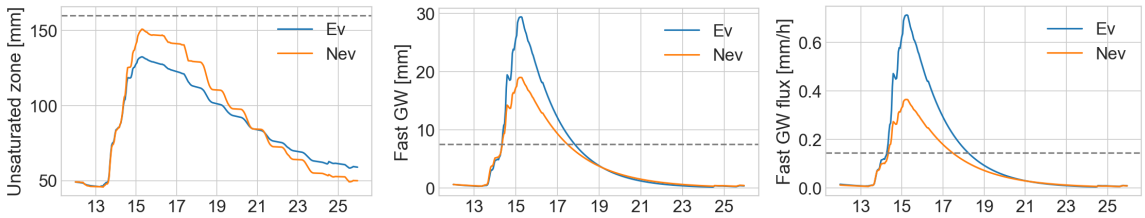
(a) Unsaturated zone *(b) Fast GW* *(c) Fast GW flux*

Figure 6.7: Water amounts in the two most relevant stores and the fast responding flux for the original (nev) and re-calibrated model (ev) in the Vesdre during the flood event. The dashed grey line shows the maximum value of the pre-event period.

The expected behaviour is confirmed by fig. 6.7 and fig. I.4 in appendix I.1 which show a fast response runoff which is higher for the re-calibrated model by a factor of 3. The changes for the re-calibrated model in the event states are similar as for the Ahr. The higher modelled discharge is mainly caused by the horizontal conductivity of the fast responding store and the stronger recession is most likely caused by the increased percolation.

Comparing the event states of the re-calibrated model with the original states in the preceding years (2016-2018) again shows much higher stores and fluxes for the event. The runoff of the fast responding store is 8 times higher than the preceding period. Such high fluxes and stores are in line with the extreme forcing of the flood event, but the increase for the fast responding store is more pronounced.

As for the other models, fig. 6.5 shows that the re-calibration of the Geul resulted in a decrease of Sumax. The increase for β is similar as for the Vesdre. The increase of Ce is bigger than for the Vesdre which will lead to even less evapotranspiration from the unsaturated store. Kf played an important role in the other catchments, and also increases for the Geul, but not significantly. On the other hand, Ks, the horizontal conductivity of the slow response store, has decreased, which will lead to a smaller and slower flux from the slow response store.



(a) Unsaturated zone *(b) Fast GW* *(c) Fast GW flux*

Figure 6.8: Water amounts in the two most relevant stores and the fast responding flux for the original (nev) and re-calibrated model (ev) in the Geul during the flood event. The dashed grey line shows the maximum value of the pre-event period.

The internal states of fig. 6.8 and fig. I.6 in appendix I.1 confirm the changes in behaviour that have been expected from the parameter changes. The evaporation has decreased, and the fast groundwater flux has increased. The required increase of the discharge is mainly caused by the decrease in Sumax and the increase of Kf.

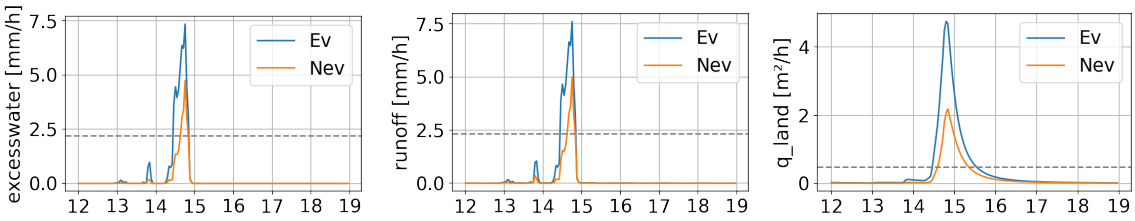
Due to the increase in Ce, the evapotranspiration from the soil is noticeably higher than for normal conditions. Similar amounts of water in the unsaturated store are common for high flow events when looking at the other high flow events for the Geul in preceding years (2019-2021). However, these high amounts usually occurred during February-March. The maximum runoff from the fast responding store is double than any maximum in the preceding years.

In general, the original models did not generate sufficient fast runoff which resulted in a low performance for the flood event. By increasing the horizontal conductivity of the fast responding store for each catchment, the ratio of runoff from the fast response increased considerably. Although all catchments had a decreased maximum capacity of the unsaturated zone and the consequential lower filling degree which would lead to more water in the fast responding store, only the Geul showed an actual increase of water in the fast responding store. Instead, the horizontal conductivity was dominant to increase the runoff of the fast responding flux. The change in other parameters for the Geul re-calibration could point to a slightly different behaviour, which is not surprising considering the catchment characteristics such as the less extreme orography and more permeable soils. The Ahr and Vesdre showed very similar behaviour for the parameter changes and the internal states. The higher flux from the fast response could indicate faster subsurface flows or overland flow in the catchments. This distinction cannot be made with the HBV model.

All catchments show a higher maximum fast response flux for the 2021 flood event than during the previous years, by a factor of 16, 8 and 5 for respectively the Ahr, Vesdre and Geul. However, the discharge of the event was also much higher than previously. To determine if the flow mechanisms are extreme, the relative changes between flow mechanisms should be considered and not only the absolute changes. Table I.1 of appendix I.3 gives the maximum ratios of fast runoff to the total runoff for the event and the previous years. The table shows that a very high ratio of fast runoff is not uncommon, values of up to 0.96 have been reached before. Since the fast runoff already shows the ability to generate almost all runoff, the ratio for the flood event can only be slightly higher, with an increase of 0.2 to 7 %. The absolute flow through the fast response was extreme, but the relative contribution to the total flow during a high flow event was not extreme since high flow events are always mainly generated by the fast responding flux.

SBM Model

The original SBM model of the Ahr underestimated the observed peak discharge. In order to increase the event peak discharge of the model to the observed value, the only sensitive parameter with sufficient influence was Kext. This extinction coefficient influences the canopy gap fraction and as such the amount of interception. Kext is originally defined as a spatially varying parameter with values between 0.6 and 0.8, which is the usual range (Imhoff et al., 2020). For the re-calibration, a factor of 0.1 has been applied to this map. Such a strong decrease can be considered as unrealistic, but it gave the best performance. It could also be an indication of an error in the Leaf Area Index, which is in the same equation as Kext. With Kext decreasing, the canopy gap fraction approaches 1 due to the exponential function, eq. (G.10). Such a high canopy gap fraction imitates the lack of vegetation and thus strongly reduces the interception.



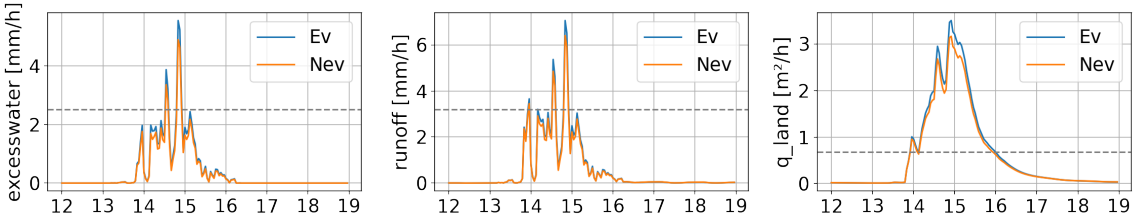
(a) excesswater *(b) runoff* *(c) q_land*

Figure 6.9: Three relevant fluxes for the original (ev) and re-calibrated model (nev) in the Ahr during the flood event. The dashed grey line shows the maximum value of the pre-event period.

Similar figures of the states as for the HBV model have been made for the SBM model and can be found in fig. 6.9 and fig. I.8 in appendix I.2. The SBM model is more complex and has more stores and fluxes. As could be expected from the reduced interception, the actual total evapotranspiration (actevap) has decreased slightly for the re-calibration. The most noticeable change is the increase of water that cannot infiltrate due to saturation (excesswater). The parameter changes also increase the total amount of water that cannot infiltrate (runoff) and consequently increases the overland runoff (h_land and q_land). Due to the minimal interception, which forms a buffer between the precipitation and the soil, more water reaches the soil faster which may lead to saturation excess. The higher peak discharge of the re-calibrated model is facilitated caused by the overland flow from saturation excess.

In fig. I.7 in appendix I.2, the states of the preceding years can be seen as a supplement to the grey lines in fig. 6.9. Due to the re-calibration, the excesswater flux is much higher than usual, as well as the consequential runoff and h_land and q_land. The amount of overland flow and saturation excess is unprecedented. Also, the amount of water in the saturated zone is unusually high for July, which is also the case for the unsaturated zone. This can indicate that the soil was indeed quite saturated before and during the event.

The re-calibration of the Vesdre catchment resulted in only one parameter change, the KsatHorFrac, which influences the horizontal subsurface flow. A higher value resembles a higher horizontal flow capacity. The original spatial distribution with values of 250 and 1000 has been replaced by a constant value of 200. Since most of the area had been identified as a KsatHorFrac of 250, this difference is only small.



(a) excesswater *(b) runoff* *(c) q_land*

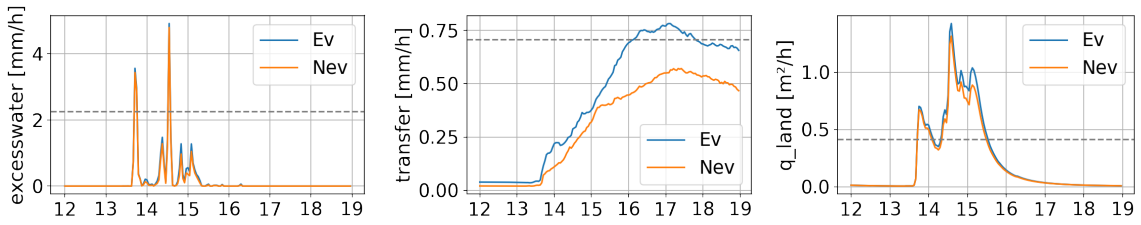
Figure 6.10: Three relevant fluxes for the original (ev) and re-calibrated model (nev) in the Vesdre during the flood event. The dashed grey line shows the maximum value of the pre-event period.

The performance of the original model was already good for the event, as could be seen from table 6.2 and fig. 6.2. As a result of the small change for calibration, fig. 6.10 visualises that the effects on the states are also small. The decrease of horizontal flow capacity can be interpreted as less permeability and since the soil type does not change for the event, this could be due to saturated soil. The high total evapotranspiration was not unusual, since the event occurred in the middle of summer.

The saturation excess was unusually high, even considering the higher amount of precipitation by accounting for their ratio. The amount of water in the unsaturated zone was higher than usual for the summer, but that was not the case for the saturated zone. Nevertheless, the high runoff caused exceptional overland flow.

Re-calibrating the Geul was complicated due to the time difference of the peak discharges. None of the mentioned sensitive parameters of Imhoff et al. (2020) caused the required shift in time for replicating the estimated hydrograph. Appendix H.7 elaborates more on this delay of the observation compared to the

model and discusses the two parameters, n and n_{river}, that were able to shift the modelled hydrograph in time. The parameter n_{river} preserved the hydrograph the most. This parameter reflects the Manning's roughness coefficient of the river cells. The parameter n reflects the roughness as well, but for the land cells. Strongly increasing n_{river} shifted the hydrograph to match the observation. To counteract the resulting decrease in peak discharge, the Ksathorfrac parameter was decreased by a factor of roughly 5. The decrease in Ksathorfrac and thus horizontal flow conductivity likely resembles less permeability and horizontal flow in the subsurface.



(a) excesswater *(b) transfer* *(c) q_land*

Figure 6.11: Three relevant fluxes for the original (ev) and re-calibrated model (nev) in the Geul during the flood event. The dashed grey line shows the maximum value of the pre-event period.

The changes in parameters for the re-calibration have little influence on most internal states of the SBM model. The biggest difference is an increase in transfer, which is the flux between the unsaturated and the saturated zone. As a result, the water in the unsaturated zone, ustoredepth, decreases and the saturated zone, satwaterdepth, increases for the re-calibration. There is also a minor increase in overland water levels and flow, h_land and q_land.

Most of the internal states reach high values, but are within the limits of occurrence in the previous years. As for the other catchments, the saturation excess, excesswater, was almost twice the previous maximum. Also the overland water level and flow are much higher than in the preceding period. This is also the case for the amount of water in the unsaturated zone. This shows that there was more overland flow than usual and more water stored in the unsaturated zone, and thus likely more fluxes throughout this layer.

For the HBV models, the relative contributions to the runoff generation have been considered to see differences in partitioning the flow between mechanisms. The same principle can be applied to the SBM model. The two horizontal fluxes of land cells are the overland flow and subsurface flow from the saturated soil layer. Table I.2 in appendix I.3 gives the ratio of the average overland flow flux to the sum of the average overland and subsurface flux for each catchment and model configuration. Although the maximum ratios are low, they increase for the 2021 flood event with a factor of 5 to 12. The overland flow is not only higher in absolute terms, but it is also considerably higher in relative terms.

For all catchments, there was a strong increase in overland waterlevel and flow for the re-calibration or a much higher peak than in the previous years, which highlights the importance of high overland flow during the flood event. The increased and high parameter values for the saturation excess show that the overland flow was likely caused by the high amount of water that could not infiltrate into the soil anymore due to saturation. This hypothesis is supported by the high amounts of water in the unsaturated zone compared to previous summers. The changes to obtain these high(er) values in overland flow and saturation excess were caused by different parameters, which could be due to the manual calibration process for the SBM model. All re-calibrated SBM models point to a high overland flow due to saturation excess.

6.4 Discussion

Both models used in this analysis have a different model structure to increase the robustness of the findings. Overall, the simple and lumped HBV model outperforms the more complex and spatially distributed model. Almost all original models had a low performance for the 2021 flood event because they underestimated the discharge. Although the structure of the HBV and SBM model differs considerably, they show similar behaviour for the re-calibration. Both models increased their modelled discharge by increasing the fluxes from fast responsive mechanisms. For the SBM model, this flux was the overland

flow due to saturation excess. For the HBV model there is no explicit overland flow flux, so fast subsurface flows could also be a possibility. Regardless of the accounting for overland flow, the decrease of Sumax and resulting higher filling degree reflects a highly saturated soil. Since the higher saturation was in an important factor for the re-calibration of both models, the high saturation is the most likely driver of the fast runoff processes. The high precipitation caused exceptional values for multiple states during the event compared to regular flows. The high exceptional states of the overland flow, fast response flux and high saturation were more extreme due to the re-calibration. The extremity and change in parameters for these processes could indicate that those fluxes and states were more exceptional for the 2021 flood event than usual.

However, there are some limitations to the derivation of the flow mechanisms in this study. Firstly, the use of two models increases the robustness but also complicates the interpretation of flow paths due to different model structure. The HBV model has a limited number of included processes. On the other hand, the SBM model is so complex that calibration has to be done manually which makes the calibration and thus the findings more subjective.

Secondly, the input data is only of a relatively short time period. Preparing the data as input of both models is work intensive and the availability of good quality data over a large area but with a small resolution is limited. Due to this short data record no validation is performed, nor is there the ability to calibrate the original model for many previous high flow events of a longer period. The exceptional flow paths for the 2021 flood event are thus only unique compared to the previous 3 to 5 years and not to preceding 15-70 years as was the case for the analysis of the forcing and response. However, longer data periods would also increase the already long run time of the model. The long run time, in combination with the work intensive calibration of the SBM model, is also the reason why only the main gauges are considered even though data on more gauges is known. This would provide more spatially detailed information on the catchments. The short input data also resulted in a warming-up period of 2 months for the Ahr and Geul. Fig. H.4 and fig. H.6 in appendix H.5 show that this is sufficient for the HBV model and SBM model, except for the Geul SBM, which would have required a warming-up period of at least 3 months. Although the warming-up periods are short, they are unlikely to have caused large errors.

Additionally, only one performance indicator is used for one characteristic: the direct hydrograph. This makes the calibration overall less robust. For the pre-event period, the use of the entire hydrograph leads to models that perform well for both low and high flow. There is no specific focus on high flow events for the pre-event calibration and the model performance of section 6.1 already showed that the peak estimates were not accurately simulated, mostly underestimated. The derived unusual flow mechanisms for the 2021 flood event could have been present for previous high flow events, but they remained undetected due to the poor simulation of these previous events. Using another characteristic than the direct hydrograph would have been complicated due to the lack of data during the event and its short time period. Other performance indicators could have been used for the calibration in theory. Due to the uncertain hydrograph estimates for the flood event, the performance indicator for the event already has some uncertainty. Combining more indicators would not decrease this uncertainty significantly.

Finally, the data set period is quite short for the Ahr and the Geul. The second sub-question of this research, to identify extreme aspects of the forcing and response, was answered by comparing the 2021 flood event and the long term data preceding the flood event. However, for the third sub-question, to analyse the extremity of the flow mechanisms, the data periods are around 3 years, which is considerably shorter than the data sets of the data analysis. The short inputs for the models could lead to unidentified flow mechanisms in previous years that actually showed similar flow mechanisms as the 2021 flood events. The conclusion that the flow mechanisms of the 2021 flood event were rare is only valid for the several years before the flood event.

Chapter 7

Conclusion and recommendations

7.1 Conclusion

The aim of this research was to analyse which hydrological aspects of the flood event of July 13-16 2021 were extreme in the catchments of the Ahr, Vesdre and Geul. These aspects were divided into the following categories: forcing, response and flow mechanisms. Forcing and response have been studied with a data analysis and flow mechanisms have been studied with models. The data analysis showed that the origin of the precipitation was extreme, as well as the return period of the 24 hour precipitation sum on July 14, the day of the event with the maximum amount of precipitation. The 30 days preceding July 14 also received extreme precipitation amounts. Such extreme forcing led to a response with extreme peak discharges. Although the response of the catchment was fast, the flood event did not qualify as a flash flood, which would have been an unseen flood type for the region. The models required a re-calibration to simulate the extreme discharges, leading to extreme fluxes and patterns in the flow mechanisms. The temporal and spatial variability of the precipitation were not extreme. The majority of the analysed hydrological aspects were extreme, which indicates that the 2021 flood event can be considered as an extreme hydrological event.

This main conclusion is the result of answering three sub-questions. The first sub-question led to a description of the forcing and response during the 2021 flood event. The stationary cold core low "Bernd" caused a precipitation event of multiple days and varying intensity which resulted in high precipitation totals, from 100 to almost 300 $[mm/5d]$. Such amounts make up 15 to 25% of the annual precipitation. The resulting high flows disabled many of the gauges, so the knowledge on the response is based on estimates. These estimates focus on the most downstream gauges and the peak discharge. For those gauges of the Ahr, Vesdre and Geul the peak discharges are estimated to be around 990, 600 and 90 $[m^3/s]$, by respectively the LFU (Landesamt für Umwelt: State Agency for Environment), (Ziemetz et al., 2021), the WL (Waterschap Limburg: Water Board of Limburg). This ranking is different from the ranking according to the total precipitation volume, since the Vesdre received more precipitation than the Ahr. The different hydrograph shapes for the three catchments are reflected by their runoff coefficients. The runoff is relatively higher and faster for the Ahr, resulting in sharper and higher hydrograph shapes compared to the Vesdre and Geul, . The main finding of the event description is the high, but spatially varying magnitude of the forcing and response.

The second sub-question concerns how extreme several aspects of the described forcing and response were. The extremity of the origin has not been quantified due to the complexity of the relevant atmospheric processes, but it can be considered as a rare event due to the combination of uncommon contributing factors, such as the large influx of humid air due to anomalies of sea surface, troughs increasing convection, orographic lifting and blocking. The resulting precipitation amount of the event was extreme. Analysing the precipitation amount of July 14 showed return periods of over 1000 years in multiple locations. Some locations showed much lower return periods such as 1 year due to the spatial variability. The spatial pattern of the precipitation during the entire event showed a positive correlation of medium strength with the climatic pattern, which means that the pattern can not be considered as extreme. The temporal variation was also not extreme since the duration for accumulating 90% of the event precipitation was not within the 90th percentile compared to other high precipitation events. A potentially aggravating

aspect of the flood event were the extreme pre-conditions, since the precipitation sums of the 30 days preceding July 14 of most gauges were in the 90th percentile compared to those of other events with high precipitation (over 30 [mm/d]). This caused a high soil moisture content before the event. The forcing was generally extreme, except for the spatial and temporal variability.

Extreme flows resulted from the extreme forcing. The estimates of the peak discharges give higher values than have ever been recorded. These peak discharges resulted in extreme return periods of over 1000 years at most locations. Historical reconstructions show that peak discharges of similar magnitude have occurred before in the Ahr. Another insight from historical documentation is that floods during the summer half year are not rare for any of the catchments.

Multiple reports (Ziemetz et al., 2021; TFFF, 2021) refer to this flood as a flash flood due to the sudden water level rise. However, the flood does not exactly match one flood type of Merz & Blöschl (2003). The lag times of 13 to 24 hours could indicate a flash flood if a broad definition is applied. However, considering the large temporal scale (a precipitation event of multiple days) and spatial scale (large, international zone of flooded areas), a more appropriate typology of this event would be a long rain event with a quick runoff response.

The third sub-question considered if the flow mechanisms were extreme and offered various interesting insights. The two models, a simple, lumped model with automatic calibration and a more complex, distributed model with manual calibration, were able to simulate the discharge well in the years before the event (NSE values of -0.17 to 0.40, with an outlier of 0.90). However, using the same parameters for the flood event resulted in a poor performance due to a strong underestimation of the peak discharge (by a factor of 2 to 5). Calibrated parameter changes that increased the fast response mechanisms, such as subsurface and overland flow, resulted in a good performance for the flood event (NSE values of 0.65 and 0.99), mostly by increasing the peak discharge. This improved performance shows that the model structures were capable of simulating the discharge of the 2021 flood events but the internal model processes functioned slightly different.

Which flow mechanisms were dominant and different for previous flows have been determined by analysing the re-calibrated parameters and the internal states of all models. Although there is a risk of equifinality due to large amount of parameters, most re-calibrated parameters enhance the fast runoff processes. The internal states reflect that as well, since the fast runoff processes have increased due to the re-calibration. The fluxes of the fast runoff processes were larger than the preceding years, as well as their relative ratios. The required re-calibration, the change in parameters and the internal states indicate that the fast flows, especially overland flow, were of rare intensity and contributed to the extreme flow.

The goal of answering the research question was to place the new understanding of the flood event in context to take appropriate measures and avoid future damage. The extreme magnitude of the forcing, response and flow mechanisms show that the catchments are capable of a fast runoff response of extreme magnitude under extreme circumstances such as high precipitation amounts during and preceding the event. Since the magnitude of the forcing, in this case the precipitation amount, was extreme, the likelihood of a similar forcing and thus flood event is low. Nevertheless, the fast and strong response capacity of the catchments could still generate (smaller) floods in case of less extreme forcing.

7.2 Recommendations

This study has a broad scope and therefore required an exploratory approach. Assessing many aspects of the flood event to get an overall understanding limits the level of detail. This is the case for the data analysis and the models. How and which aspects of this research would be interesting for further research is described in the next section. This research should be seen as a stepping stone to further research before using it in policy plans.

It would be valuable to vary the spatial focus of similar research. Focusing on one catchment allows a more detailed analysis, for example about the subcatchments. This would be especially helpful when developing specific measures or policy plans for that catchment. It could also give more insight in the flood typology since smaller catchments have shorter response times and thus could generate flash floods. This detailed analysis could show that the flood typology of the entire event was of a hybrid nature. Making an explicit comparison between catchments, which was not the goal of this research, could also provide useful insights such as specific behaviour from a specific characteristic. The influence of the

steeper hills of the Ahr catchment or the chalk soils in the Geul would be good examples. Comparative studies broaden the general understanding and could show which measures would work in one area, but not in the other. Some aspects of the flood event, such as the precipitation, could benefit from an analysis with a regional focus instead of specific catchments. Whereas the spatial variability of the precipitation in this study was only considered for gauges located in or closely to the catchments, including the entire region would provide a more complete insight. Such research could include a study on the influence of the orography on the precipitation pattern. There are still many research opportunities by varying the spatial scale.

Anthropogenic influences formed one of those differences between catchments that complicated the analysis. The influence of land cover has not been considered explicitly as a controlling factor, although the catchments characteristics have shown that they vary considerably. Another anthropogenic influence that has been touched upon shortly are the water management measures. Especially large reservoirs can have a considerable effect, such as for the Vesdre. Specific research into how these reservoirs have affected the flood event could show their potential in flood damage reduction or show opportunities to optimize their management for flood risk purposes.

In this study, the models were used as a supporting method to answer one of the sub-questions. Developing models for a higher performance in flow or even flood area simulation would be vital for a potential warning system or taking structural measures. The goal of any model must always be considered carefully. The simple lumped model outperformed the more complex distributed model, but the distributed model could be used for experimental modelling or testing hypotheses due to its inclusion of many spatially varying factors. An interesting use would be to assess the effect of summer vegetation by varying the roughness and see if a similar forcing would have had a similar response during winter. A study focused on models would benefit from a longer dataset to include more weather and flow variability. Even more important, validation would be possible with a longer observation set. In that case it might be useful to replace the multi-variate evaluation of this study with a detailed multi-variate calibration and validation, fit specifically to the model purpose. Although the models of this research have proven their capacity and value, they have even more potential.

With lag times that almost approach flash flood characteristics, the catchments have shown potential for a fast runoff response. In such cases it is important to set up a warning system for local communities as there is little time to evacuate or take other measures. Flood wave velocity has not been analysed explicitly but would be an important factor in developing warning systems. The delay of the observations of the Geul compared to the models indicate that delays can have a large impact on the course of a flood event and require further research. A more hydraulic approach would be useful to assess the effects of wider flood planes, obstacles such as bridges and increased roughness on the flood wave velocity. Due to flood wave attenuation, these factors would also change the downstream hydrograph. A more hydraulic approach for the flood wave can offer new and helpful insights.

Extreme events are rare by definition and so is their data availability. The data sparsity is increased by possible damage to measurement infrastructure. The lack of data sometimes requires the addition of estimates. Sparse data and estimates are a source of uncertainty due to the small sample and possible measurement or estimate errors. Uncertainty has not been treated in detail in this study, but it is recommended for future research. Firstly, the estimates of the peak discharges should be subjected to a critical analysis which uses multiple methods to verify these values. Finally, how we deal with these potential input errors is important. Different methods of estimating the return periods and quantifying extremity could give different results and show added uncertainty throughout this research. Error propagation is therefore worth analysing. The foremost recommendation to decrease the uncertainty for the future is to add more robust measurement devices to decrease the uncertainty of the input as a first step.

This study focused on international catchments whereas most studies only focus on catchments in one country. Despite the language barriers and more complicated data collection, the international approach shows how everyone can benefit from shared knowledge. As an example, the Ahr could benefit from knowing why the Geul had a more damped response and if they could mimic this effect. Alternatively, the Vesdre could learn from the importance of reservoirs compared to the Ahr. A more intense cooperation would create benefits for all responsible authorities. This cooperation should not only be international but also interdisciplinary. It is valuable to learn from various fields such as academic research, practical engineering, policy, social interests and involve all parties when working towards a more flood resilient region. Hopefully this research forms a first step towards a flood resilient society.

Bibliography

Abrahams, A. D., Li, G., & Parsons, A. J. (1996). Rill hydraulics on a semiarid hillslope, southern Arizona. *Earth surface processes and landforms*, *21*(1), 35–47.

Ahmed, N. (1995). Estimating soil moisture from 6· 6 ghz dual polarization, and/or satellite derived vegetation index. *International Journal of Remote Sensing*, *16*(4), 687–708.

Ailliot, P., Thompson, C., & Thomson, P. (2011). Mixed methods for fitting the GEV distribution. *Water Resources Research*, *47*(5).

Alfieri, L., Burek, P., Feyen, L., & Forzieri, G. (2015). Global warming increases the frequency of river floods in Europe. *Hydrology and Earth System Sciences*, *19*(5), 2247–2260.

Amponsah, W., Ayral, P.-A., Boudevillain, B., Bouvier, C., Braud, I., Brunet, P., ... Lebouc, L. (2018). Integrated high-resolution dataset of high-intensity European and Mediterranean flash floods. *Earth System Science Data*, *10*(4), 1783–1794.

Arcement, G. J., & Schneider, V. R. (1989). *Guide for selecting Manning's roughness coefficients for natural channels and flood plains.* US Government Printing Office Washington, DC.

Archambeau, P., Erpicum, S., Dewals, B., & Pirotton, M. (2021). *Première reconstruction des valeurs hydrologiques associées aux évènements de juillet 2021 dans la vallée de la Vesdre.*

Asselman, N., van Heeringen, K.-J., Overeem, A., Beersma, J., & Philip, S. (2022). *Watersysteemevaluatie Waterschap Limburg: Analyse overstroming Valkenburg.* Deltares.

Bachmann, D. (2021, 09). *Flood event 2021 in Germany: More research into practice!*

Bačová Mitková, V., Pekárová, P., Halmová, D., & Miklánek, P. (2018). Reconstruction and post-event analysis of a flash flood in a small ungauged basin: a case study in Slovak territory. *Natural hazards*, *92*(2), 741–760.

Barlas, Y. (1990). An autocorrelation function test for output validation. *SIMULATION*, *55*(1), 7-16. doi: 10.1177/003754979005500104

Barredo, J. I. (2007). Major flood disasters in Europe: 1950–2005. *Natural Hazards*, *42*(1), 125–148.

Barredo, J. I. (2009). Normalised flood losses in Europe: 1970–2006. *Natural hazards and earth system sciences*, *9*(1), 97–104.

Bartholomeus, V. (2021). *Zo raasde het water door Limburg.*

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*(1-2), 33–45.

Berger, H. E. J. (1992). *Flow forecasting for the river Meuse.*

Berghuijs, W., Woods, R., & Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature climate change*, *4*(7), 583–586.

Bergström, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments.*

Beven, K. (1989). Changing ideas in hydrology—the case of physically-based models. *Journal of hydrology*, *105*(1-2), 157–172.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in water resources*, *16*(1), 41–51.

Bezak, N., Šraj, M., & Mikoš, M. (2016). Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. *Journal of Hydrology*, *541*, 272–284.

Bhunya, P., Panda, S., & Goel, M. (2011). Synthetic unit hydrograph methods: a critical review. *The Open Hydrology Journal*, *5*(1).

Blanchard, B. J., McFarland, M. J., Schmugge, T. J., & Rhoades, E. (1981). Estimation of soil moisture with API algorithms and microwave emission 1. *JAWRA Journal of the American Water Resources Association*, *17*(5), 767–774.

Blöschl, G., Ardoin-Bardin, S., Bonell, M., Dorninger, M., Goodrich, D., Gutknecht, D., ... Szolgay, J. (2007). At what scales do climate variability and land cover change impact on flooding and low flows? *Hydrological Processes*, *21*(9), 1241–1247.

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.

Borga, M., Gaume, E., Creutin, J. D., & Marchi, L. (2008). Surveying flash floods: gauging the ungauged extremes. *Hydrological processes*, *22*(18), 3883.

Bormann, H., & Klaassen, K. (2008). Seasonal and land use dependent variability of soil hydraulic and soil hydrological properties of two Northern German soils. *Geoderma*, *145*(3), 295-302. (Modelling Pedogenesis) doi: https://doi.org/10.1016/j.geoderma.2008.03.017

Braud, I., Ayral, P.-A., Bouvier, C., Branger, F., Delrieu, G., Coz, J. L., ... Adamovic, M. (2014). Multi-scale hydrometeorological observation and modelling for flash flood understanding. *Hydrology and Earth System Sciences*, *18*(9), 3733–3761.

Brázdil, R., Kundzewicz, Z. W., Benito, G., Demarée, G., Macdonald, N., & Roald, L. A. (2012). Historical floods in Europe in the past millennium. *Changes in Flood Risk in Europe, edited by: Kundzewicz, ZW, IAHS Press, Wallingford*, 121–166.

Brisson, E., Demuzere, M., Kwakernaak, B., & Van Lipzig, N. (2011). Relations between atmospheric circulation and precipitation in Belgium. *Meteorology and atmospheric physics*, *111*(1), 27–39.

Brown, A. E., McMahon, T. A., Podger, G. M., & Zhang, L. (2006). A methodology to predict the impact of changes in forest cover on flow duration curves. *Science Report*, *8*(06).

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A.-C., & Seibert, J. (2017). Flood type specific construction of synthetic design hydrographs. *Water Resources Research*, *53*(2), 1390–1406.

Bryndal, T. (2015). Local flash floods in Central Europe: A case study of Poland. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography*, *69*(5), 288–298.

Budyko, M. I. (1974). *Climate and life*. Academic press, New York.

Camus, P., Haigh, I. D., Wahl, T., & Nasr, A. (2020). Potential compound flooding due to pluvial, fluvial and coastal drivers along Eastern North Atlantic and Mediterranean coasts: sensitivity analysis and spatial patterns. In *Agu fall meeting abstracts* (Vol. 2020, pp. NH024–05).

Castro, A. L. C. d. (1996). *Manual de desastres: desastres naturais*.

Chow, V. T. (1959). Open-channel hydraulics. *McGraw-Hill civil engineering series*.

Chow, V. T., Maidment, D. R., & Larry, W. (1988). Applied hydrology. *International edition, MacGraw-Hill, Inc*, *149*.

Ciervo, F., Papa, M., Medina, V., & Bateman, A. (2015). Simulation of flash floods in ungauged basins using post-event surveys and numerical modelling. *Journal of Flood Risk Management*, *8*(4), 343–355.

Contrat de rivière de Vesdre. (2022). *Le bassin de la Vesdre en quelques chiffres.* Retrieved from `https://www.crvesdre.be/territoire/le-bassin-de-la-vesdre`

Copernicus: Emergency Management Service. (2021). *The Copernicus Emergency Management Service forecasts, notifies, and monitors devastating floods in Germany, Netherlands, Belgium and Switzerland.* Retrieved from `https://emergency.copernicus.eu/mapping/ems/copernicus-emergency-management-service-forecasts-notifies-and-monitors-devastating-floods`

Cornwall, W. (2021, July, 20). Europe's deadly floods leave scientists stunned. *Science*. Retrieved 23-11-2021, from `https://www-science-org.tudelft.idm.oclc.org/content/article/europe-s-deadly-floods-leave-scientists-stunned`

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., & Smith, P. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water resources research*, *51*(7), 5531–5546.

de Almeida, G. A., Bates, P., Freer, J. E., & Souvignet, M. (2012). Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. *Water Resources Research*, *48*(5).

De Kraker, A. (2013). Storminess in the low countries, 1390-1725. *Environment and History*, *19*(2), 149–171.

Delforge, P. (2021). *Comment faire barrage à une montée des eaux violente et soudaine ?* Retrieved from `https://www.institut-destree.eu/wa_files/2021-08-10_paul-delforge_barrage-vesdre.pdf`

Deltares. (2019). *The wflow_sbm model.* Retrieved from `https://wflow.readthedocs.io/en/latest/wflow_sbm_old.html`

Diakakis, M., Deligiannakis, G., Andreadakis, E., Katsetsiadou, K. N., Spyrou, N. I., & Gogou, M. E. (2020). How different surrounding environments influence the characteristics of flash flood-mortality: The case of the 2017 extreme flood in Mandra, Greece. *Journal of Flood Risk Management*, *13*(3), e12613.

Di Baldassarre, G., & Claps, P. (2011). A hydraulic study on the applicability of flood rating curves. *Hydrology Research*, *42*(1), 10–19.

Direction des Cours d'eau non navigables. (2022). *Réseau de surveillance hydrométrique: Aqualim station : L6660 (Sippenaeken).* Retrieved from `http://aqualim.environnement.wallonie.be/Station.do?method=ficheStation&isMobile=false&time=1655667532678`

Directive 2007/60/EC of the European Parliament and of the Council on the assessment and management of flood risks. (2007).
Official Journal, L288, 27-34.

Douinot, A., Roux, H., Garambois, P.-A., Larnier, K., Labat, D., & Dartus, D. (2016). Accounting for rainfall systematic spatial variability in flash flood forecasting. *Journal of Hydrology*, *541*, 359–370.

Eeckman, J., Roux, H., Douinot, A., Bonan, B., & Albergel, C. (2021). A multi-sourced assessment of the spatiotemporal dynamics of soil moisture in the MARINE flash flood model. *Hydrology and Earth System Sciences*, *25*(3), 1425–1446.

Einfalt, T., Hatzfeld, F., Wagner, A., Seltmann, J., Castro, D., & Frerichs, S. (2009). URBAS: forecasting and management of flash floods in urban areas. *Urban Water Journal*, *6*(5), 369–374.

Eisenbies, M., Aust, W., Burger, J., & Adams, M. B. (2007). Forest operations, extreme flooding events, and considerations for hydrologic modeling in the Appalachians—a review. *Forest Ecology and Management*, *242*(2-3), 77–98.

Engeland, K., Hisdal, H., & Frigessi, A. (2004). Practical extreme value modelling of hydrological floods and droughts: a case study. *Extremes*, *7*(1), 5–30.

Fang, X., Thompson, D. B., Cleveland, T. G., Pradhan, P., & Malla, R. (2008). Time of concentration estimated using watershed parameters determined by automated and manual methods. *Journal of Irrigation and Drainage Engineering*, *134*(2), 202–211.

Feddes, R., & Zaradny, H. (1978). Model for simulating soil-water content considering evapotranspiration—comments. *Journal of Hydrology*, *37*(3-4), 393–397.

FEMA. (1981). *Design guidelines for flood damage reduction.* Federal Emergency Management Agency: Washington, DC; 102. Retrieved from `https://dnr.nebraska.gov/sites/dnr.nebraska.gov/files/doc/general-nfip-resources/fema-publications/5FEMA15.pdf`

Fu, J., & Wang, W. (2019). On the lower bound of Budyko curve: The influence of precipitation seasonality. *Journal of Hydrology*, *570*, 292–303.

Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., & Blöschl, G. (2012). Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research*, *48*(4).

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., ... Dumitrescu, A. (2009). A compilation of data on European flash floods. *Journal of Hydrology*, *367*(1-2), 70–78.

Georgakakos, K. P. (1986). On the design of national, real-time warning systems with capability for site-specific, flash-flood forecasts. *Bulletin of the American Meteorological Society*, *67*(10), 1233–1239.

Grayson, R. B., Moore, I. D., & Mcmahon, T. A. (1992). Physically based hydrologic modeling: 1. a terrain-based model for investigative purposes. *Water resources research*, *28*(10), 2639–2658.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, *377*(1-2), 80–91.

Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—i. basic concept. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(3), 219–233.

Heggen, R. J. (2001). Normalized antecedent precipitation index. *Journal of hydrologic Engineering*, *6*(5), 377–381.

Hirpa, F. A., Gebremichael, M., & Over, T. M. (2010). River flow fluctuation analysis: Effect of watershed area. *Water Resources Research*, *46*(12).

Hosking, J. R. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(1), 105–124.

Hosking, J. R. M., Wallis, J. R., & Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, *27*(3), 251–261.

Huang, S., Kumar, R., Flörke, M., Yang, T., Hundecha, Y., Kraft, P., ... Lobanova, A. (2017). Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide. *Climatic Change*, *141*(3), 381–397.

Imhoff, R., Van Verseveld, W., Van Osnabrugge, B., & Weerts, A. (2020). Scaling point-scale (pedo) transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modeling: An example for the Rhine river. *Water Resources Research*, *56*(4), e2019WR026807.

International Meuse Commission. (2005). International river basin district Meuse–characteristics, review of the environmental impact of human activity. *Economic Analysis of Water Use, International Meuse Commission, Belgium*.

IPCC. (2018). *Annex i: Glossary [matthews, j.b.r. (ed.)]. in: Global warming of 1.5°c. an ipcc special report on the impacts of global warming of 1.5°c above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.* ([Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 541-562, doi:10.1017/9781009157940.008.)

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, *81*(348), 158–171.

Johst, M., & Demuth, N. (2021). *The July flood in Germany with focus on the Ahr basin and challenges in forecasting.* Retrieved from `https://www.efas.eu/sites/default/files/2021-10/20211028_EFAS_AhrFlood_Johst.pdf`

Jonkman, S., Vrijling, J., & Vrouwenvelder, A. (2008). Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method. *Natural Hazards*, *46*(3), 353–389.

Jordans, F., & Casert, R. (2022). *'unimaginable': Germany, Belgium remember deadly 2021 floods.* Retrieved from `https://www.bloomberg.com/news/articles/2022-07-14/germany-belgium-remember-deadly-floods-1-year-later`

Jorgensen, D. W., & Gardner, T. W. (1987). Infiltration capacity of disturbed soils: Temporal change and lithologic control. *JAWRA Journal of the American Water Resources Association*, *23*(6), 1161-1172. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.1987.tb00868.x` doi: https://doi.org/10.1111/j.1752-1688.1987.tb00868.x

Junghänel, T., Bissolli, P., Daßler, J., Fleckenstein, R., Imbery, F., Janssen, W., ... Weigl, E. (2021). *Hydro-klimatologische Einordnung der Stark- und Dauerniederschläge in Teilen Deutschlands im Zusammenhang mit dem Tiefdruckgebiet „Bernd“ vom 12. bis 19. Juli 2021.* Deutscher Wetterdienst.

Kastridis, A., Kirkenidis, C., & Sapountzis, M. (2020). An integrated approach of flash flood analysis in ungauged Mediterranean watersheds using post-flood surveys and unmanned aerial vehicles. *Hydrological Processes*, *34*(25), 4920–4939.

Kess, D. (1986). *Hochwasser hinterläßt Schäden im Mündungsgebiet der Ahr.* Retrieved from `https://relaunch.kreis-ahrweiler.de/kvar/`

Kess, D. (1989). *Ahr und Rhein traten über die Ufer.* Retrieved from `https://relaunch.kreis-ahrweiler.de/kvar/`

Kirpich, Z. (1940). Time of concentration of small agricultural watersheds. *Civil engineering*, *10*(6), 362.

Klein, A. (2022). *Hydrological response of the Geul catchment to the rainfall in July 2021.*

KMI. (2021). *Eerste cijfers en duiding bij de hevige neerslag van 14 en 15 juli.* Retrieved 29-11-2021, from `https://www.meteo.be/nl/info/nieuwsoverzicht/eerste-cijfers-en-duiding-bij-de-hevige-neerslag-van-14-en-15-juli`

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. Retrieved from `https://hess.copernicus.org/articles/23/4323/2019/` doi: 10.5194/hess-23-4323-2019

Kobiyama, M., & Goerl, R. F. (2007). Quantitative method to distinguish flood and flash flood as disasters. *SUISUI Hydrological Research Letters*, *1*, 11–14.

Kohler, M. A., & Linsley, R. K. (1951). *Predicting the runoff from storm rainfall* (Vol. 30). US Department of Commerce, Weather Bureau.

Koks, E., Van Ginkel, K., Van Marle, M., & Lemnitzer, A. (2021). Brief communication: Critical infrastructure impacts of the 2021 mid-July western European flood event. *Natural Hazards and Earth System Sciences Discussions*, 1–11.

Koppa, A., & Gebremichael, M. (2017). A framework for validation of remotely sensed precipitation and evapotranspiration based on the Budyko hypothesis. *Water Resources Research*, *53*(10), 8487-8499. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020593` doi: https://doi.org/10.1002/2017WR020593

Kramer, K., & Ware, J. (2021). *Counting the cost: 2021: A year of climate breakdown.* Christian Aid.

Kreienkamp, F., Philip, S., Tradowsky, J., Kew, S., Lorenz, P., Arrighi, J., ... Wanders, N. (2021). *Rapid attribution of heavy rainfall events leading to the severe flooding in Western Europe during July 2021.* World Weather Attribution.

KULeuven. (2020). *Meadow region (Land van Herve).* Retrieved from `https://ees.kuleuven.be/soil-monoliths/herve/meadow-region.html`

Landesamt für Natur, Umwelt und Verbracherschutz Nordrhein-Westfalen. (2019). *Gsk3e_verzeichnis_gewaesser [xls].* Retrieved from `https://www.lanuv.nrw.de/fileadmin/lanuv/wasser/pdf/gsk3e_verzeichnis_gewaesser.xlsx)`

Landesamt für Umwelt, R.-P. (2022). *Übersicht des Pegels Altenahr: Hochwasserereignisse.* Retrieved from `https://www.hochwasser-rlp.de/karte/einzelpegel/flussgebiet/rhein/teilgebiet/mittelrhein/pegel/ALTENAHR`

Le Vif. (2022). *Dégâts, décès: les chiffres des inondations en vallée de la Vesdre.* Retrieved from `https://www.levif.be/belgique/degats-deces-les-chiffres-des-inondations-en-vallee-de-la-vesdre/`

Liu, D. (2020). A rational performance criterion for hydrological model. *Journal of Hydrology*, *590*, 125488.

Marchi, L., Borga, M., Preciso, E., & Gaume, E. (2010). Characterisation of selected extreme flash floods in europe and implications for flood risk management. *Journal of Hydrology*, *394*(1-2), 118–133.

Marchi, L., Cavalli, M., Amponsah, W., Borga, M., & Crema, S. (2016). Upper limits of flash flood stream power in Europe. *Geomorphology*, *272*, 68–77.

McMillan, H. K. (2021). A review of hydrologic signatures and their applications. *Wiley Interdisciplinary Reviews: Water*, *8*(1), e1499.

McPhillips, L. E., Chang, H., Chester, M. V., Depietri, Y., Friedman, E., Grimm, N. B., ... Rosi, E. J. (2018). Defining extreme events: A cross-disciplinary review. *Earth's Future*, *6*(3), 441–455.

Merckx, V. (2022). *Hoe vermijden we dat dit nog eens gebeurt? dit is er veranderd één jaar na de overstromingen in Wallonië.* Retrieved from `https://www.vrt.be/vrtnws/nl/2022/07/07/hoe-vermijden-we-dat-dit-nog-eens-gebeurt-dit-is-er-veranderd-e/#:~:text=Een%20nieuwe%20crisiscel%2C%20duidelijkere%20weersvoorspellingen,vernieling%20door%20de%20regio%20trokken.&text=39%20doden.,100.000%20slachtoffers%2C%20direct%20of%20indirect.`

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

Merz, R., & Blöschl, G. (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, *45*(1).

Meshkat, M., Amanian, N., Talebi, A., Kiani-Harchegani, M., & Rodrigo-Comino, J. (2019). Effects of roughness coefficients and complex hillslope morphology on runoff variables under laboratory conditions. *Water*, *11*(12), 2550.

Mizukami, N., Rakovec, O., Newman, A. J., Clark, M. P., Wood, A. W., Gupta, H. V., & Kumar, R. (2019). On the choice of calibration metrics for "high-flow" estimation using hydrologic models. *Hydrology and Earth System Sciences*, *23*(6), 2601–2614. Retrieved from `https://hess.copernicus.org/articles/23/2601/2019/` doi: 10.5194/hess-23-2601-2019

Moges, E., Demissie, Y., Larsen, L., & Yassin, F. (2020). *Review: Sources of hydrological model uncertainties and advances in their analysis.* s Note: MDPI stays neu-tral with regard to jurisdictional claims in ....

Moges, E., Demissie, Y., & Li, H.-Y. (2016). Hierarchical mixture of experts and diagnostic modeling approach to reduce hydrologic model structural uncertainty. *Water Resources Research*, *52*(4), 2551–2570.

Mohr, S., Ehret, U., Kunz, M., Ludwig, P., Caldas-Alvarez, A., Daniell, J. E., ... Gattke, C. (2022). A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. part 1: Event description and analysis. *Natural Hazards and Earth System Sciences Discussions*, 1–44.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Mudelsee, M. (2007). Long memory of rivers from spatial aggregation. *Water Resources Research*, *43*(1).

Nied, M., Pardowitz, T., Nissen, K., Ulbrich, U., Hundecha, Y., & Merz, B. (2014). On the relationship between hydro-meteorological patterns and flood types. *Journal of Hydrology*, *519*, 3249–3262.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., & Anquetin, S. (2008). Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. *Journal of Hydrology*, *362*(3-4), 274–290.

Noulet, J.-F. (2021, July, 29). *"ce n'était pas un raz-de-marée, mais un tsunami": quel rôle ont joué les barrages dans la brusque montée des eaux?* Retrieved 14-12-2021, from `https://www.rtbf.be/info/belgique/detail_ce-n-etait-pas-un-raz-de-maree-mais-un-tsunami-quel-role-ont-joue-les-barrages-dans-la-brusque-montee-des-eaux?id=10814073`

Overeem, A., & Leijnse, H. (2021). *Evaluatie & verbetering radar neerslagproducten casestudie extreme neerslag Zuid-Limburg 13–15 juli 2021.*

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, *49*(1), 187–201.

Paprotny, D., Sebastian, A., Morales-Napoles, O., & Jonkman, S. N. (2018, May 29). Trends in flood losses in Europe over the past 150 years. *Nature Communications*, *9*. doi: 10.1038/s41467-018-04253-1

Parsons, A. J., Abrahams, A. D., & Wainwright, J. (1996). Responses of interrill runoff and erosion rates to vegetation change in southern Arizona. *Geomorphology*, *14*(4), 311–317.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences*, *11*(5), 1633–1644.

Pelletier, J., Broxton, P., Hazenberg, P., Zeng, X., Troch, P., Niu, G., ... Gochis, D. (2016). *Global 1-km gridded thickness of soil, regolith, and sedimentary deposit layers.* ORNL Distributed Active Archive Center. Retrieved from `http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1304` doi: 10.3334/ORNLDAAC/1304

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. *Hydrological Processes*, *31*(10), 1828–1845.

Potter, W. D. (1958). Upper and lower frequency curves for peak rates of runoff. *Eos, Transactions American Geophysical Union*, *39*(1), 100–105.

Rahman, M., & Rosolem, R. (2017). Towards a simple representation of chalk hydrology in land surface modelling. *Hydrology and Earth System Sciences*, *21*(1), 459–471.

Raths, W. (1985). *Hochwasser an Ahr und Rhein.* Retrieved from `https://relaunch.kreis-ahrweiler.de/kvar/`

Raynaud, D., Thielen, J., Salamon, P., Burek, P., Anquetin, S., & Alfieri, L. (2015). A dynamic runoff co-efficient to improve flash flood early warning in Europe: evaluation on the 2013 central European floods in Germany. *Meteorological Applications*, *22*(3), 410–418.

Refsgaard, J. C., & Storm, B. (1990). *Construction, calibration and validation of hydrological models.* Springer.

Rijke, J., van Herk, S., Zevenbergen, C., & Ashley, R. (2012). Room for the River: delivering integrated river basin management in the Netherlands. *International Journal of River Basin Management*, *10*(4), 369-382. Retrieved from `https://doi.org/10.1080/15715124.2012.739173` doi: 10.1080/15715124.2012.739173

Roggenkamp, T., & Herget, J. (2014). Reconstructing peak discharges of historic floods of the river Ahr, Germany. *Erdkunde*, 49–59.

Roggenkamp, T., & Herget, J. (2015). *Historische Hochwasser der Ahr.* Retrieved from `https://www.geographie.uni-bonn.de/bilder-pressemitteilungen/RoggenkampHerget2015_Ahr_Heimatjahrbuch.pdf`

Roggenkamp, T., & Herget, J. (2022). Hochwasser der Ahr im Juli 2021 - Abflusseinschätzung und Einordnung. *Hydrologie und Wasserbewirtschaftung*, *1*, 1. 40-49.

Rosso, R., Peano, A., Becchi, I., & Bemporad, G. (1994). An introduction to spatially distributed modelling of basin response. *Advances in distributed hydrology*, *3*, 30.

Rumynin, V. G. (2015). Surface runoff generation, vertical infiltration and subsurface lateral flow. In *Overland flow dynamics and solute transport* (Vol. 26, p. 3-50). doi: 10.1007/978-3-319-21801-4_1

Sallam, O. M. (2006). Aquifers parameters estimation using well log and pumping test data, in arid regions—step in sustainable development. In *The 2nd international conference on water resources and arid environment, riyadh, saudi arabia* (pp. 26–29).

Schmid-Johanssen, J., Lang, U., & Heiliger, N. (2021, November, 24). *Datenanalyse zur Flutkatastrophe an der Ahr - aktuele Daten und Fakten.* SWR. Retrieved 2022-01-10, from `https://www.swr.de/swraktuell/rheinland-pfalz/flut-in-ahrweiler-so-gross-ist-der-schaden-104.html`

Schröter, K., Kunz, M., Elmer, F., Mühr, B., & Merz, B. (2015). What made the june 2013 flood in Germany an exceptional event? a hydro-meteorological evaluation. *Hydrology and Earth System Sciences*, *19*(1), 309–327.

Schumm, S. A. (1956). Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *Geological society of America bulletin*, *67*(5), 597–646.

Schäfer, A., Mühr, B., Daniel, J., Ehret, U., Ehmele, F., Küpfer, K., ... Kunz, M. (2021). *Hochwasser Mitteleuropa, Juli 2021 (Deutschland).* doi: 10.5445/IR/1000135730

Scofield, R. A., & Kuligowski, R. J. (2003). Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Weather and Forecasting*, *18*(6), 1037–1051.

Seel, K. (1983). Die Ahr und ihre Hochwässer in alten Quellen. *Heimatjahrbuch des Kreises Ahrweiler*.

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., ... Rahimi, M. (2012). *Changes in climate extremes and their impacts on the natural physical environment. in: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.* (Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.)

Stein, L., Pianosi, F., & Woods, R. (2020). Event-based classification for global study of river flood generating processes. *Hydrological Processes*, *34*(7), 1514–1529.

SWB. (n.d.). *Geschäftsfelder der swb regional.* Retrieved from `https://www.swb-regional.de/geschaeftsfelder/`

Syvitski, J. P. (2003). Supply and flux of sediment along hydrological pathways: research for the 21st century. *Global and Planetary Change*, *39*(1-2), 1–11.

Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., ... Schumann, A. (2019). Causative classification of river flood events. *Wiley Interdisciplinary Reviews: Water*, *6*(4), e1353.

Task Force Fact Finding hoogwater 2021. (2021). *Hoogwater 2021 Feiten en Duiding.*

Teng, W., Wang, J., & Doraiswamy, P. (1993). Relationship between satellite microwave radiometric data, antecedent precipitation index, and regional soil moisture. *International Journal of Remote Sensing*, *14*(13), 2483–2500.

Thieken, A., Kemter, M., Vorogushyn, S., Berghäuser, L., Sieg, T., Natho, S., ... Bronstert, A. (2021). *Extreme Hochwasser bleiben trotz integriertem Risikomanagement eine Herausforderung.* GFZ Potsdam, PIK Potsdam NatRiskChange (https://uni-potsdam. de/fileadmin ....

Van der Veen, R. (2021). *Topafvoeren hoogwater Maas juli 2021.* Rura-Arnhem and Agtersloot Hydraulisch Advies.

Van Dijk, A., & Bruijnzeel, L. (2001). Modelling rainfall interception by vegetation of variable density using an adapted analytical model. part 1. model description. *Journal of Hydrology*, *247*(3-4), 230–238.

Van Heeringen, K., Asselman, N., Overeem, A., Beersma, J., & S., P. (2022). *Analyse overstroming Valkenburg.* Retrieved from `https://www.waterschaplimburg.nl/publish/pages/7013/analyse_overstroming_valkenburg.pdf`

Vannier, O., Braud, I., & Anquetin, S. (2014). Regional estimation of catchment-scale soil properties by means of streamflow recession analysis for use in distributed hydrological models. *Hydrological Processes*, *28*(26), 6276–6291.

Van Osnabrugge, B., Weerts, A., & Uijlenhoet, R. (2017). genre: A method to extend gridded precipitation climatology data sets in near real-time for hydrological forecasting purposes. *Water Resources Research*, *53*(11), 9284–9303.

van Stokkom, H. T., Smits, A. J., & Leuven, R. S. (2005). Flood defense in the Netherlands. *Water International*, *30*(1), 76-87. doi: 10.1080/02508060508691839

Van Verseveld, W. J., Weerts, A. H., Visser, M., Buitink, J., Imhoff, R. O., Boisgontier, H., ... ten Velden, C. (2022). Wflow_sbm v0. 6.1, a spatially distributed hydrologic model: from global data to local applications. *Geoscientific Model Development Discussions*, 1–52.

Verkade, J. (2021). *July 2021 floods: EFAS performance.* Retrieved from `https://www.efas.eu/sites/default/files/2021-10/July2021flood_EFASperformance.pdf`

Vertessy, R. A., & Elsenbeer, H. (1999). Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization. *Water Resources Research*, *35*(7), 2173–2187.

Viessman, W., & Lewis, G. L. (1996). *Introduction to hydrology.* Harper Collins College.

Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M., & Blöschl, G. (2010). Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, *394*(1-2), 213–229.

Villarini, G., Smith, J. A., Baeck, M. L., Vitolo, R., Stephenson, D. B., & Krajewski, W. F. (2011). On the frequency of heavy rainfall for the Midwest of the United States. *Journal of Hydrology*, *400*(1), 103-120. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0022169411000527` doi: https://doi.org/10.1016/j.jhydrol.2011.01.027

Waterschap Limburg. (2021). *(Gesloten) liveblog - watersnood Limburg.* Retrieved 14-12-2021, from `https://www.waterschaplimburg.nl/@6426/liveblog-watersnood/`

WML. (2021). *De bronnen van het drinkwater uit voerendaal.* Retrieved from `https://limburgsdrinkwater.nl/gemeente/voerendaal/detail`

Ziemetz, F., Launay, M., Bourqui, P., Calixte, E., Fallon, C., & Teller, J. (2021). *Analyse indépendante sur la gestion des voies hydrauliques lors des intempéries de la semaine du 12 juillet 2021. lot 1 - factualisation.* Stucky Gruner.

Appendix A

Scopes of flash flood research

The purpose of this appendix is to provide more insight in the scopes of flash flood studies. The studies have been found with search terms such as "flash floods" and "Europe" during the literature research for flash floods in this thesis. The main search engine for this search was Google Scholar. The search engine of the library of TUDelft and Web of Science have been used as well. In general, the studies either focused on specific aspects of flash floods within a singular catchment or multiple catchments within one country or they analysed international flash floods in a general way.

Table A.1: Different scopes for flash flood research

Research paper	**In-depth analysis in one country**	**General analysis in multiple countries**
Alfieri et al. (2015)		x
Amponsah et al. (2018)		x
Barredo (2007)		x
Barredo (2009)		x
Bačová Mitková et al. (2018)	x	
Bezak et al. (2016)	x	
Borga et al. (2011)		x
Bryndal (2015)	x	
Ciervo et al. (2015)	x	
Diakakis et al. (2020)	x	
Douinot et al. (2016)	x	
Einfalt et al. (2009)	x	
Gaume et al. (2009)		x
Kastridis et al. (2020)	x	
Marchi et al. (2010)		x
Marchi et al. (2016)		x
Merz & Blöschl (2003)	x	
Paprotny et al. (2018)		x
Raynaud et al. (2015)	x	

Table A.1 shows that most flash flood studies perform an in-depth analysis of one catchment or multiple catchments in one country. There are also numerous studies that performed a general analysis of various flash flood events that occurred in multiple countries. Since no study has focused on a single flash flood event in international catchments, this research on a flood event referenced to as a flash flood has a unique approach.

Appendix B

Catchment characteristics

This section provides more detailed information and figures for the description of the catchments. The discussed catchment characteristics are: the elongation and relief ratio, catchment slopes, lithology, measurement locations, population density, the Budyko plot, annual precipitation and flow duration curves.

B.1 Elongation and relief ratio

Schumm (1956) has developed the elongation and relief ratio to quantify the relative shape and elevation differences in catchments. The elongation ratio is defined as the ratio between the diameter of a circle with the same area as the catchment and the maximum catchment length, see eq. (B.1). This parameter approaches 1 when the form of the catchment becomes similar to a circle.

$$\text{elongation ratio} = \frac{2 \cdot \sqrt{\frac{A_{catchment}}{\pi}}}{L_{max}} \tag{B.1}$$

The relief ratio is defined as the ratio between total elevation difference and the maximum catchment length, see eq. (B.2). The total elevation difference is calculated from the points with the highest and lowest elevation within the catchment. A value of 0 would indicate a completely flat catchment while a high value indicates a high relatively strong relief.

$$\text{relief ratio} = \frac{h_{max} - h_{min}}{L_{max}} \tag{B.2}$$

Both ratios depend on the maximum length of the catchment, which is defined as the maximum straight distance within the catchment that is parallel to the principal drainage line. These lengths are shown in fig. B.1. The maximum catchments lengths are only an estimation since the elevation and relief ratio are used as indication and the definition of the maximum length is not precise.

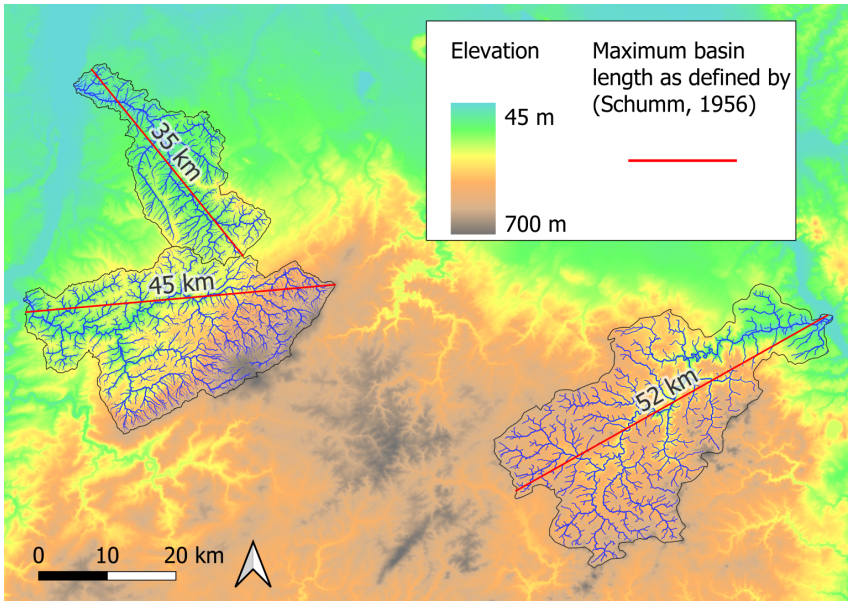


Figure B.1: Maximum catchment length for the three catchments as defined by Schumm (1956), with elevation as background and the river system indicated in blue

The elevation differences are based on the EU-DEM v1.1 topography map, resulting in maximum elevation differences of 635, 644, 326 $[m]$ for respectively the Ahr, Vesdre and Geul. For the catchment areas, the values of table 2.1 are used. For the relief ratio, the maximum catchment length is used in the unit of $[m]$ to ensure that the units drop out and the ratio is dimensionless. For the elongation ratio, this is already the case. The resulting ratios are given in table 2.1.

B.2 Slope

In fig. B.2, fig. B.3 and fig. B.4, the darker colors indicate a relatively flat surface, while the lighter colors indicate steep slopes. Although the river network is not indicated in these figures, the pattern of the river network can be recognized in the slope figures, showing that each river has carved its network through the landscape. The overall brightness of slope figure for the Ahr catchment indicates a steeper relief compared to the Vesdre and Geul. Additionally, the figure of the Ahr contains almost no dark areas, which indicate plateaus if they are situated in upstream regions. These plateaus are visible for the Vesdre and the Geul. Of all the catchments, the Geul has the darkest colors and thus the least steep relief. The valleys also seems slightly wider for the Geul than for the Ahr and the Vesdre, although all valleys can be considered as narrow. The relief varies in all catchments and is the strongest in the Ahr, followed by the Vesdre and the Geul.

The conclusions from these slope figures are confirmed by observations from field trips to the catchments.

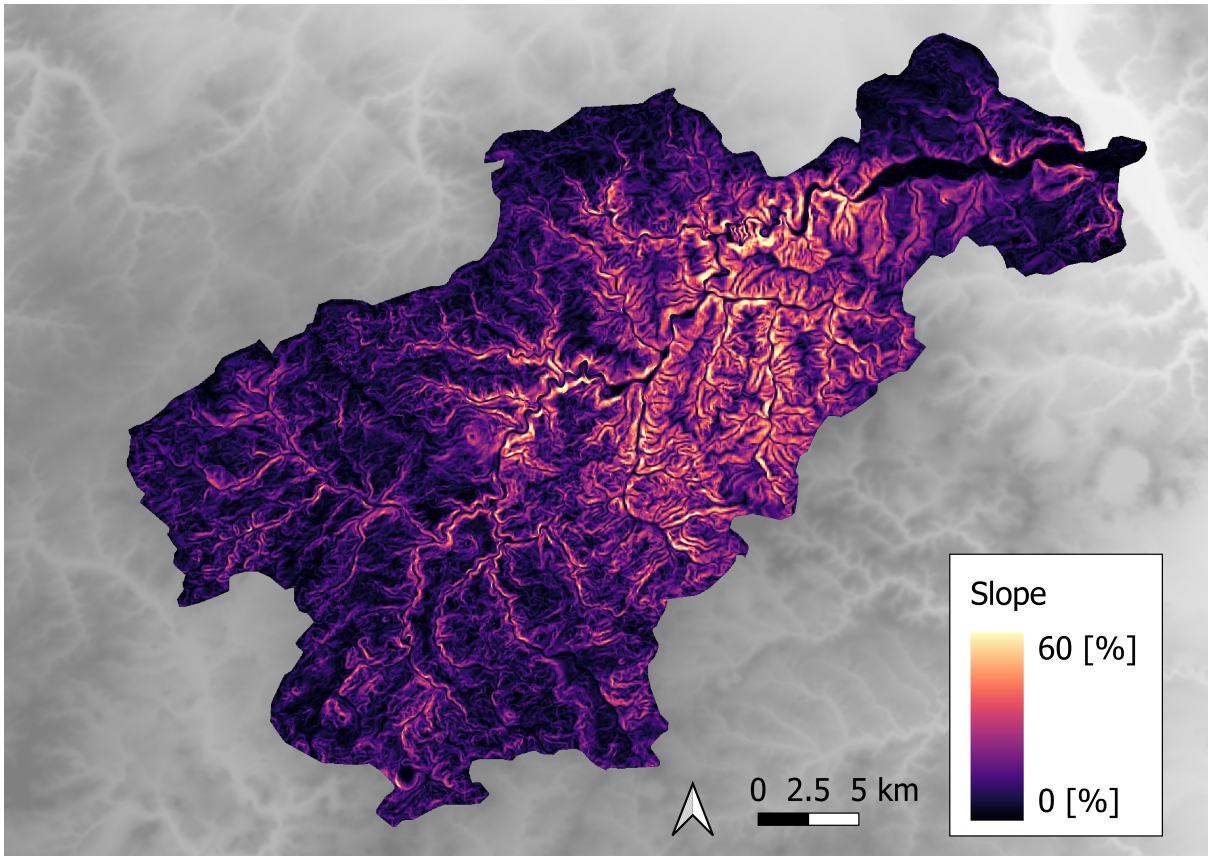


Figure B.2: Slope of the Ahr catchment, derived from EU-DEM v1.1

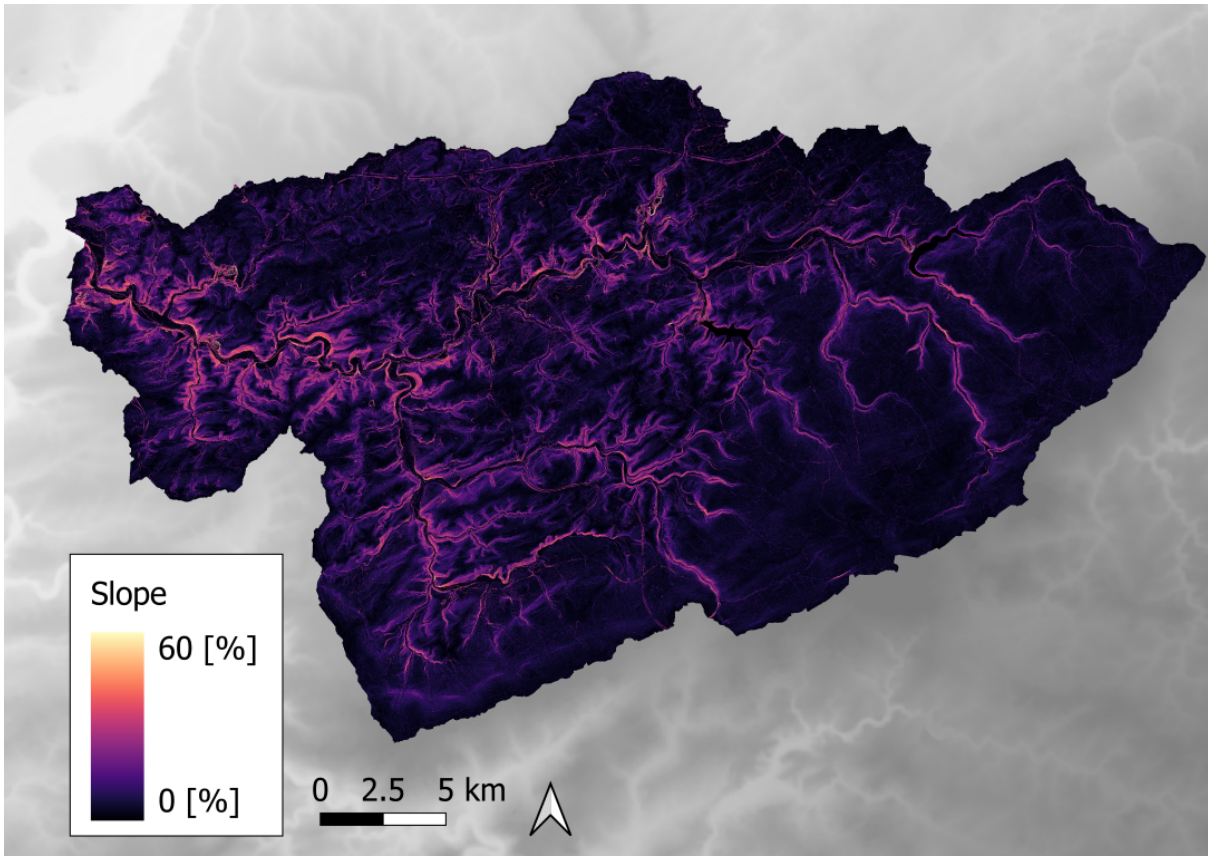


Figure B.3: Slope of the Vesdre catchment, derived from Modèle Numérique de Terrain (MNT) de la Wallonie 2013-2014

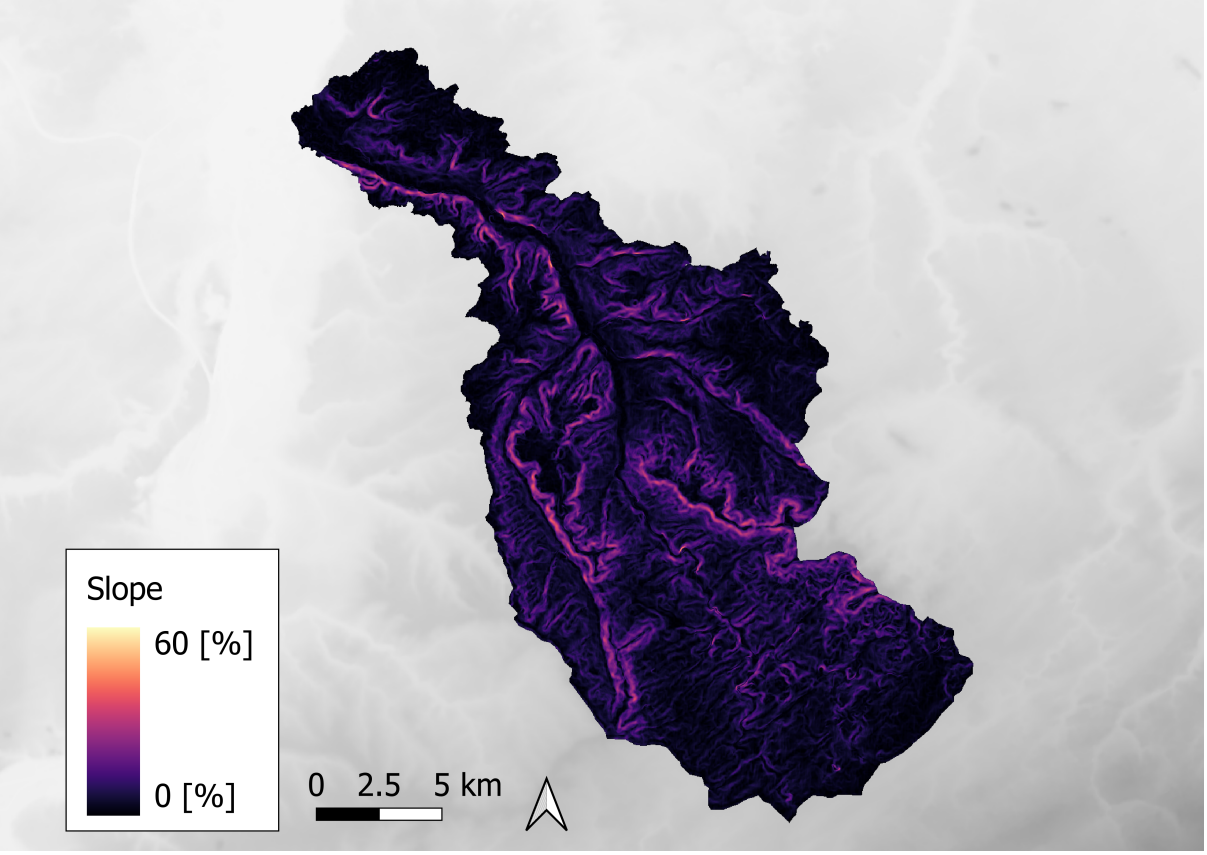


Figure B.4: Slope of the Geul catchment, derived from EU-DEM v1.1

B.3 Lithology

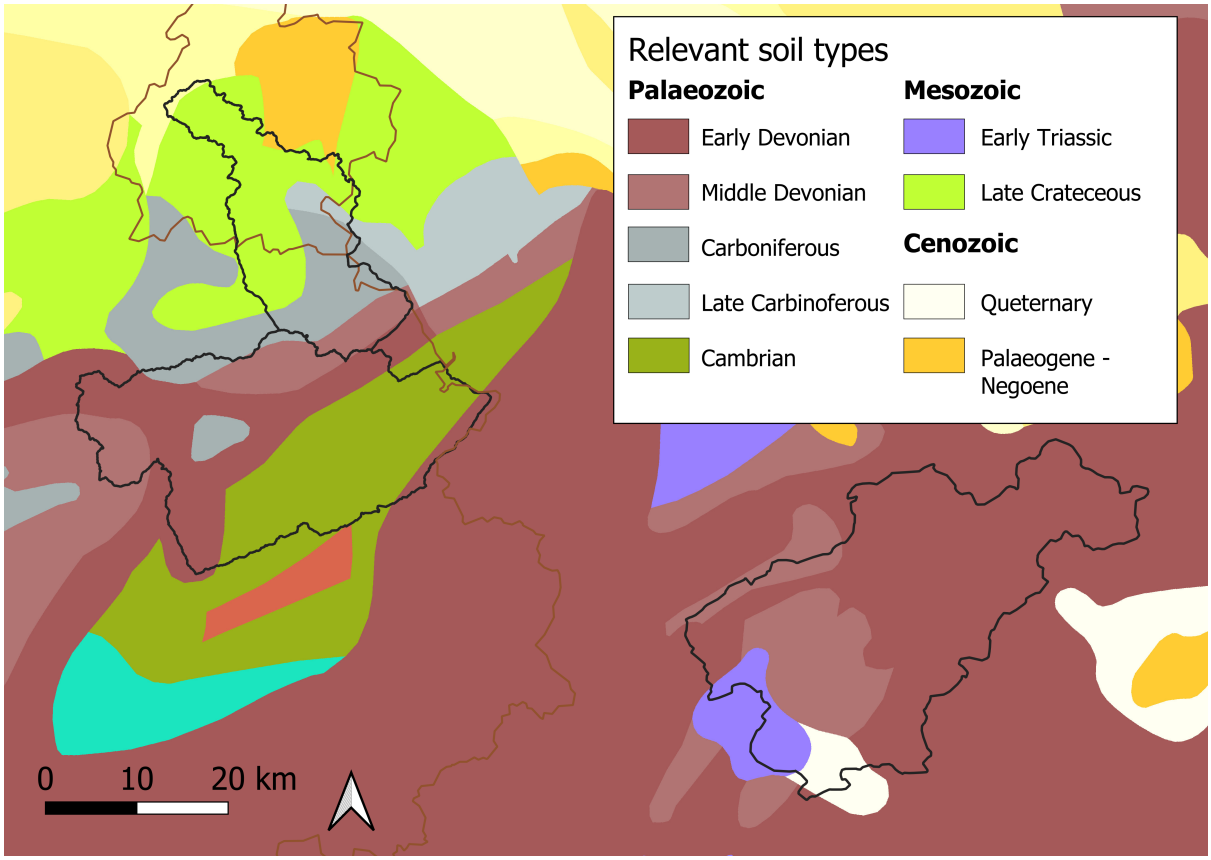


Figure B.5: Pre-Quarternary geology of the study region, data source: IGME5000 provided by EGDI Metadata

Fig. B.5 shows the 1:5 Million International Geological Map of Europe and Adjacent Areas (IGME5000). It visualises the formation periods of the different soils in the study region. The majority of the soils originate from the palaeozoic period, especially for the Ahr and the Vesdre. The northern part of the Geul has soils originating from the mesozoic, specifically late crateceous period. The geologic age of this region is similar since it is largely part of the Rhenish Massif. This common geology will likely lead to similar geo-hydrologic properties.

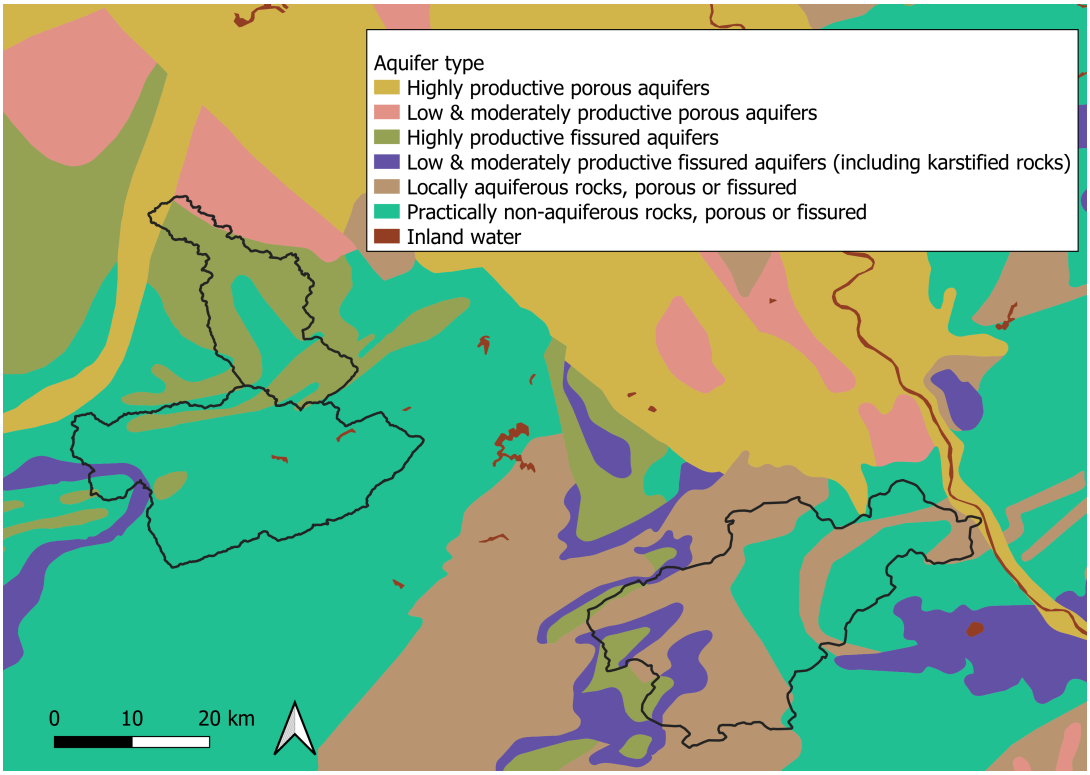


Figure B.6: IHME1500 showing the aquifer types for all catchments.

The deducted geohydrological properties from fig. B.5 can be compared with the IHME1500 dataset (International Hydrogeological Map of Europe 1:1,500,000), which indicates the aquifers and the lithology of certain layers and is visualised in fig. B.6. For the Ahr, the western part is indicated as locally aquiferous rocks, which could be due to the limestone and dolomite in that region. The rest of the Ahr catchment, as well as the Vesdre catchment is classified as practically non-aquiferous rocks which confirms the low permeability soil in that region. The Geul has both highly productive fissured aquifers due to the sand and chalk, as well as a non-aquiferous part, probably due to the silt. Subsurface flow is expected to be largest for the Geul while they may be limited for the Vesdre and the Ahr. Thereby, fig. B.6 confirms the findings from fig. 2.2.

B.4 Measurement station information

More information is given for the measurement stations. Types include precipitation (P), discharge (Q) and water levels (H). The abbreviations of the countries are given as well (column CTRY). The coordinates (X, Y) are given in the form of the EPSG:4326 coordinate reference system. The frequencies of time steps of the data (Freq.) include hourly (H), daily (D) and per minute (T).

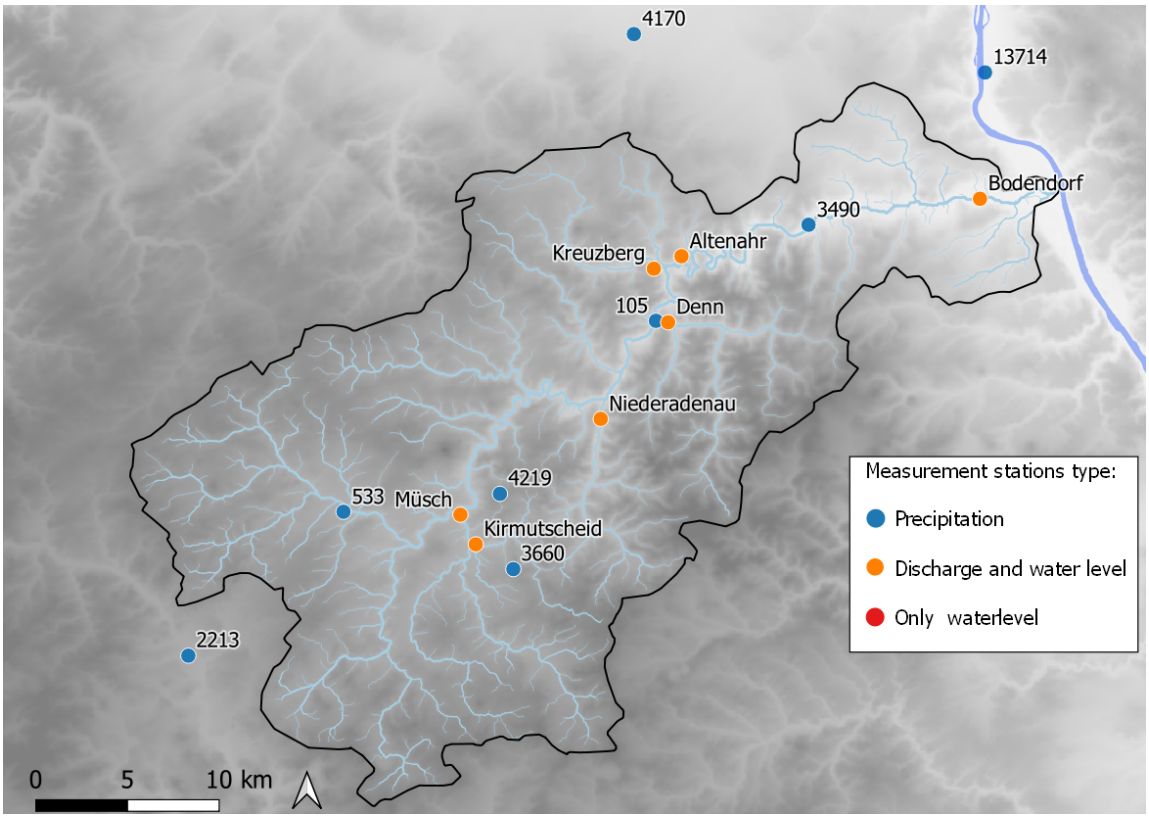


Figure B.7: Measurement stations for the Ahr catchment.

Table B.1: Information of measurement stations for the Ahr catchment.

Type	Station	Town	CTRY	X	Y	Start data	End data	Freq.	Source	River
P	105	Ahrbrück	DE	6.971023	50.484485	1941	13/07/2021	D	DWD	-
P	533	Blankenheim-Ahrhütte	DE	6.737474	50.384408	1961	05/12/2021	D	DWD	-
P	2213	Lissendorf	DE	6.623578	50.310518	1931	07/12/2021	D	DWD	-
P	2213	Lissendorf	DE	6.623578	50.310518	01/11/2002	07/12/2021	H	DWD	-
P	3225	Mehren	DE	6.876649	50.18203	1971	31/12/2018	D	DWD	-
P	3490	Neuenahr	DE	7.085337	50.534561	1931	06/12/2021	D	DWD	-
P	3490	Neuenahr	DE	7.085337	50.534561	01/09/2004	06/12/2021	H	DWD	-
P	4170	Rheinbach	DE	6.94496	50.624435	1931	31/10/2020	D	DWD	-
P	4219	Rodder	DE	6.856804	50.396548	1931	11/01/2022	D	DWD	-
P	4219	Rodder	DE	6.856804	50.396548	01/10/2004	11/01/2022	H	DWD	-
P	3660	Nurburg-Barweiler	DE	6.869540184	50.35998604	1995	11/01/2022	D	DWD	-
P	3660	Nurburg-Barweiler	DE	6.869540184	50.35998604	01/09/1995	11/01/2022	D	DWD	-
P	13714	Linz-Unkel	DE	7.216695684	50.61250819	2007-10	11/01/2022	H	DWD	-
Q	Müsch	Müsch	DE	6.827178805	50.00	11/11/1972	31/12/2019	D	Lfu	Ahr
Q	Altenahr	Altenahr	DE	6.988616779	50.51663107	11/11/1946	31/12/2021	D	Lfu	Ahr
Q	Kirmutscheid	Kirmutscheid	DE	6.839956721	50.37132175	11/11/1972	31/12/2021	D	Lfu	Trierbach
Q	Denn	Denn	DE	6.980423464	50.48398444	11/11/1977	31/12/2021	D	Lfu	Kesselingerbach
Q	Niederadenaeu	Niederadenaeu	DE	6.931711081	50.43534653	31/12/2013	31/12/2021	D	Lfu	Adenauerbach
Q	Kreuzberg	Kreuzberg	DE	6.967765535	50.5099822	11/11/1972	31/12/2021	D	Lfu	Sahrbach
Q	Bodendorf	Bodendorf	DE	7.216457007	50.5505509	05/02/2020	31/12/2021	D	Lfu	Ahr
H	Müsch	Müsch	DE	6.827178805	50.38548103	11/11/1972	31/12/2019	D	Lfu	Ahr
H	Altenahr	Altenahr	DE	6.988616779	50.51663107	11/11/1946	31/12/2020	D	Lfu	Ahr
H	Kirmutscheid	Kirmutscheid	DE	6.839956721	50.37132175	11/11/1972	31/12/2020	D	Lfu	Trierbach
H	Denn	Denn	DE	6.980423464	50.48398444	11/11/1977	31/12/2020	D	Lfu	Kesselingerbach
H	Niederadenaeu	Niederadenaeu	DE	6.931711081	50.43534653	31/12/2013	31/12/2020	D	Lfu	Adenauerbach
H	Kreuzberg	Kreuzberg	DE	6.967765535	50.5099822	11/11/1972	31/12/2020	D	Lfu	Sahrbach
H	Bodendorf	Bodendorf	DE	7.216457007	50.5505509	05/02/2020	31/12/2021	D	Lfu	Ahr

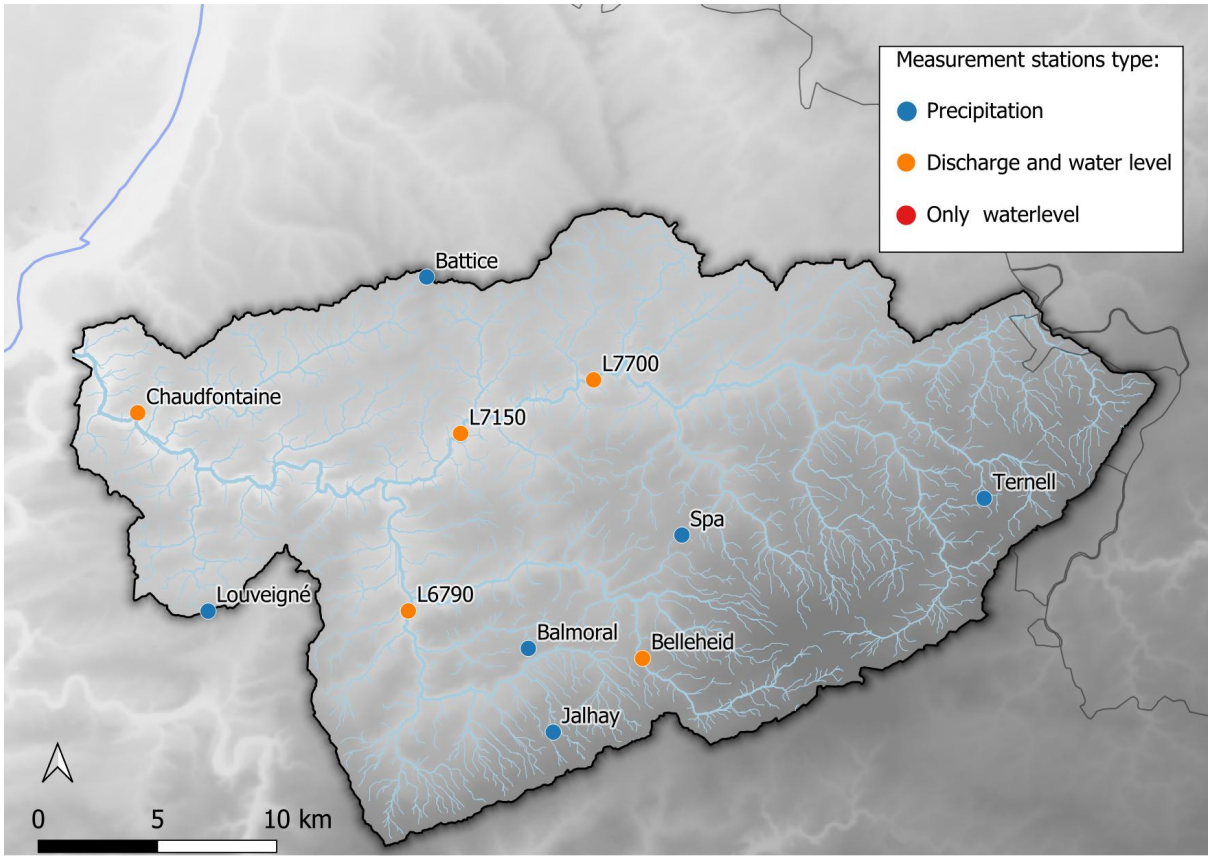


Figure B.8: Measurement stations for the Vesdre catchment.

Table B.2: Information of measurement stations for the Vesdre catchment.

Type	Station	Town	CTRY	X	Y	Start data	End data	Freq.	Source	River
P	Ternell	Ternell	BE	6.157033	50.57528	01/01/2002	31/12/2021	D	SPW	-
P	Ternell	Ternell	BE	6.157033	50.57528	01/01/2002	31/12/2021	H	SPW	-
P	Spa	Spa	BE	5.979466	50.55537	01/01/2017	31/12/2021	D	SPW	-
P	Spa	Spa	BE	5.979466	50.55537	01/01/2017	31/12/2021	H	SPW	-
P	Jalhay	Jalhay	BE	5.909887	50.47854	01/01/2002	31/12/2021	D	SPW	-
P	Jalhay	Jalhay	BE	5.909887	50.47854	01/01/2002	31/12/2021	H	SPW	-
P	Battice	Battice	BE	5.819987	50.64707	01/01/2002	31/12/2021	D	SPW	-
P	Battice	Battice	BE	5.819987	50.64707	01/01/2002	31/12/2021	H	SPW	-
P	Balmoral	Balmoral	BE	5.892437	50.50951	01/01/2002	31/9/2016	D	SPW	-
P	Louveigné	Louveigné	BE	5.730506	51.50277	01/01/2002	31/12/2021	D	SPW	-
P	Louveigné	Louveigné	BE	5.730506	51.50277	01/01/2002	31/12/2021	H	SPW	-
Q	L7700	Bellevaux	BE	5.922198	50.61192	01/09/2011	31/12/2021	H	SPW	Vesdre
Q	L7150	Verviers	BE	5.845297	50.58893	17/08/2004	31/12/2021	H	SPW	Vesdre
Q	L6790	Spixhe	BE	5.820336	50.52106	29/03/2002	14/07/2021	H	SPW	Wayai
Q	Chaudfontaine	Chaudfontaine	BE	5.653621	50.58964	01/01/1992	31/07/2021	H	SPW	Vesdre
Q	Belleheid	Belleheid	BE	5.960228	50.50814	01/01/1994	31/07/2021	H	SPW	Hoëgne
H	L7700	Bellevaux	BE	5.922198	50.61192	01/09/2011	31/12/2021	H	SPW	Vesdre
H	L7150	Verviers	BE	5.845297	50.58893	17/08/2004	31/12/2021	H	SPW	Vesdre
H	L6790	Spixhe	BE	5.820336	50.52106	29/03/2002	14/07/2021	H	SPW	Wayai
H	Chaudfontaine	Chaudfontaine	BE	5.653621	50.58964	01/01/1992	31/07/2021	H	SPW	Vesdre
H	Belleheid	Belleheid	BE	5.960228	50.50814	01/01/1994	30/04/2020	H	SPW	Hoëgne

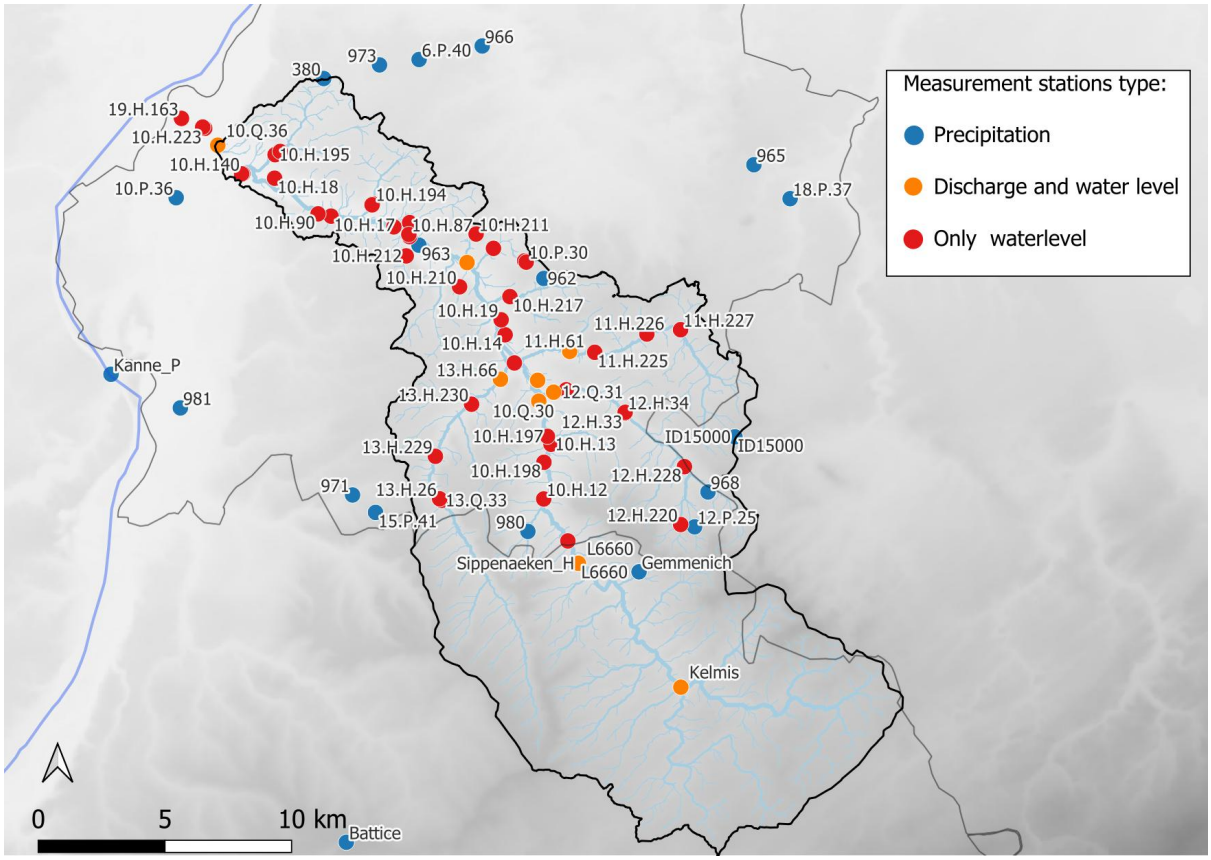


Figure B.9: Measurement stations for the Geul catchment

Table B.3: Information of measurement stations for the Geul catchment.

Type	Station	Town	CTRY	X	Y	Start data	End data	Freq.	Source	River
P	963	Valkenburg	NL	5.841672	50.85993	02/01/1951	20/11/2021	D	KNMI	-
P	965	Schaesberg	NL	6.026789	50.89501	02/01/1951	20/11/2021	D	KNMI	-
P	981	Oost Maarland	NL	5.713411	50.79744	01/01/1986	20/11/2021	D	KNMI	-
P	980	Epen	NL	5.911562	50.76075	01/08/1980	20/11/2021	D	KNMI	-
P	971	Noorbeek	NL	5.812443	50.77015	02/01/1951	20/11/2021	D	KNMI	-
P	966	Schinnen	NL	5.870827	50.93172	02/01/1951	20/11/2021	D	KNMI	-
P	973	Beek	NL	5.813848	50.92296	02/01/1951	20/11/2021	D	KNMI	-
P	962	Ubachsberg	NL	5.912579	50.85064	01/11/1952	20/11/2021	D	KNMI	-
P	968	Vaals	NL	6.011002	50.77824	02/01/1951	20/11/2021	D	KNMI	-
P	380	Maastricht	NL	5.782993	50.9170	01/01/1958	31/12/2021	H	KNMI	-
P	1.P.2	Roermond	NL	5.986756	51.21979	13/07/2021	15/07/2021	5T	WL	-
P	6.P.40	Spaubeek	NL	5.835861	50.92567	13/07/2021	15/07/2021	5T	WL	-
P	10.P.30	Ransdaal	NL	5.902261	50.85623	13/07/2021	15/07/2021	5T	WL	-
P	10.P.36	Maastricht	NL	5.704189	50.87177	13/07/2021	15/07/2021	5T	WL	-
P	12.P.25	Vaals	NL	6.004582	50.76564	13/07/2021	15/07/2021	5T	WL	-
P	15.P.41	Noorbeek	NL	5.825855	50.76445	13/07/2021	15/07/2021	5T	WL	-
P	18.P.37	Kerkrade	NL	6.048089	50.8837	13/07/2021	15/07/2021	5T	WL	-
P	Kanne_P	Kanne	BE	5.673533	50.8079	01/06/2021	17/07/2021	15T	VMM	-
P	Gemmenich	Gemmenich	BE	5.974898	50.74865	01/07/2021	31/07/2021	H	SPW	-
P	Battice	Battice	BE	5.819985	50.64707	01/01/2002	31/12/2021	H	SPW	-
P	ID15000	Aachen	DE	6.0244	50.7983	01/04/2011	22/02/2022	H	DWD	-

Type	Station	Town	CTRY	X	Y	Start data	End data	Freq.	Source	River
Q	10.Q.30	Hommerich	NL	5.913613	50.80709	01/01/1970	25/07/2011	H	WL	Geul
Q	11.Q.32	Eys	NL	5.929379	50.82533	17/07/1991	25/07/2011	H	WL	Eyserbeek
Q	12.Q.31	Partij	NL	5.921602	50.81058	06/08/1991	25/07/2011	H	WL	Selzerbeek
Q	12.Q.46	Wittem	NL	5.912339	50.81446	01/10/1994	25/07/2011	H	WL	Selzerbeek
Q	13.Q.34	Gulpen	NL	5.891506	50.81409	16/09/1973	25/07/2011	H	WL	Gulp
Q	L6660	Sippenaeken	BE	5.940761	50.75047	13/06/1996	31/12/2021	H	SPW	Geul
Q	10.Q.36	Meerssen	NL	5.725796	50.89121	15/02/2012	18/07/2021	H	WL	Geul
Q	10.Q.63	Schin op Geul	NL	5.869142	50.85476	17/02/2016	14/07/2021	H	WL	Geul
Q	Kelmis	Kelmis	BE	6.001745	50.708599	01/01/2009	31/12/2021	H	SPW	Geul
H	10.H.12	Epen	NL	5.919392	50.77257	01/02/1995	12/11/2021	H	W-WRO	Geul
H	10.H.13	Mechelen	NL	5.921758	50.79208	01/07/1993	12/11/2021	H	W-WRO	Geul
H	10.H.14	Wijlre	NL	5.892715	50.82993	01/10/1994	12/11/2021	H	W-WRO	Geul
H	10.H.140	Meerssen	NL	5.740717	50.88166	07/07/2013	12/11/2021	H	W-WRO	Geulke
H	10.H.141	Meerssen	NL	5.740717	50.88165	07/07/2013	12/11/2021	H	W-WRO	Kanjel
H	10.H.142	Meerssen	NL	5.740077	50.88157	07/07/2013	12/11/2021	H	W-WRO	Kanjel
H	10.H.16	Valkenburg	NL	5.83609	50.86278	17/07/1991	12/11/2021	H	W-WRO	Geul
H	10.H.17	Geulhem	NL	5.791541	50.86843	25/11/1994	12/11/2021	H	W-WRO	Geul
H	10.H.18	Meerssen	NL	5.758698	50.8807	01/06/1993	12/11/2021	H	W-WRO	Geul
H	10.H.19	Wijlre	NL	5.890018	50.83515	04/01/2008	12/11/2021	H	W-WRO	Geul
H	10.H.194	Valkenburg	NL	5.814572	50.87361	01/12/2015	12/11/2021	H	W-WRO	-
H	10.H.195	Meerssen	NL	5.758211	50.88905	20/07/2016	10/09/2021	H	W-WRO	-
H	10.H.196	Hommerich	NL	5.898767	50.82015	04/12/2015	05/11/2021	H	W-WRO	Geul
H	10.H.197	Mechelen	NL	5.919534	50.7948	17/11/2015	08/11/2021	H	W-WRO	Geul
H	10.H.198	Mechelen	NL	5.918379	50.78556	04/11/2015	12/11/2021	H	W-WRO	Geul
H	10.H.199	Schin op Geul	NL	5.869142	50.85476	17/02/2016	14/10/2021	H	W-WRO	Geul
H	10.H.209	Valkenburg	NL	5.814143	50.87326	30/05/2016	12/11/2021	H	W-WRO	Buffer
H	10.H.210	Valkenburg	NL	5.865729	50.84602	27/05/2016	12/11/2021	H	W-WRO	Buffer
H	10.H.211	Klimmen	NL	5.873312	50.86502	01/12/2015	16/07/2021	H	W-WRO	Buffer
H	10.H.212	Valkenburg	NL	5.834887	50.85587	05/11/2015	12/11/2021	H	W-WRO	Buffer
H	10.H.213	Ransdaal	NL	5.901326	50.85647	21/11/2015	11/11/2021	H	W-WRO	Buffer
H	10.H.214	Ransdaal	NL	5.883488	50.8603	02/02/2016	31/10/2021	H	W-WRO	Buffer
H	10.H.215	Meerssen	NL	5.760978	50.89029	13/11/2015	12/11/2021	H	W-WRO	Buffer
H	10.H.216	Valkenburg	NL	5.835726	50.86773	22/02/2016	10/08/2021	H	W-WRO	Buffer
H	10.H.217	Etenaken	NL	5.894212	50.84355	27/05/2016	15/07/2021	H	W-WRO	Buffer
H	10.H.222	Valkenburg	NL	5.827043	50.86591	04/11/2015	04/11/2021	H	W-WRO	Geul
H	10.H.223	Bunde	NL	5.717921	50.8969	07/11/2015	12/11/2021	H	W-WRO	Geul
H	10.H.224	Bunde	NL	5.716773	50.89731	21/02/2016	16/07/2021	H	W-WRO	Geul
H	10.H.56	Cottessen	NL	5.934192	50.75818	01/08/1991	12/11/2021	H	W-WRO	Geul
H	10.H.58	Hommerich	NL	5.91366	50.80711	27/01/1970	12/11/2021	H	W-WRO	Geul
H	10.H.59	Ransdaal	NL	5.902259	50.85608	22/09/2011	12/11/2021	H	W-WRO	Geul
H	10.H.60	Meerssen	NL	5.725796	50.89121	01/04/1994	12/11/2021	H	W-WRO	Geul
H	10.H.87	Valkenburg	NL	5.835671	50.86353	03/11/2010	12/11/2021	H	W-WRO	Geul
H	10.H.90	Berg en Terblijt	NL	5.784173	50.86897	28/10/2010	12/11/2021	H	W-WRO	Geul
H	11.H.225	Eys	NL	5.943403	50.82552	04/11/2015	05/10/2021	H	W-WRO	Eyserbeek
H	11.H.226	Eys	NL	5.971995	50.8330	04/11/2015	12/11/2021	H	W-WRO	Eyserbeek
H	11.H.227	Eys	NL	5.990759	50.83517	13/07/2016	12/11/2021	H	W-WRO	Eyserbeek
H	11.H.61	Eys	NL	5.92935	50.8253	01/07/1991	12/11/2021	H	W-WRO	Eyserbeek
H	12.H.219	Wittem	NL	5.928694	50.81167	04/11/2015	12/11/2021	H	W-WRO	Buffer
H	12.H.220	Vaals	NL	5.996723	50.7662	30/11/2015	11/11/2021	H	W-WRO	Buffer
H	12.H.228	Lemiers	NL	5.997064	50.78677	04/11/2015	29/10/2021	H	W-WRO	Selzerbeek
H	12.H.33	Wittem	NL	5.962344	50.80455	01/08/2006	12/11/2021	H	W-WRO	Selzerbeek
H	12.H.34	Wittem	NL	5.962249	50.80485	01/08/2006	12/11/2021	H	W-WRO	Selzerbeek
H	12.H.62	Partij	NL	5.921602	50.81062	28/08/1991	12/11/2021	H	W-WRO	Selzerbeek
H	12.H.63	Wittem	NL	5.912296	50.81444	01/10/1994	12/11/2021	H	W-WRO	Selzerbeek
H	12.H.66	Wittem	NL	5.928609	50.81169	10/01/2014	07/10/2021	H	W-WRO	Selzerbeek
H	13.H.229	Slenaken	NL	5.857533	50.78551	23/11/2015	12/11/2021	H	W-WRO	Gulp
H	13.H.230	Euverem	NL	5.876112	50.80468	17/11/2015	11/11/2021	H	W-WRO	Gulp
H	13.H.25	Slenaken	NL	5.86238	50.77003	01/01/1978	31/08/2021	H	W-WRO	Gulp
H	13.H.26	Slenaken	NL	5.861237	50.77055	14/12/2005	31/08/2021	H	W-WRO	Gulp
H	13.H.66	Gulpen	NL	5.891478	50.81409	02/12/1999	12/11/2021	H	W-WRO	Gulp
H	L6660	Sippenaeken	BE	5.940761	50.75047	27/11/2012	04/05/2018	H	SPW	Geul

B.5 Population density

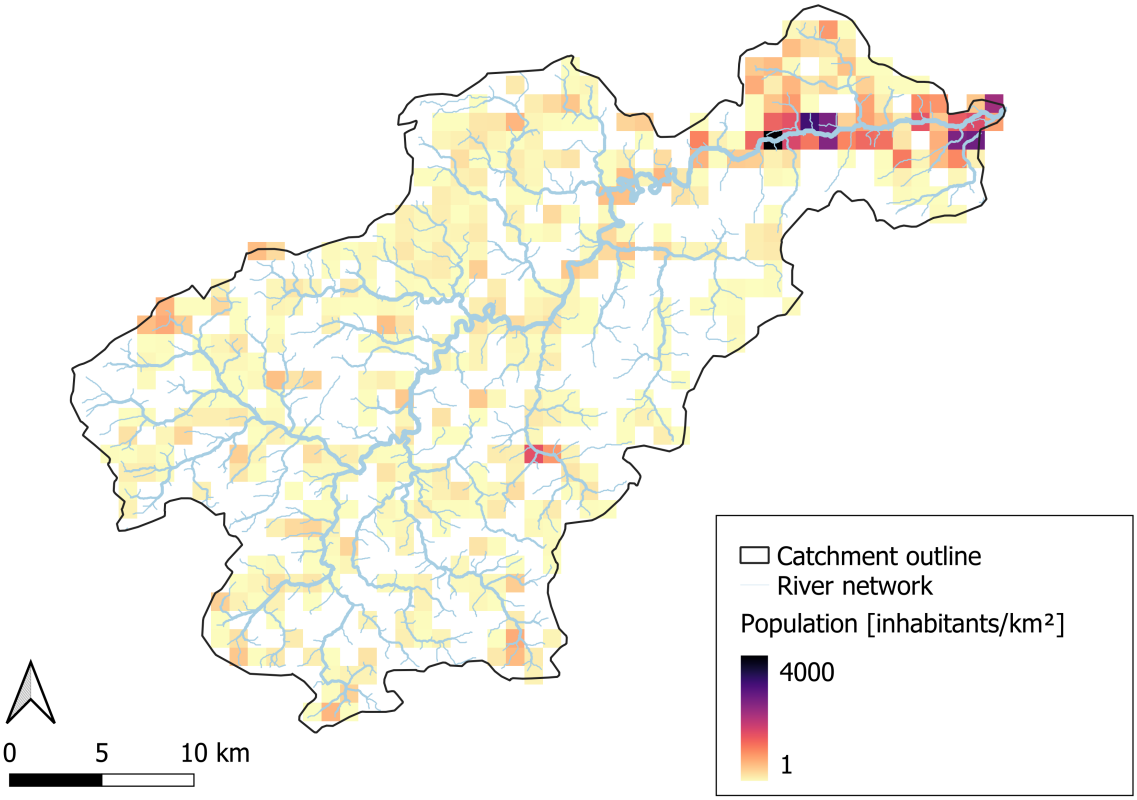


Figure B.10: Population density [$inhabitants/km^2$] of the Ahr catchment with the river network as overlay.

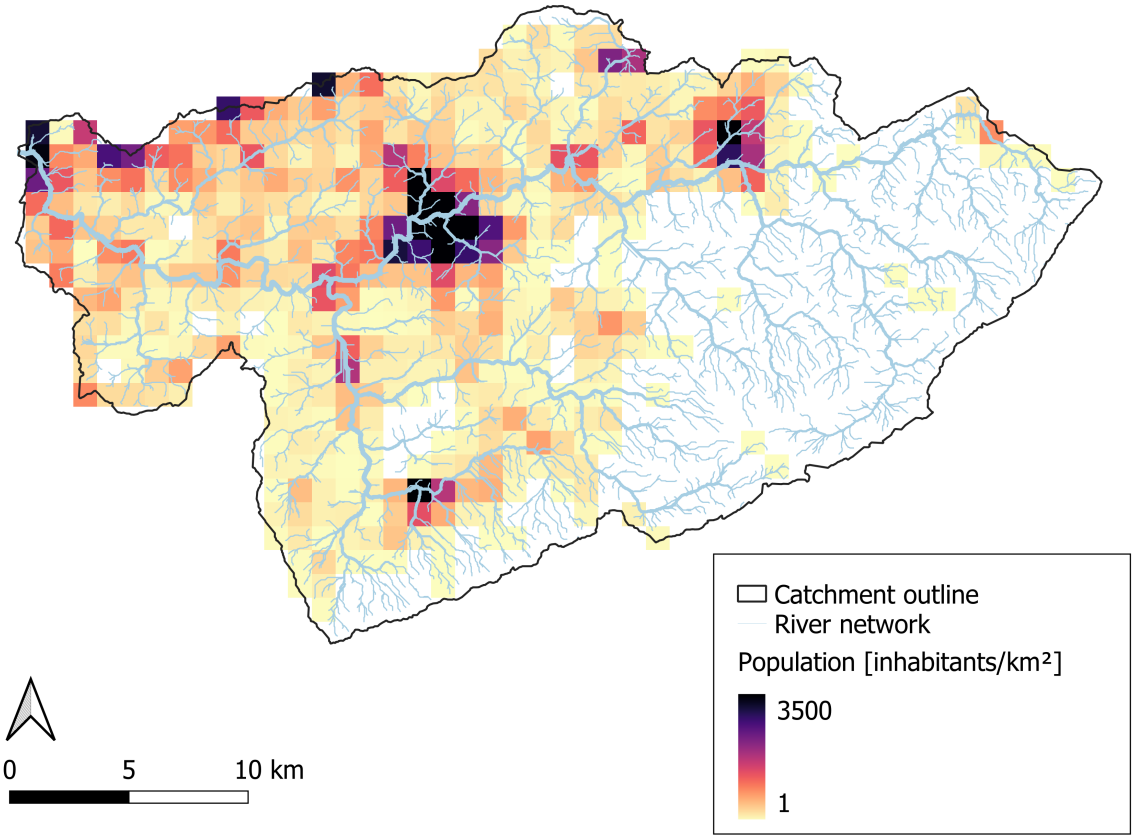


Figure B.11: Population density [inhabitants/km^2] of the Vesdre catchment with the river network as overlay.

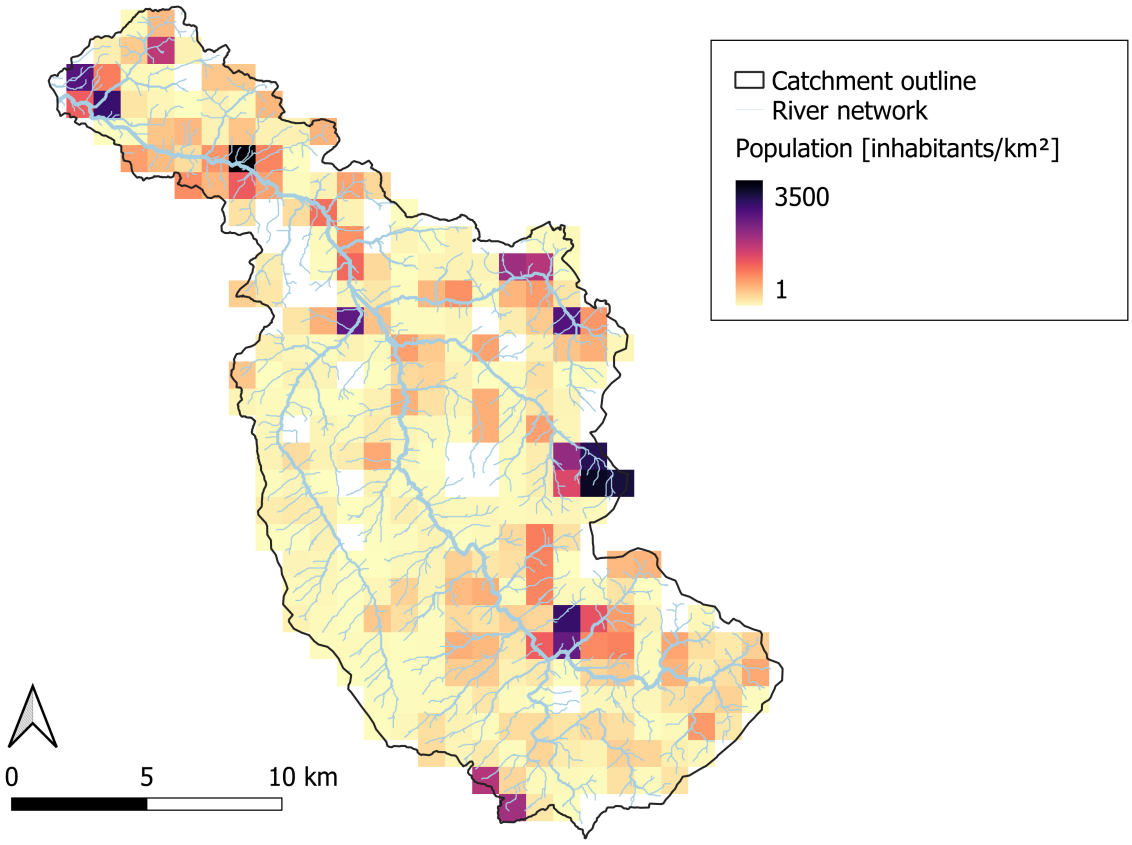


Figure B.12: Population density [$inhabitants/km^2$] of the Geul catchment with the river network as overlay.

Fig. B.10, fig. B.11 and fig. B.12 show how the most densely populated regions are often located near the river. Higher densities of population can especially be found near the mouth of the river. This proximity between the population and the river has effects in both directions. The urban areas require room for housing, transportation, et cetera. A wide river profile and curving path of an undisturbed river would decrease the available space and so, the available room of the river is often limited in favor of urbanization. On the other hand, the proximity of the population also increases their exposure to potential floods. This is especially the case for some of the narrow valleys in the catchments. Due to limited area suited for construction, buildings are often located near the river, or even in the winter bed of the river. This high exposure and limited flow area, leading to higher water levels, increase the flood risk in the catchments.

B.6 Budyko plot

Budyko (1974) proposed a simple supply-demand framework to visualise the long-term coupled water-energy balance. It simplifies the water and energy balance equations into two ratios that can be calculated and plotted for each subcatchment in the Budyko plot. On the horizontal axis, the mean annual climate aridity index is shown, which is the ratio of the potential evaporation and the precipitation. The vertical axis shows the mean annual evaporation ratio, the ratio of annual actual evaporation and the precipitation. Using the water balance equation, eq. B.3, with the assumption that the yearly change in storage is zero and that there are no other in- or outfluxes, the actual evaporation is equal to the precipitation minus the discharge and the mean annual evaporation ratio can be rewritten as eq. B.4.

$$\frac{dS}{dt} = P - Q - E_a \tag{B.3}$$

$$\frac{E_a}{P} = 1 - \frac{Q}{P} \tag{B.4}$$

The Budyko plot can be used for many purposes (Fu & Wang, 2019) and one of its main applications is to provide climatic context for catchments. The Budyko plot includes two physical limitations and the relevant limit indicates a wet or arid climate. The actual evaporation can be limited by the water supply or by the energy input. In the water-limiting condition, the aridity index is larger than 1, so the potential evaporation is higher than the precipitation. The actual evaporation of the catchment is then limited to the amount of precipitation, indicating an arid climate. In energy-limiting conditions, the aridity index is smaller than 1 and the actual evaporation is limited by the amount of energy, a characteristic of wet environments. As the study region of all three catchments in Western Europe has a temperate oceanic climate according to the Köppen-Geiger classification system, and is thus not arid, an energy-limited condition is expected with the according aridity index of less than 1.

The Budyko plot can also be used for data validation (Koppa & Gebremichael, 2017). If a catchment is plotted above either of the two limits, this can indicate a data error. Alternatively, it can indicate an incorrect assumption about the catchment. As the actual evaporation of the catchments is unknown, eq. B.4 replaced the actual evaporation with the discharge. The required assumptions of net-zero change in storage and no fluxes except discharge, precipitation and evaporation, can be invalid in case of high storage variability or additional fluxes. A potential extra flux in all three catchments is water extraction for the public drinking water supply (SWB, n.d.; Ziemetz et al., 2021; WML, 2021), as well as for commercial products such as bottled water (Apollinaris and Ahr Quelle from the Ahr, Spa and Chaudfontaine from the Vesdre) and beer (Brand from the Geul). Although the water is primarily extracted from deeper groundwater layers and the extraction rate is difficult to estimate, they can disturb the assumed water balance with no additional fluxes. Agriculture can be another source of water extraction, one that uses surface water. Additionally for the Vesdre, the water storage flux is not necessarily zero due to the large reservoirs. For validating the data input with the Budyko plot, the catchments should be plotted below the two limits but data errors and additional fluxes or changes in storage could cause deviations.

Plotting each subcatchment on the Budyko plot requires three values: the precipitation sum, potential evaporation sum and the discharge sum. Due to the importance of spatial detail for the subcatchments, the radar data has been used for the precipitation and potential evaporation data for the entire data period. The same delineated catchments and their derived data values have been used as for the HBV model. Although the Budyko plot is normally based on annual values, the complete radar dataset lengths are used. The resulting Budyko plot for all subcatchments with the total aridity index on the horizontal axis and evaporation ratio on the vertical axis is shown in fig. B.13.

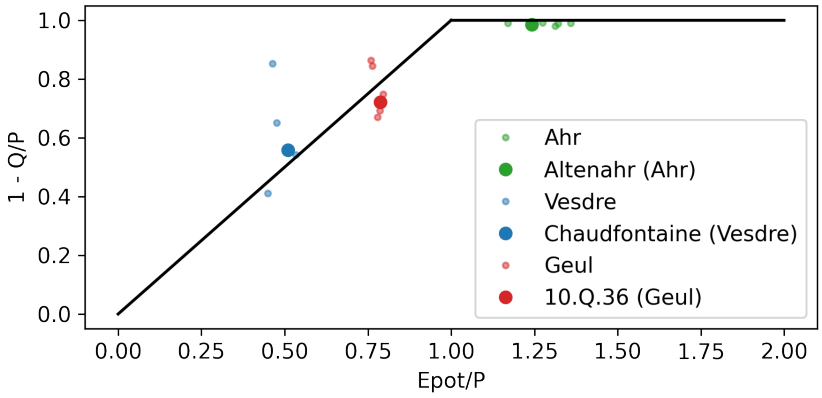


Figure B.13: Budyko plot for all stations of the Ahr, Vesdre and Geul. The main gauges are highlighted.

Figure B.13 shows that the Geul and Vesdre are indeed part of a wet climate due to the low aridity index which indicates them as energy-limited. There are some subcatchments that exceed the energy limit, which could point to data errors. For the Vesdre, the three catchments that lie above the limit are all downstream of the two reservoirs and their distance to the limit in fig. B.13 increases with decreasing distance to the reservoirs in reality. Since the subcatchment with the relatively largest influence of the reservoirs has the largest deviation, it is likely that the reservoirs are the cause of the deviation of the three stations. Coming back to the assumptions for the Budyko plot here, both the possible change in storage and extraction for drinking water could be the reason behind this influence of the reservoirs. For the Geul, the two subcatchments exceeding the limit are the Eyserbeek and Selzerbeek. Drinking water

is extracted from the Eyserbeek subcatchment, but this is only a small drinking water extraction zone. The Dutch part of the Geul has a large spread of artificial buffers that have been constructed in context of national climate adaptation plans. The total volume of water that these buffers can store is 781.000 $[m^3]$. Almost 60% of this total buffer volume is situated in these two subcatchments with deviating values (36% for Eyserbeek and 23 % for Selzerbeek) which creates a total buffer volume of almost 400.000 $[m^3]$. However, the total buffer volume is only 0.3% of the annual precipitation in the catchment and therefore, the influence of these buffers is expected to be limited. Considering all influences of reservoirs, unknown water extraction and buffers, which could explain the Budyko deviations, there is no clear evidence of data errors for the Vesdre or the Geul.

The Ahr is shown to be arid according to the Budyko plot, with a high aridity index and also a high evaporation ratio. A possibly incorrect delineation can not be the cause, since it would then overestimate precipitation and potential evapotranspiration, conserving its ratio. Furthermore, the delineated subcatchments have been checked for mistakes. Although it is difficult to see the lush, green hills of the Ahr as arid, the region receives with roughly 700 [mm/y] less rain than the Vesdre (1100 [mm/y]) and the Geul (900 [mm/y]). The ratios of those precipitation amounts also match with the distances on the aridity index axis between the catchments. Since the temperatures are similar on such small distances and the elevation difference is also not large, it can be assumed that the potential evaporation are similar for all catchments. With a lower precipitation amount, possibly due to orographic shading, the Ahr indeed is more arid than the other catchments and it is possible that for the Budyko plot, this is considered as arid. Although the arid characteristic of the Ahr catchments of the Budyko plot cannot be noticed in real life and thus could indicate a data error, the relatively low precipitation amount of the Ahr could explain this phenomenon. Again, there is no clear evidence of data errors from the Budyko plot.

B.7 Spatial interpolation of annual precipitation

This section gives a more detailed explanation of the mean annual precipitation amounts for the stations with over 10 years of measurements and its spatial interpolation in fig. 2.6. The annual precipitation sums for years with less than 10 days of missing data have been averaged to compute the mean annual precipitation amount of each measurement station. The Thiessen polygons have been determined by applying the standard Vector geometry tool of “Voronoi Polygons” in QGIS to the stations. Dividing the area of each Thiessen polygon by the total catchment area gives the areal fraction of each station. The locations of the stations on the map can be found in fig. B.14. More information about the stations and their Thiessen polygons can be found in table B.4, including the number of valid years for the average and the areal fractions.

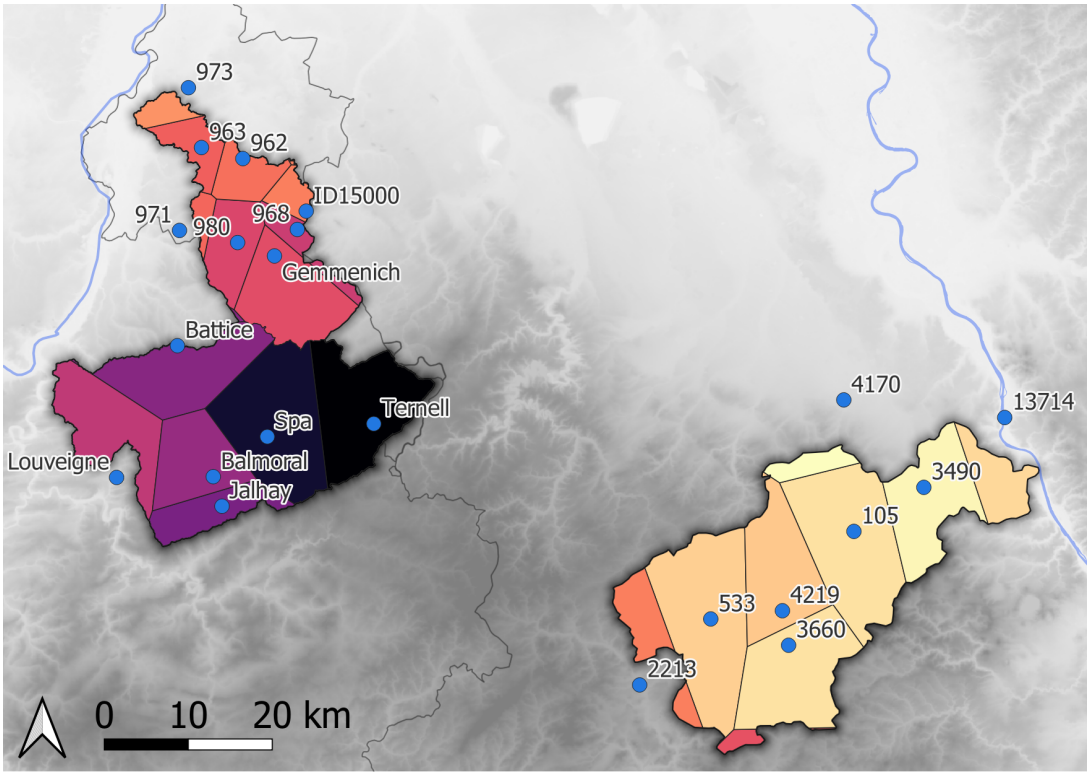


Figure B.14: Supplement of fig. 2.6 to show the locations and names of the stations on which the Thiessen polygons are based.

Table B.4: Information on the stations of fig. B.14, including the catchment, town, x-coordinate (X), y-coordinate (Y), starting year of the dataset (Start), final year of the dataset (End), average annual precipitation (Average), number of years with sufficient data (Years) and the fraction of the area of the Thiessen polygon to the total area (Areal fraction).

Catchment	Station	Town	X	Y	Start	End	Average	Years	Areal fraction
Ahr	105	Ahrbrück	6.971023	50.48449	1941	2021	675	76	0.21
Ahr	533	Blankenheim	6.737474	50.38441	1961	2021	699	58	0.22
Ahr	2213	Lissendorf	6.623578	50.31052	1931	2021	808	86	0.06
Ahr	3225	Mehren	6.876649	50.18203	1971	2018	879	47	0.01
Ahr	3490	Neuenahr	7.085337	50.53456	1931	2021	647	84	0.13
Ahr	4170	Rheinbach	6.94496	50.62444	1931	2020	635	87	0.03
Ahr	4219	Rodder	6.856804	50.39655	1931	2021	709	88	0.13
Ahr	3660	Nurburg-Barweiler	6.86954	50.35999	1995	2021	673	25	0.16
Ahr	13714	Linz-Unkel	7.216696	50.61251	2007	2021	686	13	0.06
Geul	980	Epen	5.911562	50.76075	1980	2021	902	41	0.18
Geul	Gemmenich	Gemmenich	5.974898	50.74865	2021	2021	889	20	0.35
Geul	968	Vaals	6.011002	50.77824	1951	2021	929	71	0.08
Geul	ID15000	Aachen	6.02440	50.7983	2011	2021	808	10	0.06
Geul	962	Ubachsberg	5.912579	50.85064	1952	2021	830	66	0.12
Geul	971	Noorbeek	5.812443	50.77015	1951	2021	843	71	0.04
Geul	963	Valkenburg	5.841672	50.85993	1951	2021	854	71	0.11
Geul	973	Beek	5.813848	50.92296	1951	2021	782	71	0.06
Vesdre	Ternell	Ternell	6.157033	50.57528	2002	2021	1286	20	0.21
Vesdre	Jalhay	Jalhay	5.909887	50.47854	2002	2021	1228	20	0.25
Vesdre	Spa	Spa	5.979466	50.55537	2017	2021	1057	5	0.08
Vesdre	Battice	Battice	5.819987	50.64707	2002	2021	1034	20	0.17
Vesdre	Louveigné	Louveigné	5.730506	50.50277	2002	2021	946	20	0.16
Vesdre	Balmoral	Balmoral	5.892437	50.50951	2002	2016	1010	14	0.13

B.8 Flow duration curves

The flow duration curves of fig. 2.10 are the average curves of the annual flow duration curves. Fig. B.15, fig. B.16 and fig. B.17 show these annual flow duration curve for the Ahr, Vesdre and Geul. Only the years with more than 85% of data for that year are included. For each year, the daily discharge is sorted in descending order. The probability of exceedance is then calculated for each sorted discharge, resulting in the figures below.

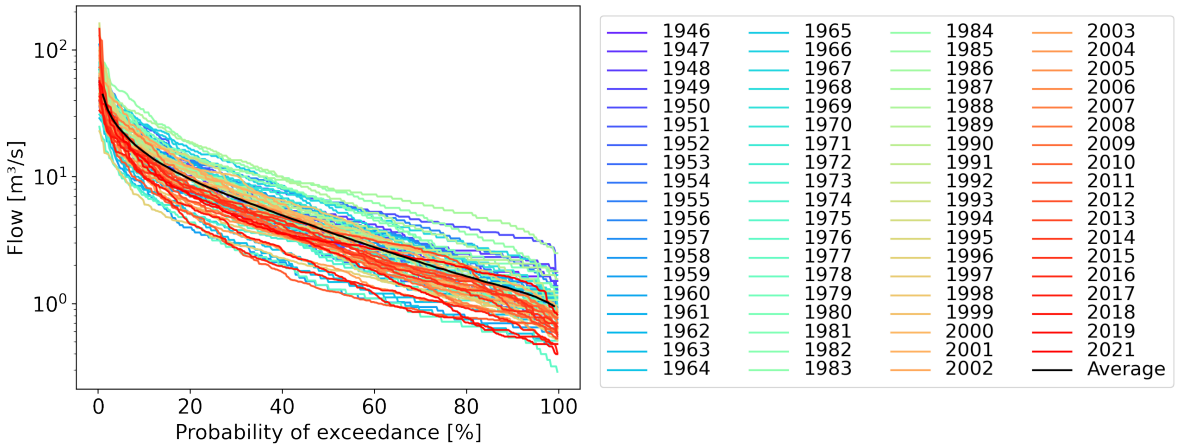


Figure B.15: Flow duration curves for the Altenahr station of the Ahr for the years with a sufficient number of measurements. The black lines indicates the average, as shown in fig. 2.10

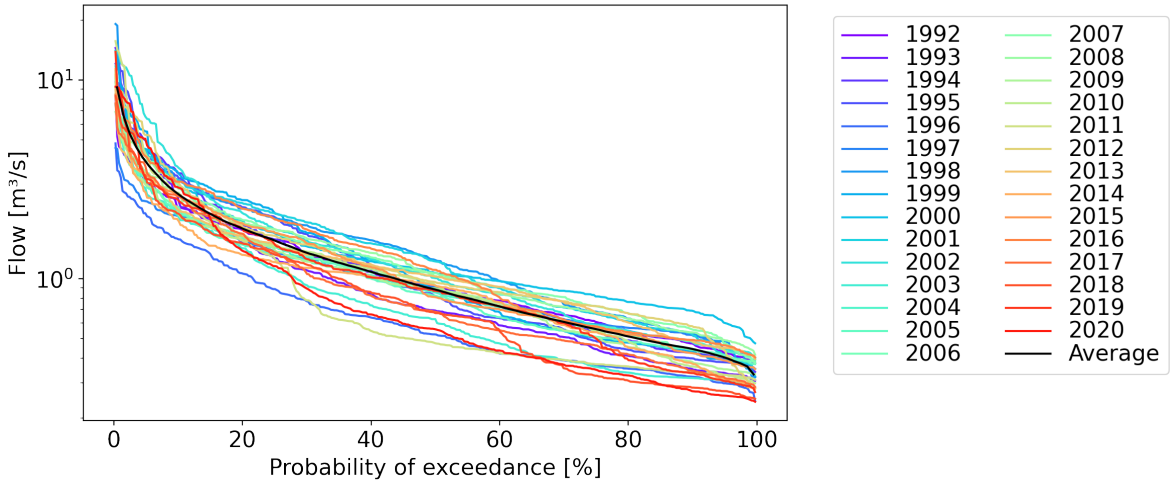


Figure B.16: Flow duration curves for the Chaudfontaine station of the Vesdre for the years with a sufficient number of measurements. The black lines indicates the average, as shown in fig. 2.10

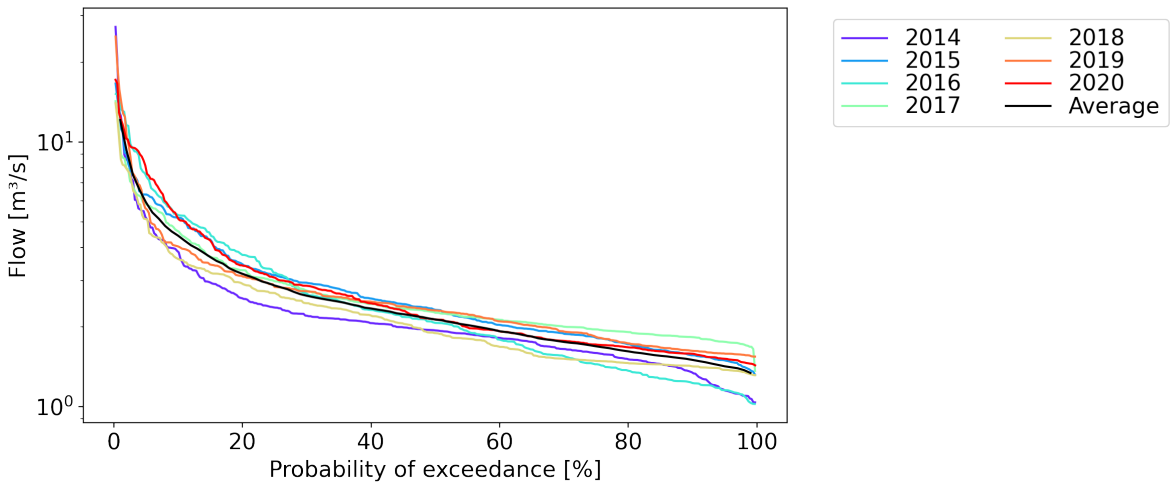


Figure B.17: Flow duration curves for the Meerssen (10.Q.36) station of the Geul for the years with a sufficient number of measurements. The black lines indicates the average, as shown in fig. 2.10

The yearly flow duration curves show that there is quite some variation in the shape of the curve for each year. This is expected since some years are drier or wetter than others. There seems to be no clear climatic pattern in this variation. The Meerssen only has a short data period, and so it is difficult to detect any pattern. Since the flow duration curves do not have a large spread, it can be assumed that the short data period does not lead to an inaccurate flow duration curve for the Geul.

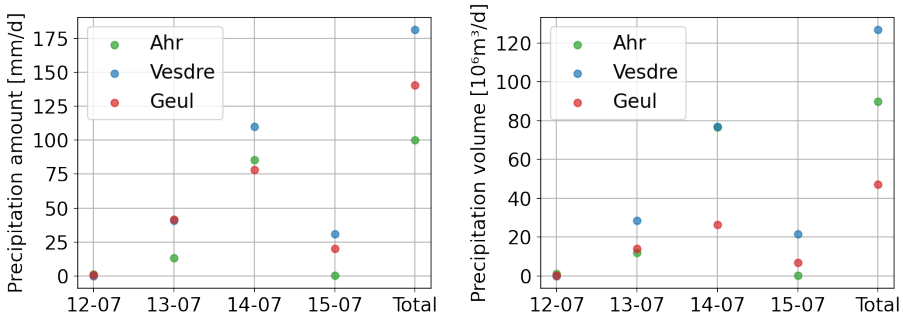
Appendix C

Precipitation characteristics

This appendix discusses several precipitation characteristics, both for the 2021 flood event itself and the extremity of that event. The precipitation totals of the 2021 flood events are analysed. The potential extreme aspects of the temporal variation, spatial variation and pre-conditions are discussed in more depth, as well as how they have been derived.

C.1 Total precipitation per catchment for the 2021 flood event

The total event precipitation for a catchment can be derived from radar products or through interpolation of station measurements. The radar products that are available for the study regions are described and checked in section 5.4. Since the radar products have a better spatial resolution than a station interpolation would have, they are the preferred method. The precipitation amounts per catchment are clipped from the radar products with the hydromt Python package and the catchment boundaries.



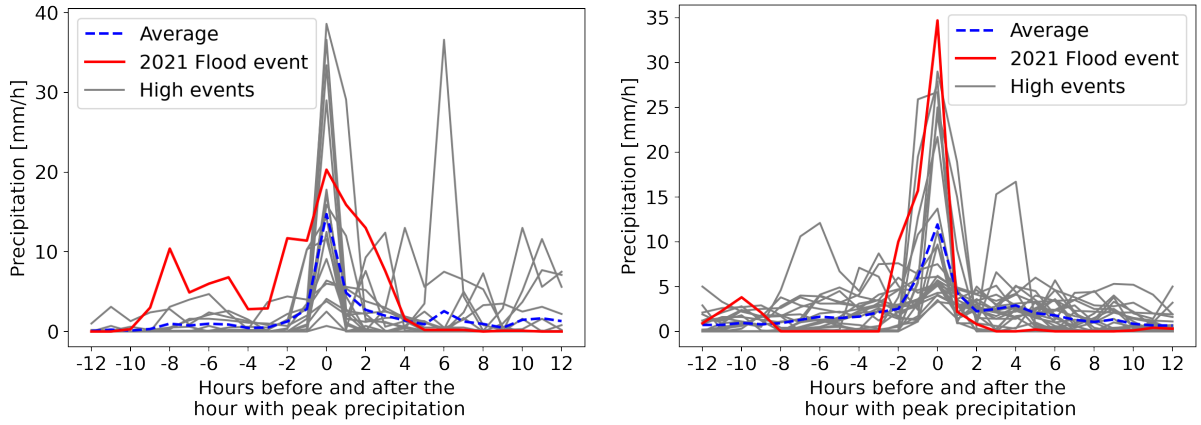
(a) Precipitation amount *(b) Precipitation volume*

Figure C.1: Precipitation volumes for the entire catchments of the Ahr, Vesdre and Geul per day and in total, based on the radar data discussed in section 5.4. The units for the total amounts are not per day, as indicated on the vertical axes, but per 4 days.

The precipitation totals can be expressed in the units of $[mm/d]$, indicating the average precipitation height in the catchment, or in the units of $[m^3/day]$, indicating the total volume. The spatial pattern of the precipitation from fig. 3.6 shows that the rain on 13 July was mostly situated near the Geul and Vesdre, and that the rain on 14 July was mostly situated near the Ahr and Vesdre. This pattern can also be seen in the precipitation amounts of fig. C.1. The precipitation amounts on 13 and 14 July of fig. 3.6 showed that the precipitation of the Geul mostly fell in the Belgian part. Although the Dutch part received less rain than any other catchment area according to that same figure, the total precipitation amount of the Geul is still higher than for the Ahr. The high total precipitation in the Geul thus strongly originated from the Belgian part. Due to the differences in catchment size, the total precipitation volume is higher for the Ahr than the Geul despite the lower precipitation amount. The Vesdre has the highest precipitation volume for the entire event, followed by the Ahr and finally the Geul.

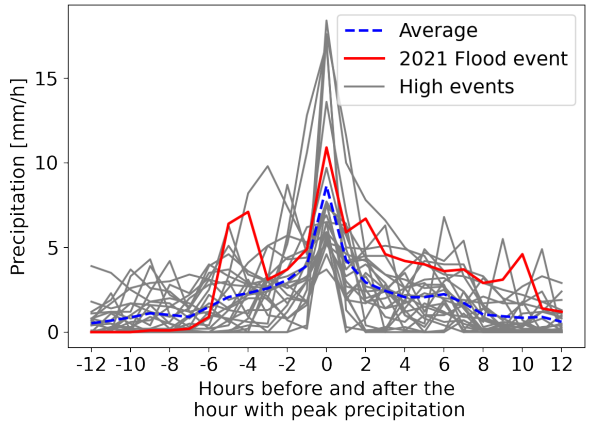
C.2 Temporal variation

The temporal variations of different stations show different patterns, but there is no clear temporal pattern for the 2021 flood event for all stations or a clear deviation from usual patterns.

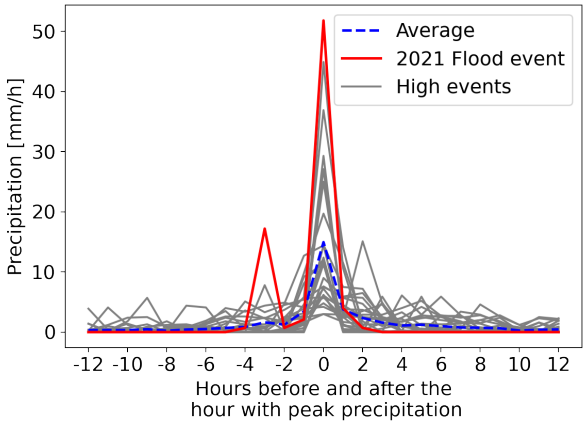


(a) Ahr (2213)

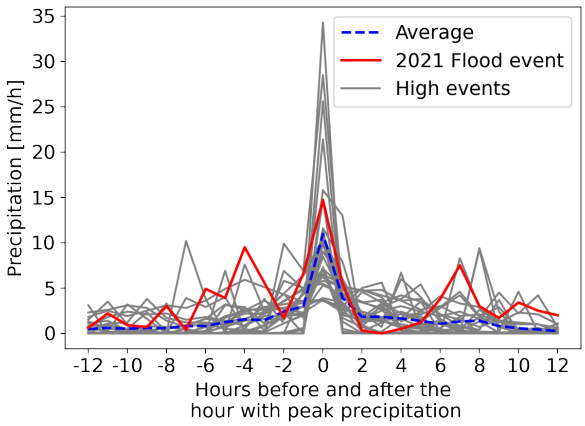
(b) Vesdre (Ternell)



(c) Vesdre (Jalhay)



(d) Vesdre (Louveigné)



(e) Geul (Gemmenich)

Figure C.2: Hyetographs for the stations not included in fig. 4.3. The hyetograph is given in grey for the 30 days with the maximum daily precipitation amount of the record period of that gauge. The hourly maximum of that day is central with 12 hours before and after that hour to show the temporal variation of an extreme rain event.

C.3 Spatial variation

Two spatial data sets can be compared in a quantitative manner with the use of correlation coefficients. In this case, Pearson's correlation coefficient, r, is used to calculate the correlation between the average annual precipitation and total precipitation amount of the 2021 flood event. This section provides the quantification of fig. 4.2.

The average annual and event precipitation amounts are compared for stations in all catchments. Most stations that measured the full event formed the base stations for deriving the Thiessen polygons in fig. 2.6 and fig. 4.2. The event and annual precipitation amount could be compared directly. Several stations in the Geul catchment could not be included in the long-term pattern of the Thiessen polygons as their data length was limited to the flood event. Their event precipitation amount is compared to the average annual precipitation amount of the Thiessen polygon they are located in. In case the station is located outside the catchment and thus the Thiessen polygon, they are included in the comparison if a nearby station (distance of less than 2 km) forms the base of a Thiessen polygon. With those requirements, 24 stations remained for the quantitative comparison.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{C.1}$$

The Pearson correlation coefficient for a sample is given in eq. (C.1). The equation for a sample is used since the stations resemble only part of the precipitation locations, hence the use of a sample equation and not the population equation which would represent all locations. The mean of the average annual precipitation does not include multiple stations twice, although some are included twice as the reference for several stations that had not been included in the Thiessen polygons, as mentioned earlier. Filling in eq. (C.1) with the annual average and event precipitation amounts for each station and the means, see table C.1, results in a correlation coefficient of 0.356.

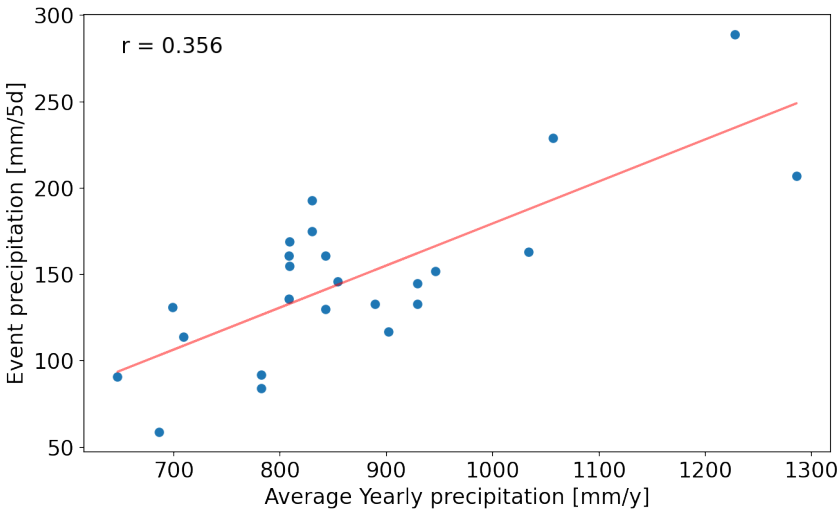


Figure C.3: Scatter plot of the yearly average precipitation and the total precipitation amount during the 2021 flood event. The trend line of the scatter plot is given in red and the figure includes the value of the Pearson correlation coefficient, r.

A correlation coefficient between 0.3 and 0.5 shows a medium positive correlation. Fig C.3 visualises this correlation through a scatter plot and the trend line which show that the average annual and event precipitation amount are indeed positively correlated. High event precipitation amounts are mostly found in stations with high average annual precipitation. Considering that the correlation coefficient can be between -1 for a perfect negative correlation and +1 for a perfect positive correlation, the medium positive correlation of 0.356 shows that the spatial pattern of the event precipitation was not extreme, since this would require a very small positive or even negative correlation.

Table C.1: Annual average and event precipitation amounts for all stations used in the computation of the correlation coefficient. The averages are given at the bottom of the table. The values between brackets are not used in the average but they are used for comparing x_i and y_i

Station for x_i	**Station for y_i**	**Annual average precipitation (x_i) [mm/y]**	**Event precipitation (y_i) [mm/5d]**
533	533	699	131
2213	2213	808	136
3490	3490	647	91
4219	4219	709	114
13714	13714	686	59
Ternell	Ternell	1286	207
Jalhay	Jalhay	1228	289
Spa	Spa	1057	229
Battice	Battice	1034	163
Louveigné	Louveigné	946	152
980	980	902	117
Gemmenich	Gemmenich	889	133
968	968	929	133
968	12.P.25	(929)	145
ID15000	ID15000	808	161
962	962	830	193
962	10.P.30	(830)	175
971	971	843	161
971	15.P.41	(843)	130
963	963	854	146
973	973	782	92
973	6.P.40	(782)	84
965	965	809	169
965	18.P.37	(809)	155
	Mean	**881**	**149**

C.4 Pre-conditions

This appendix explains in more detail how the pre-conditions have been derived, which limitations are applied and how varying these limitations influences the results. For the latter, each limitation is varied individually while conserving the original values of the other limitations.

The pre-conditions of fig. 4.4 consist of the rain during the day of a high precipitation event and the accumulated rain in the 30 days preceding that event. High precipitation events were defined by a daily precipitation amount of more than 30 $[mm/d]$. These events were selected for all relevant stations of the Ahr, Vesdre and Geul. The stations were classified as relevant when they had daily precipitation measurements for more than 10 years and contained measurements for the 2021 flood event. This led to the inclusion of the following stations:

- For the Ahr (4): 3490, 533, 2213, 13714
- For the Vesdre (4): Battice, Ternell, Louveigné, Jalhay
- For the Geul (12): P_962, P_963, P_965, P_966, P_968, P_969, P_971, P_973, P_980, P_981, Aachen_D, Gemmenich

The high precipitation events of all stations in the same catchment were combined into one data set. A single precipitation event can thus be included from multiple stations. For all events in each catchment, the 90th percentile of preceding precipitation has been computed, indicated by the dashed line in fig. 4.4. If the precipitation amount of the 2021 flood event for a station exceeded this 90th percentile, it is

classified as extreme. If the majority of the stations show extreme pre-conditions, the pre-conditions of the entire catchment can be considered as extreme.

Since there can be two consecutive days with precipitation amounts over 30 $[mm/d]$, only the first day has been accounted for in fig. 4.4, since the first day indicates the start of the event. However, this method can lead to the exclusion of the day with the highest precipitation during the 2021 flood event. If the precipitation amount of July 13 already exceeded the limit of 30 $[mm/d]$, the higher precipitation amount of July 14 is not included in the pre-conditions.

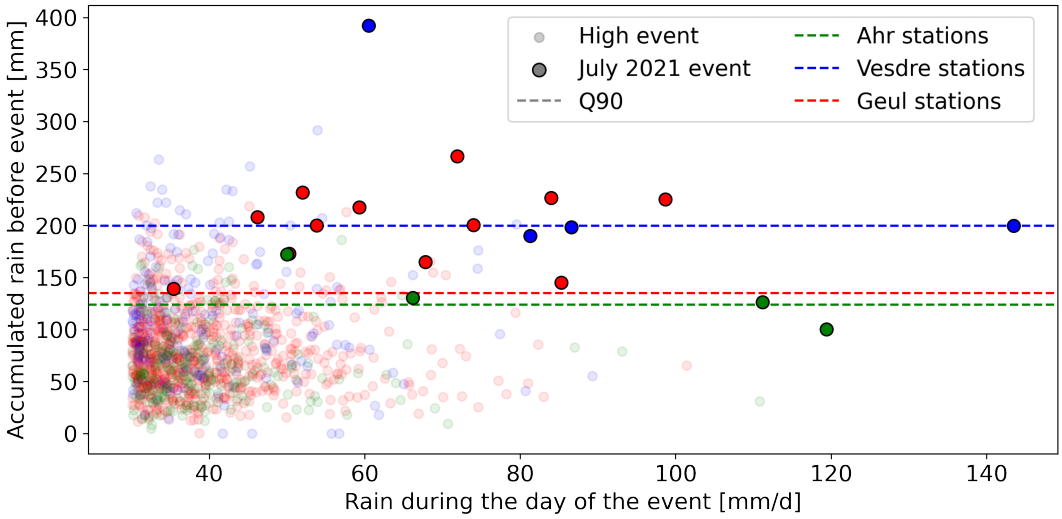


Figure C.4: Pre-conditions for the relevant stations of the Ahr (green), Vesdre (blue) and Geul (red), similarly to fig. 4.4. Instead of using the first day of consecutive high precipitation events, this figure is based on the use of the second day. Except for this change, the method is identical.

The original figure of fig. 4.4 showed that 75% of the stations of the Ahr and the Geul exceeded the 90th percentile limit. None of the Vesdre stations did. Fig. C.4 accounts for the second day of two consecutive events, which shows an overall increase of stations exceeding the limit. For the Geul, the number of stations exceeding the limit goes from 75% to 100%, so all stations had extreme pre-conditions. For the Vesdre, 50% of the stations exceed the limit instead of none. For the Ahr, there are no changes. Accounting for the second day of consecutive events leads to more stations exceeding the 90th percentile limit and thus more stations with extreme pre-conditions.

There are three additional limitations, or design decisions, that were applied for fig. 4.4 which influence the results. Firstly, the 90th percentile was used. This shows the limit which exceeds 90% of all values. There is no exact limit in literature to account for extreme pre-conditions. Some researchers or policy makers could already consider 70% as an extreme limit. Such a lower limit which would likely result in more stations with extreme pre-conditions. Varying the percentile limit for the definition of extreme pre-conditions could give additional insights.

Fig. C.5 confirms the expectation that a lower percentile limit results in more stations with extreme pre-conditions. If the 80th percentile is used as a limit, all stations for the Ahr and Geul, as well as 75% of the Vesdre, show extreme pre-conditions. This decreases strongly for higher percentile limits. The original figure of the pre-conditions would lead to the conclusion that the Vesdre had no extreme pre-conditions since none of the stations exceeded the 90th percentile. However, the lower 80th percentile results in extreme pre-conditions for 3 of 4 stations. This shows that the Vesdre may not have had extreme pre-conditions according to the 90th percentile limit, but they were still very high. The catchments of the Ahr and Geul also show a stronger extremity for the lower percentile limits. It can be concluded that maybe not all stations show extreme pre-conditions according to the 90th percentile limit, but they are still unusually high.

The Vesdre shows an interesting pattern with a sudden jump from 0 to 75% of extreme stations when the extrimity limit increases from the 80th to the 85th percentile. This is due to the fact that three out of four stations of the Vesdre are closely located near these limits, as can be seen in fig. C.4. The similar behaviour of 75% of the stations explains the suddenly large changes in extreme stations. This is also

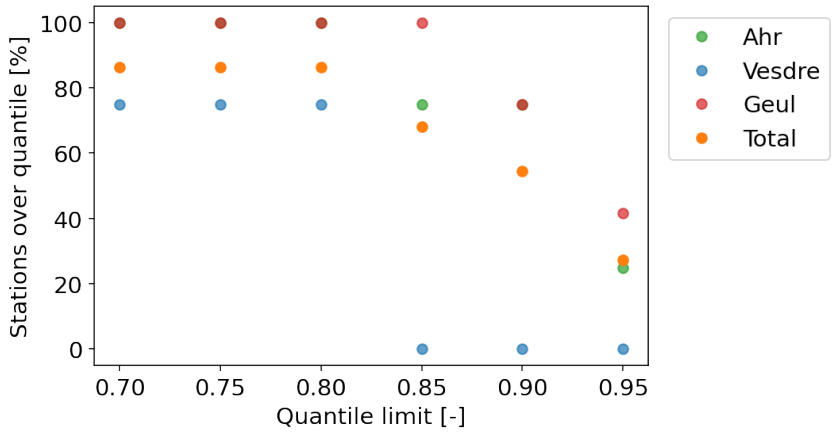


Figure C.5: Influence of the percentile limit on the extreme pre-conditions. The variable percentile limit (ranging from 0.7 to 0.95) is given on the horizontal axis. The percentage of stations that exceed the variable percentile limit is given on the vertical axis. The percentages are given for each catchment and the all catchments (Total). The markers are slightly transparent to visualise overlapping markers.

the case for the following analyses.

Another limitation is that the pre-conditions are considered for 30 days preceding the event. Although the consideration of 30 days is applied by multiple studies (Heggen, 2001; Schröter et al., 2015; Mohr et al., 2022), it is still important to know if the pre-conditions were also extreme if the preceding period were longer or shorter.

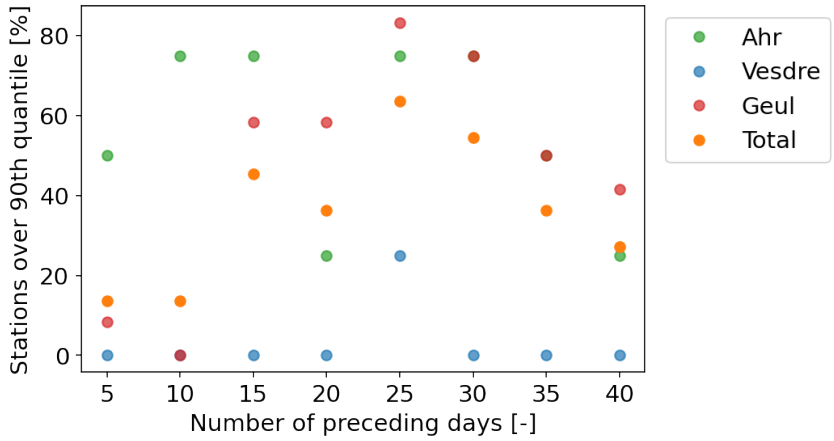


Figure C.6: Influence of the duration of the preceding precipitation on the extreme pre-conditions. The variable duration (ranging from 5 to 40 days) is given on the horizontal axis. The percentage of stations that exceed the 90th percentile limit is given on the vertical axis. The percentages are given for each catchment and the all catchments (Total). The markers are slightly transparent to visualise overlapping markers.

Fig. C.6 indicates that the Ahr had extreme pre-conditions for many stations for a shorter period, starting from 10 days. For the Geul, the majority of stations with extreme pre-conditions started after using a period of 15 days. So even a shorter period than 30 days results in extreme pre-conditions for the Ahr and the Geul. For a period longer than 30 days, the amount of stations with extreme pre-conditions started to decrease again. The choice of 30 days is representative for the overall results.

In the original figure, high precipitation events were considered if the daily precipitation amount exceeded 30 $[mm/d]$. A lower limit would include more events, but taking lower precipitation events into account also leads to less influence of the high precipitation events on the percentile limit. A higher limit would lead to less events and maybe even too little events to compute a representative 90th percentile.

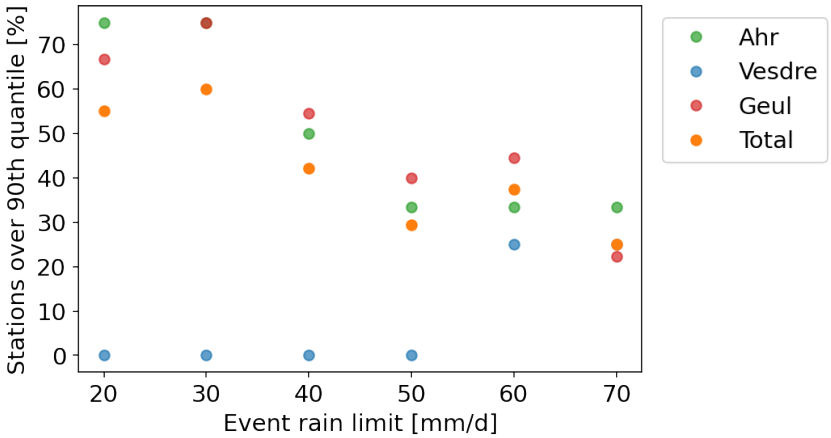


Figure C.7: Influence of the rain event limit on the extreme pre-conditions. The variable rain event limit (ranging from 20 to 70 $[mm/d]$) is given on the horizontal axis. The percentage of stations that exceed the 90th percentile limit is given on the vertical axis. The percentages are given for each catchment and the all catchments (Total). The markers are slightly transparent to visualise overlapping markers.

Fig. C.7 shows a decrease in stations exceeding the 90th percentile for an increasing event rain limit. This decrease can be explained by the decrease in the number of events for each catchments when the event rain limit is increased, as shown in fig. C.8. The event rain limit of 30 $[mm/d]$ seems fitting since each catchments has more than 100 events and no more than 2000. A higher rain limit and less events would lead to a sample smaller than 100 to derive a 90th percentile, which is considered as too small.

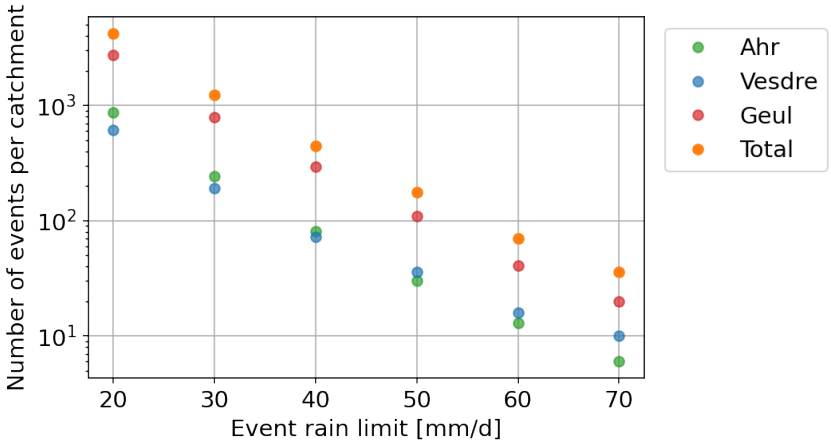


Figure C.8: The number of events that are included for each catchment and all catchments when the event rain limit is changed. The horizontal axis indicates the event rain limit (ranging from 20 to 70 $[mm/d]$). The vertical axis shows the number of events included for the variable rain limits.

The original figure of the pre-conditions, fig. 4.4, showed that 75% percent of the Ahr and Geul stations had preceding precipitation over the 90th percentile of each catchment. With the majority of the stations showing extreme pre-conditions, both the Ahr and Geul catchment can be considered as having extreme

pre-conditions. If the second day of consecutive events is included instead of the first day, 50% of the Vesdre stations have extreme pre-conditions instead of none. If the 90th percentile limit is lowered to the 80th percentile, the Vesdre also has a majority of the stations exceeding the limit. Even if not all stations are above the 90th percentile and thus extreme, their pre-conditions are still unusually high. Changing the duration of the preceding precipitation from 30 days, results in similar results for 15 to 30 days. The 30 day assumption can thus be considered as representative. The rain event limit of 30 $[mm/d]$ is also representative, since it balances the amount of included events. Although fig. 4.4 used several limitations, the conclusion of extreme pre-conditions for the Ahr and Geul still holds and can be extended by the fact the pre-conditions were unusually high for all catchments.

Appendix D

Return periods

This appendix discusses the return periods of the forcing and response magnitude of the 2021 flood event. Firstly, it explains how the return periods are derived with the use of the GEV distribution. The return period of the forcing magnitude is quantified by the return period of the 1-day precipitation amount of 14 July. The uncertainty ranges of the return periods are also provided, as well as the return periods of the 2-day and 3-day precipitation amounts. For the response, the return periods of the peak discharge are given, as well as the uncertainty range.

D.1 Deriving return periods with the GEV distribution

The GEV distribution has been computed according to the method of the course CIE5450, Hydrology of Catchments, Rivers and Landscapes. This method has been applied to both the precipitation and the discharge data sets. First, each data set is transformed to a data set of the daily maximum value. These maximum values are then processed by using functions of the "mevpy" Python package, developed by Enrico Zorzetto (https://github.com/EnricoZorzetto/mevpy). The final result is the GEV distribution and its return period estimates.

Detailed information on the mevpy package and its functions can be found at the link above. The package also includes the MEV (Metastatistical Extreme Value) distribution, but initial test results indicated poor performance of replicating return periods from the observations. The GEV distribution can be derived with multiple methods. The most common methods have been applied, as was done in CIE5450. A short explanation is given here. For selecting the extreme values from the transformed dataset, the method of block maxima is selected, which takes the maximum value of each year into account. An alternative would be to use the peak-over-threshold method, which selects all values of more than a certain threshold value, but this method is more suited for the Pareto distribution (Engeland et al., 2004). The distribution of the extreme values that have been selected by the block maxima can be approximated with the GEV distribution by estimating its three parameters.

The two most common methods to derive the parameters of the GEV distribution are the method of maximum likelihood and L-moments (Ailliot et al., 2011). J. R. M. Hosking et al. (1985) have shown that the method of maximum likelihood is outperformed by the use of probability-weighted moments (PWMs), especially since it led to a smaller variance. Several years later, the use of PWMs formed the base of the method of L-moments (J. R. Hosking, 1990). Although more recent parameter estimators show improvements to the original L-moments (Engeland et al., 2004), they are not included in the mevpy package and therefore, are not applied here. Due to its slightly better performance, the method of L-moments was chosen for deriving the GEV parameters.

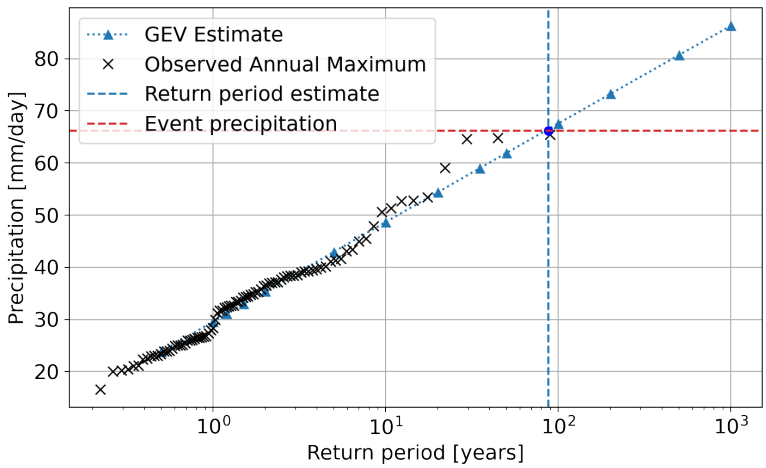


Figure D.1: Plot of return periods of the annual maxima and the GEV distribution for the precipitation of station 3490 in the Ahr. The dashed lines give the precipitation and its return period for the 2021 flood event.

Fig. D.1 gives an example of the resulting return period for the 2021 flood event. The return periods of the GEV distribution follow the observations well. The return periods of the highest observed maxima seem overestimated by the GEV estimate. This is caused by the limitation of the maximum return period of the observations, which can only be as high as the length of the record period. The GEV distribution allows for extrapolation beyond these years and therefore gives a higher return period estimate for the higher precipitation events. The intersection of the precipitation of the 2021 flood event and the GEV estimate results in a return period of 88 years.

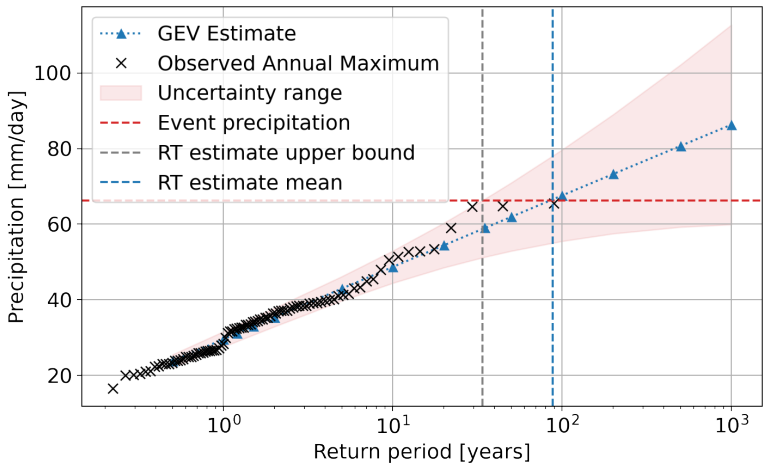


Figure D.2: Plot of return periods of the annual maxima and the GEV distribution for the precipitation of station 3490 in the Ahr. The uncertainty bounds of the 95th percentile have been added. The dashed lines give the precipitation and the return period for the 2021 flood event for the mean GEV estimate (RT estimate mean) and the upper bound (RT estimate upper bound).

The derived return period of 88 years from fig. D.1 is not a certain value. The 95% confidence intervals have been added in fig. D.2. These bounds indicate the range of values in which the true mean of the sample lies with 95% confidence. The upper bound results in a return period of 34 years, which is lower than the mean estimate of 88 years. The lower bound results in a higher return period of more than 1000 years.

The uncertainty bounds are computed with the mevpy package. A bootstrap method randomly samples 1000 data sets of the original annual maxima and computes their GEV distribution parameters. The covariance matrix of the 1000 estimates of each of the 3 parameters is computed. A final function, gev_quant, uses this covariance matrix of the parameter estimates to compute the 95% confidence intervals of the return periods.

The minimum and maximum ranges of the return periods are provided for the precipitation and discharge in table D.2 and table D.6. The method described above is applied in the same way to both parameters for each catchment. All data sets have an hourly timestep, so the daily maximum value is the maximum hourly value of that day. The only exception is the discharge of the Ahr, which has a daily timestep. The daily maximum value is then simply the daily average value. All GEV distributions are estimated from the data period that starts at the beginning of the observations and ends at June 2021. This data period excludes the event, to identify the extremity of the 2021 flood event compared to the previous conditions. The number of years with significant data gaps are negligible or at the start of the observations, which are then not included.

D.2 Return period estimates of the precipitation

This section provides the table of return period values for table D.1 with additional information, the 95% confidence intervals of those values and the return periods for the 2-day and 3-day precipitation sum.

Table D.1: The return periods for relevant stations of all catchments are provided with the peak value of the precipitation amount of 14 July. The RT_ev column has been computed by including the flood event in the statistics. In RT_nev the event is excluded. The last column is the main return period estimate and is used for fig. 4.1. The number of years for which the GEV distribution is computed is also given for each station.

Catchment	Station	Town	Years	Peak value	RT_ev	RT_nev
Ahr	533	Blankenheim-Ahrhütte	61	111	450	>1000
Ahr	2213	Lissendorf	89	119	249	730
Ahr	3490	Neuenahr	89	66	63	88
Ahr	13714	Linz-Unkel	15	50	4	4
Vesdre	Ternell	Ternell	20	144	82	>1000
Vesdre	Jalhay	Jalhay	20	172	74	>1000
Vesdre	Battice	Battice	20	81	26	43
Vesdre	Louveigné	Louveigné	20	87	23	40
Geul	380	Maastricht	64	32	2	2
Geul	962	Ubachsberg	70	98	858	>1000
Geul	963	Valkenburg	71	80	90	156
Geul	965	Schaesberg	71	86	350	>1000
Geul	966	Schinnen	71	36	2	2
Geul	968	Vaals	71	51	6	7
Geul	971	Noorbeek	71	88	58	80
Geul	973	Beek	71	36	2	2
Geul	980	Epen	42	42	3	4
Geul	981	Oost Maarland	24	55	30	64
Geul	Kanne	Kanne	15	35	1	2
Geul	Gemmenich	Gemmenich	20	85	47	278

The values of the return periods in fig. 4.1 are given in table D.1. The return periods of the precipitation amount for 14 July are calculated for two reference periods: one that includes the 2021 flood event (RT_ev) and one that does not include the event (RT_nev). Since the inclusive reference period has an additional high precipitation event, the return period from this reference period would be lower. This is indeed the case as can be seen from fig. 4.1. Fig. 4.1 is based on the reference period excluding the event, column RT_nev, to compare the flood event to the precipitation characteristics preceding the event.

Table D.2 gives the upper and lower bounds of the 95% confidence intervals of the return periods for the 1-day precipitation. The upper bounds of nearly all stations are over 1000 years, except for the mean return periods that were below 10 years. These exceptionally low return period estimates remain low within the uncertainty ranges. This means that it can be stated with confidence that several stations of the Geul catchment had very low return periods, on the condition that the measurements were correct. As the data has been collected from the authority, this condition can be assumed to be satisfied. The lower bounds are considerably lower than the mean, even by an order of magnitude in some cases. Even several lower bounds of this large confidence interval still indicate return periods from 300 to over 1000

Table D.2: 95% confidence intervals of the precipitation return periods. The mean value is the same as for table D.1. The "Min" and "Max" value are the minimum and maximum return period according to the confidence interval.

Catchment	**Station**	**Min**	**Mean**	**Max**
Ahr	533	>1000	>1000	>1000
Ahr	2213	117	730	>1000
Ahr	3490	33	88	>1000
Ahr	13714	2	4	>1000
Vesdre	Ternell	>1000	>1000	>1000
Vesdre	Jalhay	361	>1000	>1000
Vesdre	Battice	14	43	>1000
Vesdre	Louveigné	19	40	>1000
Geul	380	1	2	3
Geul	962	879	>1000	>1000
Geul	963	47	156	>1000
Geul	965	338	>1000	>1000
Geul	966	1	2	3
Geul	968	4	7	12
Geul	971	30	80	>1000
Geul	973	1	2	3
Geul	980	2	4	7
Geul	981	17	64	>1000
Geul	Gemmenich	37	278	>1000

years. The wide confidence interval shows that some stations are certainly of an extreme return periods and some are not at all of an extreme occurrence.

The method for the return period of the 2-day and 3-day precipitation sums is similar as for the 1-day precipitation amount of table D.1. The difference is that the block maxima are not the maximum precipitation in 1 day per year, but for 2 and 3 consecutive days. The 2021 flood event value is the maximum precipitation for 2 and 3 days between 12 July and 15 July. As the return periods of table D.3 are not based on hourly data, but daily data, they should not be confused with the return period of the 48h-return period. Nevertheless, they give an indication of how extreme the precipitation magnitude was for multiple days.

Table D.3: The return periods for relevant stations of all catchments are given for the precipitation amount in 1 day (RT_1day), 2 days (RT_2day) and 3 days (RT_3day), based on the daily precipitation data sets. Green cells indicate an increase, red cells a decrease.

Catchment	Station	RT_1day	RT_2day	RT_3day
Ahr	533	>1000	>1000	>1000
Ahr	2213	730	356	655
Ahr	3490	88	57	55
Ahr	13714	4	3	4
Vesdre	Ternell	>1000	>1000	>1000
Vesdre	Jalhay	>1000	992	>1000
Vesdre	Battice	43	434	>1000
Vesdre	Louveigné	40	>1000	>1000
Geul	380	2	18	13
Geul	962	>1000	>1000	>1000
Geul	963	156	749	298
Geul	965	>1000	>1000	>1000
Geul	966	2	36	27
Geul	968	7	724	180
Geul	971	80	461	272
Geul	973	2	31	22
Geul	980	4	69	35
Geul	981	64	157	108
Geul	Kanne	2	29	19
Geul	Gemmenich	278	>1000	113

Fig. D.3 shows that the return periods increase for most stations when the precipitation amount is considered for a longer period. For the return periods that are already over 1000 years, no change can be determined. It is noticeable how mostly the values for the Ahr show a decrease in return periods for the longer time period. This could indicate that the overall extremity of the precipitation event was not caused by the long duration, but rather the high precipitation amount on 1 day. For the Vesdre, the return periods are the highest for the 3-day sum and for the Geul, for the 2-day sum. Overall, the return periods for the Vesdre and Geul increase when considering a multi-day precipitation sum instead of the precipitation amount of 1 day. So despite the fact that the 1-day precipitation sum was already an extreme event, the fact that the event lasted for several days with such a total precipitation amount was even more extreme.

The mean return periods of the precipitation are extreme for most stations, and even the lower bounds of a wide confidence interval are extreme for some stations. There are several Geul stations that show no extremity at all. The return periods of the 2-day and 3-day precipitations sums show that also for this longer time period the precipitation event was extreme, even more so than for the 1-day precipitation amount.

D.3 Return periods of discharge

The return periods of the peak discharge measurements and estimates, as shown in fig. 4.5, are given in table D.4. As for the precipitation, the return periods have been estimated for a reference period including and excluding the 2021 flood event. Again, the return periods are higher when the event is excluded, as was to be expected.

Since the Ahr is the only station with daily values during the reference period on which the GEV distribution is based, the peak discharge is also a daily value. The peak discharge occurred during the night and therefore, the maximum discharge averaged over 24 hours is split over 2 days. For fair comparison, table D.4 used the true daily average, but table D.5 gives the return periods of the true maximum discharge averaged over 24 hours. As could be expected from using the maximum 24 hour average, the return periods increase considerably compared to fig. 4.5 and are over 1000 years for 3 out of 5 stations and over 100 years for the 2 remaining stations.

Table D.4: The return periods for relevant stations of all catchments are provided with the peak value of the discharge (daily for the Ahr, hourly for the Vesdre and Geul). The RT_ev column has been computed by including the flood event in the statistics. In RT_nev the event is excluded. The last column is the main return period estimate and is used for fig. 4.1. The number of years for which the GEV distribution is computed is also given for each station.

Catchment	**Station**	**Town**	**Years**	**Peak value**	**RT_ev**	**RT_nev**
Ahr	Müsch	Müsch	48	168	487	487
Ahr	Altenahr	Altenahr	77	466	>1000	>1000
Ahr	Kirmutscheid	Kirmutscheid	50	42	60	84
Ahr	Denn	Denn	45	38	34	53
Ahr	Kreuzberg	Kreuzberg	50	42	167	>1000
Vesdre	Chaudfontaine	Chaudfontaine	30	600	>1000	>1000
Vesdre	Belleheid	Belleheid	28	60	243	>1000
Geul	L6660	Sippenaeken	26	53	120	>1000
Geul	10.Q.36	Meerssen	12	88	>1000	>1000
Geul	Kelmis	Kelmis	13	56	64	515

Table D.5: Return periods of the Ahr discharge stations in case the maximum 24h average discharge is used for the 2021 flood event instead of the discharge of the maximum calendar day.

Station	**Town**	**Peak value**	**RT_ev**	**RT_nev**
Müsch	Müsch	271	>1000	>1000
Altenahr	Altenahr	556	>1000	>1000
Kirmutscheid	Kirmutscheid	68	256	420
Denn	Denn	46	61	101
Kreuzberg	Kreuzberg	69	421	>1000

Table D.6: 95% confidence intervals of the discharge return periods. The mean value is the same as for table D.4. The "Min" and "Max" value are the minimum and maximum return period according to the confidence interval.

Catchment	**Station**	**Min**	**Mean**	**Max**
Ahr	Müsch	181	487	>1000
Ahr	Altenahr	>1000	>1000	>1000
Ahr	Kirmutscheid	36	84	725
Ahr	Denn	24	53	>1000
Ahr	Kreuzberg	>1000	>1000	>1000
Vesdre	Chaudfontaine	>1000	>1000	>1000
Vesdre	Belleheid	>1000	>1000	>1000
Geul	L6660	152	>1000	>1000
Geul	10.Q.36	>1000	>1000	>1000
Geul	Kelmis	82	515	>1000

Table D.6 shows that for the upper bounds of the 95% confidence interval all stations have a return period of over 1000 years, with one exception of 725 years. Even for the lower bounds, half of the stations have a return period of over 1000 years. The other half have much lower return periods of less than 200 years. It can be concluded that even for a wide confidence interval, most stations show an extreme peak discharge magnitude from the high return periods.

Appendix E

Discharge estimates

This appendix gives more information on the estimates of peak discharges. This includes the derivation of historical discharges in the Ahr from Roggenkamp & Herget (2014) and the application of that same method for the 2021 flood event, as done by Roggenkamp & Herget (2022). This appendix also includes the peak discharge estimates for the 2021 flood event for the Vesdre, based on the findings of Ziemetz et al. (2021).

E.1 Deriving historical discharges for the Ahr

Roggenkamp & Herget (2014) have reconstructed five pre-instrumental floods in the Ahr catchment for which there was sufficient information. Their general method was to select clear floodmarks (on buildings or from pictures) and compute the mean flow velocity with the empirical Manning formula (Chow, 1959). By multiplying the mean velocity with the flow area, the peak discharge can be calculated according to eq. (E.1). The peak discharge is defined as the discharge when the water level is at its maximum.

$$Q_p = A_p \cdot R_p^{2/3} \cdot S^{1/2} \cdot n^{-1} \tag{E.1}$$

The cross-section during the highest water level, A_p, is determined using the height of the flood mark and the local topography. As these flood marks were indicated in towns and their lay-out has changed considerably in the recent centuries, the researches used old maps to reconstruct the cross-section as accurately as possible. The same goes for the hydraulic radius, R_p. The slope, S, is related to the slope of the water level instead of the energy line as this is common practice in paleohydrology. Since no engineering measures have been implemented (Roggenkamp & Herget, 2015), the current value of 0.0035 $[mm^{-1}]$ has been used.

The Manning roughness coefficient, n, as used in eq. (E.1) is calculated in eq. (E.2) (Chow, 1959). The physical significance of each variable is given in table E.1.

$$n = (n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 + n_8 + n_9) \cdot m \tag{E.2}$$

Variable	Significance
n_1	Surface roughness
n_2	Vegetation
n_3	Channel irregularity
n_4	Channel alignment
n_5	Silting and scouring
n_6	Obstruction
n_7	Variations of stage and discharge
n_8	Sediment load (density of water)
n_9	Seasonal changes
m	Correction factor for meandering of the channel

Table E.1: Physical significance of variables of eq. (E.2)

For each historic flood event, the start of the calculation is the water level. This can be derived from a flood mark on a structure that has clearly been labelled in time and for which the structure has undergone no considerable changes in elevation (or appearance). Alternatively, pictures can be used to derive water levels using remaining structures as reference. The water level is only relevant for the cross-section in that location. Therefore, the peak discharge is only valid for the location of the flood mark. The flood marks have mostly been found in Dernau and Altenahr, with Dernau being located two kilometers upstream of Altenahr. As this distance is relatively small and with no significant changes in the river system, such as confluences, no difference is made between the locations from here on and Altenahr is considered as the location of these peak discharges as there is modern gauge at this location.

For each cross-section of the discharge derivation, Roggenkamp & Herget (2014) have determined different areas based on their roughness characteristics, such as flood plains, urban centers and the river channel. The Manning roughness coefficient variables of eq. (E.2) could then be derived using pictures and assumptions in combination with roughness tables of Arcement & Schneider (1989). Several variables were determined to be irrelevant due to the circumstances. Since the entire flooded area is considered, alignment effects, n_4, are not relevant. Silting and scouring, n_5, are not considered for the floodplains as these are overgrown surfaces. For the river channel, this variable has also be left out, as it is of minor importance according to scale and surface. There were no considerable obstructions, n_6, for the locations in question. Stage and discharge effects, n_7, are considered unimportant as the cross-section area is divided into different elements. The sediment load, n_8, could not be quantified directly and is thus assumed to be of minor importance. Seasonal changes, n_9, are irrelevant for the summer flood events (1804, 1888, 1910). Finally, the river channel is straight in the urban areas (which are the locations of the flood marks), so the correction factor for meandering, m, is considered to be 1. As the roughness variables are not completely certain, Roggenkamp & Herget (2014) provides a range of possible discharges between a minimum and maximum value, as well as a mean value which can be considered as the main estimate. These values can be found in table E.2.

Peak discharge	1804	1888	1910	1918	1920
Q_{max}	1600	360	630	300	220
Q_{mean}	1210	280	500	240	170
Q_{min}	940	230	410	190	140

Table E.2: Peak discharge estimates $[m^3/s]$ *(Roggenkamp & Herget, 2014)*

E.2 Discharge estimates for the 2021 flood event in the Ahr

Roggenkamp & Herget (2022) have estimated several peak discharges throughout the Ahr catchment with a similar method as for the historical floods. Applying this method is likely results in better estimate for recent floods since the topography and thus potential flow area can be known accurately. Furthermore, there are many flood marks present, which allows for making estimates at locations with minimal backwater effects. While the historical flood marks were mostly present in urban areas, which cause a narrowing of the flow area and so backwater effects, the estimates for the 2021 flood event were deliberately made outside urban areas.

Erosion and accumulation in the channels or floodplains have not been considered as relevant. However, considerable erosion has been observed (Koks et al., 2021) and if part of this erosion occurred after the peak discharge, which is likely, the assumption of no erosion could have caused an error. Since the estimate locations have been chosen to avoid backwater effects, the water level gradient has been read from continuous water level indicators along the bank or from the current water level. The cross-section has been divided into relatively homogeneous sub-areas and for each area the roughness coefficient is determined. Three roughness coefficient estimates have been made for each sub-area: the minimum, maximum and most plausible. Adding all parameter values into the manning equation, eq. (E.1), results in a minimum, maximum and most plausible discharge peak for each location.

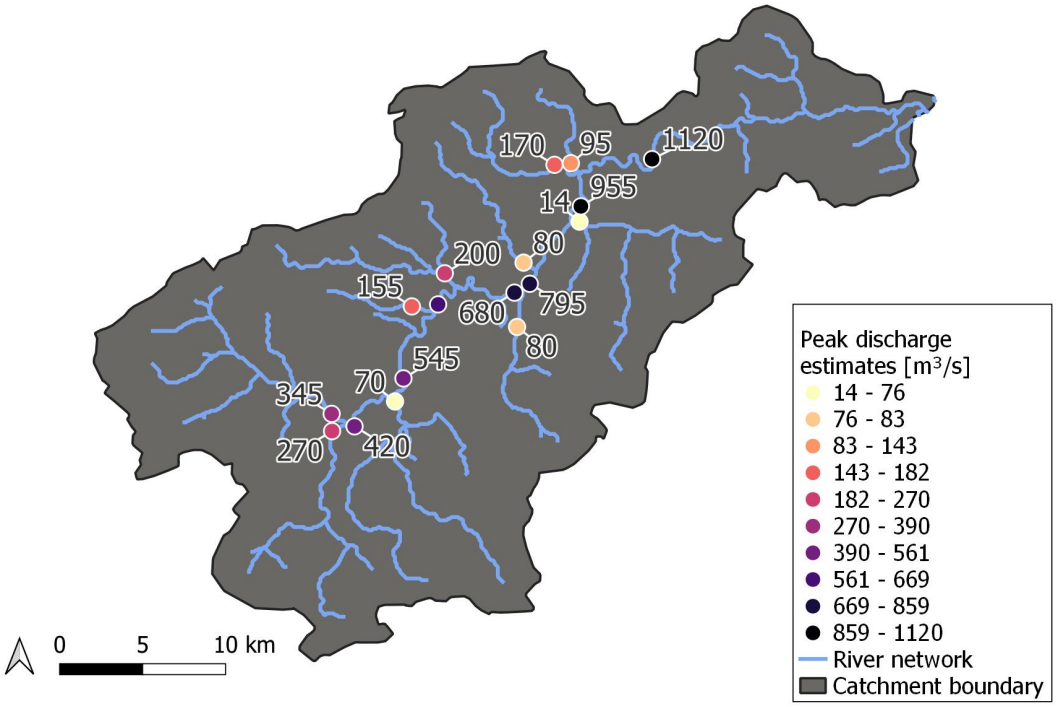


Figure E.1: Peak discharge estimates (most plausible) and their location from Roggenkamp & Herget (2022)

The most plausible peak discharge estimates of fig. E.1 show the high contribution of the north and western tributaries to the discharge in the main river. The contribution of the eastern tributaries (as they have smaller drainage areas) are much smaller. The peak discharges increases in downstream direction and has a value of 1120 $[m^3/s]$ just downstream of the Altenahr station. The minimum estimate is 1030 $[m^3/s]$ and the maximum estimate is 1230 $[m^3/s]$, which is a range of 8-10% from the most plausible estimate. The exact derivation for these values can be found in Roggenkamp & Herget (2022) as the station relevant for Altenahr is worked as an example.

E.3 Discharge estimates for the 2021 flood event in the Vesdre

Commissioned by the Wallonian Ministry of Climate, Energy, Mobility and Infrastructures, led by Philippe Henry, the Swiss engineering company Stucky Gruner has performed two analyses of the 2021 flood. The first part is the factualization of the event (Ziemetz et al., 2021) which includes several discharge estimates. Using water level marks and estimates of the Manning Strickler coefficient, 18 peak discharges have been estimated. This method assumes uniform flow and is dependent on the Strickler coefficient estimate, which are two considerable limitations. The results are visualised in fig. E.2. The report recommends to interpret these value with an confidence interval of 30%, which means that they are confident of the range of values that are 30% lower and higher than their mean estimate.

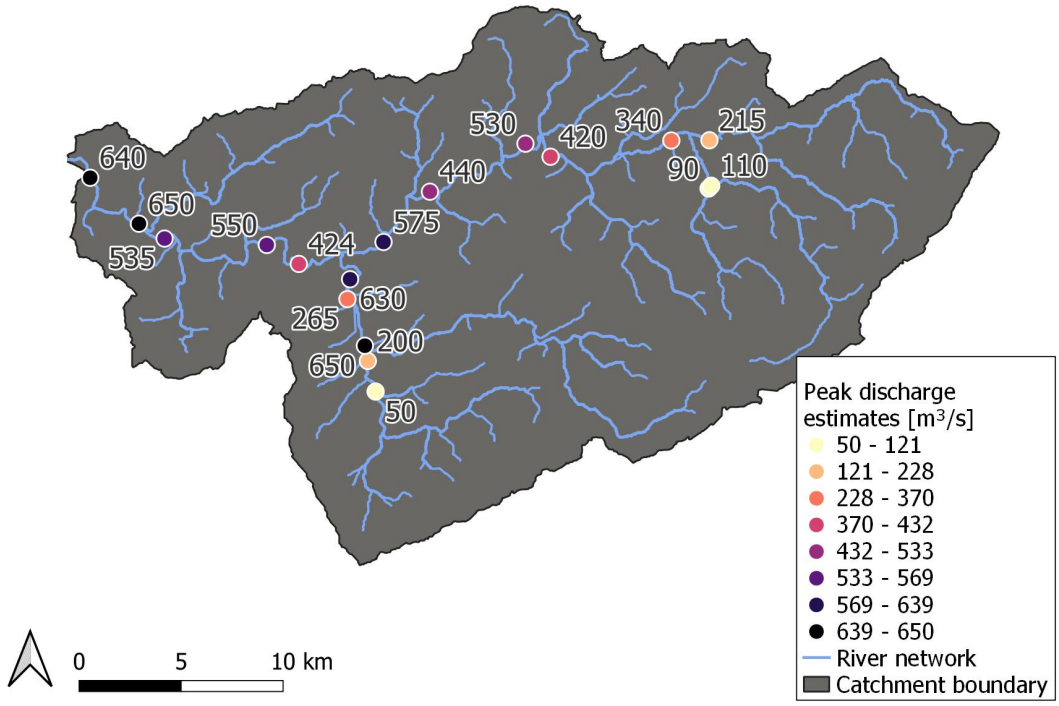


Figure E.2: Discharge estimates for the 18 points in the Vesdre catchment, based on values from Ziemetz et al. (2021)

The Direction de la Gestion Hydrologique (DGH, Board of Hydrologic Policy) and the Direction des Cours d'eau non navigables (DCENN, Board of non-navigable waterways) have made several manual measurements on July 14. Due to the high waterlevels and consequent danger, these measurements were mostly taken in the morning and noon. At Chaudfontaine, a discharge of 150 $[m^3/s]$ was recorded at 07:37 h. More upstream, just before the confluence with the Hoëgne at Verviers, a discharge of 115 $[m^3/s]$ was recorded at 13:15 h. At the Hoëgne, in Theux, a discharge of 170 $[m^3/s]$ was recorded at 11:00 h. More upstream, along the tributary Wayai of the Hoëgne, in Spixhe, a discharge of 92 $[m^3/s]$ was recorded at 08:40 h. The gauge of Pepinster, also along the Hoëgne, was the only one with a measurement at the evening, at 20:00 h, with a discharge of 390 $[m^3/s]$.

There are also other methods of estimating peak discharges. Firstly, the rational method can be applied which is a simple, semi-empirical method, based on the rainfall intensity, i, the drainage area, A, and the runoff coefficient, C. The runoff coefficient dictates the percentage of the rainfall that becomes runoff and is based on land cover maps. These parameters are combined in eq. (E.3).

$$Q = i \cdot C \cdot A \tag{E.3}$$

Table E.3 gives the measurements and estimates of peak discharges throughout the Vesdre catchments. all measurements are given for the gauge locations. Not all of these locations are included in the station overview in fig. B.8 as some were not publicly available. The table indicates the measured peak discharge, although this data is missing for almost all stations. The reconstructed measurements are derived from measurements with a small gap during the peak with complete measurements from nearby gauges. The flood mark estimates were already given in fig. E.2 but are given here again. For several locations there have been nearby estimates, resulting in a range of possible discharges. However, in the list on which fig. E.2 was based, only the most probable value was included. The rational method was explained above and is a rather crude simplification. Finally, the results from Archambeau et al. (2021) have also been included in this table.

These peak discharges vary considerably between the different methods for all stations, which show that the peak discharges throughout the catchment are quite uncertain. This report only continues with the discharges of Belleheid and Chaudfontaine due to the available discharge records preceding the 2021 flood event. Although there are measurements available for the station of Belleheid, Ziemetz et al. (2021)

Table E.3: Discharge estimates from multiple techniques for the Vesdre catchment (Ziemetz et al., 2021)

Flood marks	**River**	**Discharge** $[m^3/s]$				
		Reconstructed measurement	**Flood marks**	**Rational method**	**(Archambeau et al., 2021)**	**Average**
Outflow reservoir Eupen	Vesdre	196			190	193
Polleur	Hoëgne	228		100	145	157
Belleheid	Hoëgne	106		130	60	99
Theux	Hoëgne	367	265-650	200	260	348
Bellevaux	Vesdre		420-530	50		333
Trooz	Vesdre		535	400		468
Verviers	Vesdre		440-575		410	475
Chaudfontaine	Vesdre		650		580-660	630
Spixhe	Wayai		50-200	150	140	135

has calculated higher values than the measurements indicate. It is possible that the measurement is an underestimation of the true peak discharge. However, since Archambeau et al. (2021) have an estimate close to the measured peak discharge, the measurements are used in the rest of this study. The peak discharge estimates for Chaudfontaine vary between 580 and 650 $[m^3/s]$ and the report uses a value of 600 $[m^3/s$ in its further analyses. With a likely scenario that the true peak discharge was 3% lower or 8% higher, the overall estimate of 600$[m^3/s]$ can be considered as a balanced estimate without too much variation.

The discharge in the Vesdre is strongly influenced by two major reservoirs, as shown in section 2.4. The reservoir of Gileppe stored almost all of its inflow, 7.6 million $[m^3]$ during the event. Only in the morning of July 15 did it start releasing 10 $[m^3/s]$ in accordance with the authorities (DGH and CRC-W). However, the other reservoir, the one of Eupen, released considerable amounts of water during the flood event.

Ziemetz et al. (2021) have reconstructed the in- and outflow of the Eupen reservoir, as can be seen in fig. E.3. The inflow has a peak discharge of 268 $[m^3/s]$, while the outflow has a peak discharge of 196 $[m^3/s]$ at 01:30h on July 15. This high discharge can be seen as the start of the second discharge peak travelling downstream through the Vesdre catchment. The sudden release of water was required to ensure the structural safety of the dam as the crucial water level had been exceeded (full, blue line in fig. E.3. The "note de manutention" (reservoir management plan) was followed, but was overwritten. This is allowed in risk of structural failure. Nevertheless, 6 million $[m^3]$ of water was stored during the flood event by the reservoir.

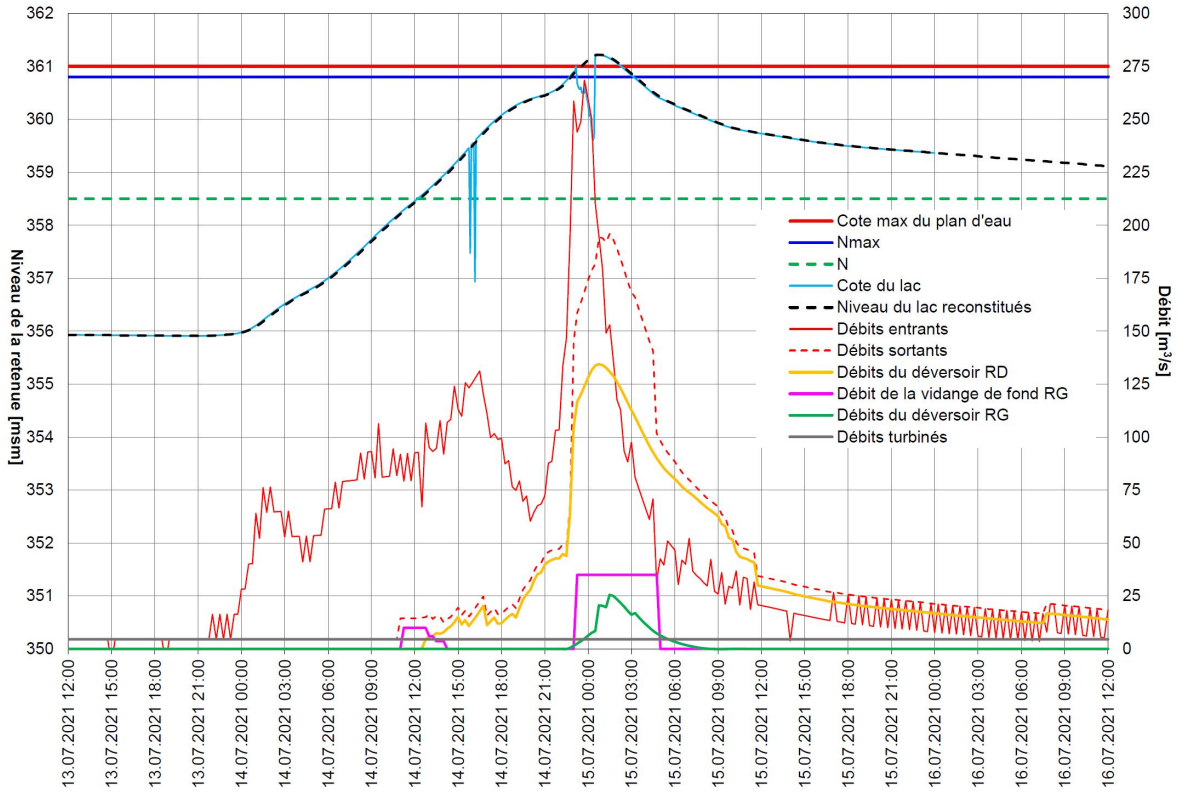


Figure E.3: Reconstructed hydrographs of the Eupen reservoir from (Ziemetz et al., 2021). The in- and outflow of the reservoir are given in red in red. The dotted line is the outflow and the full line is the inflow.

In accordance with the reservoir management plan, the critical flow scenario took effect on July 13 at 23:07 h as the expected volume to enter the reservoir was larger than the available storage. The discharge through the turbines is then maximised and the water level in the reservoir is carefully monitored. The left drain was opened to equalise the in- and outflow of the reservoir. If there is a risk of flooding between the dam and the confluence of the Vesdre and the Helle, the tunnel between the Helle and the reservoir must be closed. This final action was not done successfully. Additionally, the upstream limnographs that record the inflow to the reservoir were not operational due to maintenance. Instead, manual measurements were done at 13:06 h, 15:10 h and 5:00 h.

The water release was halted when a crucial water level in Pepinster was exceeded and the drain was closed again at July 14, 04:17 h. With increasing water levels and the consequent worry of structural safety, 10 $[m^3/s]$ started to be released at 14:57 h. At 12:35 h the previously malfunctioning weir gate on the right side was repaired and its opening allowed the closure of the bottom drain. However, at 22:43 h the critical height, Nmax, of 360.8 $[m]$ was exceeded. At 00:28 h the structural height maximum was also exceeded, which resulted in the re-opening of the bottom valve and also the spillway of the left bank. This led to a high discharge from the reservoir. If this opening was the cause of a second discharge peak in the Vesdre is beyond the scope of this research.

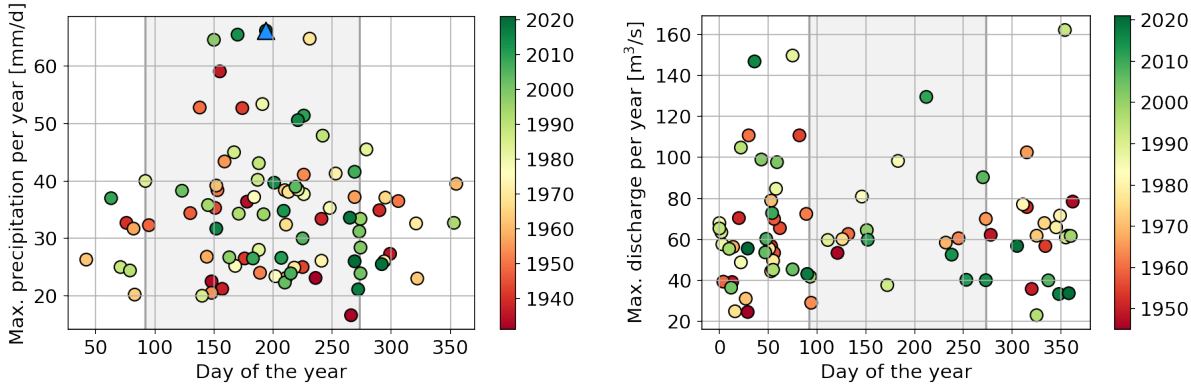
Appendix F

Seasonality

This appendix provides more information concerning the seasonality of the 2021 flood event. The seasonality of the yearly precipitation and discharge maxima are given for all catchments. This appendix also includes a more detailed elaboration on the historical flood data for the Geul.

F.1 Seasonality of high precipitation and discharge

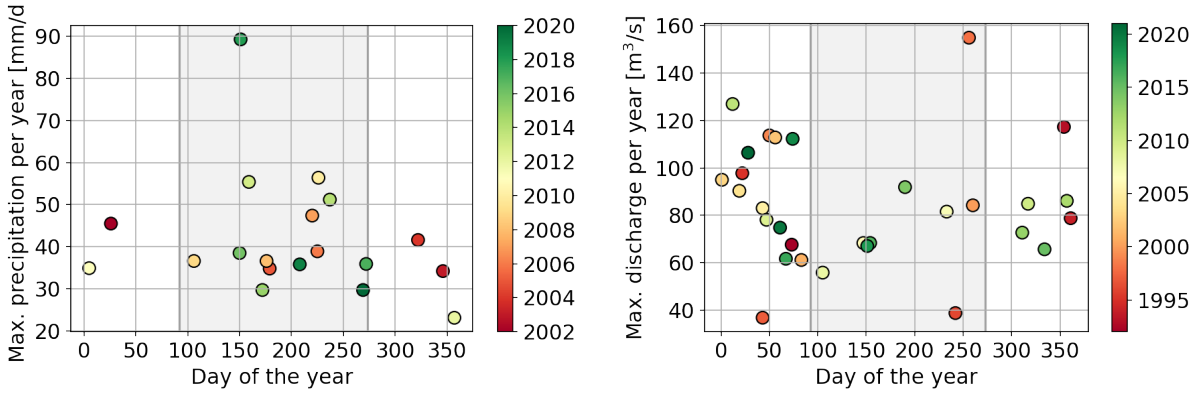
To get insight in the seasonality of extreme events, fig. F.1, fig. F.2 and fig. F.3 show the day of occurrence of the maximum precipitation and discharge of each year for the Ahr, Vesdre and Geul. The stations have been chosen to have a relatively high proximity and a similar record period. A specific focus is on the summer half-year as it was said that the 2021 flood event was rare to have occurred during summer. Since the moment of occurrence of historical floods is often known for only the month, the summer half-year is applied here (April-September), instead of the true summer period.



(a) Precipitation from station 3490

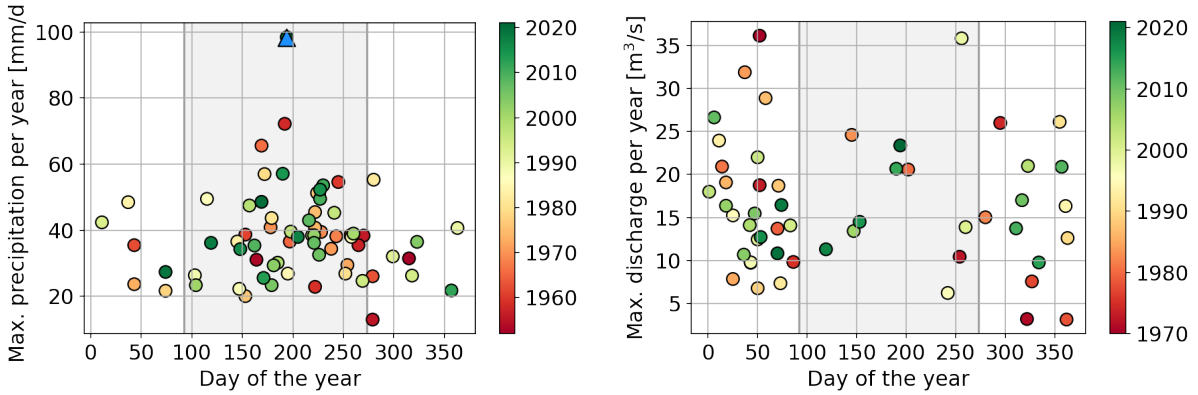
(b) Discharge from station Altenahr

Figure F.1: Day of occurrence and magnitude of the yearly maximum precipitation (daily) and discharge (daily) in the Ahr. A colorbar indicates the year in which the maximum day occurred. The summer half year is indicated by the grey zone.



(a) Precipitation from station Battice *(b) Discharge from station Chaudfontaine*

Figure F.2: Day of occurrence and magnitude of the yearly maximum precipitation (daily) and discharge (daily) in the Vesdre. A colorbar indicates the year in which the maximum day occurred. The summer half year is indicated by the grey zone.



(a) Precipitation of station P_962 *(b) Discharge of station 10.Q.30*

Figure F.3: Day of occurrence and magnitude of the yearly maximum precipitation (daily) and discharge (daily) in the Geul. A colorbar indicates the year in which the maximum day occurred. The summer half year is indicated by the grey zone.

The percentages of the maximum precipitation day within the summer half-year are 74, 74 and 79% for the Ahr, Vesdre and Geul. For the discharge, these percentages are only 22, 30 and 21% for the respective catchments. This quantification confirms the pattern that can be seen in all figures: a lot of rain events during the summer half-year, but very little high discharge events in that same period. The reason why the high precipitation and discharge events seem to have an opposing seasonality is outside the scope of the research. Nevertheless, it is an interesting observation since the discharge is dependent on the precipitation.

F.2 Historical floods Geul

This section includes the list of historical floods in the Geul and how it was made. The tool that has been used to find flood events in old newspapers is Delpher. This is a Dutch database with digitized texts from newspapers, books, and magazines from collections of scientific organisations, libraries and heritage institutions. With Optical Character Recognition (OCR), the texts can be searched for specific terms. For this search the boolean search terms are the following: "overstroming OR watersnood OR vloed AND Geul". Using synonyms for a flood event allows a wide search. Filtering the search to only newspaper articles, 3 426 matching results were found. Sorting these results to relevance, followed by a manual scan of titles, publication location and publisher, allowed for a quick method to filter out relevant flood event. A flood is here defined as an exceedance of the natural river course, or as often mentioned

in the articles: “buiten de oevers treden”. Only 1 article is used as reference in table F.1 although some events were mentioned by multiple articles. The articles are referenced with a URL-link for convenience.

Table F.1: All historical floods of the Geul and their moment of occurrence found using Delpher.

Year	Month	Day	Newspaper	Date	Source
1643	1		Limburgs Dagblad	01/10/1963	https://resolver.kb.nl/resolve?urn=ddd:010526077:mpeg21:p002
1909	2		De Maasbode	11/01/1914	https://resolver.kb.nl/resolve?urn=MMKB04:000189036:mpeg21:p002
1914	1	11	De Maasbode	12/01/1914	https://resolver.kb.nl/resolve?urn=MMKB04:000189036:mpeg21:p002
1917	2	6	De grondwet	06/02/1917	https://resolver.kb.nl/resolve?urn=ddd:110621570:mpeg21:p010
1926	5	20	Overijsselsch dagblad	20/05/1926	https://resolver.kb.nl/resolve?urn=MMKB23:001317095:mpeg21:p00001
1926	2		Nieuwe Tilburgsche Courant	22/02/1926	https://resolver.kb.nl/resolve?urn=ddd:010234720:mpeg21:p005
1939	1	4	De Maasbode	04/01/1939	https://resolver.kb.nl/resolve?urn=MMKB04:000195388:mpeg21:p009
1940	2	6	Limburger koerier: provinciaal dagblad	06/02/1940	https://resolver.kb.nl/resolve?urn=ddd:010327230:mpeg21:p005
1947	3	5	Gazet van Limburg	07/03/1947	https://resolver.kb.nl/resolve?urn=MMCC01:048011041:mpeg21:p00001
1952	12	20	Eindhovensch Dagblad	23/12/1952	https://resolver.kb.nl/resolve?urn=MMRHCE02:163606071:mpeg21:p00001
1956	3	6	Provinciale Overijsselsche en Zwolsche courant	05/03/1956	https://resolver.kb.nl/resolve?urn=MMHCO02:163872055:mpeg21:p00001
1958	2	28	Nieuwsblad voor de Hoeksche Waard en Ijselmonde	28/02/1958	https://resolver.kb.nl/resolve?urn=MMMHW01:001115025:mpeg21:p00001
1960	12	6	Limburgs Dagblad	06/12/1960	https://resolver.kb.nl/resolve?urn=MMKB23:001934056:mpeg21:p00009
1960	5		De Volkskrant	14/05/1960	https://resolver.kb.nl/resolve?urn=ABCDDD:010875559:mpeg21:p001
1962	2	12	Limburgs Dagblad	15/02/1962	https://resolver.kb.nl/resolve?urn=ddd:011028049:mpeg21:p019
1965	1	10	Limburgs Dagblad	14/01/1965	https://resolver.kb.nl/resolve?urn=ddd:010526167:mpeg21:p023
1966	12	1	De Tijd De Maasbode	27/12/1996	https://resolver.kb.nl/resolve?urn=ddd:011237786:mpeg21:p005
1970	2	22	Limburgsch Dagblad	12/05/1970	https://resolver.kb.nl/resolve?urn=ddd:010541164:mpeg21:p013
1973	2	23	Limburgsch Dagblad	23/02/1973	https://resolver.kb.nl/resolve?urn=ddd:010556266:mpeg21:p015
1978	5	7	Limburgsch Dagblad	08/05/1978	https://resolver.kb.nl/resolve?urn=ddd:010563255:mpeg21:p005
1980	7	20	Limburgs Dagblad	21/07/1980	https://resolver.kb.nl/resolve?urn=ddd:010570822:mpeg21:p007
1981	6	30	Limburgs Dagblad	30/06/1981	https://resolver.kb.nl/resolve?urn=ddd:010570374:mpeg21:p011
1984	2	7	Limburgs Dagblad	09/02/1984	https://resolver.kb.nl/resolve?urn=ddd:010593217:mpeg21:p001
1987	2	31	Limburgsch Dagblad	02/03/1987	https://resolver.kb.nl/resolve?urn=ddd:010611204:mpeg21:p001
1987	1	3	Nederlands dagblad: gereformeerd gezinsblad	03/01/1987	https://resolver.kb.nl/resolve?urn=ddd:010562334:mpeg21:p003
1990	12	31	Limburgsch Dagblad	31/12/1990	https://resolver.kb.nl/resolve?urn=ddd:010624173:mpeg21:p011
1994	12	30	NRC Handelsblad	30/12/1994	https://resolver.kb.nl/resolve?urn=KBNRC01:000030833:mpeg21:p002
2012	6		Limburg 1	29/07/2012	https://l1.nl/hoogste-piek-rivier-de-gulp-is-gulpen-gepasseerd-50813

Appendix G

Model descriptions

This appendix describes the lumped HBV model and distributed SBM model with their processes, governing equations and parameters.

G.1 Lumped HBV model

The lumped model based on the HBV model of the course CIE4431 is applied here. Its input is the precipitation and potential evaporation in the unit of $[mm/h]$. The model has an hourly timestep and no spatial variation. The input data of precipitation and potential evaporation, described in section 5.4, have been summed for all grid cells in the drainage area of the relevant station with the use of multiple Python packages (rasterio for reading in the catchment elevation, pyflwdir for deriving the drainage areas and xarray for processing the data of the NETCDF input files). The derived drainage areas have been checked to see if they match the expected drainage area of each station visually. With the prepared input of precipitation and potential evaporation and discharge, the model can be run. The discharge is used in the calibration.

Fluxes and stores

The lumped HBV model consists of stores through which the water travels (in fluxes) during each timestep. In fig. G.1, the scheme of these stores and fluxes are give. It shows the four stores: interception (Si), unsaturated store (Su), fast-responding groundwater store (Sf) and slow-responding groundwater store (Ss). The stores are connected by fluxes, for which the precise equations are given below. The discharges from the groundwater stores are delayed with a weighing function (triangular transfer function with base Tlag), which is explained in more detail in the parameter subsection.

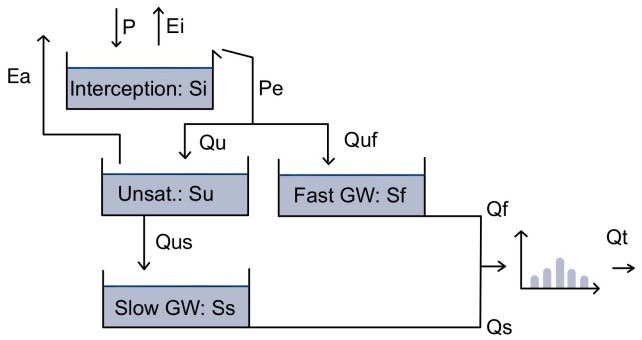


Figure G.1: Schematization of the lumped model with the stores and fluxes, as used in the equations.

The precipitation input enters the first storage unit which is interception. With the added precipitation input, the amount of water in the interception store is Si. From this store, the available water can evaporate as the flux of Ei, which is limited by the potential evaporation Ep, see eq. G.1. In this model, water evaporation only occurs in the timesteps without precipitation.

$$Ei = min(Ep, Si) \tag{G.1}$$

The maximum capacity of the interception storage, $Imax$ determines how much water can move to the next two storage units. They both represent water stores in the soil, so the next flux can be considered as infiltration. Overland flow is not included explicitly in this model. The amount of water that can infiltrate, Pe, is the excess of the interception store, see eq. G.2. In other words, the interception store must be full before water moves to the next stores.

$$Pe = max(0, Si - Imax) \tag{G.2}$$

The water flux to the next two stores, the unsaturated zone and fast responding groundwater zone, is partitioned by the partitioning factor Cr given in eq. G.3. It depends on the filling degree of the unsaturated store, since Su is the amount of water stored in that zone and Su_max is the maximum capacity of that zone. The ratio has an exponent β.

$$Cr = (\frac{Su}{Su\_max})^{\beta} \tag{G.3}$$

The infiltration to the fast responding groundwater zone, Quf, is defined by eq. G.4. So if the partitioning factor Cr is high, the percentage of the infiltration, Pe, flowing to the fast groundwater zone is also higher. The infiltration to the unsaturated zone is the remaining part of the flux Pe.

$$Quf = Cr \cdot Pe \tag{G.4}$$

(G.5)

Water can also exit the unsaturated zone with the flux Ea to reflect the natural processes of transpiration and soil evaporation. The parameter Ep_r is the residual potential evaporation, so it is zero if all potential evaporation has been evaporated in the interception zone. The flux Ea depends on the filling degree of the unsaturated zone and an additional factor in the denominator, Ce, as shown in eq. G.6.

$$\begin{aligned} Ea &= Ep_r \cdot \frac{Su}{Su\_max \cdot Ce} \\ Ea &= min(Ea, Su) \end{aligned} \tag{G.6}$$

The flux from the unsaturated zone to the slow responding groundwater zone, Qus reflects the process of percolation. As shown in eq. G.7, it depends on the filling degree of the unsaturated zone and a factor $Pmax$.

$$Qus = Pmax \cdot \frac{Su}{Su\_max} \tag{G.7}$$

The only two stores that generate the runoff are the fast and slow responding groundwater stores. Those fluxes, respectively Qf and Qs, depend on the horizontal conductivity factors, Kf and Ks, as shown in eq. G.8.

$$\begin{aligned} Qf &= Kf \cdot Sf \\ Qs &= Ks \cdot Ss \end{aligned} \tag{G.8}$$

Parameters
The equations of the fluxes contained several parameters that are user defined (manually or through calibration). They are given in table G.1 and explained here in more detail. It is important to understand the effect of these parameters since the calibration of the parameters is used to detect changes in flow mechanisms and

Table G.1: Parameters of the HBV Model with their meaning and range

Parameter	Meaning	Value range
Imax	Maximum interception storage	0 - 4
Sumax	Maximum storage of unsaturated zone	40 - 700
Beta	Exponent for unsaturated/fast groundwater flow	0.5 - 2
Ce	Transpiration parameter	0.1 - 0.9
Pmax	Factor of percolation	0.001 - 0.5
Kf	Horizontal conductivity of fast reservoir to river discharge	0.01 - 0.7
Ks	Horizontal conductivity of slow reservoir to river discharge	0.0001 - 0.01
Tlag	Time delay of river discharge	0-10

Imax:
When the maximum storage capacity of the interception zone is higher, a lot of water can be stored before running off. One could see it as a region with a lot of vegetation that intercepts the rain. Alternatively, it can be considered as implicit surface water storage due to depressions in the surface or a low infiltration capacity since more water is stored before it infiltrates.

Su_max:
The maximum storage of the unsaturated zone reflects the thickness and permeability of the soil. With a high maximum storage, a lot of water can infiltrate and remain in the unsaturated zone, before flowing to groundwater layers. The maximum storage is important for the filling degree of the unsaturated zone, which is defined as the ratio of the actual storage of the unsaturated zone and the maximum storage. When this maximum storage is very large, the filling degree will likely be smaller. Since this degree influences the partitioning to the fast groundwater zone, as well as the flow to the slow groundwater, a high Sumax would result in more flow through the unsaturated and slow groundwater stores, thus slowing down the flow response.

β:
The exponent β influences the partitioning of water between the unsaturated zone and the fast-responding groundwater zone through the factor Cr, see eq. G.3. It can be considered as a factor of how much water flows from the unsaturated zone immediately to the fast-responding groundwater zone. This partitioning depends on the filling degree of the unsaturated zone. When there is more water in the unsaturated zone, more rainwater flows immediately to the fast-responding groundwater zone. When β is equal to 1, that ratio fully determines the partitioning. By varying β above and below 1, this partitioning can be tweaked.

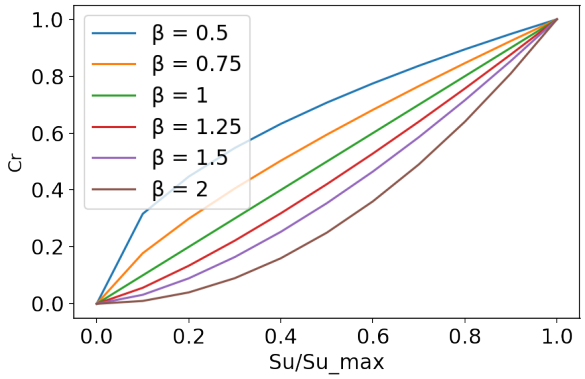


Figure G.2: Plot of ratio between Su and Su_max for varying values of β.

When β is lower than 1, relatively more water flows to the fast-responding groundwater zone. When β is higher than 1, more water flows to the unsaturated zone. Both affect the lower ratio more than the high ratio, since β is an exponent in eq. G.3.

Ce:

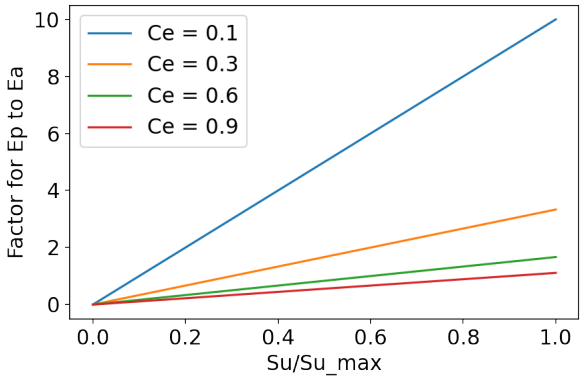


Figure G.3: Plot of ratio between Su and Su_max for varying values of Ce.

Similar as for the partitioning of the rain between the unsaturated and fast-responding groundwater zone, the transpiration from the unsaturated zone depends on the ratio between the amount of actual water in the unsaturated zone and the maximum storage. In contrast with β, the parameter is no exponent but a simple factor in the denominator, see eq. G.6. When this value is higher, less water evaporates or transpires. One can compare it to little vegetation and little potential for soil evaporation, so for example paved surface. When Ce is low, more water of the remaining potential evapotranspiration is released.

Pmax:

Whereas Ce was a factor in the denominator, the percolation factor $Pmax$ is a factor in the numerator, see eq. G.7. So when the factor is lower, relatively less water flows to the slow groundwater zone than by the ratio of the amount of water in the unsaturated store and its maximum capacity. This effect can be seen in fig. G.4. It can be interpreted as a low vertical hydraulic conductivity of the soil, slowing the process of percolation. As a consequence, a low flow to the slow groundwater flow leads to more fast groundwater flow as the amount of water in the unsaturated store remains high.

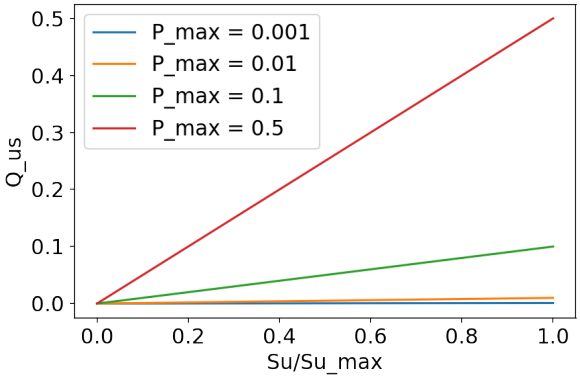


Figure G.4: Plot of ratio between Su and Su_max for varying values of Pmax.

Kf:

This parameter has a similar function as the horizontal conductivity. It determines how much of the storage in the fast-responding groundwater zone becomes river discharge. A high value reflects a higher conductivity, thus more water will runoff.

Ks:

This parameter works the same as Kf, but now for the slow-responding groundwater zone. A high value reflects a higher conductivity, thus more water will runoff.

Tlag:

All processes occur in one time step, but in reality these processes take much longer. Therefore, the flows from the fast- and slow-responding groundwater store are delayed with a weighing function. The parameter Tlag determines the length of this delay. For the duration of Tlag, which expresses a number of timesteps, a weight is given to each hour to form a symmetrical triangle with an area of 1 to assure conservation of the flow. A convolution function combines the weighing values with the instantaneous discharge time series.

G.2 Distributed SBM Model

General description

Wflow is a modelling platform from Deltares that includes multiple fully distributed hydrological models. It can account for precipitation, interception, infiltration, evapotranspiration, subsurface water and even reservoirs. One of those models within the modelling environment of wflow is the wflow_sbm model. Its vertical processes are based on the Topog_SBM model from Vertessy & Elsenbeer (1999). The original model is a simple bucket model with a one-dimensional kinematic wave overland flow scheme, and a contour-based element network for routing surface and subsurface flows, hence the topography reference in the name. The wflow_sbm model is more detailed than the original model and includes much more parameters and processes. These changes allow the wflow_sbm model to be applied more widely than the Topog_SBM model, which was developed to simulate fast runoff processes in small catchments. Some of the main differences are the fact that the unsaturated zone can consist of multiple layers, evapotranspiration losses are accounted for, as well as capillary rise. The routing is also using a D8 network instead of contour lines. This D8 network will be discussed in the section on lateral processes. Detailed information on the wflow_sbm model can be found here .

The model is especially useful for steep catchments with thin soils (<2 m) due to the lateral ground flow following topography instead of the true hydraulic head. This makes the model also less reliable in non-steep areas. Furthermore, it takes into account many parameters that are derived from look-up tables and pedotransfer functions (Imhoff et al., 2020). This gives the model a strong physical basis, which makes it useful for model experimentation. A possible danger of the high number of parameters is equifinality.

The basic functioning of this fully distributed model is quite simple. It updates each cell for each time step based on vertical processes and lateral processes. With the defined grid, cells are created, which are divided in river and land cells. The vertical processes and lateral processes described in the next chapters form the update of 1 time step for 1 cell.

The main vertical processes that are included in the wflow_sbm model are shown in fig. G.5.

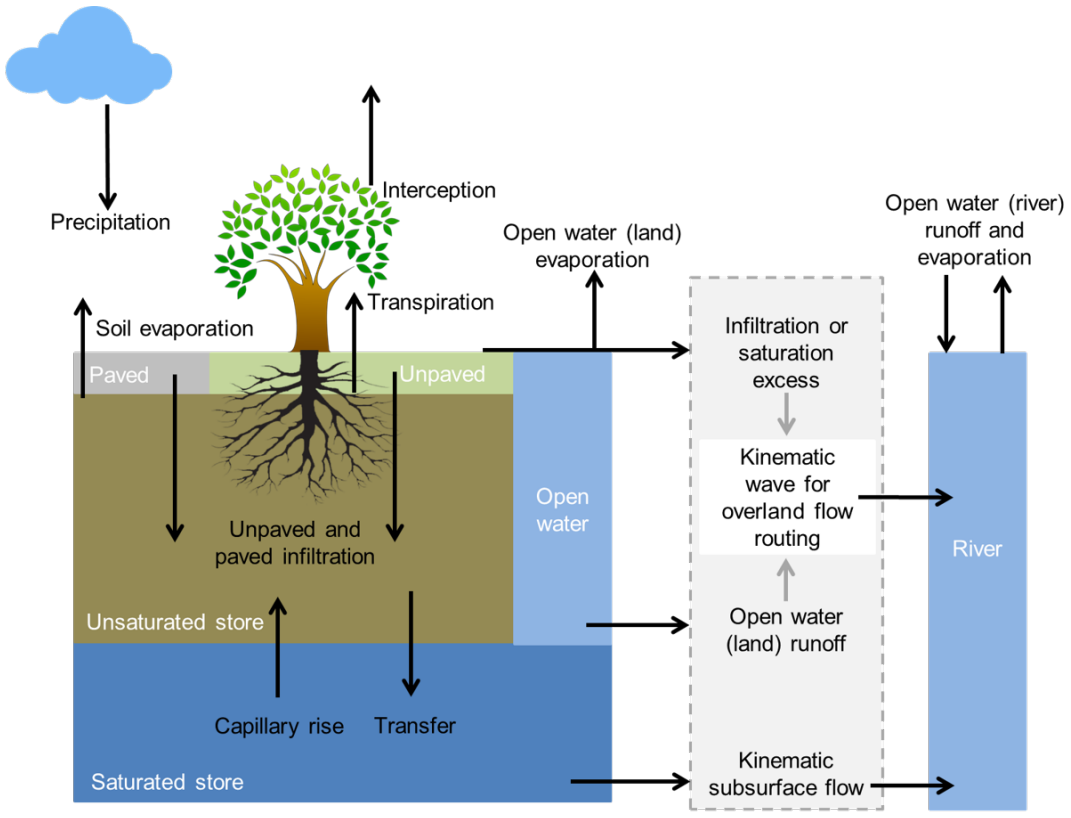


Figure G.5: Overview of the different processes and fluxes in the wflow_sbm model. Source: Deltares

This document discusses the parameters that are used in the model, as well as its vertical and lateral processes. The latter are discussed in less detail, as the main focus of the wflow_sbm model is on the vertical processes. Snow and glaciers are not included in this document but can be accounted for in the model. In order to use this document as a guide for reading the code and understanding the details, the parameter names of the model are used and not the names of the official equations. The final chapter provides some help to bridge this guide and the detailed code of the model, for example where certain elements can be found.

Table G.2: Parameters of wflow_sbm

Parameter	Meaning	Unit
User defined parameter		
Δt	Model time step	s
maxlayers	Maximum number of soil layers	-
θ_s	Saturated water content (porosity)	-
θ_r	Residual water content	-
kv_0	Vertical hydraulic conductivity at soil surface	$mm\Delta t^{-1}$
soilthickness	Soil thickness	mm
infiltcappath	Infiltration capacity of compacted areas	$mm\Delta t^{-1}$
infiltcapsoil	Infiltration capacity of non-compacted areas (soil)	$mm\Delta t^{-1}$
maxleakage	Maximum leakage from saturated zone	$mm\Delta t^{-1}$
waterfrac	Fraction of open water (excl. rivers)	-
pathfrac	Fraction of compacted area	-
rootingdepth	Rooting depth	mm
rootdistpar	Controls how roots are linked to water table	-
c	Brooks-Corey power coefficient for each soil layer	-
f	Scaling parameter controlling exponential decline of kv_0	mm^{-1}
kext	Extinction coefficient to calculate canopy gap fraction	-
leaf_area_index	Leaf area index	m^2m^{-2}
e_r	Gash interception model parameter	$mm\Delta t^{-1}$
cfmax	Degree-day factor	$mm^\circ C^{-1}\Delta t^{-1}$
tt	Threshold temperature for snowfall	$^\circ C$
tti	Threshold temperature interval length	$^\circ C$
ttm	Threshold temperature for snowmelt	$^\circ C$
whc	Water holding capacity as fraction of current snow pack	-
w_soil	Soil temperature smooth factor	-
cf_soil	Controls soil infiltration reduction factor when soil is frozen	-
g_tt	Threshold temperature for snowfall above glacier	$^\circ C$
g_cfmax	Degree-day factor for glacier	$mm^\circ C^{-1}t^{-1}$
g_sifrac	Fraction of the snowpack on top of the glacier converted into ice	-
glacierfrac	Fraction covered by a glacier	-
Forcing parameters		
precipitation	Precipitation	$mm\Delta t^{-1}$
temperature	Temperature	$^\circ C$
potential_evaporation	Potential evapotranspiration	$mm\Delta t^{-1}$
Internal parameters		
n	number of cells	-
nlayers	Number of soil layers	-
n_unsatlayers	Number of unsaturated soil layers	-
riverfrac	Fraction of river	-
kvfrac	Multiplication factor applied to kv_z	-
hb	Air entry pressure of soil (Brooks-Corey)	cm
act_tickl	Actual thickness of soil layers	mm
sumlayers	Cumulative sum of soil layers, starting at soil surface	mm
cap_hmax	Parameter controlling capillary rise	mm
cap_n	Coefficient controlling capillary rise	-
et_reftopot	Multiplication factor to correct the input of potential evapotranspiration	-
soilwatercapacity	Capacity of the entire soil	mm
cmax	Maximum canopy storage	mm
sl	Specific leaf storage	mm
swood	Storage woody part of vegetation	mm

Table G.3: Parameters of wflow_sbm

Parameter	Meaning	Unit
Model state		
ustorelayerdepth	Amount of water in the unsaturated zone, per layer	mm
satwaterdepth	Amount of water in the saturated zone	mm
ustoredepth	Amount of available water in unsaturated zone	mm
zi	Pseudo-water table depth (top saturated zone)	mm
canopystorage	Canopy storage	mm
canopygapfraction	Canopy gap fraction	-
avail_forinfilt	Water available for infiltration	$mm\Delta t^{-1}$
waterlevel_land	Water level land	mm
waterlevel_river	Water level river	mm
rainfallplusmelt	Snow melt + precipitation as rainfall	mm
vwc	Volumetric water content per soil layer	-
vwc_perc	Volumetric water content per soil layer	%
rootstore	Root water storage in unsaturated and saturated zone	mm
vwc_root	Volumetric water content in root zone	-
vwc_percroot	Volumetric water content in root zone	%
tsoil	Top soil temperature	$°C$
snow	Snow storage	mm
snowwater	Liquid water content in the snow pack	mm
glacierstore	Water within the glacier	mm
Model flux		
stemflow	Stemflow	$mm\Delta t^{-1}$
throughfall	Throughfall	$mm\Delta t^{-1}$
pottrans_soil	Potential transpiration, open water, river and soil evaporation (after substracting interception)	$mm\Delta t^{-1}$
transpiration	Transpiration	$mm/\Delta t^{-1}$
ae_ustore	Actual evaporation from unsaturated store	$mm\Delta t^{-1}$
interception	Interception	$mm\Delta t^{-1}$
soilevap	Soil evaporation from unsaturated zone	$mm\Delta t^{-1}$
soilevapsat	Soil evaporation from saturated zone	$mm\Delta t^{-1}$
actcapflux	Actual capillary rise	$mm\Delta t^{-1}$
actevapsat	Actual transpiration from saturated zone	$mm\Delta t^{-1}$
actevapustore	Actual transpiration from the unsaturated zone	$mm\Delta t^{-1}$
actevap	Total actual evapotranspiration	$mm\Delta t^{-1}$
runoff_river	Runoff from river based on riverfrac	$mm\Delta t^{-1}$
runoff_land	Runoff from land based on waterfrac	$mm\Delta t^{-1}$
ae_openw_l	Actual evaporation from land	$mm\Delta t^{-1}$
ae_openw_r	Actual evaporation from river	$mm\Delta t^{-1}$
net_runoff_river	Net runoff from river	$mm\Delta t^{-1}$
actinfilt	Actual infiltration into the unsaturated zone	$mm\Delta t^{-1}$
actinfiltsoil	Actual infiltration non-compacted fraction	$mm\Delta t^{-1}$
actinfiltpath	Actual infiltration compacted fraction	$mm\Delta t^{-1}$
infiltsoilpath	Actual infiltration (compacted and non-compacted areas)	$mm\Delta t^{-1}$
infiltexcess	Infiltration excess water	$mm\Delta t^{-1}$
excesswater	Water that cannot infiltrate due to saturated soil (saturation excess)	$mm\Delta t^{-1}$
exfiltsatwater	Water exfiltrating during saturation excess conditions	$mm\Delta t^{-1}$
exfiltustore	Water exfiltrating from unsaturated store because of change in water table	$mm\Delta t^{-1}$
excesswatersoil	Excess water for non-compacted fraction	$mm\Delta t^{-1}$
excesswaterpath	Excess water for compacted fraction	$mm\Delta t^{-1}$
runoff	Total surface runoff from infiltration and saturation excess	$mm\Delta t^{-1}$
transfer	Downward flux from unsaturated to saturated zone	$mm\Delta t^{-1}$
recharge	Net recharge to saturated zone	$mm\Delta t^{-1}$
actleakage	Actual leakage from saturated store	$mm\Delta t^{-1}$

Vertical processes per cell

Many vertical hydrological processes are included in this model, as can be seen in fig. G.6. A short note on a detail must be made to avoid confusion further on. Two unsaturated soil layers are visualised to show the possibility of multiple unsaturated soil layers. The same goes for evaporation from the saturated soil. However, the combination of multiple soil layers and evaporation from the saturated soil soil is not possible.

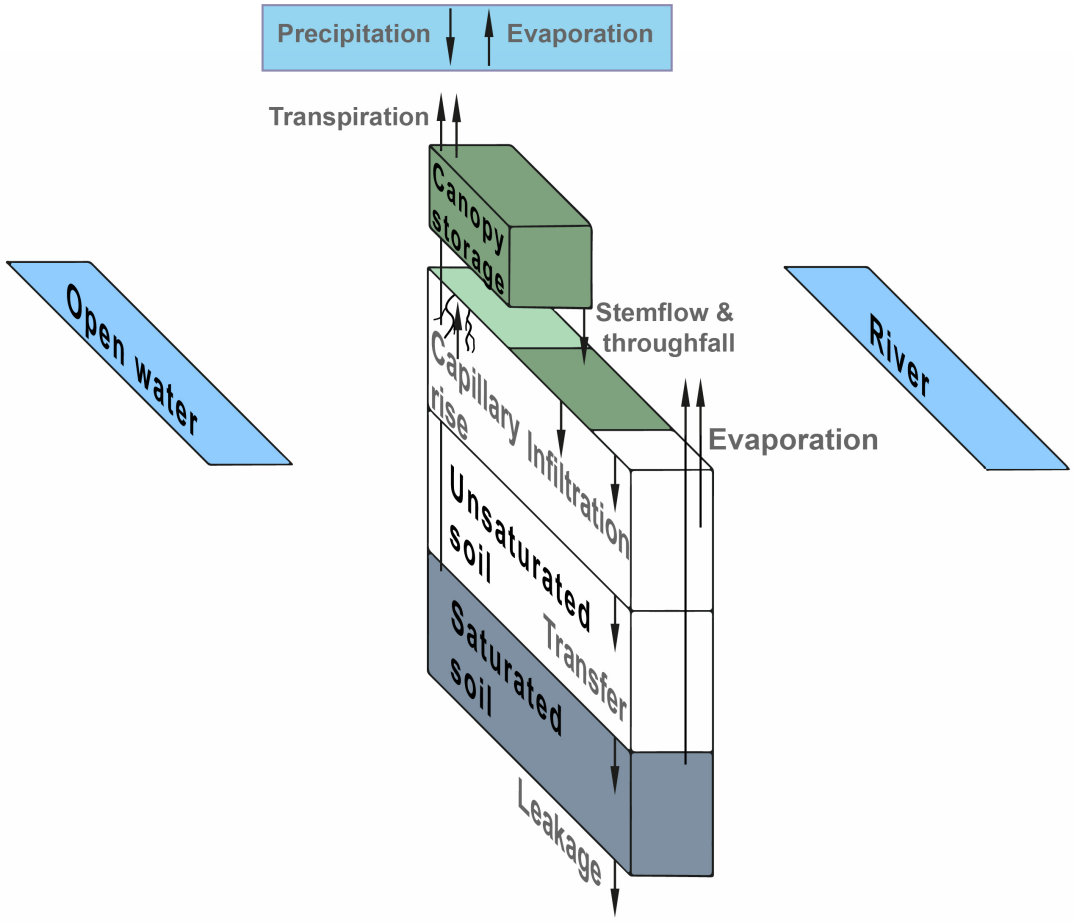


Figure G.6: Vertical processes of the wflow_sbm model

The order of these processes matters as some parameters are updated before they are used in another process. The order of these processes are given below. This list forms the structure of the following subsections in which each process is explained in detail. This explanation focuses on the physical processes and their equations, a link between this guide and the code is given in the final chapter. For more detailed information or questions about the code, please contact the author.

1. Vegetation parameters
2. Potential evaporation
3. Interception
4. Snow conditions
5. Rainfall plus snowmelt
6. Water available for infiltration
7. Open water evaporation
8. Infiltration

9. Defining the soil layers
10. Soil layers flow
11. Soil layers evaporation
12. Transpiration
13. Excess water
14. Capillary rise
15. Leakage and recharge
16. Exfiltration from unsaturated zone
17. Runoff

Vegetation parameters

Within the model there are 2 options of setting the vegetation parameters. These parameters are the maximum canopy storage, *cmax*, the canopy gap fraction, *canopygapfraction*, and the Gash interception model parameter, *e_r*. They can be defined directly in the input file of the static parameters by the user, but they can also be derived from other parameters if the leaf area index climatology is provided. In the latter case, the parameters are defined as follows.

The maximum canopy storage, *cmax*, is the maximum amount of water that can be stored in the canopy. Its quantification is split into a part for leaves and wood and depends on the specific leaf storage, *sl*, the leaf area index, *LAI* and the storage of the woody part of the vegetation, *swood*. This is combined in eq. G.9. The leaf area index, *LAI*, is an input map of the model with monthly values. The specific leaf storage is derived from a lookup table which depends on the land cover. This is also the case for the storage of the woody part of the vegetation.

$$cmax = sl \cdot LAI + swood \tag{G.9}$$

The canopy gap fraction, $canopygapfraction$, represents the fraction of gaps in the canopy. In other words, how much of the area is not underneath any form of canopy. In eq. G.10 the leaf area index is combined with an extinction coefficient in an exponential function (Van Dijk & Bruijnzeel, 2001). This coefficient depends on leaf distribution and the inclination angle and the values range between 0.6 and 0.8 and are dependent on the land cover.

$$canopygapfraction = e^{-kext \cdot LAI} \tag{G.10}$$

As an intermediate step the parameter *ewet* is introduced in eq. G.11. It represents the fraction of the cell with canopy cover multiplied with the potential evapotranspiration.

$$ewet = (1 - e^{-kext \cdot LAI}) \cdot potential\_evaporation \tag{G.11}$$

This parameter is then used in the computation of the Gash interception model parameter, e_r, in eq. G.12. There is an upper limit of 0.25 for e_r.

$$\begin{cases} e\_r = \frac{ewet}{P} & \text{if } P > 0.0001 \\ e\_r = 0 & \text{else} \end{cases} \tag{G.12}$$

Potential evaporation

One of the three forcing parameters is the potential evapotranspiration. A factor, $et\_reftopot$ can be added to convert reference grass evaporation (from the input) to potential evaporation , see eq. G.13.

$$potential\_evaporation = potential\_evaporation \cdot et\_reftopot \tag{G.13}$$

Interception

For a sub-daily timestep, the model uses a simplified Rutter model, which calculates the net interception, throughfall, stemflow and canopy storage for the next time step. The mechanism used for a daily time step is not explained in this document as most users will probably use an hourly time step.

The amount of stemflow, p_t, is taken as a fraction of 10% of the canopy gap fraction. With these parameters the interception steps of the model can be described. First the canopy storage is drained, eq. G.14, so not there is not more water in the canopy store than possible.

$$\begin{cases} DD = canopystorage - cmax & \text{if } canopystorage > cmax \\ DD = 0 & \text{else} \end{cases} \tag{G.14}$$

$$canopystorage = canopystorage - DD$$

The amount of rain that falls on the canopy, P_{frac}, is calculated and added to the storage with eq. G.15. This rain fraction has a lower limit of 0.

$$\begin{aligned} P_{frac} &= (1 - canopygapfraction - p_t) \cdot precipitation \\ canopystorage &= canopystorage + P_{frac} \end{aligned} \tag{G.15}$$

The evaporation is subtracted from the canopy storage, eq. G.16. However, the evaporation cannot be higher than the canopy storage.

$$\begin{aligned} dc &= -1 \cdot min(canopystorage, potential\_evaporation) \\ canopystorage &= canopystorage + dc \end{aligned} \tag{G.16}$$

The amount of evaporation that is not used is the leftover and is quantified according to eq. G.17. Keep in mind that the evaporation from the canopy, dc, is defined as negative. The leftover has a lower limit of 0 and is used for other evapotranspiration processes in the model in the next steps.

$$leftover = potential\_evaporation + dc \tag{G.17}$$

The canopy storage is drained again, eq. G.18.

$$\begin{cases} D = canopystorage - cmax & \text{if } canopystorage > cmax \\ D = 0 & \text{else} \end{cases} \tag{G.18}$$

$$canopystorage = canopystorage - D$$

The throughfall is the calculated using eq. G.19.

$$throughfall = DD + D + canopygapfraction \cdot precipitation \tag{G.19}$$

The stemflow is computed using eq. G.20.

$$stemflow = p_t \cdot precipitation \tag{G.20}$$

Finally the net interception is computed using eq. G.21.

$$Netinterception = precipitation - throughfall - stemflow \tag{G.21}$$

The net interception is then used in the rest of the model as simply the interception.

Snow conditions

When the snow is accounted for in the model, the parameter of *modelsnow* must be set to true. Then additional steps are added, which return the snow, snowwater, snowmelt, rainfallplusmelt and snowfall.

Rainfall plus snowmelt

The precipitation amount that reaches the soil surface is defined as the $rainfallplusmelt$ parameter. In case the $modelsnow$ parameter is true, the previous section has calculated the $rainfallplusmelt$ parameter already. In the case that there is also a glacier present (the parameter of $modelglacier$ is true), glaciermelt is added.

When snow is not specifically accounted for, the $rainfallplusmelt$ is taken as the sum of the stemflow and throughfall, eq. G.22, thus the remainder of the precipitation that has not been intercepted.

$$rainfallplusmelt = throughfall + stemflow \tag{G.22}$$

Water available for infiltration

The water that is available for infiltration, $avail\_forinfilt$, is set equal to the $rainfallplusmelt$. Part of this water falls on a river or any open water body. For this reason, two parameters are introduced: $runoff\_river$ and $runoff\_land$. The water that they collect depends on their areal fraction, see eq. G.23.

$$\begin{aligned} runoff\_river &= riverfrac \cdot avail\_forinfilt \\ runoff\_land &= waterfrac \cdot avail\_forinfilt \end{aligned} \tag{G.23}$$

The remaining water for infiltration then is calculated with eq. G.24.

$$avail\_infilt = avail\_infilt - runoff\_river - runoff\_land \tag{G.24}$$

Open water evaporation

The water that fell on rivers or open water is thus not available for infiltration, but it is still available for evaporation. The actual evaporation from the rivers, ae_openw_r, and open water, ae_openw_l, is thus taken as their areal fraction multiplied by the leftover of the potential evapotranspiration from the interception process, as is shown in eq. G.25. The actual evaporation has a maximum of the waterlevel multiplied by the areal fraction as there cannot be more evaporation than the amount of water present in the water body.

$$\begin{aligned} ae\_openw\_r &= riverfrac \cdot leftover \\ ae\_openw\_l &= waterfrac \cdot leftover \end{aligned} \tag{G.25}$$

The remaining potential evapotranspiration, $restevap$, is given in eq. G.26, where it is then divided for the bare soil, $potsoilevap$, and the vegetated areas, $pottrans$, using the canopy gap fraction.

$$\begin{aligned} restevap &= leftover - ae\_openw\_r - ae\_openw\_l \\ potsoilevap &= restevap \cdot canopygapfraction \\ pottrans &= restevap \cdot (1 - canopygapfraction) \end{aligned} \tag{G.26}$$

Infiltration

Before calculating the infiltration into the soil column, the initial capacity of the unsaturated store, $ustorecapacity$, is calculated with eq. G.27. It is equal to the storage capacity of the entire soil layer without the initial amount of water in the unsaturated, $ustoredepth$, and saturated store, $satwaterdepth$. It may be confusing that the remaining part is the capacity of only the unsaturated store, but be aware that the store of the saturated zone is by definition always full and thus the $satwaterdepth$ is also equal to the capacity of the saturated store.

$$ustorecapacity = soilwatercapacity - satwaterdepth - ustoredepth \tag{G.27}$$

The infiltration scheme of wflow_sbm can be visualised with the following schematization and steps in fig. G.7. The next page explains these steps in detail. Elements of frost are not considered in this description but are included in the model.

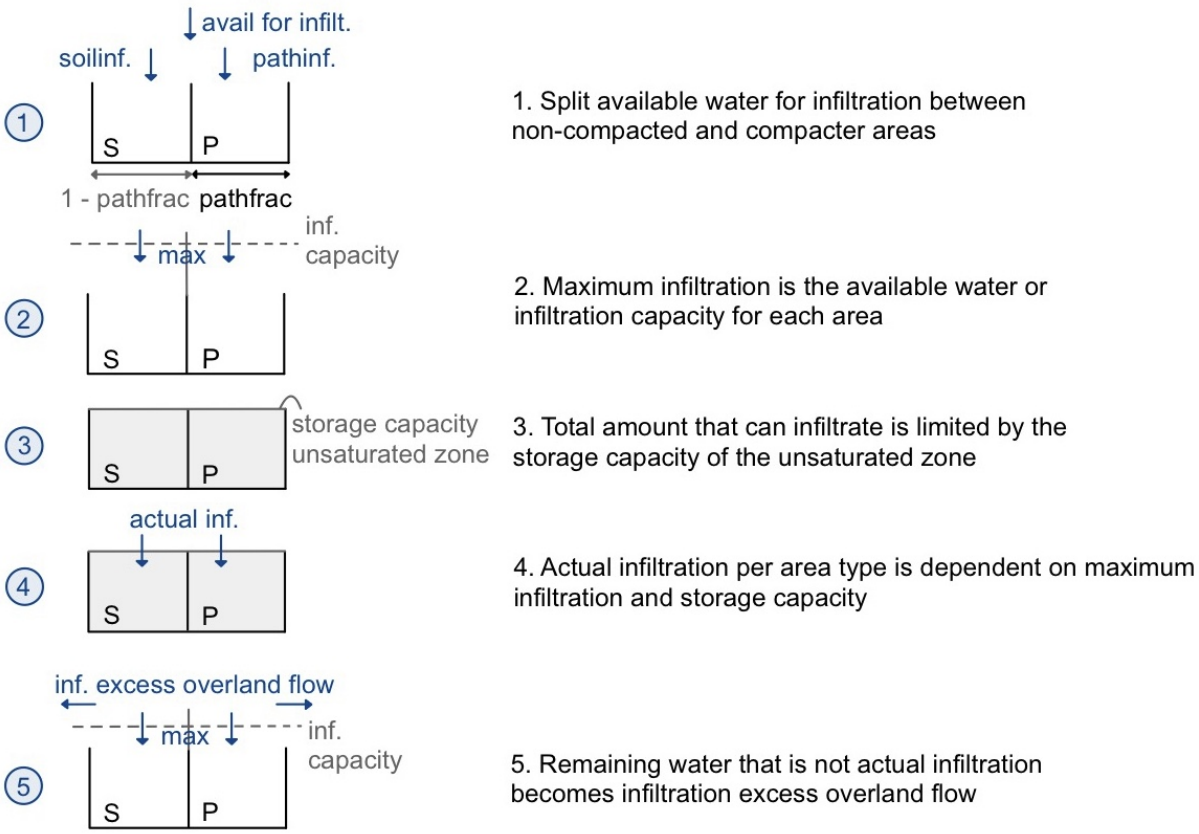


Figure G.7: Infiltration scheme in wflow_sbm

Detailed model description

1. Available water
The wflow_sbm model uses two area types for the infiltration process. It makes a difference between compacted area, called path, and non-compacted area, referred to as soil. Firstly, the available water for infiltration, $avail\_forinfilt$, is split over these areas dependent on the fraction of the surface that is compacted, indicated by $pathfrac$. This water division is given in eq. G.28.

$$\begin{aligned} pathinf &= avail\_forinfilt \cdot pathfrac \\ soilinf &= avail\_forinfilt \cdot (1 - pathfrac) \end{aligned} \tag{G.28}$$

2. Maximum infiltration for each area type
The maximum infiltration is equal to the amount of water available for that area, as was just determined in step 1. However, it is limited by the infiltration capacity of the soil. In other words, the maximum infiltration is equal to the available amount of water unless the infiltration capacity is lower. The maximum infiltration can thus be described with eq. G.29.

$$\begin{aligned} max\_infiltsoil &= min(infiltcapsoil, soilinf) \\ max\_infiltpath &= min(infiltcappath, pathinf) \end{aligned} \tag{G.29}$$

3. Actual infiltration capacity
The actual total infiltration capacity, $infiltsoilpath$, is equal to the sum of the maximum infiltration for the non-compacted and compacted areas. However, this sum cannot be larger than the storage of the unsaturated zone, $ustorecapacity$. You can compare it to the overflowing of a bucket. The formulation is shown in eq. G.30.

$$infiltsoilpath = min(max\_infiltpath + max\_infiltsoil, ustorecapacity) \tag{G.30}$$

4. Actual infiltration

A similar approach can be taken for the actual infiltration in the separate areas. This is quantified in eq. G.31. The model code ensures that the storage of the storage of the unsaturated zone is positive by taking a lower limit of 0. The fraction has an upper limit of 1, for the same reason as the limit in step 3 (the maximum total infiltration is limited by the storage capacity in the unsaturated zone). So when the storage is larger then the sum of the maximum infiltration values, the actual infiltration simply becomes the maximum infiltration from the infiltration capacity.

$$infiltsoil = max\_infiltsoil \cdot m \frac{ustorecapacity}{max\_infiltpath + max\_infiltsoil} \tag{G.31}$$
$$infiltpath = max\_infiltpath \cdot \frac{ustorecapacity}{max\_infiltpath + max\_infilsoil}$$

5. Infiltration excess overland flow
The infiltration excess overland flow, $infiltexcess$, is then quantified as the amount of water available for infiltration minus the maximum infiltration for both areas, as is shown in eq. G.32

$$infiltexcess = (soilinf - max\_infiltsoil) + (pathinf - max\_infiltpath) \tag{G.32}$$

Defining the soil layers

The soil in wflow_sbm can be divided into multiple layers by the user. The parameter $thicknesslayers$ is an array in which multiple thicknesses of layers can be entered. A standard for this is: [100,300,800], with a unit of millimeters as seen in fig. G.8.

However, these layers are not always the same as the total depth of these layers cannot be larger than the actual thickness of the soil. For this purpose, the function "set_layerthickness" is developed. The input is the reference or maximum depth, the cumulative soil depth and the pre-defined thickness per layer. The output is the actual thickness of the layers and the number of them. In the function, a loop iterates through each defined layer and determines its thickness and depth, if this is possible for the given soil thickness. This iteration is done with the following three options, which are checked by the loop in this order.

Given per iteration:

- Soil thickness (entire soil)
- Sum of pre-defined layer depths, starting at 0
- Layer thickness of layer i

Option 1:
Soil thickness > bottom depth of layer i:
Actual thickness = layer thickness

Option 2:
Soil thickness > upper depth of layer i:
Actual thickness = soil thickness - upper depth of layer i

Option 3: If both option 1 and 2 do not apply, no layer can be defined:
Actual thickness gets no value

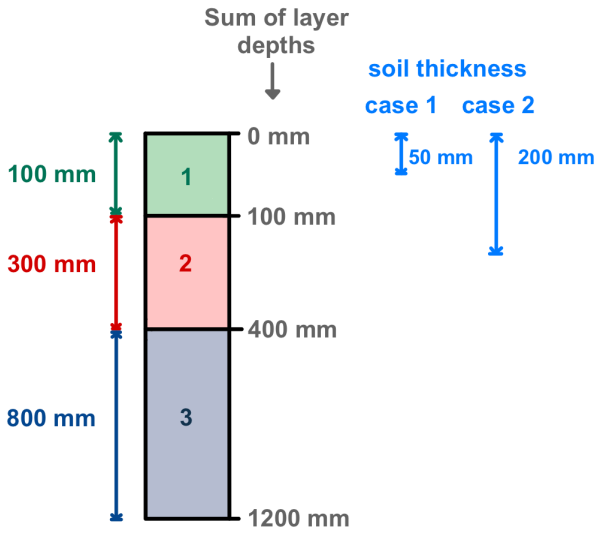


Figure G.8: Example of the "set_layerthickness" function

The example of fig. G.8 may clarify the function. It uses the standard definition of the *layerthickness* parameter of [100,300,800]. This results in the *sumlayers* parameter of [0,100,400, 1200]. For *sumlayers* the model simply cumulatively sums the thicknesses of the layers and adds 0 to the front of the array, to get the top and bottom depth of each layer.

For case 1, the soil thickness is set to 50 mm. For layer 1, with a thickness of 100 mm, option 1 does not apply, since the soil thickness (50 mm) is not larger than the bottom depth (100 mm). Option 2 applies as the soil thickness (50 mm) is larger than the upper depth of the layer (0 mm). The thickness of the first layer is thus equal to the soil thickness (50 mm) - minus the top depth of the layer (0m), so the thickness is 50 mm, equal to the maximum soil depth. The other values fall under option 3 and thus get no value.

For case 2, the soil thickness (200 mm) is larger than the bottom depth of the first layer (100 mm), so the actual thickness of this layer is taken as its pred-defined thickness (100 mm). For the second layer, the soil thickness is limiting. Option 1 does not apply, since the soil thickness (200 mm) is not larger than the bottom depth (400 mm). Option 2 applies, as the soil thickness (200 mm) is larger than the top depth of layer 2 (100 mm), so the actual thickness of layer 2 is the soil thickness (200 mm) minus the top depth (100 mm), resulting in an actual layer thickness of 100 mm.

These soil layers can now be split into unsaturated layers and saturated layers. Whereas the previous distinction was made based on the soil thickness as a reference depth, this split uses the pseudo-water table depth, zi in the "set_layerthickness" function. It works exactly the same, but the reference depth is thus different, as well as the user defined layer thicknesses which have been replaced by the actual layer thicknesses that have just been determined. The result is the layer thicknesses of the unsaturated layers, usl, as well as the amount of layers, n_usl. The total depth of the unsaturated layers is defined as z. The parameter $usld$ is equal to the amount of water in the unsaturated zone per layer, also described by the working parameter $ustorelayerdepth$.

Soil layers flow

Before looking at the evapotranspiration processes from the soil, the infiltrated water is distributed over the different layers that have just been determined. The following processes differ for the number of layers. All processes are based on only gravity based flow.

The first case is when there is only 1 layer set by the $thicknesslayers$ parameter and the $transfermethod$ option in the configuration file is set to true. Then the model uses the original Topog_SBM vertical transfer formulation, which is explained below.

1. Add infiltrated water to unsaturated store
In eq. G.33 the preceding amount of water in the first layer of the unsaturated zone is increased with

the total, actual infiltration.

$$ustorelayerdepth = ustorelayerdepth + infiltsoilpath \tag{G.33}$$

2. Calculate the vertical saturated hydraulic conductivity
The vertical saturated hydraulic conductivity, kv_z, is calculated using eq. G.34. It is equal to the standard equation of the saturated hydraulic conductivity, but with an added multiplication factor, $kvfrac$. The saturated hydraulic conductivity depends on the saturated conductivity at the soil surface, kv_0, the depth of the pseudo water table, zi (which is also the bottom depth of this layer) and a scaling parameter f. The saturated conductivity declines exponentially with soil depth and the smoothness is determined by the scaling parameter. For a lower value of this parameter, the decrease is more gradual, while a high parameter value indicates a sharp transition. This is shown in fig. G.9.

$$kv\_z = kvfrac \cdot kv_0 \cdot exp^{-f \cdot zi} \tag{G.34}$$

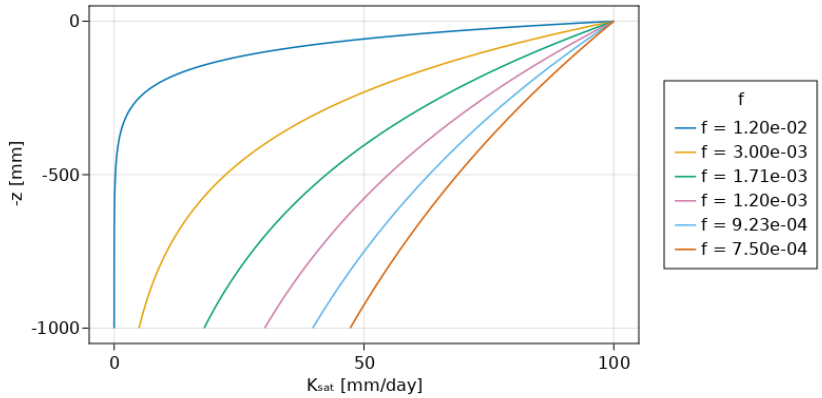


Figure G.9: The relation between soil depth z and saturated hydraulic conductivity for varying values of the scaling parameter, f

3. Calculate the saturation deficit.
The saturation deficit, sd, is defined as the difference between the capacity of the saturated zone, $soilwatercapacity$, and the amount of water in the saturated zone, $satwaterdepth$, see eq. G.35.

$$sd = soilwatercapacity - satwaterdepth \tag{G.35}$$

4. Calculate the flux to the saturated zone
The flux to the saturated zone, ast, is non-zero on the condition that the saturation deficit is larger than 0.00001. In other words, the storage capacity of the unsaturated zone must be larger than 0. When this condition is satisfied, the flux is based on eq. G.36.

$$ast = kv\_z \cdot (\frac{usl \cdot (\theta_s - \theta_r)}{sd}) \tag{G.36}$$

The vertical saturated hydraulic conductivity, kv_z, is multiplied by the fraction of the unsaturated store and the saturation deficit. This unsaturated store is here defined as the product of the unsaturated zone thickness, usl, and the difference of the saturated, θ_s, and residual soil water content, θ_r. This fraction has an upper limit, which is the amount of water in the unsaturated zone. This upper limit also applies to the entire flux. In other words, the flux of the unsaturated zone to the saturated zone cannot be larger than the amount of water in the unsaturated zone.

5. Update the storage in the unsaturated zone
Finally, the flux is subtracted from the amount of water in the unsaturated zone as this has flowed into the saturated zone. This is then updated in the $usld$ parameter.

$$ustorelayerdepth = ustorelayerdepth - ast \tag{G.37}$$

There is also the possibility that the number of unsaturated layers is not 1, then there is also flow between the unsaturated layers themselves. Then for each layer the following is done.

1. Calculate the available water amount in the unsaturated layer
The available water amount in the unsaturated layer, *l_sat*, is calculated based on the thickness of that layer, *usl*, and the saturated and residual soil water content, see eq. G.38.

$$l\_sat = usl \cdot (\theta_s - \theta_r) \tag{G.38}$$

2. Calculate the vertical saturated hydraulic conductivity
This calculation is the same as for the case of just 1 layer, see eq. G.34.

3. Update the amount of water in each layer
For the first layer, the same is done as for the 1 layer case, so the total, actual infiltration amount is added there. For the other layers, the flux, *ast*, of the previous layer is added.

4. Compute the flux from gravity to the next layer and the remaining amount of water
Assuming a unit head gradient, the transfer of water from an unsaturated store layer, *ustorelayerdepth*, is controlled by the vertical saturated hydraulic conductivity, *kv_z*, (at the bottom layer or water table), the effective saturation degree of the layer (which is the ratio of *ustorelayerdepth* and *l_sat*), and a Brooks-Corey power coefficient, *c*. When the unsaturated store for the layer is empty, nothing happens, as there is no water to flow to a next layer. Otherwise the following process is started. The flux, *st* is calculated with eq. G.39. The fraction to the power *c* has an upper limit of 1.

$$st = kv\_z \cdot \frac{ustorelayerdepth^{\,c}}{l\_sat} \tag{G.39}$$

The flux has a upper limit, *st_sat*, that is defined by eq. G.40 and that has a lower limit of 0. It is the difference of the water stored in the layer and the available water. To reduce overshooting, a loop is introduced in the function to limit the change in soil water per iteration step to a maximum 0.2 mm.

$$st\_sat = ustorelayerdepth - l\_sat \tag{G.40}$$

The unsaturated store of the layer is then updated by subtracting the flux. The function returns the new unsaturated store of the layer, and the flux to the next layer. This is repeated for each unsaturated soil layer, resulting in their updated amount of water, and a flux to the saturated zone, but this one is not yet transferred.

Soil layers evaporation

A considerable part of the initial potential evapotranspiration is already taken up by interception and open water evaporation, while the remaining part is split into potential soil evaporation and potential transpiration. First the evaporation of all soil layers is determined.

The saturation deficit, *sd*, is calculated again as the difference of the soil water capacity and the amount of water in the saturated zone, see eq. G.35.

If there is only 1 layer defined in the *thicknesslayers* parameter, the evaporation from the unsaturated soil, *soilevapunsat* is calculated with eq. G.41. The parameter *potsoilevap* was the remaining potential evaporation for the bare soil after the open water evaporation.

$$soilevapunsat = potsoilevap \cdot \frac{sd}{soilwatercapacity} \tag{G.41}$$

If the number of unsaturated soil layers is 1, the evaporation from the unsaturated soil is calculated with eq. G.42. The fraction has an upper limit of 1 as there cannot suddenly be more evaporation than the potential. The actual evaporation from the soil layer is thus dependent on the amount of water present, the water table depth and the water content.

$$soilevapunsat = potsoilevap \cdot \frac{ustorelayerdepth}{zi \cdot (\theta_s - \theta_r)} \tag{G.42}$$

If the number of unsaturated soil layers is higher than 1, a similar equation (eq. G.43) is used, but the water table depth is replaced by the thickness of the first unsaturated soil layer, *usl*.

$$soilevapunsat = potsoilevap \cdot \frac{ustorelayerdepth}{usl \cdot (\theta_s - \theta_r)} \tag{G.43}$$

So the only evaporation from the unsaturated zone originates from the first unsaturated soil layer and this is limited by the available amount of water in that layer. This evaporation is then subtracted from the initial remaining potential evaporation of the bare soil, which is used next for the saturated soil. The amount of water in the first unsaturated soil layer is updated by subtracting its evaporation.

The evaporation from the saturated soil is only non-zero when the number of unsaturated soil layers is 0 or 1. In other words, there are more than 1 layers defined and a saturated layer is at the surface or below 1 unsaturated layer. When this is the case, the evaporation of the saturated soil, *soilevapsat*, is calculated with eq. G.44. The actual evaporation is thus dependent on the actual thickness of the first layer and the water table depth.

$$soilevapsat = potsoilevap \cdot \frac{act\_thickl - zi}{act\_thickl} \tag{G.44}$$

There is also an upper limit of the evaporation of the saturated soil, eq. G.45 to ensure that there is not more evaporation than possible.

$$soilevapsat\_max = (act\_thickl - zi) \cdot (\theta_s - \theta_s) \tag{G.45}$$

Finally, the total evaporation, *soilevap*, can be computed with eq. G.46.

$$soilevap = soilevapunsat + soilevapsat \tag{G.46}$$

The amount of water in the saturated zone is then updated by subtracting the evaporation of the saturated zone, see eq. G.47.

$$satwaterdepth = satwaterdepth - soilevapsat \tag{G.47}$$

Transpiration

Plants take up water from the soil and transpire this as water vapour to the air. As roots can extract water from deeper soil layers, this can play an important role in the hydrological cycle. In wflow_sbm the uptake of water occurs in the saturated and unsaturated zone.

For the saturated zone, the fraction of wet roots is determined for an assumed rooting depth and the water table depth using a sigmoid curve for which the sharpness can be changed with the model parameter "rootdistpar".

The transpiration from the saturated zone, *actevapsat* (although a confusing term) is calculated with eq. G.48 and has an upper limit of the amount of water in the saturated zone, *satwaterdepth*. The transpiration is thus simply the amount of water remaining from the potential evapotranspiration for the canopy covered areas multiplied with the fraction of wet roots.

$$actevapsat = pottrans \cdot wetroots \tag{G.48}$$

The amount of water in the saturated zone is then updated by subtracting the transpiration. The same goes for the remaining potential evapotranspiration. This rest amount can be transpired from the unsaturated zone. The unsaturated soil zone can be divided into multiple layers but for simplicity, this section only discusses the zone as 1 layer. In case of multiple layers, this process is simply repeated for each layer. The maximum extraction depends on the fraction of the unsaturated zone depth that is above the rooting depth. This fraction, *availcap* is derived from eq. G.49, where *rootingdepth* is the depth of the maximum depth of the roots below the surface and *sumlayer* is the depth of the upper boundary unsaturated layer. An example is shown in fig. G.10. The allowed value of *availcap* lies between 0 and 1.

$$availcap = \frac{rootingdepth - sumlayer}{usl} \tag{G.49}$$

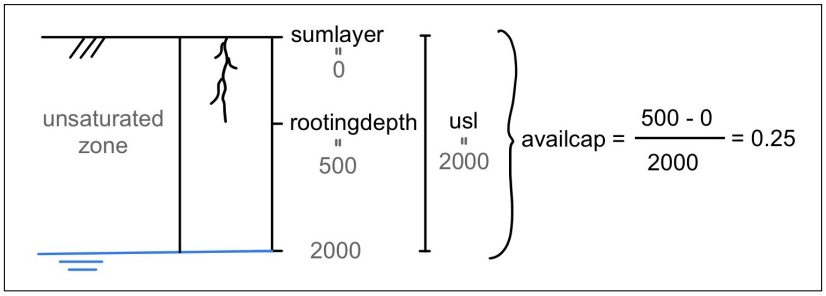


Figure G.10: Figure visualising eq. G.49

The maximum water extraction from the unsaturated zone can be calculated from eq. G.50. It is simply the amount of water stored in the unsaturated zone multiplied with the fraction of the unsaturated zone that is above the rooting depth and thus can reach water for transpiration.

$$maxextr = availcap \cdot ustorelayerdepth \tag{G.50}$$

A reduction coefficient is introduced since the transpiration is also dependent on the soil water pressure. This has been quantified by Feddes & Zaradny (1978) in what is now known as the Feddes curve which depends on soil water pressure. When the soil water pressure is lower than the wilting point, no water can be taken up by the roots and so the reduction coefficient is 0. This increases linearly to 1 when the soil water pressure is at least higher than the critical soil moisture content. This is visualised in fig. G.11

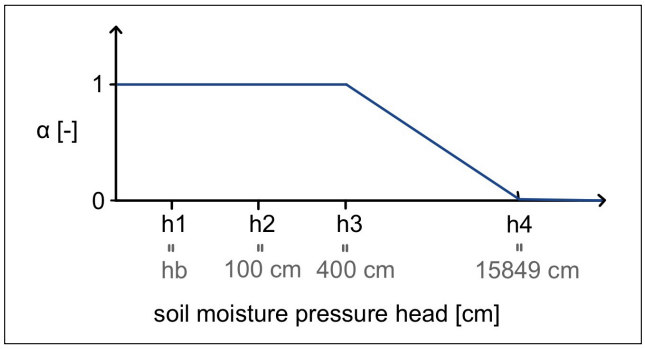


Figure G.11: Feddes curve as defined in the wflow_sbm model

The four identifying parameters of the soil moisture pressure head are given in table G.4

Table G.4: Values and meaning of the parameters of the Feddes curve in the wflow_sbm model

Symbol	**Value [cm]**	**Meaning**
h_1	h_b	air entry pressure
h_2	100	pF2 for field capacity
h_3	400	pF3 for critical point
h_4	15849	pF4.2 for wilting point

The volumetric water content, *vwc*, is defined by eq. G.51.

$$vwc = \frac{ustorelayerdepth}{usl} \tag{G.51}$$

The head is then calculated with eq. G.52.

$$head = \frac{h_b}{(\frac{vwc}{\theta_s - \theta_r})^{1/\lambda}} \tag{G.52}$$

$$\text{with } \lambda = \frac{2}{c-3}$$

Using the calculated head and the Feddes curve, the reduction coefficient, α, can be derived. The amount of transpiration is then quantified as the multiplication of the reduction coefficient and the maximal extraction. However, the maximal extraction can not be higher than the amount of water in the unsaturated zone nor the rest of the evaporation, as was just computed after the transpiration from the saturated zone. The resulting actual transpiration is then subtracted from the rest evaporation and the storage in the unsaturated zone is decreased with the transpiration flux.

Excess water

The amount of water in the unsaturated soil layers is updated from the bottom to the top, by checking the water balance and updating the amount of water in the layer if necessary. This is done with the imaginary flux, du, which is calculated with eq. G.53 and has a lower limit of 0. It is equal to the difference between the calculated water in the layer, $usld$ and the actual amount of water that can be stored or extracted, which depends on its thickness, usl, and water content, θ.

$$du = usld - usl \cdot (\theta_s - \theta_r) \tag{G.53}$$

The amount of water in the unsaturated layer is then updated by subtracting this imaginary flux from the current layer and adding it to layer above.

The remaining excess of water in the soil layers, so the du value of the top soil layer, is considered as a decrease to the actual infiltration, $actinfilt$, in eq. G.54. Remember that $infiltsoilpath$ is the infiltration amount from the infiltration process, which was only limited by the infiltration capacity.

$$actinfilt = infiltsoilpath - du \tag{G.54}$$

The total of excess water, $excesswater$, then becomes with eq. G.55

$$excesswater = avail\_forinfilt - infiltsoilpath - infiltexcess + du \tag{G.55}$$

The excess water is split up in the soil and path, based on the canopy gap fraction, as was also done before for the potential evaporation with eq. G.26.

Capillary rise

The actual capillary rise, $actcapflux$, is calculated only if there are unsaturated soil layers. When this is the case, the vertical saturated hydraulic conductivity is calculated again, using eq. G.34.

The initial capacity of the unsaturated store, $ustorecapacity$, is again calculated with eq. G.27 as the total capacity of the soil minus the amount of water present in the saturated and unsaturated zone.

The maximum capillary rise, $maxcapflux$, is then the minimum of the vertical hydraulic conductivity, kv_z, the actual transpiration from the unsaturated zone, $actevapustore$, the initial capacity of the unsaturated store, $ustorecapacity$, which has just been calculated, and the amount of water in the saturated zone, $satwaterdepth$.

Capillary rise is only possible when the water table depth is larger than the rooting depth. In this case, the net capillary rise, $netcapflux$, is calculated with eq. G.56. The value for the water table depth, zi, has an upper limit of the controlling parameter, cap_hmax. Another controlling parameter is the coefficient, cap_n.

$$netcapflux = maxcapflux \cdot (1 - \frac{zi}{cap\_hmax})^{cap\_n} \tag{G.56}$$

For each layer, from the bottom to the top, an added factor, $toadd$, is determined which is the $netcapflux$ parameter or its maximum, $netcapfluxmax$, calculated with eq. G.57.

$$netcapflux\_max = usl \cdot (\theta_s - \theta_r) - usld \tag{G.57}$$

The unsaturated store of that layer is then increased with this added factor, $toadd$. At the same time, the $netcapflux$ for the next iteration step is decreased with the added factor and the actual capillary rise is increased with it.

Leakage and recharge

Using the unit head gradient, gravity based flow has been determined through all soil layers. This flow also occurs at the bottom of the soil. As there is not yet a groundwater store, this is seen as leakage and

is calculated with eq. G.58. The maximum is the amount of water in the saturated zone, *satwaterdepth* and a user-defined value of the maximum leakage, *maxleakage*. The leakage cannot be negative.

$$actleakage = kv_0 \cdot exp^{-f \cdot soilthickness} \tag{G.58}$$

The recharge to the saturated zone is then calculated with eq. G.59. The actual amount of water transferring to the saturated zone is thus equal to the transfer of water from the unsaturated to the saturated zone, while substracting the capillary rise from the unsaturated zone (which had not yet been taken into account), the leakage (from the saturated soil to below), the transpiration and evaporation from the saturated zone.

$$recharge = ast - actcapflux - actleakage - actevapsat - soilevapsat \tag{G.59}$$

The total transpiration is then determined by eq. G.60, by taking the transpiration of the saturated and unsaturated store.

$$transpiration = actevapsat + actevapustore \tag{G.60}$$

The actual evapotranspiration is determined by eq. G.61 and consists of the total soil evaporation, the total transpiration and the river and open water transpiration.

$$actevap = soilevap + transpiration + ae\_openw\_r + ae\_openw\_l \tag{G.61}$$

Exfiltration from unsaturated zone

After the lateral flows have been processed in the model, the water table depth has been updated. A possible result is that the water table could be higher than the soil surface for some cells. Especially in steep, river valleys this may be the case. This additional step accounts for this exfiltration from the unsaturated zone.

First the layer thicknesses are set again with the "set_layerthickness" function, with the updated water table depth as reference depth. A loop is run through for each original unsaturated soil layer with a storage resulting from the previous steps of vertical processes. Two options exist for each layer. If the original unsaturated soil layer is still unsaturated, the exfiltration, $exfiltustore$, is calculated with eq. G.62, and is thus dependent on the original amount of water in the layer, its thickness and the water content.

$$exfiltustore = usld - usl \cdot (\theta_s - \theta_r) \tag{G.62}$$

If this is not the case, so the unsaturated layer has become a saturated layer due to the rise in the water table, the exfiltration is calculated with eq. G.63 and is simply equal to the amount of water in the layer.

$$exfiltustore = usld \tag{G.63}$$

The amount of water in the layer is updated by subtracting the exfiltration and the layer above it is replenished with that amount of water.

With the new water table depth, the amount of water in the saturated zone, *satwaterdepth*, should also be updated. With eq. G.64, it becomes the following.

$$satwaterdepth = (soilthickness - zi) \cdot (\theta_s - \theta_r) \tag{G.64}$$

Runoff

The total runoff from the land cells is then determined by eq. G.65, which is the exfiltration from the unsaturated zone plus the water exfiltrating during saturation excess conditions plus the water that cannot infiltrate due to saturated soil plus the runoff of open water plus the infiltration capacity excess minus the evaporation of the open water. The runoff and evaporation of the river is left out as the cells are only land cells and not river cells.

$$\begin{aligned} runoff = {} & exfiltustore + exfiltsatwater + excesswater + runoff\_land \\ & + infiltexcess - ae\_openw\_l \end{aligned} \tag{G.65}$$

Lateral processes

There are three major flow paths of the water in the lateral direction, as shown in fig. G.12. There is the flow between the river cells and the flow between the land cells on the surface and in the subsurface. Firstly, this chapter will focus on the surface flows, both for the river and overland flow, which are referred to as river and land routing. The surface flows are modelled with two possible routing schemes: the kinematic wave and the local inertial approximation of shallow water flow. Both will be discussed, although wflow_sbm uses the kinematic wave by standard. Finally, this chapter describes the subsurface flows which uses a kinematic wave approximation. In contrast with the previous chapter, this description is not made by a combined study of the documentation and the model code, but by only the documentation provided by Deltares.

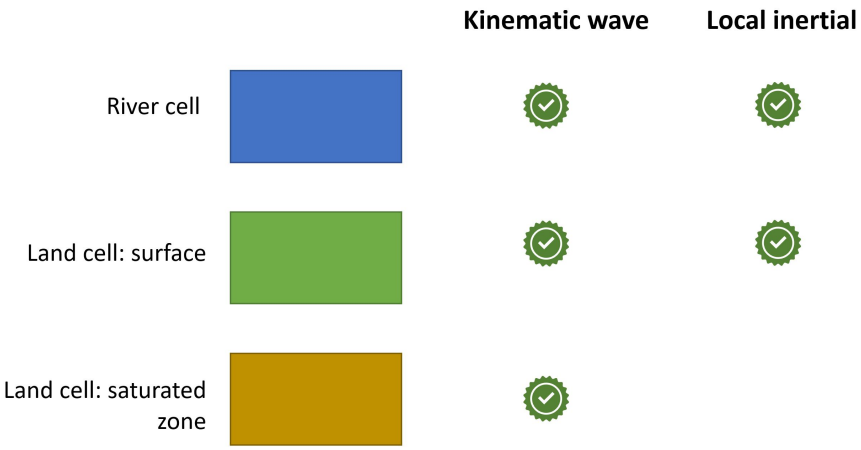


Figure G.12: Lateral flow paths and their shallow water flow approximation

Kinematic wave

As the wflow_sbm model is based on the fact that the topography controls the water flow, the kinematic wave approach is applied by standard to both the river flow and the overland flow. The kinematic wave equations by Chow et al. (1988) can be combined to 1 equation, see eq. G.66. Now the equation is only a function of stream flow, or rather the surface runoff in the kinematic wave, Q. Other parameters are the length of the runoff pathway, x, the integration time step, t, the coefficients, α and β. The cross-section area of the runoff pathway, A has been eliminated in the combined equation by using its own definition.

$$\frac{dQ}{dx} + \alpha\beta Q^{\beta-1}\frac{dQ}{dt} = q \text{ from } \begin{cases} \frac{dQ}{dx} + \frac{dA}{dt} = q \\ A = \alpha Q^{\beta} \end{cases} \tag{G.66}$$

The equation is solved with a nonlinear scheme using Newton's method. This method includes an iteration until the solution has stabilised. The α parameter depends on the wetted parameter and therefore, depends on the water level. Due to this dependence, α parameter is also iteratively solved with eq. G.66. For large models, the number of iterations can be fixed to limited computation times.

The surface flows may not be correctly routed with the kinematic wave approach for less steep terrain as the approach neglects pressure differences and inertial momentum, two characteristics that become more relevant for less steep terrain.

Local inertial

Whereas the kinematic wave approximation of shallow water flow also ignored the pressure differences and inertial momentum, the local inertial approximation only neglects the convective acceleration term in the Saint-Venant momentum conservation equation. Bates et al. (2010) developed the numerical solution, eq. G.67, of this approximation for a staggered grid. The current time step of a variable is indicated by the subscript t, while the next time step is indicated with $t + \Delta t$. The variables in this equation are the river flow Q, the gravity constant, g, the cross-section flow area, A, the hydraulic radius, R, the water surface slope, S, and the Manning's roughness coefficient, n. This momentum equation is applied to each link between two river grid cells.

$$Q_{t+\Delta t} = \frac{Q_t - gA_t\Delta tS_t}{(1 + g\Delta tn^2|Q_t|/R_t^{4/3}A_t)} \tag{G.67}$$

This momentum equation may become unstable for steeper slopes. Therefore, a default limit is applied for the Froude number. This value remains lower than 1, thus keeping subcritical flow conditions, by reducing the water flow (and thus velocity) if this limit is about to be exceeded.

With the momentum equation defined, the continuity equation completes the shallow water flow description. In contrast with the momentum equation, the continuity equation, eq. G.68, is only applied on each river cell itself. The variables are the water depth, $h[m]$, the river area, $A[m^2]$, and the up-, Q_{src}, and downstream river flow link, Q_{dst}, of the cell.

$$h^{t+\Delta t} = h^t + \Delta t \cdot \frac{Q_{src}^{t+\Delta t} - Q_{dst}^{t+\Delta t}}{A} \tag{G.68}$$

A variable that plays an important role in both equations, is the model time step of the local inertial model, Δt. This is not the same time step as the overall model time step which can be daily or hourly or even smaller. The local inertial time step is estimated by the Courant-Friedrichs-Lewy (CFL) condition, eq. G.69. This condition ensures a stable solution for the numerical model. The water depth and gravity constant are combined to the wave celerity of the river cell i, $\sqrt{(gh_i)}$. The river length of the river cell i is indicated by $\Delta x[m]$ and finally a coefficient α is added to enhance the stability of the simulation. Its values usually lie between 0.2 and 0.7.

$$\Delta t = min(\alpha \cdot \frac{\Delta x_i}{\sqrt{(gh_i)}}) \tag{G.69}$$

Whereas the kinematic wave application uses the same scheme for the land and river cells, the local inertial approximation for the river cells differs from the approximation of the land cells, so the overland flow. For 2D overland flow on a staggered grid, the numerical scheme of de Almeida et al. (2012) is applied. Written as the explicit solution of water discharge between two cells in the x-direction, one gets eq. G.70.

$$Q_{i-1/2}^{n+1} = \frac{[\theta Q_{i-1/2}^n + \frac{(1-\theta)}{2}(Q_{i-3/2}^n + Q_{i+1/2}^n] - gh_f\frac{\Delta t}{\Delta x}(\eta_i^n - \eta_{i-1}^n)\Delta y}{1 + g\Delta tn^2|Q_{i-1/2}^n|/(h_f^{7/3}\Delta y)} \tag{G.70}$$

This numerical scheme is visualised in fig. G.13. From this, one can see that the superscript n refers to the space index. So be aware that this is an index and not a power. The subscript is still the time index, but $t + \Delta t$ has been replaced by $i + 1$, so forward and backward steps can be taken. The water flow, Q is defined at the link between two nodes, while the water surface elevation, η, is defined at the centre of the nodes. The water depth, h_f, is defined between cells. The distance between the centres of the nodes is Δx, while Δy is the flow width.

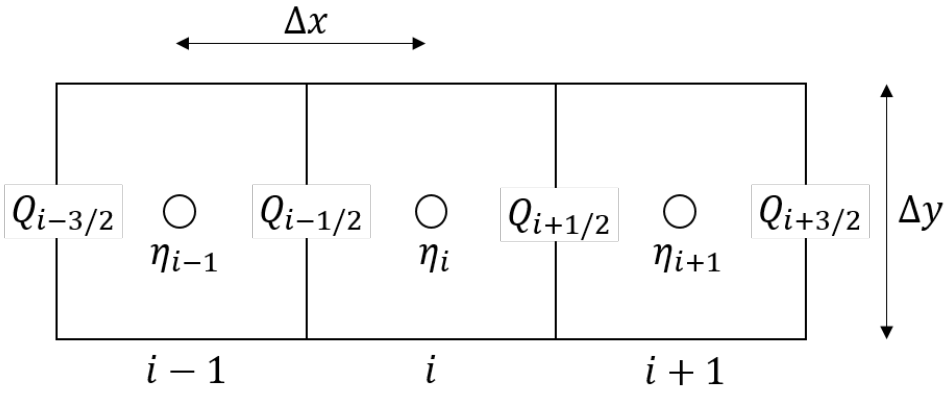


Figure G.13: Visualisation of the numerical scheme of eq. G.70

In wflow, a D8 and D4 scheme is used to determine the flow direction, as shown in fig. G.14. The water flows to the lowest of the surrounding cells when this is lower than the water level of the cell in question. As the name suggests a D8 routing method allows the water to flow to all 8 directions surrounding the cell, while a D4 method only allows flow to fully adjacent cells, as is indicated in green in fig. G.14. A D8 scheme is used for the river routing and a D4 scheme is used for the land routing in combination with the previously described local inertial schemes. This means that the river flow is modelled in 8 directions and the land flow is in 4 directions, which may lead to slower flow times on land.

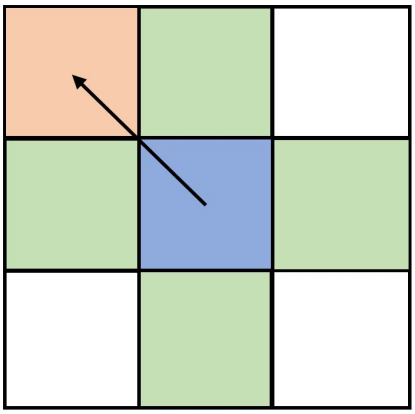


Figure G.14: D8 scheme and D4 scheme (green)

Subsurface flow routing

The lateral subsurface flow of wflow_sbm is always modelled with the kinematic wave approach. The water flows laterally from the saturated store by saturated downslope subsurface flow per unit width of slope w, as is defined in eq. G.71. This subsurface flow per unit width, q, is dependent on multiple parameters that have also been mentioned in the section on vertical processes. We have already seen the saturated hydraulic conductivity at the soil surface, K_0, the water table depth, z_i, the total soil depth, z_t, and the scaling parameter, f. Added is the element slope angle, β [°].

$$q = \frac{K_0 tan(\beta)}{f} \cdot (e^{-f \cdot z_i} - e^{-f \cdot z_t}) \tag{G.71}$$

This can be combined with the continuity equation, eq. G.72. As for previous equations, the saturated, θ_s, and residual soil moisture content, θ_r, play a role. As well as the water table height, h, the distance downslope, x, and the net input rate to the saturated store, r $[m/t]$

$$(\theta_s - \theta_r)\frac{\delta h}{\delta t} = -w\frac{\delta q}{\delta x} + wr \tag{G.72}$$

Rewriting eq. G.72 to use the subsurface flow instead of the water table height using ($h\frac{\delta q}{\delta h}$), results in eq. G.73, which is then again iteratively solved using Newton's method.

$$w\frac{\delta q}{\delta t} = -cw\frac{\delta q}{\delta t} + cwr \tag{G.73}$$

The celerity, c, is given in eq. G.74.

$$c = \frac{K_0 tan(\beta)}{\theta_s - \theta_r} \cdot e^{-f \cdot z_i} \tag{G.74}$$

For the subsurface flow, the use of the kinematic wave approach may not always be appropriate as the hydraulic gradient is not equal to the surface slope in less step terrain, an assumption that the kinematic wave approach relies on.

Additional features

External water can be supplied or abstracted for surface water routing, both as a cyclic parameter or as part of forcing. Similarly, reservoirs and lakes can be defined.

Wflow inner workings in coding

The source code of wflow is available here on Github. The relevant documents for the wflow_sbm model are the following files. The vertical processes per cell can be found in the sbm.jl file, in combination with several functions of the vertical_process.jl file. The steps of the vertical procceses in appendix G.2 can be found in the following functions. Items 1 and 4 fall under the function of *update_until_snow*. Items 5 to 15 are described in *update_until_recharge*. Finally, items 16 to 17 are described in *update_after_subsurfaceflow*. It updates all parameters of the wflow_sbm model for each cell for 1 timestep. The complete functioning of the model, including the lateral processes, is described in the sbm_model.jl file, which uses several functions of the horizontal_process.jl file.

The model code often works with max() and min() functions. These are not written in the formulas as they overcomplicate the equations. Instead, these maximum and minimum functions have been included in the text by describing upper and lower limits.

The use of wflow_sbm is not a topic of this document. However, a short note is required on the input file, as it is mentioned several times. Wflow uses an executable containing the model that requires a configuration file (with the .toml extension) to function. In this configuration file, the user can describe the model properties, such as the timesteps, period, output names, and so on. The configuration file also forms the link between other input files that are required, such as the forcing. Another input file is the staticmaps.nc file. This is a NETCDF that contains the user-defined static parameters. As it is a NETCDF, these parameters can vary spatially. So when this document mentions the input of static parameters, these values come from the staticmaps.nc file. For calibration, the parameter values can be changed directly in the staticmaps.nc file or in the configuration file.

Appendix H

Model calibration

This appendix gives additional information for the model calibration. This concerns the set-up of the calibration, including the calibration periods, the data check, the ranges of the HBV parameters. The appendix also includes the results of the calibration: the calibrated values of the HBV parameters, the performance for the non-event period and event period and the evaluation of the results for the non-event period. The Geul required additional further analysis for the calibration since a delay disturbed the initial results.

H.1 Data check

This section provides figures, fig. H.1, fig. H.2, fig. H.3, to compare the precipitation for the 2021 flood event of station measurements and of the point precipitation extracted from the radar products. This provides more context than the NSE values of fig. 5.4.

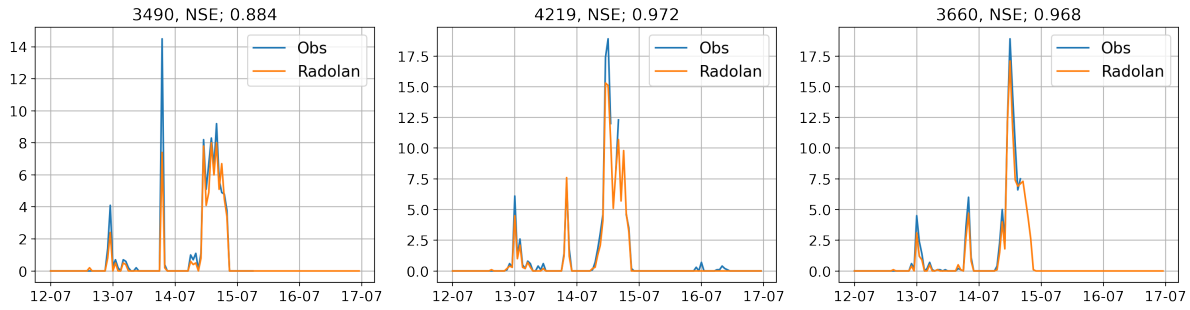


Figure H.1: Comparison of the precipitation between 12-07-2021 and 17-07-2021 for the relevant stations in the Ahr catchment. The station name and NSE value are given at the top of the figure.

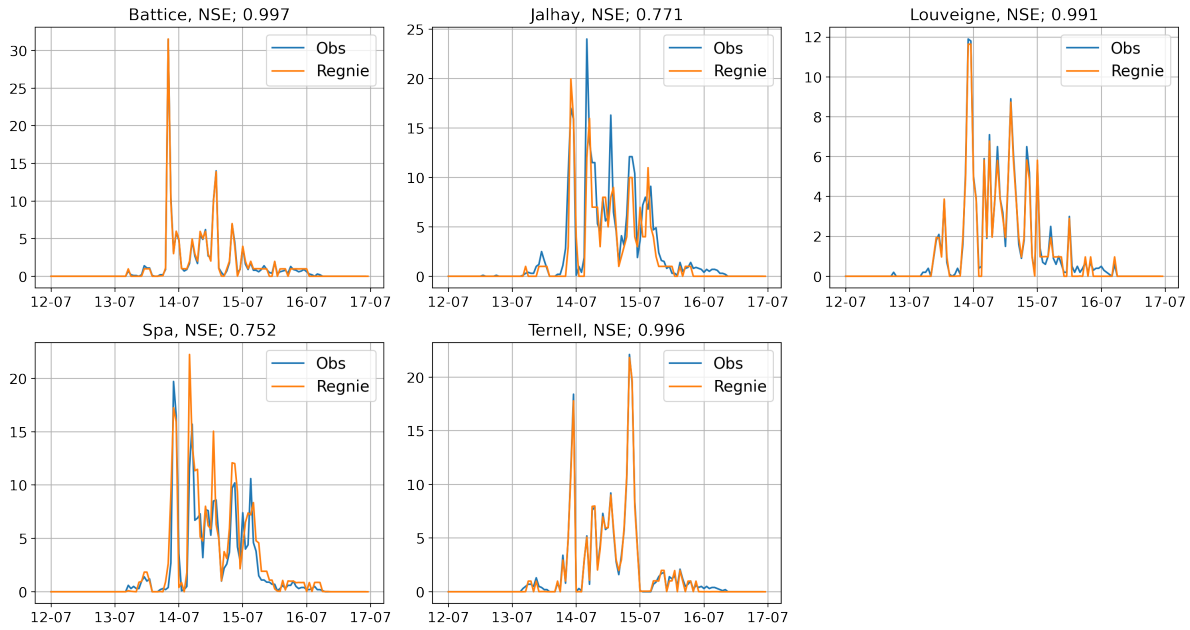


Figure H.2: Comparison of the precipitation between 12-07-2021 and 17-07-2021 for the relevant stations in the Vesdre catchment. The station name and NSE value are given at the top of the figure.

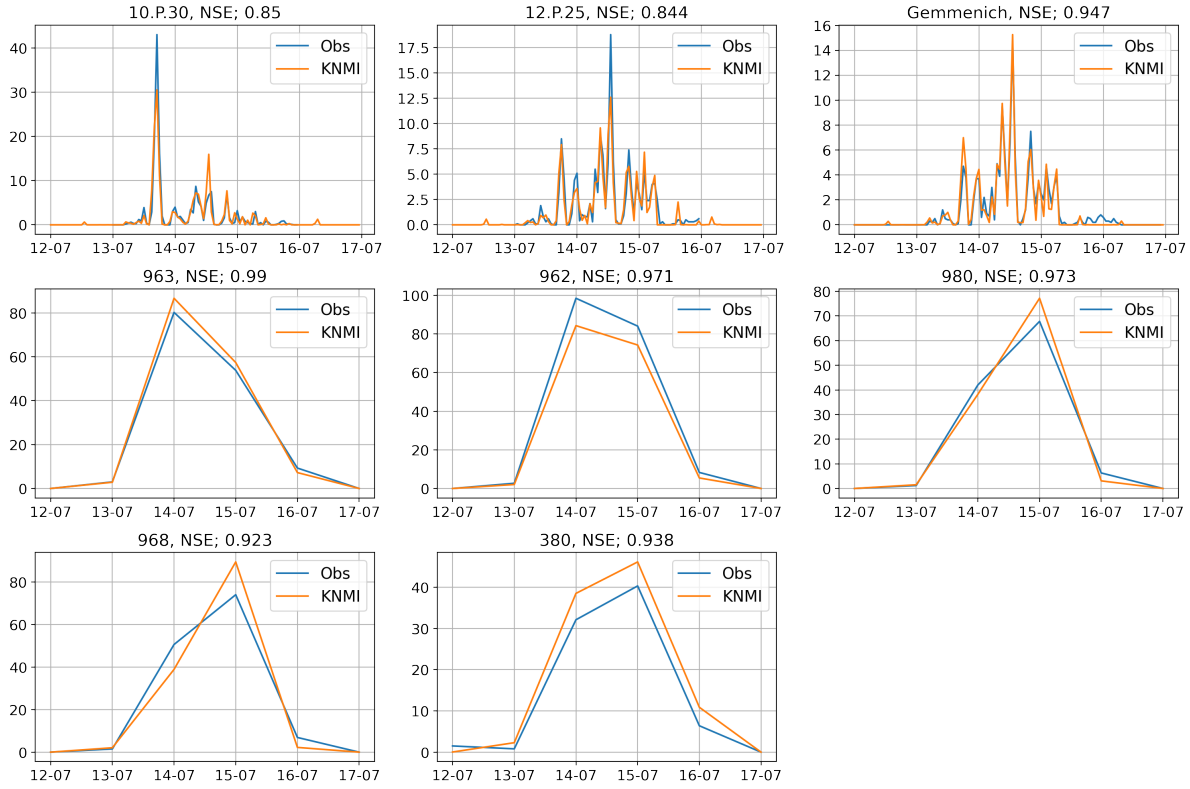


Figure H.3: Comparison of the precipitation between 12-07-2021 and 17-07-2021 for the relevant stations in the Geul catchment. The station name and NSE value are given at the top of the figure.

The good match between the data sets as indicated by the high NSE value is confirmed by the figures.

H.2 Calibration periods

Table H.1 provides the exact calibration periods for the models (HBV and SBM) for each catchment. They are referenced as the Nev period (or non-event period). The data period is slightly longer, as given in section 5.4, but some time is required for the warming-up which is not included in the calibration period.

Table H.1: Calibration periods for the non-event calibration.

	Start	**End**
Ahr	2018-03-01	2021-07-11
Vesdre	2016-01-01	2021-02-21
Geul	2019-03-01	2021-07-11

H.3 HBV parameter ranges for calibration

This section provides the value ranges for the HBV parameters for the Monte-Carlo calibration. The values are taken from the course CIE4431, but slightly adapted as recommended by Markus Hrachowitz. These values are given in table H.2 and have been used for the Ahr and Vesdre. Calibrating the Geul model with these parameter ranges resulted in a poor performance. The value ranges have been adapted in accordance with scatter plots and heatmaps from iterative runs indicating parameter ranges with good performance.

Table H.2: Parameter ranges for the HBV model calibration for the Ahr and Vesdre

	Minimal value	**Maximum value**
Imax	0	4
Ce	0.1	0.9
Sumax	40	700
β	0.5	2
Pmax	0.001	0.5
Tlag	0	10
Kf	0.01	0.7
Ks	0.0001	0.01

Table H.3: Parameter ranges for the HBV model calibration for the Geul

	Minimal value	**Maximum value**
Imax	0	4
Ce	0.05	0.6
Sumax	150	800
β	0.5	2.5
Pmax	0.001	0.3
Tlag	0	410
Kf	0.01	0.2
Ks	0.0001	0.01

H.4 Calibrated parameters of the HBV model

The calibrated parameters of the HBV model for both the non-event and event calibration periods can be found in table H.4, table H.5 and table H.6 for the Ahr, Vesdre and Geul. The factor between the original and re-calibrated value is given as well to show the degree in change. A change of twice or half

Table H.4: Parameter values for the HBV model calibrated for the pre-event period (Nev) and the flood event (Ev) for the Ahr. The factor between the original value and re-calibrated value is also provided. If the value at least doubled or halved, the factor is seen as significant and thus marked as bolt. Increases are given in green, decreases in red.

	Nev	**Ev**	**Factor**
Imax	1.29	1.04	0.81
Ce	0.78	0.69	0.88
Sumax	113.38	68.36	0.60
β	1.61	0.79	**0.49**
Pmax	0.01	0.45	**45.00**
Tlag	0.52	9.66	**18.58**
Kf	0.01	0.07	**7.00**
Ks	0.006	0.006	1.00

of the original value is considered as significant. This section also includes their expected effects on the model results.

According to table H.4, the following parameters for the HBV model of the Ahr show significant changes after the re-calibration for the 2021 flood event: β, Pmax, Tlag and Kf. The most significant increase is Pmax, which increases by a factor of 45. Parameter Pmax affects the percolation, as shown in eq. (G.7) in appendix G.1. This equation shows that the percolation flux to the slow responding store is equal to the filling degree of the unsaturated zone, multiplied by Pmax. A higher value, especially at a factor of 45, strongly increases the flux from the unsaturated zone to the slow responding store. This drains the unsaturated zone faster and leads to a higher slow groundwater flux. A quickly decreasing filling degree of the unsaturated zone could then also lead to a less flow to the fast responding store.

The parameter of Tlag increases from 0.5 h to 10 h. This will lead to a later peak discharge. Another strong increase can be found in parameter Kf, which influences the runoff from the fast responding store to the total runoff (before the convolution function). This flux is equal to the parameter Kf multiplied by the amount of water in the fast responding flux, see eq. (G.8). A higher value for Kf will result in a higher runoff from the fast responding store. Since more water runs off from this store per timestep, the runoff of the fast responding store also becomes faster.

The parameter β is halved, which is not a very big change compared to the other parameters. A decrease in β leads to an increased flux from the unsaturated store to the fast responding store in case of the same filling degree.The filling degree is determined by the amount of water in the unsaturated zone and its maximum capacity Sumax. Although the change in the parameter Sumax is not significant, the filling degree of the unsaturated store has a direct influence in half of all fluxes, and so the small change in Sumax still has considerable influence. Its decrease results in relatively higher filling degree for the same amount of water. A higher filling degree results in more water partitioning to the fast responding flux and also faster percolation from the unsaturated zone to the slow responding store. It represents a high saturation of the soil and the faster responding process is similar to overland flow, although it is not included explicitly in the model.

The parameter changes of the Ahr are expected to lead to a decline of the amount of water in the unsaturated zone due to increased percolation. The increased Tlag will lead to a delay of the peak discharge. The decrease in Kf will likely lead to higher and faster runoff from the fast responding store. The decrease in β and Sumax can lead strengthen this effect.

Table H.5: Parameter values for the HBV model calibrated for the pre-event period (Nev) and the flood event (Ev) for the Vesdre.

	Nev	**Ev**	**Factor**
Imax	2.07	2.80	1.35
Ce	0.11	0.43	**3.91**
Sumax	514.45	143.08	**0.28**
β	0.80	1.88	**2.35**
Pmax	0.12	0.26	**2.17**
Tlag	6.73	6.95	1.03
Kf	0.02	0.07	**3.50**
Ks	0.008	0.008	1.00

According to table H.5, the HBV model of the Vesdre has many parameter with significant changes: Ce, Sumax, β, Pmax and Kf. The increase for parameter Pmax is much lower than for the Ahr (2 instead of 45), but this increase will still lead to a higher percolation and thus quicker draining of the unsaturated store (and consequently the fast response store). The decrease of Sumax is larger than for the Ahr (0.32 instead of 0.6) and this lower Sumax will likely lead to more water partitioning to the fast responding store. For the Vesdre, β increases with a factor of 2.35 which would lead to less water partitioning to the fast responding store, as explained in appendix G.1 and so opposes the effect of a lower Sumax. However, since β is only an exponent to the filling degree, the overall decrease of Sumax is expected to have a larger effect. The increase in Kf will lead to a higher and faster response from the fast responding store, similarly to the Ahr. A new significant parameter is Ce, which influences the evapotranspiration from the unsaturated zone, see eq. (G.6). The increase in this parameter will lead to a lower evapotranspiration. This leads to more water remaining in the unsaturated zone and thus a higher filling degree. The effect is similar to a lower Sumax. In conclusion, the parameter changes in the Vesdre are expected to lead to a higher and faster response.

Table H.6: Parameter values for the HBV model calibrated for the pre-event period (Nev) and the flood event (Ev) for the Geul.

	Nev	**Ev**	**Factor**
Imax	2.27	1.14	0.50
Ce	0.07	0.42	**6.46**
Sumax	514.39	163.94	**0.32**
β	0.98	2.31	**2.36**
Pmax	0.18	0.22	1.22
Tlag	6.03	6.33	1.05
Kf	0.019	0.024	1.26
Ks	0.0004	0.0001	**0.25**

Table H.6 shows that the significant changes for the HBV Geul model occur for the following parameters: Ce, Sumax, β and Ks. The factors and effects of Sumax and β are similar as for the Vesdre, thus a relatively higher filling degree with a higher partitioning to the fast response flux, which is slightly damped by β. The effect of Ce is also similar, but the effect will be larger, thus leading to even less water evaporating or transpiring from the unsaturated soil and thus increasing the filling degree. As for the Ahr and Vesdre, the factor Kf increases but not significantly. Nevertheless, it will still lead to more and faster runoff from the fast responding store. A parameter that remains equal for the Ahr and the Vesdre, but not for the Geul, is Ks. It has a similar effect as Kf, but for the slow responding store. Its decrease will thus result in a lower and slower response from the slow responding store. In conclusion, the fast responding flux is expected to increase and the slow responding flux is expected to decrease.

H.5 Model performance pre-event

This section contains the hydrographs (both direct and cumulative) for the pre-event periods of the observation, HBV model and SBM model.

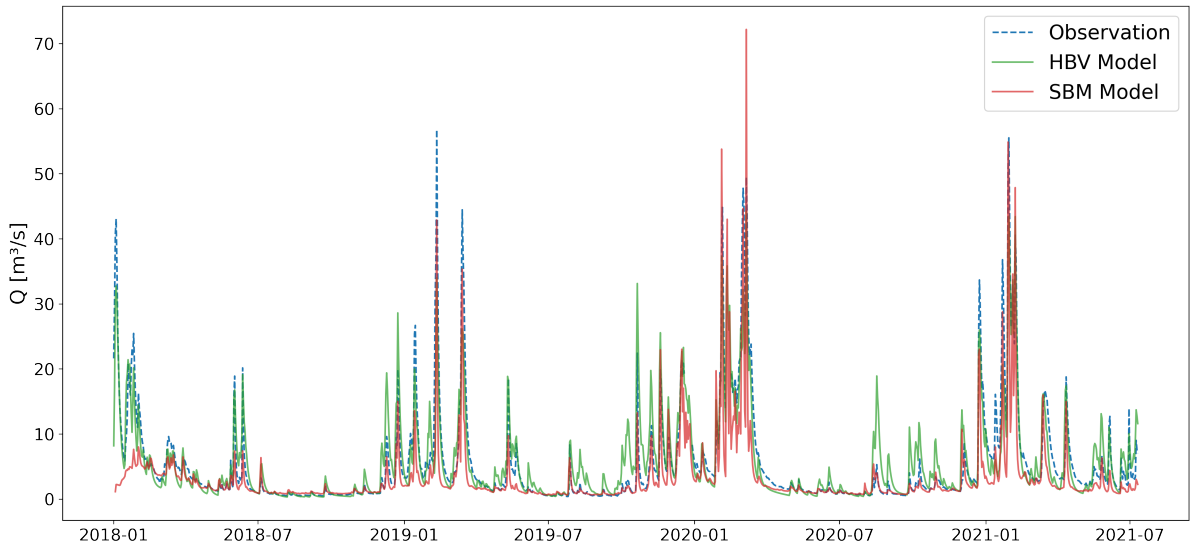


Figure H.4: Observed and modelled hydrographs (HBV and SBM model) of the pre-event period for the Ahr (Altenahr station). The warming-up period of 60 days is included in this figure, but not in the calibration.

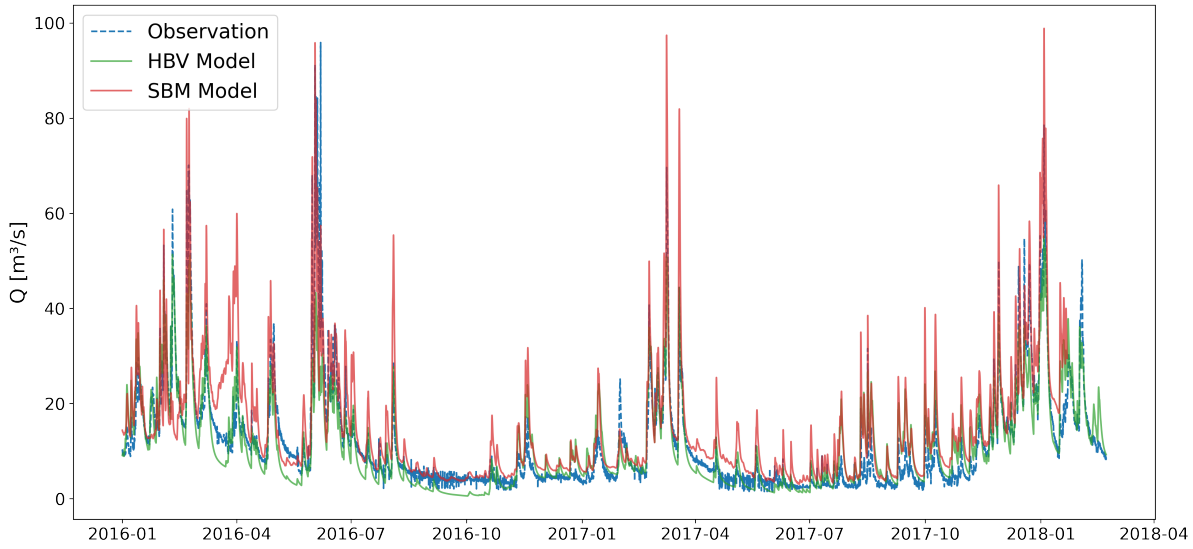


Figure H.5: Observed and modelled hydrographs (HBV and SBM model) of the pre-event period for the Vesdre (Chaudfontaine station).The warming-up period of 1 year is not included in this figure.

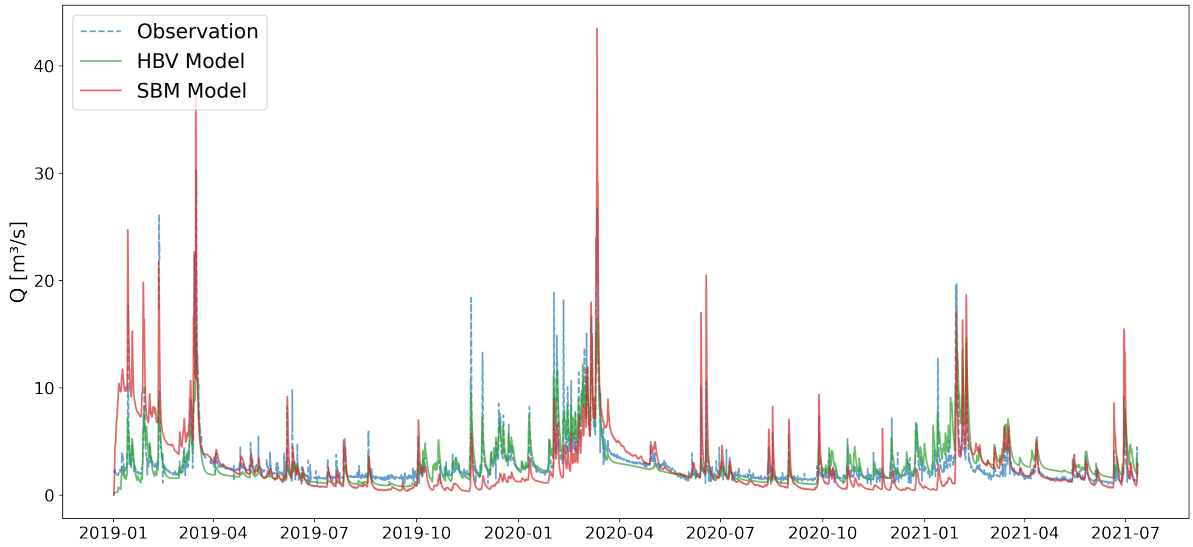


Figure H.6: Observed and modelled hydrographs (HBV and SBM model) of the pre-event period for the Geul (10.Q.36, Meerssen station). The warming-up period of 60 days is included in this figure, but not in the calibration.

H.6 Evaluation performance

This section provides the figures for the evaluation of the original models for other frequency domains compared to the calibration method. This includes visualisation of overestimation and underestimation for varying discharge ranges. The varying frequency domains are the cumulative hydrograph, the autocorrelation function and the flow frequency curves. The method of deriving these figures has been described in chapter 6.

Overestimation and underestimation visualisation for sorted discharge
The average daily discharge for each catchment has been sorted from high to low (so day zero of the figures has the highest discharge). The normalised difference of the modelled discharge and the observed discharge is then calculated for each day to detect overestimation or underestimation patterns for the different discharge ranges.

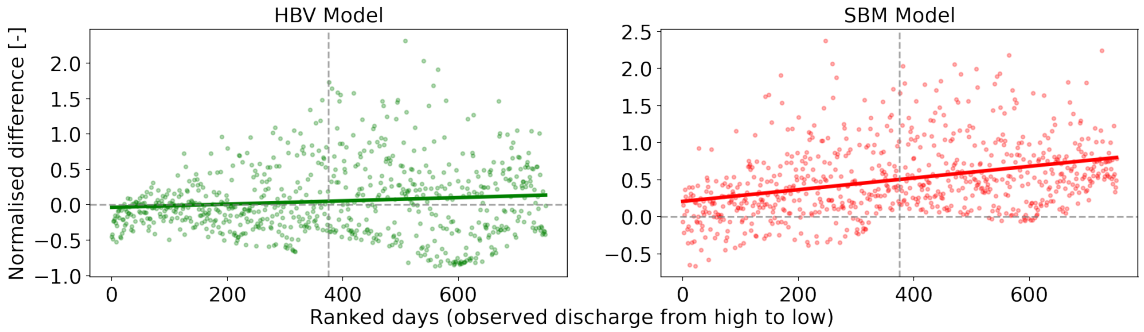


Figure H.7: Normalised differences between the simulated and observed discharge of the Vesdre (Chaudfontaine station) for each day, ranked from high to low discharge values. Both the HBV model (left) and SBM model (right) discharges are shown. The hourly discharges of the observations and simulation have been averaged for each day.

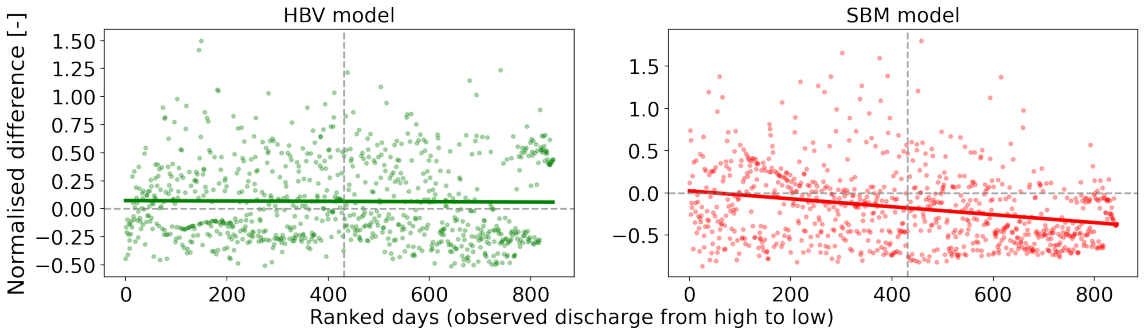


Figure H.8: Normalised differences between the simulated and observed discharge of the Geul (10.Q.36, Meerssen station) for each day, ranked from high to low discharge values. Both the HBV model (left) and SBM model (right) discharges are shown. The hourly discharges of the observations and simulation have been averaged for each day.

All models for the Vesdre and Geul show a wide spread of overestimation and underestimation. There is no clear trend for either HBV model. The SBM model for the Vesdre shows an increasing overestimation for lower discharges, confirming the previous statement that the model is incapable of replicating the varying baseflow of the Vesdre. For the Geul, there is a general underestimation which increases for lower discharges. These figures, along with fig. 6.4 confirm the statements of section 6.1.

Cumulative hydrographs

The cumulative hydrographs are an indication if the model has a good representation of the water balance. To accentuate this importance, the precipitation has also been added to the figures. This is the mean precipitation for the entire catchment, derived from the radar products.

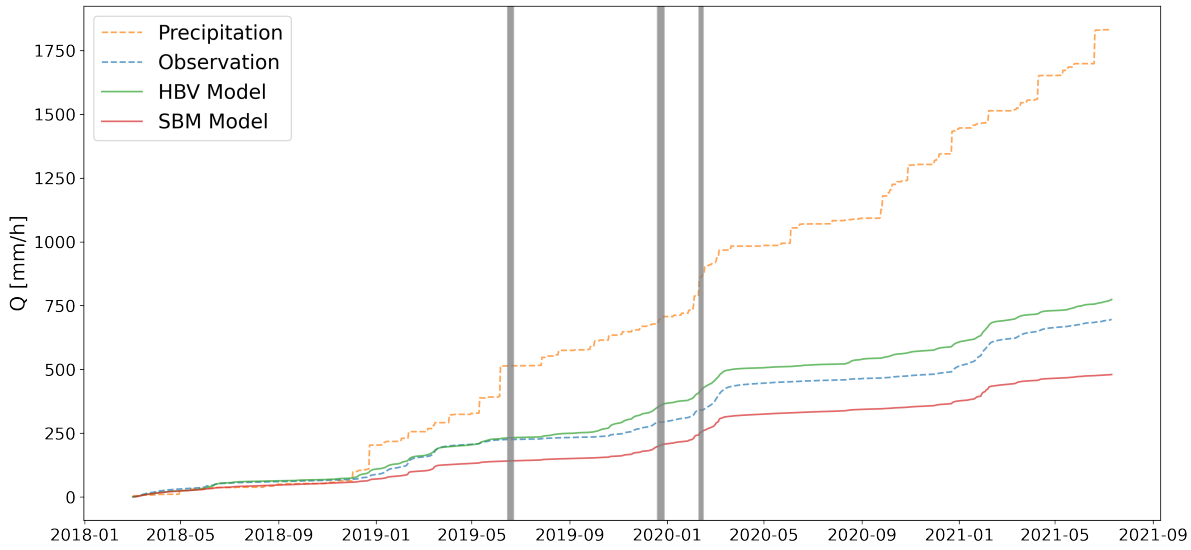


Figure H.9: Observed and modelled cumulative hydrographs (HBV and SBM model) of the pre-event period for the Ahr (Altenahr station). The precipitation has been is added for more context.

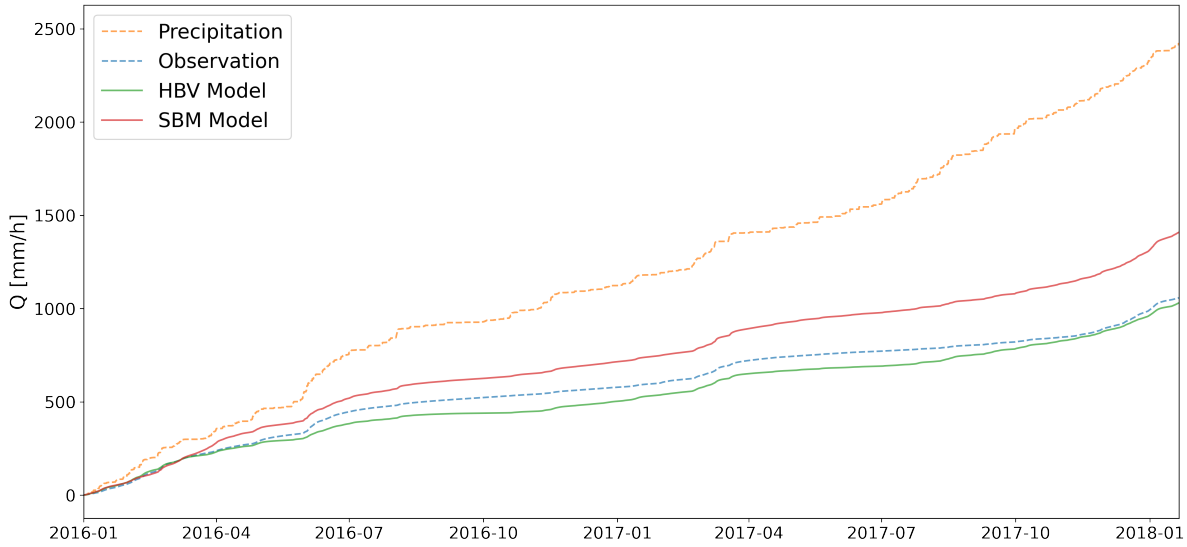


Figure H.10: Observed and modelled cumulative hydrographs (HBV and SBM model) of the pre-event period for the Vesdre (Chaudfontaine station). The precipitation has been is added for more context.

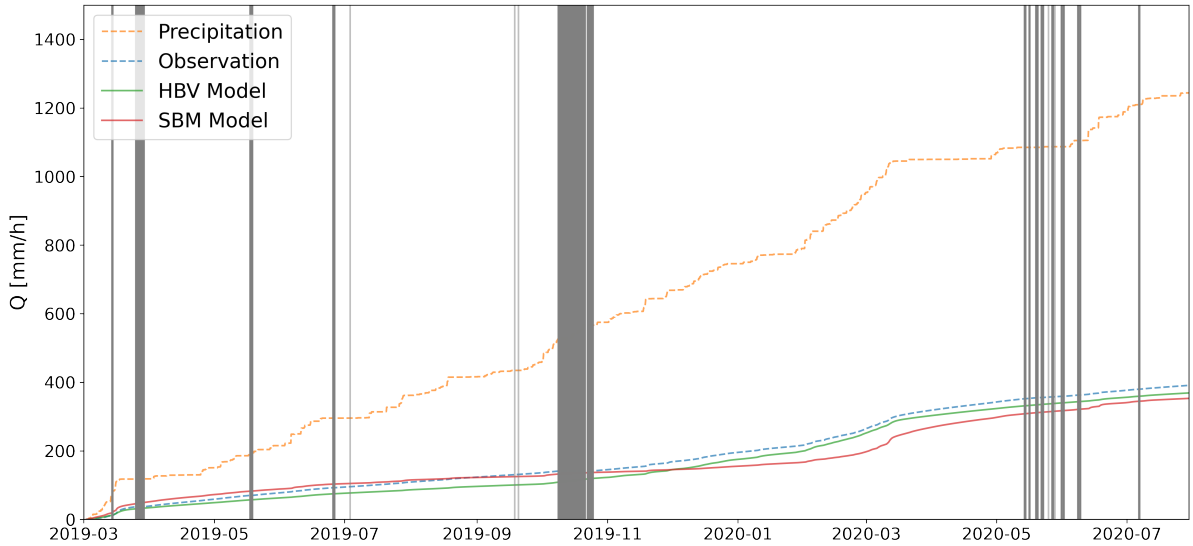


Figure H.11: Observed and modelled cumulative hydrographs (HBV and SBM model) of the pre-event period for the Geul (10.Q.36, Meerssen station). The precipitation has been is added for more context.

Due to some data gaps, it is not possible to make a completely accurate comparison, but it is still clear that the HBV model is better at representing the cumulative hydrograph than the SBM model.

Autocorrelation functions

The autocorrelation function compares the original data set with a copy of itself which shifts 1 day at each step. So for the figures below, 30 daily shifts have been applied. For each shift, a correlation coefficient can be calculated between the original data set with no shift and the shifted data set. For the modelled discharges, each shifted data set is compared to the observed data set with no shift. The autocorrelation function is usually applied to quantify the "memory" of the river A quickly decreasing correlation coefficient indicates a low memory, since a short time shift already leads to low correlation.

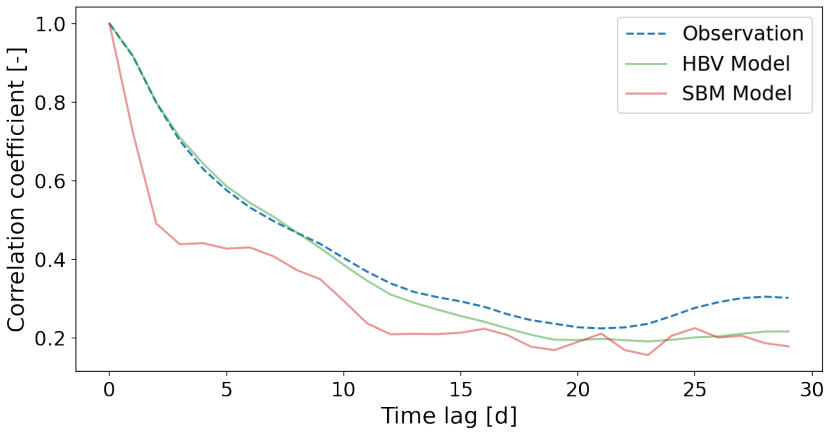


Figure H.12: Autocorrelation function and its correlation coefficient for time steps of 30 days for the discharge of the observation, HBV model and SBM model of the Ahr.

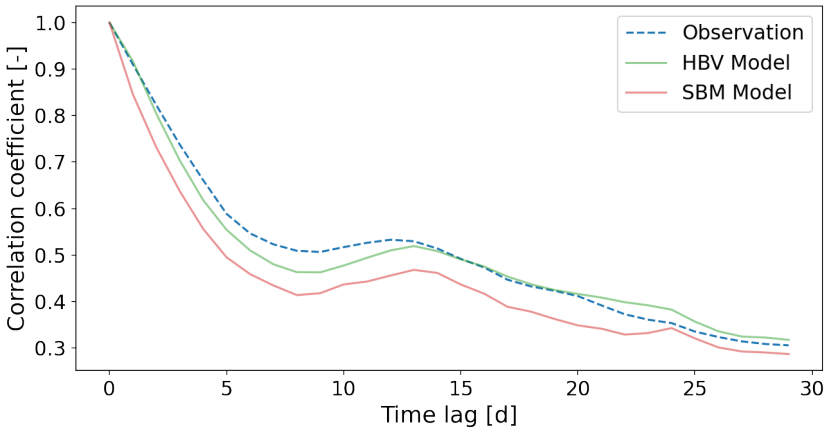


Figure H.13: Autocorrelation function and its correlation coefficient for time steps of 30 days for the discharge of the observation, HBV model and SBM model of the Vesdre.

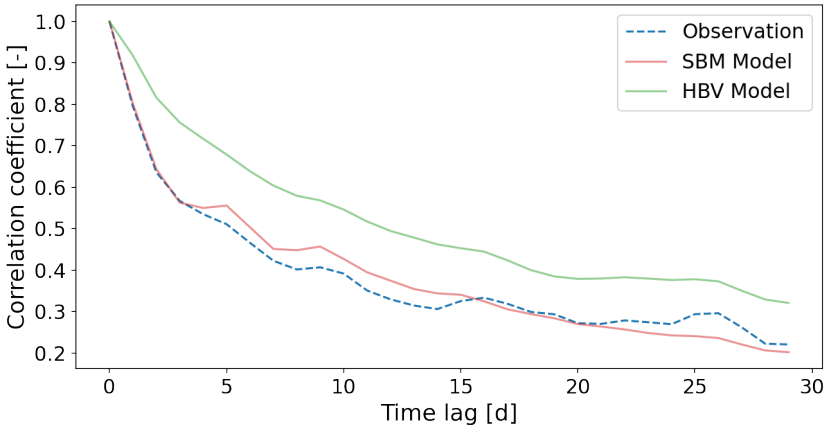


Figure H.14: Autocorrelation function and its correlation coefficient for time steps of 30 days for the discharge of the observation, HBV model and SBM model of the Geul.

The HBV model approximates the autocorrelation function better for the Ahr and Vesdre catchment. Both SBM models show a shorter memory of the river. For the Geul catchment, the SBM model has a much better approximation, while the HBV model expects a longer memory of the river. Considering the observed autocorrelation function, the Ahr has the longest memory, followed by the Vesdre. The Geul has the shortest memory. Although it was previously stated that the Geul has the most dampened and least variable response, one could expect that it would therefore have a long memory. However, the memory is strongly influenced by the drainage area (Hirpa et al., 2010; Mudelsee, 2007). This would explain why the ranking of decreasing memory is the same as of decreasing area.

Flow duration curves

The flow duration curves of the long-term observations have been given in fig. 2.10 in chapter 2 and are explained in more detail in appendix B.8. For the observations and models in this section, the flow duration curves are not created per year and then averaged. Here, the discharge data of all years has been sorted simultaneously, followed by the calculation of the probability of exceedance. For the Ahr, the daily discharge is used and for the Vesdre and Geul, the hourly discharge.

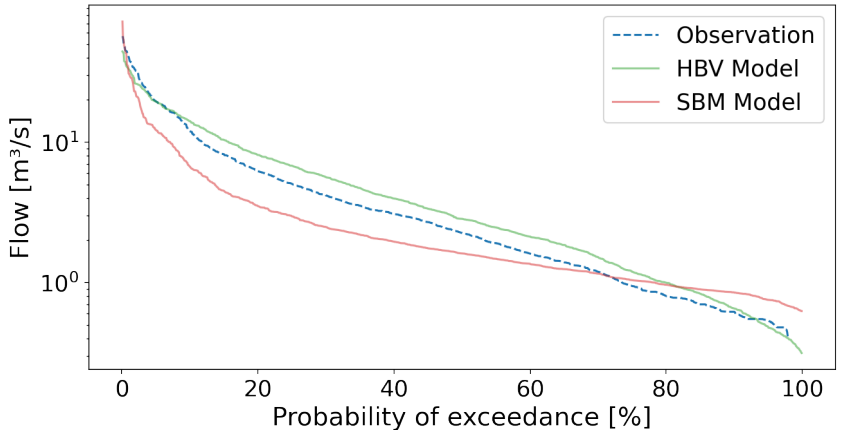


Figure H.15: Flow duration curves for the discharge of the observation, HBV model and SBM model of the Ahr.

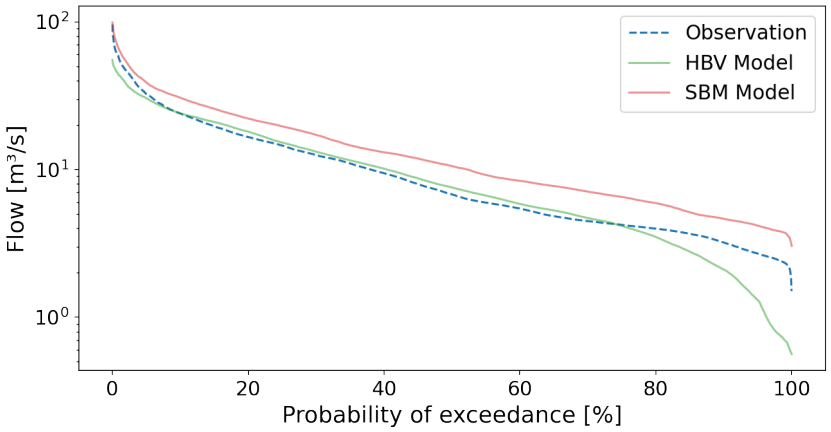


Figure H.16: Flow duration curves for the discharge of the observation, HBV model and SBM model of the Vesdre.

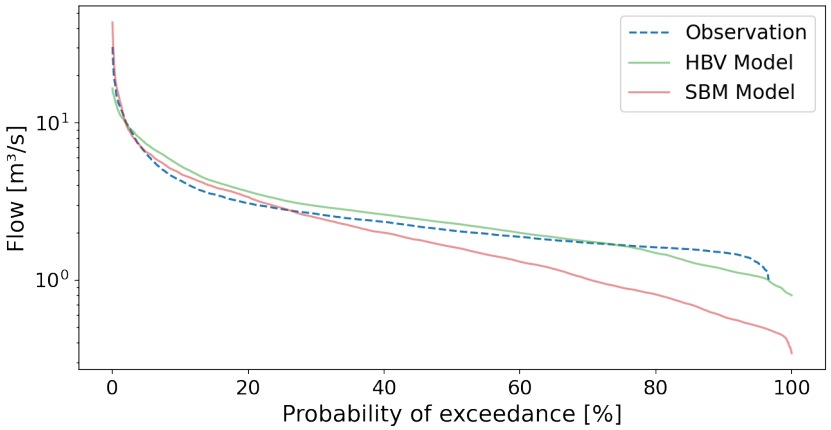


Figure H.17: Flow duration curves for the discharge of the observation, HBV model and SBM model of the Geul.

For the Ahr, the HBV model approximates the observed flow duration curves the best since the SBM model shows a flow duration curve of a more constant river. For the Vesdre, the HBV model has a better approximation for most of the curve, but for the lowest discharges, the SBM model has a more similar shape. For the Geul, the HBV model again has the best approximation. In contrast with the Ahr, the SBM model now shows a more variable river pattern than the observation. These figures confirm the better approximation of the HBV model as indicated in table 6.4. The SBM model is not completely off, but it has no good performance for this metric.

H.7 Flow delay in the Geul

The original SBM model for the Geul had a peak discharge with a moment of occurrence that preceded the estimated peak discharge by 1 day and 5 hours. Such a time difference is a significant error of the model. Additionally, the typical sensitive parameter of Imhoff et al. (2020) did not decrease this time shift and as such, complicated the analysis. Finally, the potential cause for this time difference could indicate a different flow behaviour during the flood. As this time difference is relevant on multiple aspects, this section analysis it in more depth.

One of the most likely reasons for the time difference would be the added buffering from flooded areas in the catchment. The SBM model does not account for flood added area and its hydraulic influences, which would explain why the SBM model has issues with the timing. Therefore, it is important to know if this delay is also present during other high flow events that did not cause floods. If this is the case, the delay is not a special characteristic of the flood event but rather a flaw in the original model calibration. Another interest for these events would be how the delay evolves upstream. Is it only present at the most downstream gauge? If the time difference slowly grows in downstream direction, flood wave attenuation could be indicate a pattern.

Three high flow events have been selected. These are the only high flow events with a discharge of more than 20 $[m^3/s]$ at the downstream gauge. The hydrographs of four gauges along the main river channel of the Geul are considered for this analysis (from upstream to downstream): L6660 at Sippenaeken, 10.Q.30 at Hommerich, 10.Q.63 at Schin op Geul and 10.Q.36 at Meerssen.

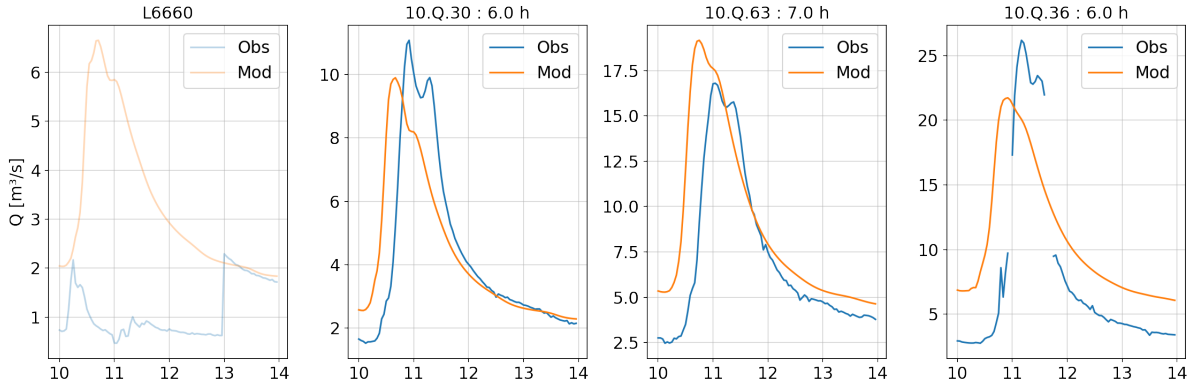


Figure H.18: Hydrographs from the observation (Obs) and original SBM model (Mod) for the high flow event of February 2019. The time differences for the peak discharges are given for each gauge.

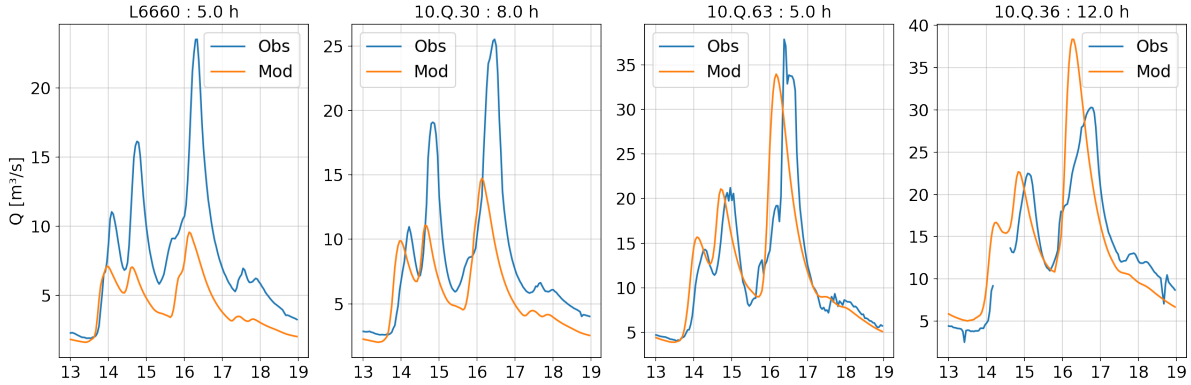


Figure H.19: Hydrographs from the observation (Obs) and original SBM model (Mod) for the high flow event of March 2019. The time differences for the peak discharges are given for each gauge.

Fig. H.18, fig. H.19 and fig. H.20 show that the time differences are present for all high flow events, although these differences are still smaller than for the 2021 flood event. The buffering of the flooded areas are thus not the dominant cause of this time difference.

The storage volume of the Geul due to artificial buffer in the Dutch region is 781.000 $[m^3]$. However, with a percentage of 0.3% this is only a small part of the average annual precipitation volume. A significant influence of the buffers would also show that the delay would increase from the upstream gauges to the downstream gauges since there are more buffers in the downstream regions. The figures above show that there is no clear increasing pattern in time difference for more downstream gauges.

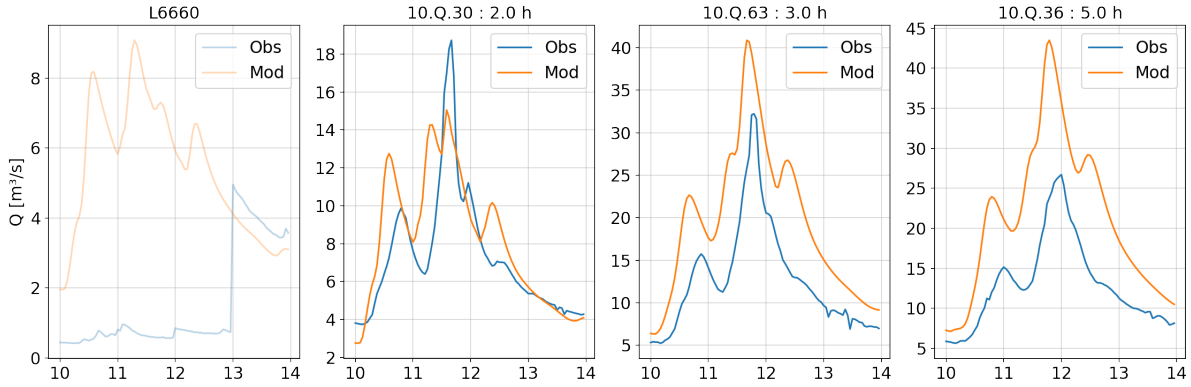


Figure H.20: Hydrographs from the observation (Obs) and original SBM model (Mod) for the high flow event of February 2020. The time differences for the peak discharges are given for each gauge.

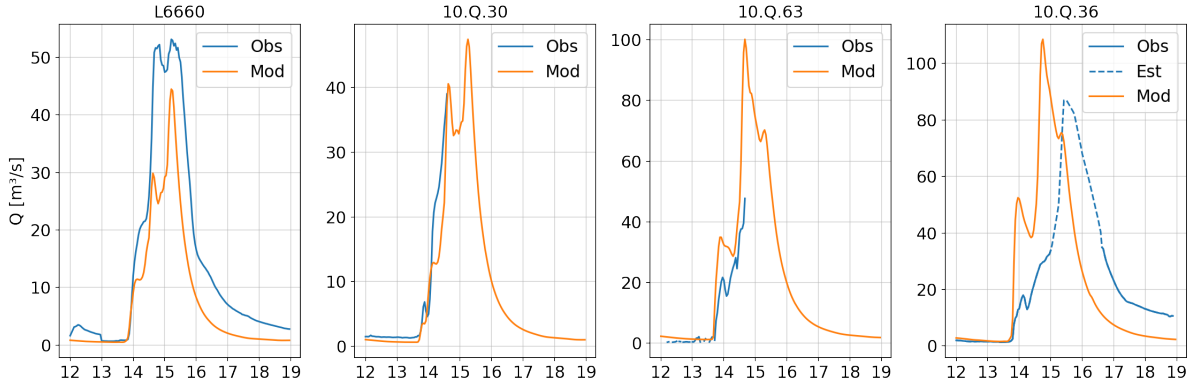


Figure H.21: Hydrographs from the observation (Obs), estimation (Est) and original SBM model (Mod) for the high flow event of July 2021.

This time difference for all gauges for the other high flow events is interesting because the time difference is not clear for upstream gauges during the 2021 flood event, as can be seen in fig. H.21. The lack of peak discharges of upstream gauges makes it difficult to draw strong conclusions for the spatial variance in time difference for the 2021 flood event.

Despite the lack of a clear cause of the time difference, it must still be changed for the re-calibration. Two parameters that provided useful for this goal are the roughness coefficients, n and n_river, respectively for land and river cells. By increasing their value, the moment of occurrence of the peak discharge could be delayed.

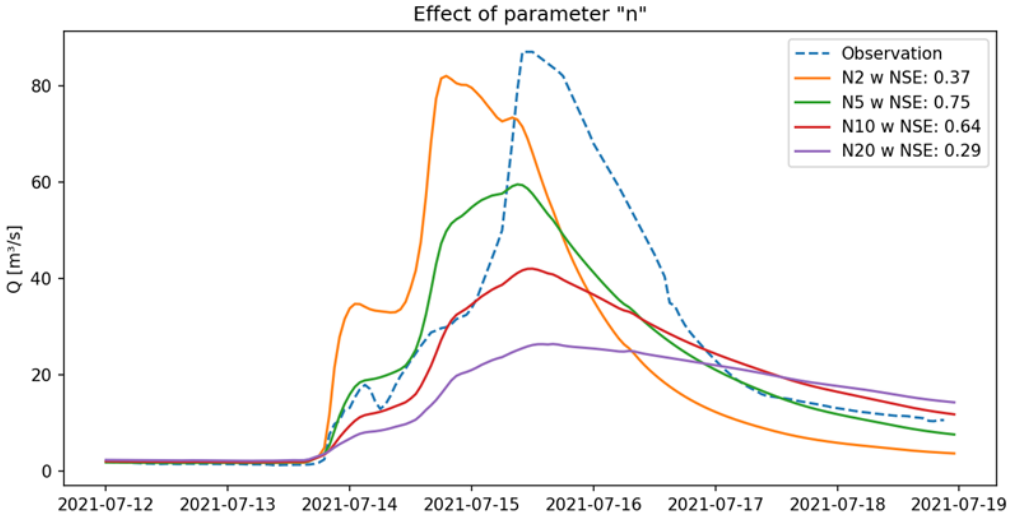


Figure H.22: Hydrographs for the 2021 flood event by calibrating the original SBM model with varying value of the n parameter.

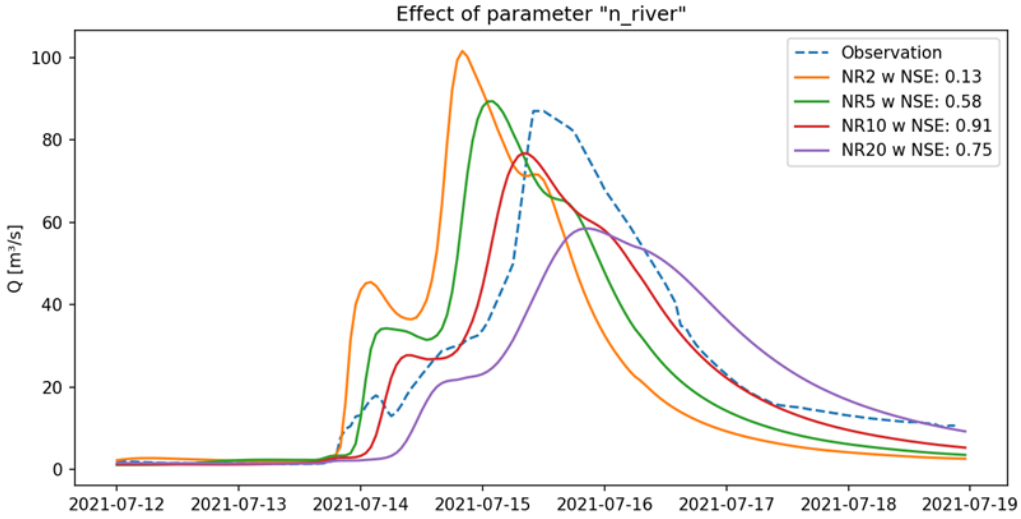


Figure H.23: Hydrographs for the 2021 flood event by calibrating the original SBM model with varying value of the n_river parameter.

Fig. H.22 and fig. H.23 show that both parameters are effective in decreasing the time difference. The parameter cause a stronger deformation of the hydrograph shape than the parameter n_river. Since the hydrograph shape is an important aspect of a good calibration, the parameter n_river has been selected for the re-calibration.

Appendix I

Model states and fluxes

This appendix contains figures of the model states of the HBV and SBM models. Both for the pre-event period and event period. For the latter, the results are given for the original and re-calibrated model. Finally, the ratios of the notable flow mechanisms are compared for the pre-event and event period.

I.1 HBV states

For the states of the HBV model the most relevant states and fluxes are shown. These include the modelled discharge, the four storage units, the evaporation and the two runoff fluxes (from the fast and slow responding groundwater zone). For each catchment, the changes for the re-calibrated model are described and how these can be linked to the parameter changes of appendix H.3 through the processes of the HBV model described in appendix G.1.

[some general char.]The slow responding zone and its flux are characterised by the slowly varying values and thus smooth curve, while the fast responding zone and flux are much more variable. The interception store and the evaporation show a very erratic pattern. The evaporation and unsaturated zone have a seasonal pattern. This is expected since the winter is characterised by lower temperatures and thus less evaporation, as well as typically higher groundwater levels.

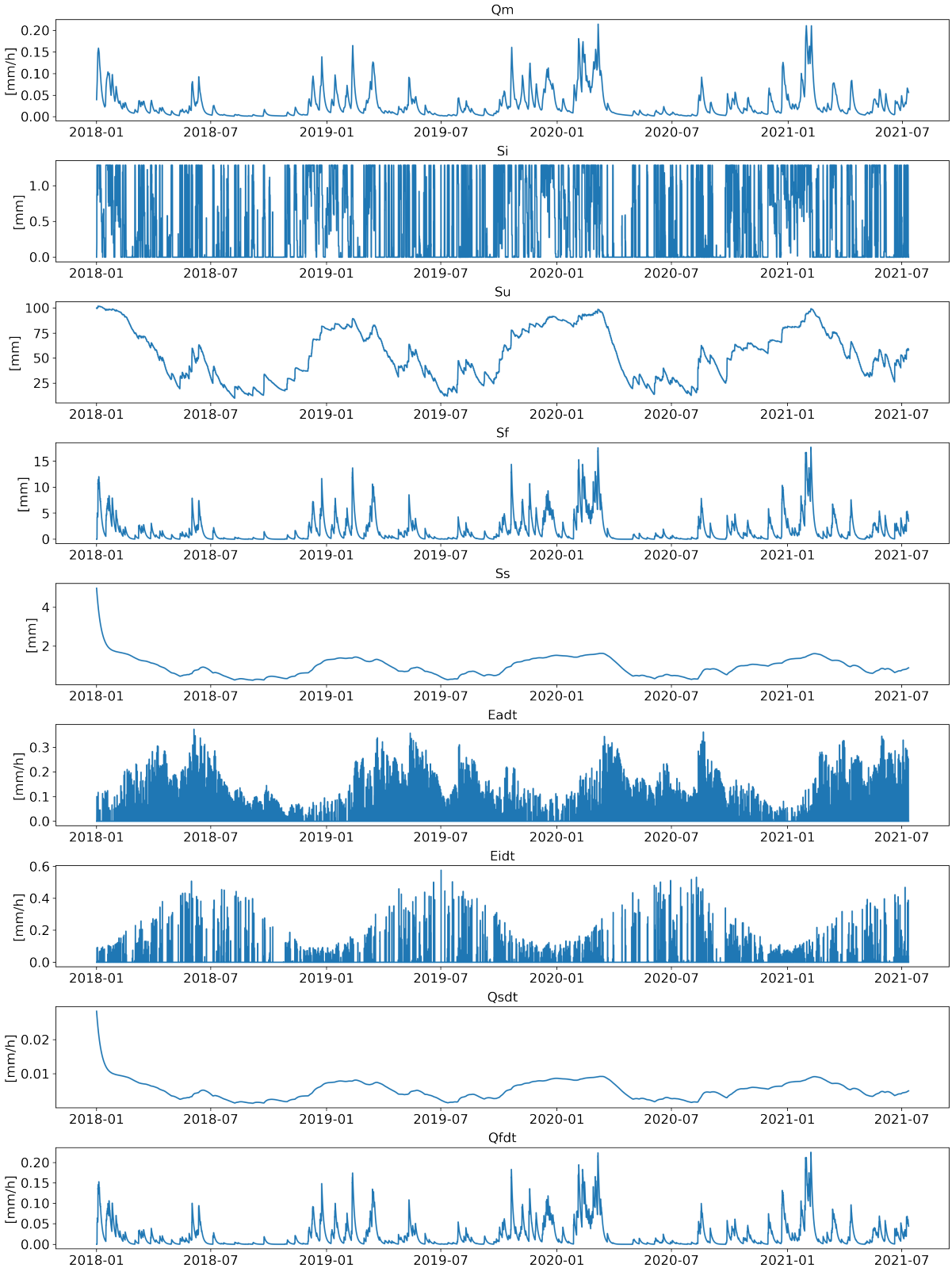


Figure I.1: States of the original HBV model before the 2021 flood event in the Ahr.

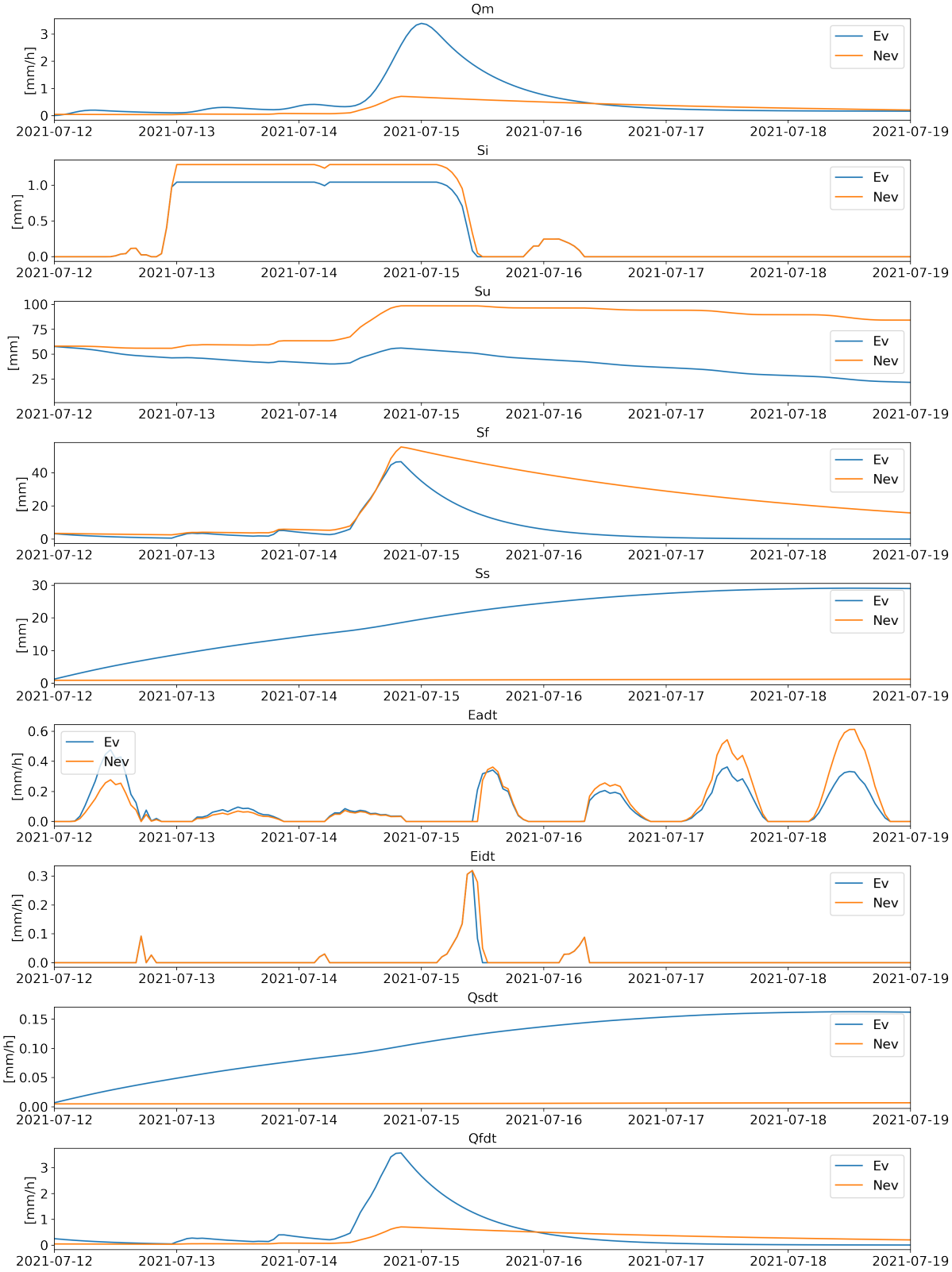


Figure I.2: States of the original (Nev) and re-calibrated (Ev) HBV model during the 2021 flood event in the Ahr.

The internal states for the Ahr model during the event, fig. I.2 show that the initial simulation of the event hydrograph with the original model resulted in a hydrograph that was very low and with a very slow recession, see the "Qm" (or modelled discharge) graph. The water in the interception store, Si, has decreased slightly. This is due to the similar reduction of the interception capacity, Imax. A larger difference can be seen for the amount of water in the unsaturated zone, Su, which has decreased strongly for the event. This is caused by the lower maximum capacity of that store, Sumax. The filling degree is still 0.7 at its peak (not shown in this figure), but it decreases faster after its peak for the re-calibration. The amount of water in the fast responding store has decreased slightly at its peak and has a faster recession. The decrease of the peak discharge could have been caused by a combination of multiple parameter changes such as the lower interception capacity or higher percolation. The fast recession is likely caused by the higher percolation which leads to a rapid decline of the filling degree and thus to a rapid decline of water partitioning to the fast responding store. This higher percolation is also the reason for the filling of the slow water store, which was empty for the original calibration. The evapotranspiration has decreased slightly, but this is a minor effect.

The runoff fluxes from the slow and fast responding store, Qsdt and Qfdt, are the most important for this analysis. The flux from the slow responding store is non-zero for the re-calibrated model. The flux of the fast responding store is still considerably higher and will have a much higher contribution to the total runoff. This flux has also increased considerably for the re-calibration. The slight decrease of water in the fast responding store is compensated by a much higher conductivity, Kf, which increased by a factor of 7. The fast recession was already visible from the water in the store. The higher modelled flow is largely created by the higher Kf parameter and the steeper recession curve is created by the higher Pmax parameter.

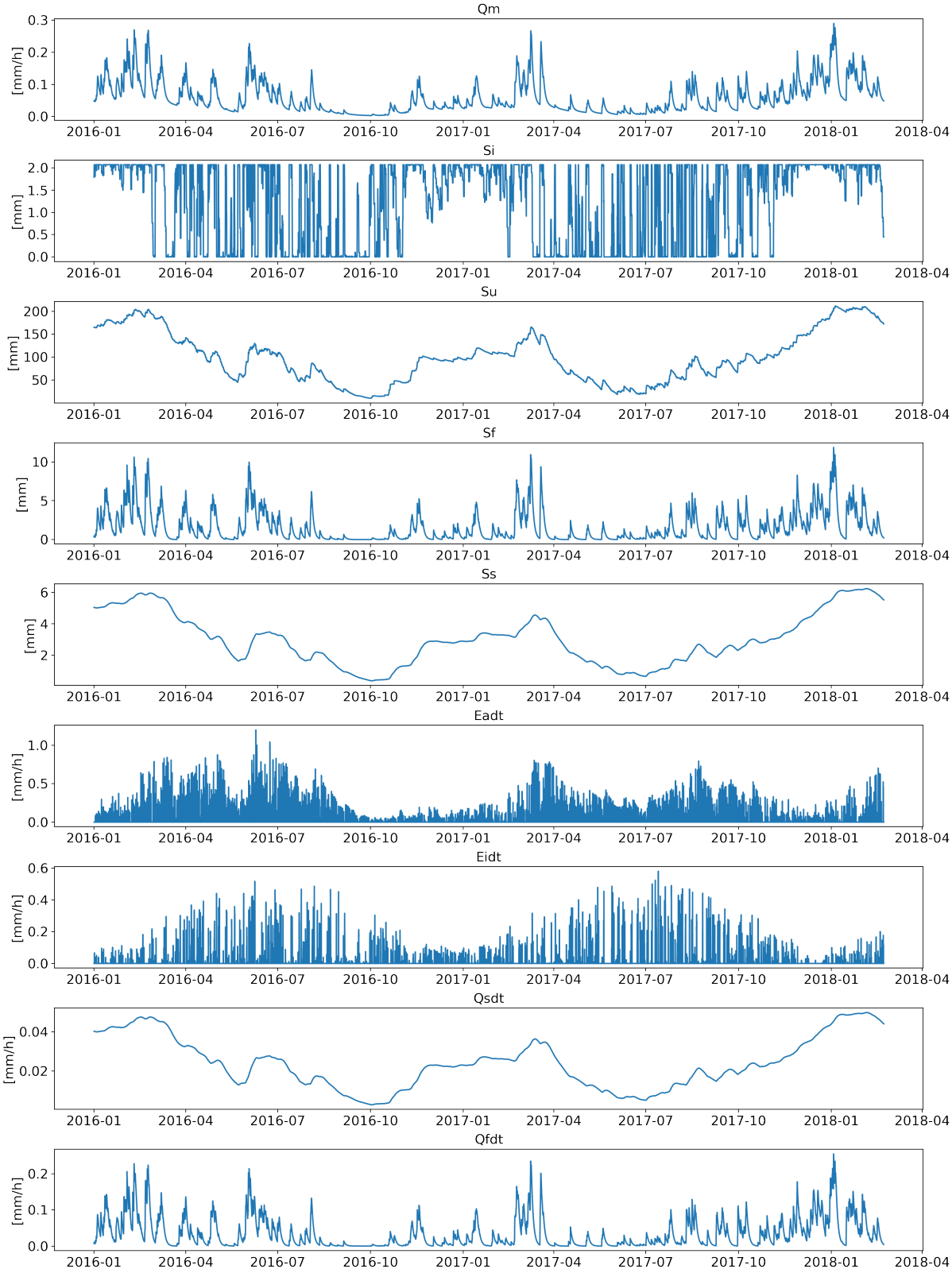


Figure I.3: States of the original HBV model before the 2021 flood event in the Vesdre.

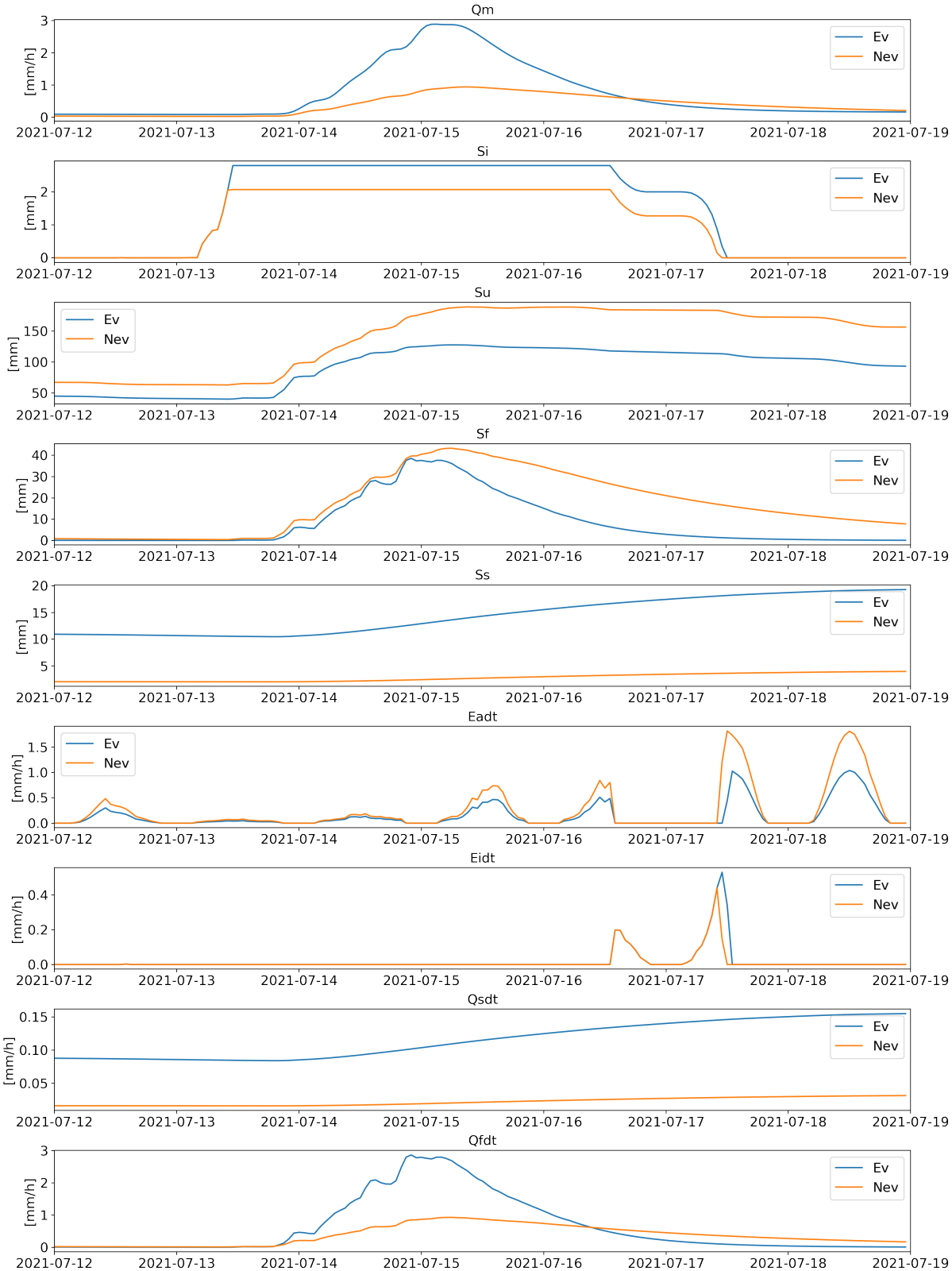


Figure I.4: States of the original (Nev) and re-calibrated (Ev) HBV model during the 2021 flood event in the Vesdre.

The internal states for the event in the Vesdre of fig. I.9 show that the re-calibration leads to a higher peak discharge and a stronger recession. The water in the interception zone has increased due to the increased interception capacity. The amount of water in the unsaturated zone has strongly decreased due to the decreased Sumax. However, the maximum filling degree has increased from 0.4 to 0.9, which is a large increase that has significant consequences. The amount of water in the fast responding store has decreased slightly and has a stronger recession. This effect is similar to what could already be noticed for the Ahr. Another similar change is the amount of water in the slow responding store which is not empty anymore for the re-calibration due to the increased percolation. The evapotranspiration has also decreased.

The changes for the two runoff fluxes are very similar to the Ahr. The runoff from the slow responding store is not zero anymore, but it is still much lower than the contribution of the fast responding store. The runoff of the latter has increased considerably (with a factor of 3.1). Additionally, there is a stronger recession of the runoff. The higher generated discharge is largely due to the increased horizontal conductivity of the fast responding store and the stronger recession is due to the increased percolation.

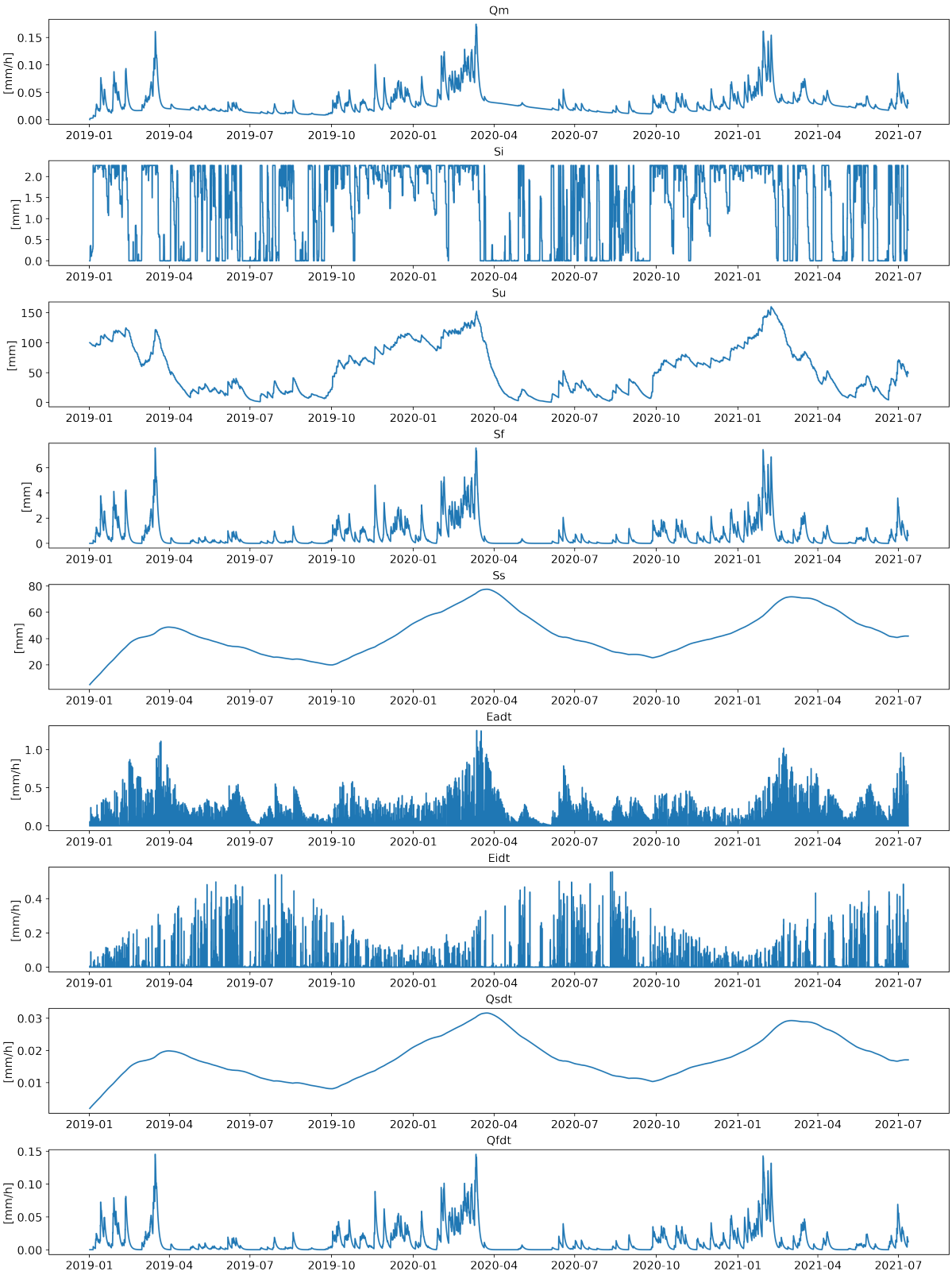


Figure I.5: States of the original HBV model before the 2021 flood event in the Geul.

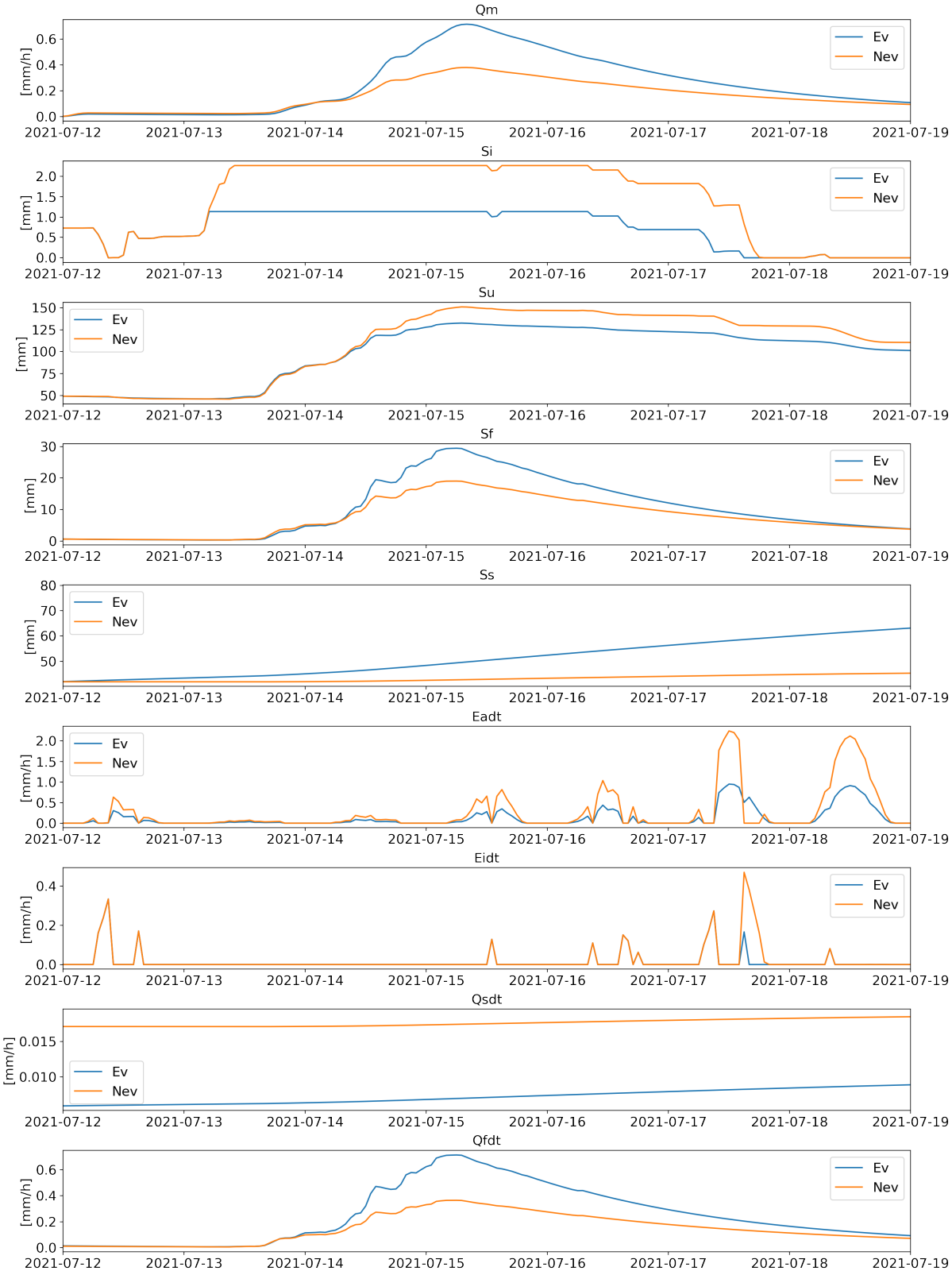


Figure I.6: States of the original (Nev) and re-calibrated (Ev) HBV model during the 2021 flood event in the Geul.

An important difference between the modelled hydrograph of the Geul and of the other catchments is that the underestimation of the peak discharge is smaller for the Geul. The changes of the internal states for the re-calibration are therefore likely to be smaller. Fig. I.6 shows that the re-calibration resulted in a higher peak discharge. The interception has halved which is in line with the decrease of the interception capacity. The water in the unsaturated zone remains nearly equal, but with the strong decrease of Sumax, the filling degree has increased considerably (from a maximum of 0.3 to 0.8). This higher filling degree has likely resulted in the higher amount of water in the fast responding store. The amount of water in the slow responding store has increased, but mostly after the peak discharge occurred. The evapotranspiration has decreased.

The increase for the amount of water in the slow responding store is less visible in its runoff due to the decreased hydraulic conductivity. The higher conductivity for the fast responding store leads to a considerable higher runoff, which makes the runoff of the slow response almost negligible. The increase of the modelled discharge seems caused by the decrease in Sumax (and thus the considerably higher filling degree of the unsaturated zone) and the increase of the hydraulic conductivity of the fast responding store.

I.2 SBM States

The most important and significant model states and fluxes of the SBM model are given in this section. There are many more states and fluxes, but the overarching ones are chosen here.

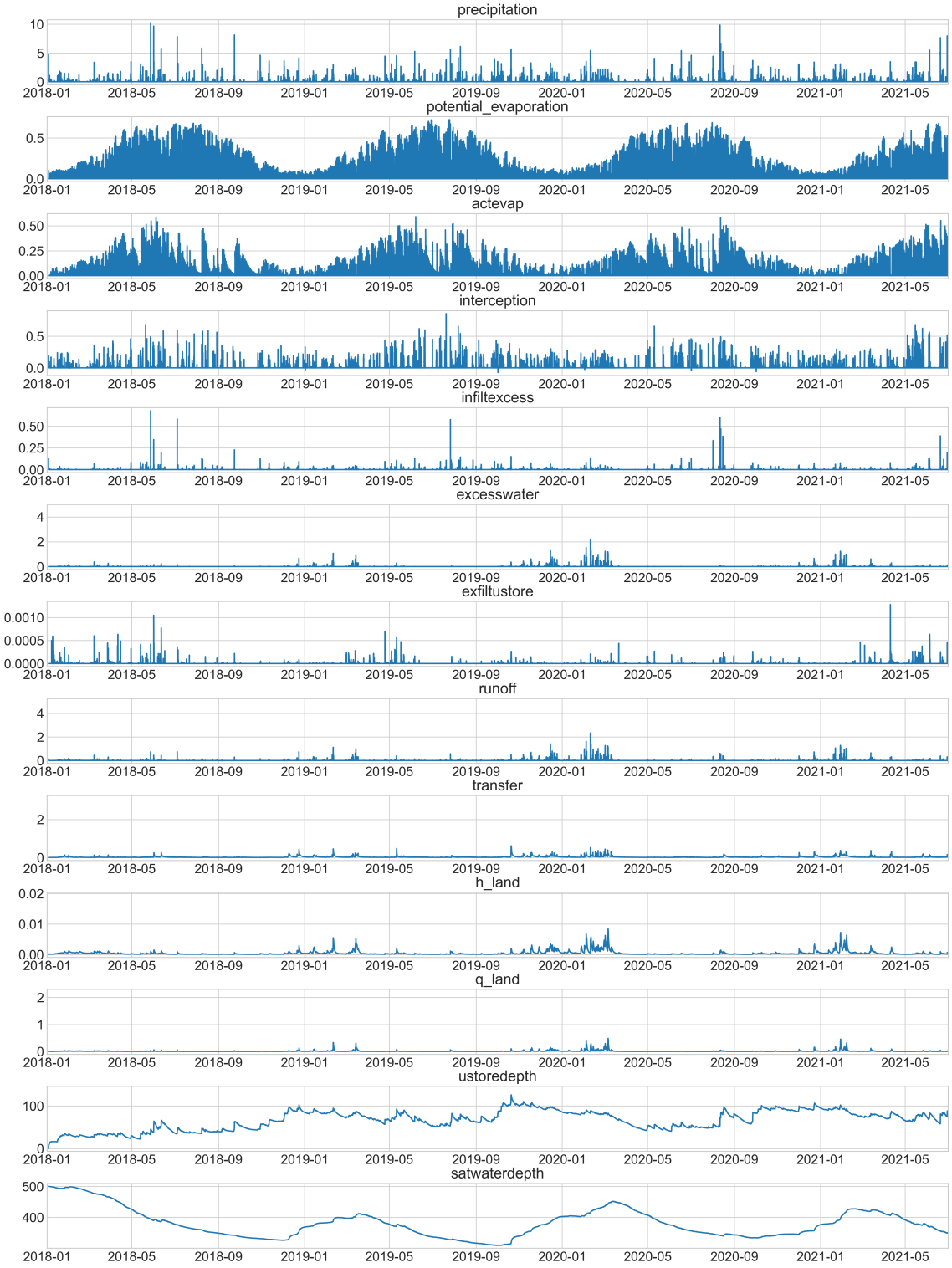


Figure I.7: States of the original SBM model during the pre-event period for the Ahr.

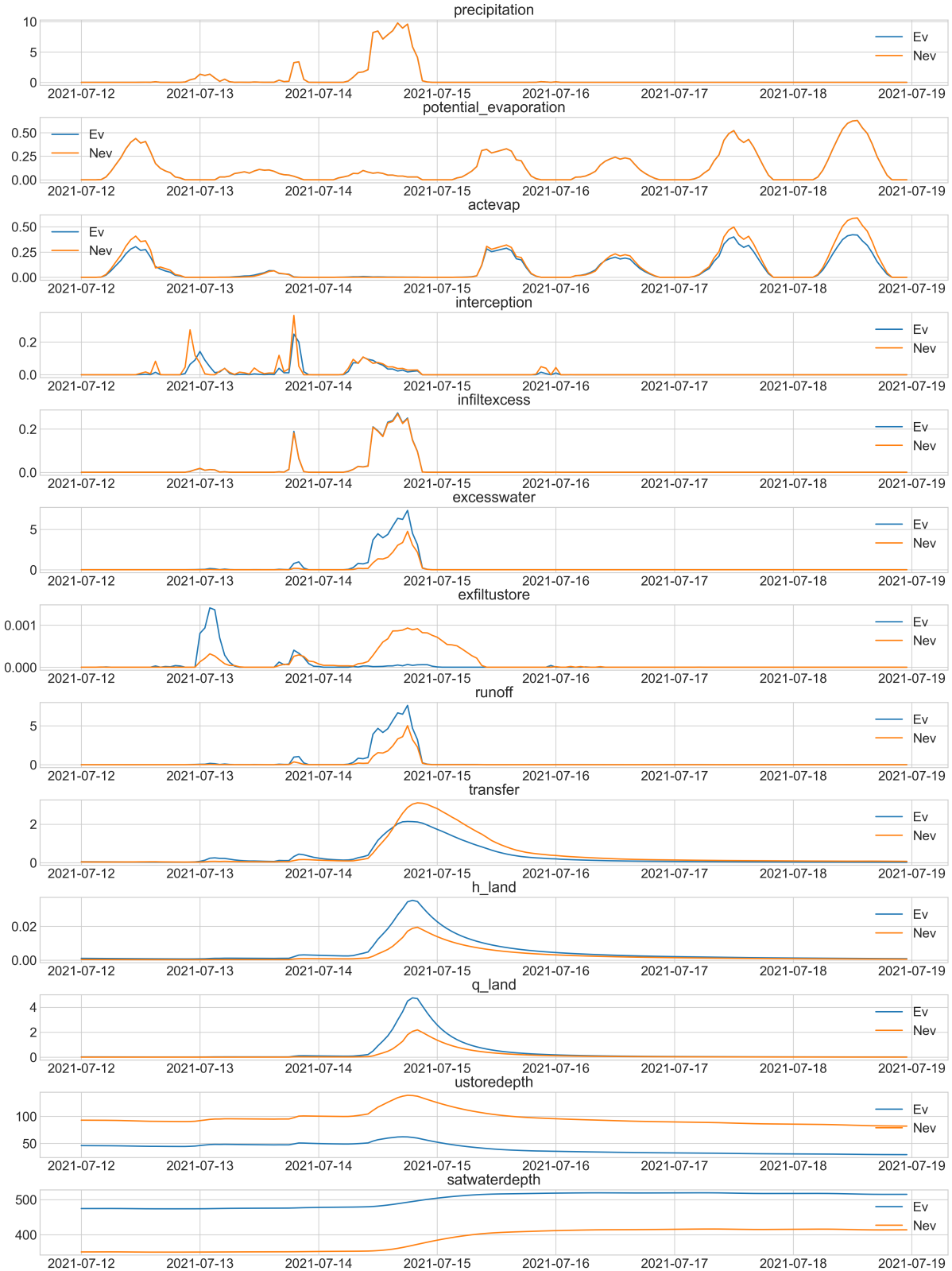


Figure I.8: States of the original SBM model during the 2021 flood event for the Ahr.

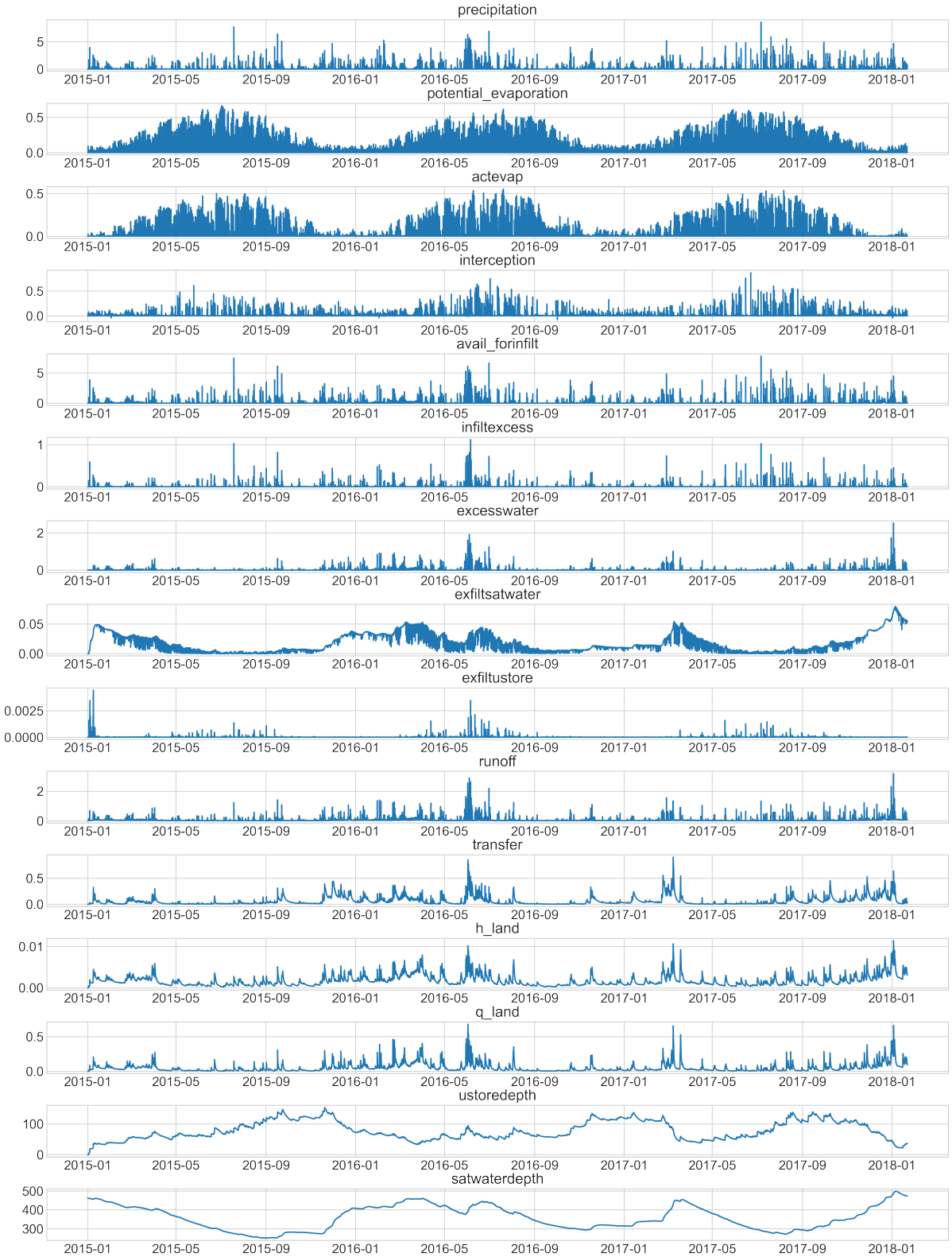


Figure I.9: States of the original SBM model during the pre-event period for the Vesdre.

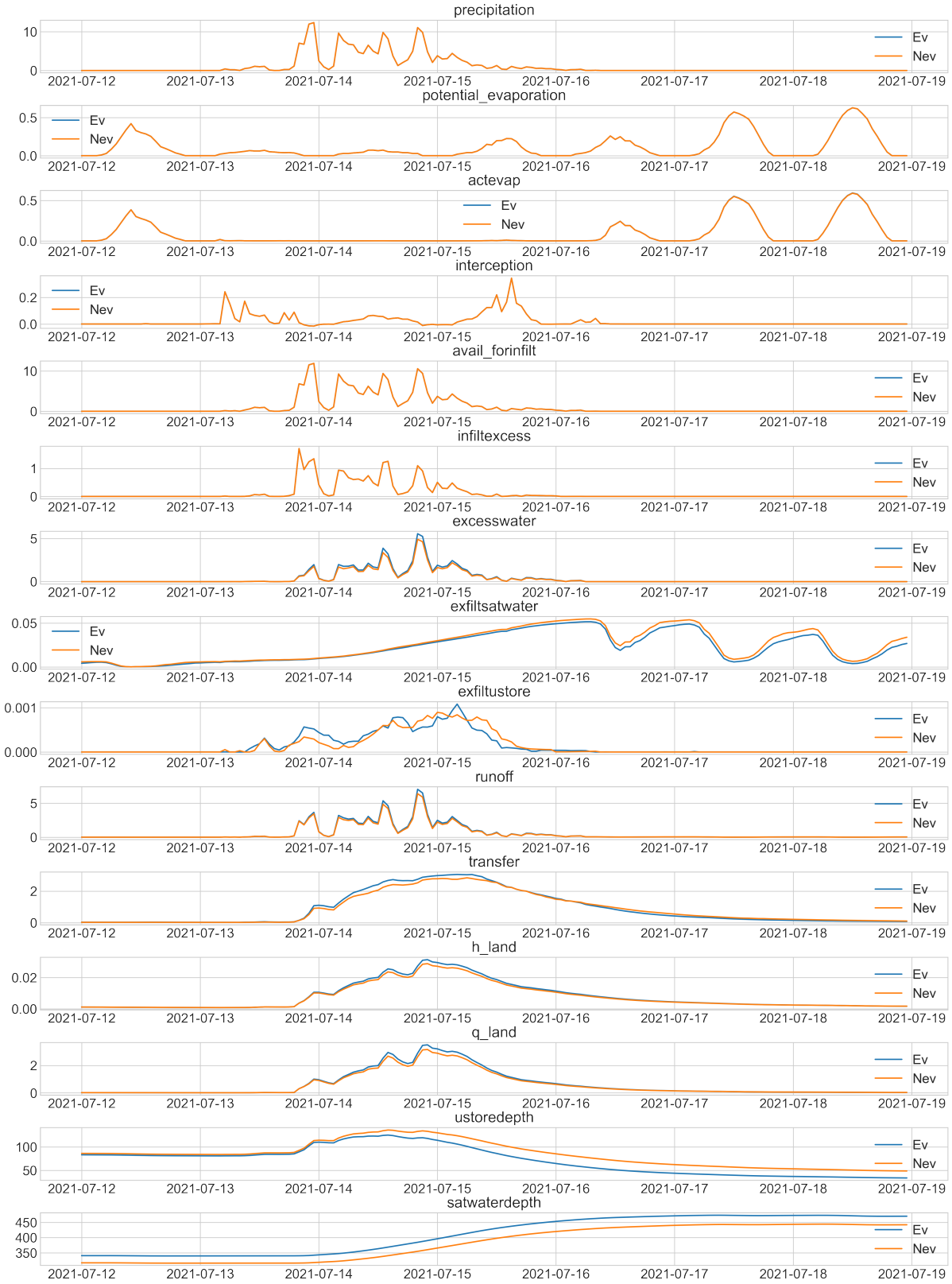


Figure I.10: States of the original SBM model during the 2021 flood event for the Vesdre.

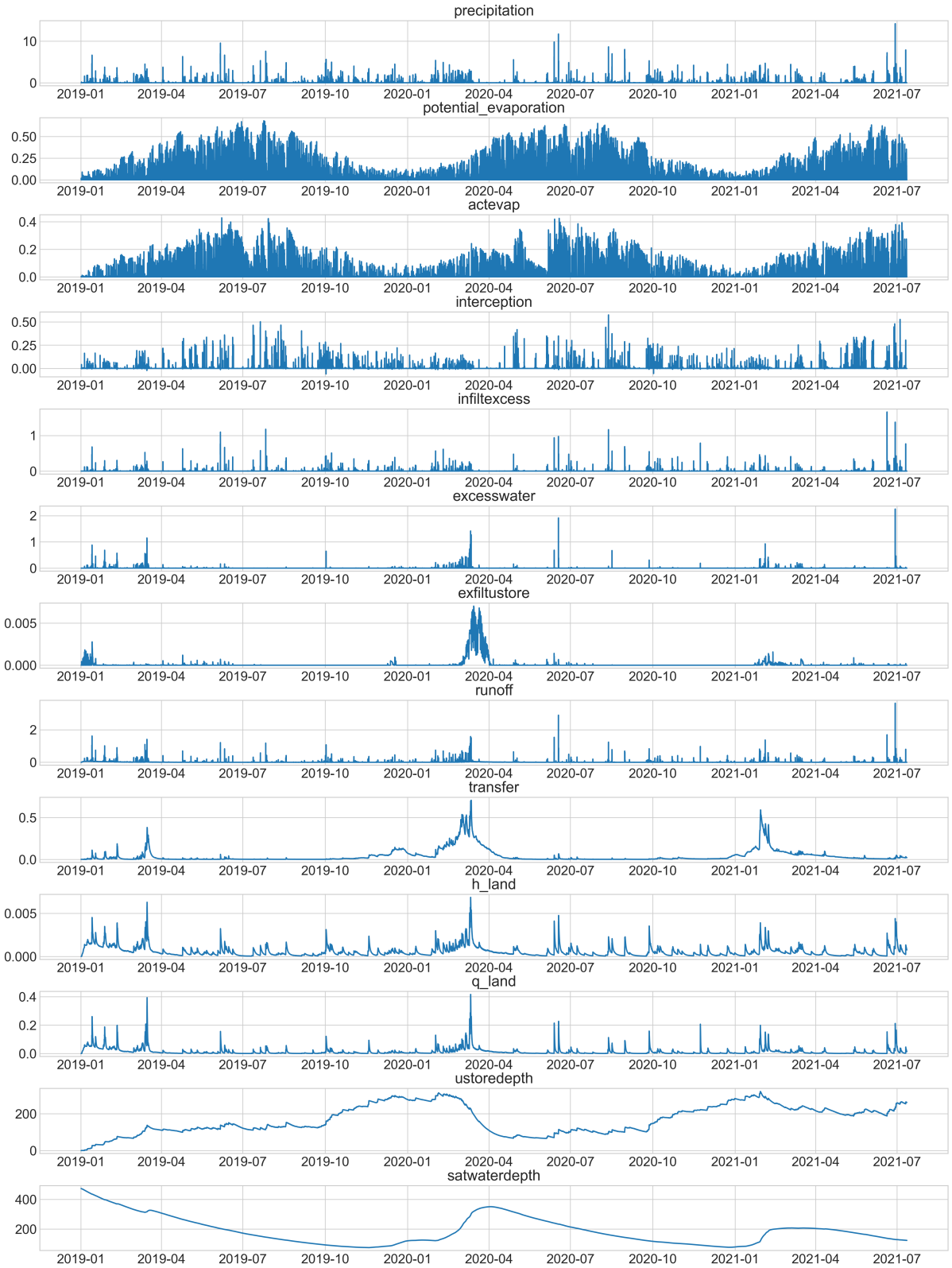


Figure I.11: States of the original SBM model during the pre-event period for the Geul.

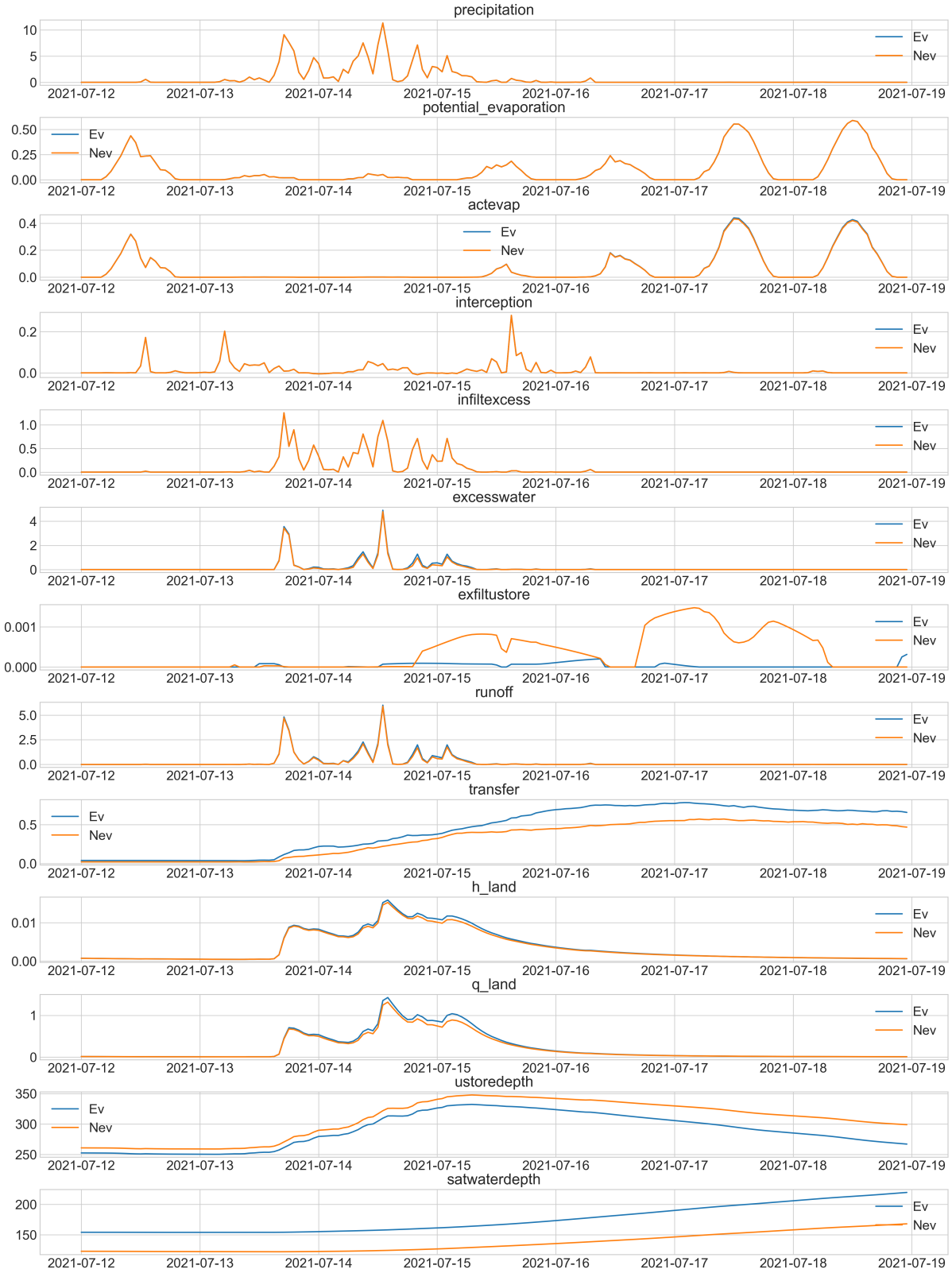


Figure I.12: States of the original SBM model during the 2021 flood event for the Geul.

I.3 Flow mechanism ratios

This section includes the maximum ratios of the fluxes reflecting the fast and slow runoff processes to compare how they vary for the pre-event period and the 2021 flood event.

For the HBV model, these fast and slow fluxes are easily recognizable since there are only two fluxes contributing to the runoff. The ratio to detect fast flow mechanisms is based on the flux from the fast responding store, indicated as Qfdt, compared to the total runoff fluxes, which adds the flux from the slow responding store, Qsdt, to Qfdt.

Table I.1: Maximum ratio of the fast responding flux to the total flux (fast and slow responding flux) of the HBV model. The results are given for the pre-event period (original model) and the 2021 flood event (of the original model, "Uncal. Event", and the re-calibrated model, "Cal. Event"). The increase of the ratios for the 2021 flood event compared to the pre-event period is given in percentages.

	Ratio			**Increase from pre-event**	
	Pre-Event	**Cal. Event**	**Uncal. Event**	**Cal. Event**	**Uncal. Event**
Ahr	0.97	0.972	0.992	0.2%	2%
Vesdre	0.959	0.966	0.979	1%	2%
Geul	0.928	0.990	0.954	7%	3%

Table I.1 show that the increases of the ratio for the re-calibration are low. However, they could not have been much higher since the previous ratio was already close to its maximum of 1. It could be surprising that the ratios are even higher for the original model during the flood event for the Ahr and the Vesdre since the faster response was a strong characteristic of the re-calibration. For the Ahr and Vesdre, the percolation parameter had been increased to allow a faster recession of the hydrograph similar to the observation. This also increased the flow from the slow responding store. Due to this higher (although small) contribution of the slow responding runoff to the total runoff, the ratio of the fast responding flux to the total runoff is lower for the re-calibrated models than for the original models.

Finding parameters to compare fast flow mechanisms ratios is more difficult for the SBM model due to the many parameters, as well as the vertical and lateral processes. Since the two runoff fluxes of the HBV model, Qfdt and Qsdt, can be considered as lateral fluxes (they are the only contributions to the "horizontal" discharge), the ratio for the SBM model is also based on two horizontal fluxes: q_land represents the lateral overland flow and ssf represents the lateral subsurface flow. Although it is not explicitly mentioned in the documentation or model source, the units of both parameters are assumed to be equal, $[m^2/h]$. However, as this is not entirely certain, their absolute ratio will not be analysed further, only its difference for the pre-event and event period.

Table I.2: Maximum ratios of the surface runoff and the sum of the surface runoff and the subsurface flow from the SBM model. The results are given for the pre-event period (original model) and the 2021 flood event (of the original model, "Uncal. Event", and the re-calibrated model, "Cal. Event"). The increase of the ratios for the 2021 flood event compared to the pre-event period is given as a factor.

	Ratio			**Increase from pre-event**	
	Pre-event	**Cal. Event**	**Uncal. Event**	**Cal. Event**	**Uncal. Event**
Ahr	0.0167	0.0928	0.0877	5.556886	5.251497
Vesdre	0.0058	0.0351	0.0271	6.051724	4.672414
Geul	0.0126	0.1611	0.052	12.78571	4.126984

Table I.2 show a strong increase of the ratio, indicating that the overland flow has strongly increased, even relative to the subsurface flow.

List of Figures

List of Tables