An aerial photograph of a wide river meandering through a dense, green forest. The river is the central focus, with its banks covered in thick vegetation. The lighting suggests a bright day, with some reflections on the water's surface.

Identifying the relative contribution of discharge components to improve discharge predictions

Rinske de Ronde

June 2024

MSc thesis

Hydrology and Environmental Hydraulics Group

Wageningen University

Source cover photo: Uliasz, M. (2019). Retrieved October 2022.

Name: RH (Rinske) de Ronde
Student number: 1028474

Master thesis Earth and Environment
Course code: HWM80436
Wageningen University and Research (WUR)

First supervisor: dr.ir. AJ (Ryan) Teuling (HWM)
Second Supervisor: LA (Lieke) Melsen (HWM)
Examiner: dr.ir. CC (Claudia) Brauer (HWM)
June 2024

Abstract

Modelling discharge is a trade-off between using simple concepts and increasing model complexity to simulate additional processes. Currently, there is a lack of understanding of the processes taking place inside a catchment that lead to the formation of discharge after a precipitation event. This research aims to create a model that takes into account the three flow processes that are responsible for the total discharge (deep and shallow groundwater based discharge and overland flow) to simulate discharge based on evapotranspiration and precipitation data, while keeping the model simple with a limited number of parameters. Starting with a baseflow separation, deep groundwater based discharge is simulated assuming a linear relationship between water in storage and discharge. Consequently, shallow groundwater based flow is modeled following the approach of Kirchner (2009), in which discharge is assumed to solely depend on water storage. Lastly, overland flow is modeled as a fraction of precipitation, where the fraction depends on catchment wetness, represented by discharge. Applied to thirteen catchments in Europe comprising a variety of response types, the model has proven to be successful in determining the relative contribution of the three discharge components based on the hydrograph. Additionally, the model is able to use knowledge on the component distribution to provide satisfactory discharge predictions for most catchments. As a predictive tool, the model did not show to outperform other rainfall-runoff models. However, the model is a valuable addition to the already-existing conceptual rainfall-runoff models, as it provides the opportunity to determine the distribution of discharge components in a relatively simple manner.

Contents

1	Introduction	1
1.1	Problem Description	1
1.2	Research Objective & Questions	2
1.3	Thesis Outline	3
2	Field Site & Data	5
3	Methods	9
3.1	Data Preparation & Analysis	9
3.2	Model Development	10
3.3	Model Quality Assessment	14
4	Results & Interpretation	17
4.1	Data Preparation & Analysis	17
4.2	Model Development	18
4.3	Model Quality Assessment	32
5	Discussion	37
5.1	Model Structure	37
5.2	Broader Model Discussion	38
5.3	Model Prospect	39
6	Conclusion	41
	Acknowledgements	43
	Bibliography	45
A	List of Variables	49
B	Additional Figures	51

1 | Introduction

1.1 Problem Description

Rainfall-runoff models are used for water resource management and have flood risk prediction as their most important application (Seibert and Bergström, 2022; Todini, 2011). The urge for high quality discharge prediction rises as extreme rainfall events will occur more frequent in the future, leading to an increase in river floodings (Arnell and Gosling, 2016). Therefore, scientists focus on improving models to predict river discharge as accurately as possible. Over the years, numerous hydrological models have been developed to predict discharge, evapotranspiration and water storage based on precipitation data (Rosbjerg and Rodda, 2019).

Peel and McMahon (2020) discussed the historical development of rainfall-runoff models, reporting 279 models with different model structures. These models are a trade-off between using simple concepts and increasing model complexity to simulate additional processes. Presently, there is no general classification of catchments that can be used as a guide to model discharge with a predefined model concept (Wagener et al., 2007). Choosing the most suitable model and parameters is therefore done by the modeller (Kalantari et al., 2015) and outputs have, accordingly, shown to depend strongly on personal judgement (Holländer et al., 2009).

There are several approaches to modelling discharge, for example in a conceptual way (Todini, 2011; Peel and McMahon, 2020). An example of a conceptual model is the HBV (Hydrologiska Byråns Vattenbalansavdelning) model, which dates back to the 1970s (Seibert and Bergström, 2022). The model consists of multiple routines; one representing snowfall and melt (based on the degree-day method) and one representing water storage through a soil box (Uhlenbrook et al., 1999). Additionally, the model represents groundwater flow based on flow equations (Driessen et al., 2010). The model was originally developed for Scandinavian catchments (Seibert, 1996), but over the years it has been used to model discharge all over the world (Nonki et al., 2021; Li et al., 2013; Driessen et al., 2010).

Model complexity is known to go along with uncertainty and therefore, less complex models might lead to better performance (Snowling and Kramer, 2001). Research has been done into increasing model complexity and the corresponding uncertainty, to make a better

consideration whether increasing complexity matches the purpose of the model (Puy et al., 2022). Besides, some scientists follow the theory of Occam's razor. The principle, dating back to the fourteenth century, states that the simplest explanation of a phenomenon is always the best (Braithwaite, 2007). This highlights that models should not be more complex than necessary, showing the importance of simple conceptual models.

An example of such a simple conceptual mode is the Simple Dynamical Systems (SDS) approach as proposed by Kirchner (2009). Despite the simplicity of the model and the minimal number of parameters, the approach has shown to perform well. In this model the assumption is made that discharge solely depends on subsurface water storage in a catchment. Without the need for base-flow separation (Eckhardt, 2008), a storage-discharge relationship can be identified which describes how changes in storage affect discharge. This relationship, also known as the discharge sensitivity function, can be inferred from streamflow recession analysis and allows for discharge modelling using only precipitation and evapotranspiration data. Kirchner (2009) demonstrated that the SDS approach yields good results for the Severn and Wye river at Plynlimon, Wales, with Nash-Sutcliffe efficiencies (NSE) up to 0.93 with parameters obtained from recession plots for individual years and NSE values up to 0.95 with model parameters calibrated to discharge time series.

Corresponding, others have shown that the SDS approach yields good results for different catchments. Teuling et al. (2010) tested whether the assumed hydraulic connectivity between saturated and unsaturated stores in a catchment is valid, by applying the SDS approach to the Swiss Rietholzbach catchment. The catchment has, similar to the Plynlimon catchment, a humid climate, but receives less precipitation and has more extreme droughts observed in the past. Teuling et al. (2010) found that the method works generally well, especially during wet periods, but failed to show good model performance during dry summers.

Additionally, Brauer et al. (2013) tested the method in the Dutch lowland Hupsel Brook catchment, which is less humid than both the Plynlimon and Rietholzbach catchment and has a considerably smaller runoff ratio. Results showed that the discharge of the Hupsel Brook catchment cannot always be modelled correctly follow-

ing Kirchner's approach, especially not during summer. Adamovic et al. (2015) found corresponding results for the Ardèche catchment in France, for which the SDS approach performed well during wet conditions, but failed to provide good discharge predictions during dry periods.

Often, rainfall-runoff models are lumped per catchment, as catchments are natural units that act as their own system in the global water balance (Wainwright and Mulligan, 2005). The hydrograph can reflect the result of catchment-wide hydrological processes, such as water routing. Catchment discharge is relatively easy to observe through, for example, stream velocity measurements (Speir et al., 2023; Gore and Banning, 2017). However, numerous processes that are not easy to observe, take place inside a catchment before precipitation becomes discharge. Consequently, developing models that accurately represent these processes and perform well at predicting discharge remains a challenge. We pose that a better understanding of flow routes and their relative contribution to discharge is needed to further improve discharge predictions.

Although the SDS approach performs well at predicting discharge, best results were obtained during wet conditions (Adamovic et al., 2015; Brauer et al., 2013; Teuling et al., 2010). Model errors could result from the fact that total discharge is composed of a combination of flow routes (van der Velde et al., 2010). Catchments dominated by overland flow are not represented well by the SDS approach (Kirchner, 2009). Rusjan and Mikoš (2015) therefore proposed a combination of quickly responding discharge resulting from shallow groundwater flow, as modeled by the SDS approach, and overland flow (very quickly responding discharge) for the Padez stream catchment in Slovakia. The catchment is characterized by rainfall-runoff that bypasses subsurface storage, resulting from extreme rainfall events. Overland flow can be caused by several mechanisms; (i) rainfall intensity can exceed the infiltration capacity of the soil, known as Hortonian overland flow, and (ii) rainfall causes overland flow when the soil is saturated and thus no more water can infiltrate, known as Durnian overland flow (de Lima, 1989; Maier and van Meerveld, 2021).

A discharge component that neither Kirchner (2009) nor Rusjan and Mikoš (2015) consider, is deep groundwater flow. At locations where the water table is deep, precipitation can infiltrate towards the deep groundwater (recharge). This usually happens in areas situated higher in the landscape, whereas water reaches the surface again in lower areas (Batelaan et al., 2003).

Deep groundwater has a longer residence time than shallow groundwater, as the residence time is determined by the depth of the aquifer (Hare et al., 2021). Figure 1.1 shows a schematic overview of the three discussed pathways from precipitation to discharge.

Even though rainfall-runoff models aim to represent catchment wide hydrological processes and flow routes through discharge, they are often validated based on the discharge at the catchment outlet (Hattermann et al., 2004). As most simple models focus on one specific process that is applicable for a specific type of catchment (border cases), those models give a good output for given catchments. However, there are only few catchments representing only one process and many catchments for which the discharge is compiled of multiple discharge components. Therefore, there is a need for a simple conceptual rainfall-runoff model that takes into account knowledge on the relative distribution of discharge components to predict discharge from precipitation more accurately for various types of catchments.

1.2 Research Objective & Questions

The aim of this research is to improve discharge predictions by creating a model that uses information on the relative contribution of the three discharge generating mechanisms to model discharge based on precipitation and evapotranspiration data, while keeping the model simple with a limited number of parameters. The possibility of using existing discharge datasets on potential groundwater recharge to predict the contribution of deep groundwater based discharge in a certain catchment will be explored, in order to reduce future calibration efforts. The quality of the discharge predictions will be evaluated by comparing the output of the new model to output of more conventional methods (SDS approach and HBV) and to observations. To reach this aim, the following set of sub-research questions will be addressed in this thesis:

- *How can relative contributions of discharge components be determined from the hydrograph?*
- *Can the relative contribution of discharge components be predicted based on already existing discharge datasets?*
- *How do discharge predictions based on three discharge components compare to conventional model predictions?*

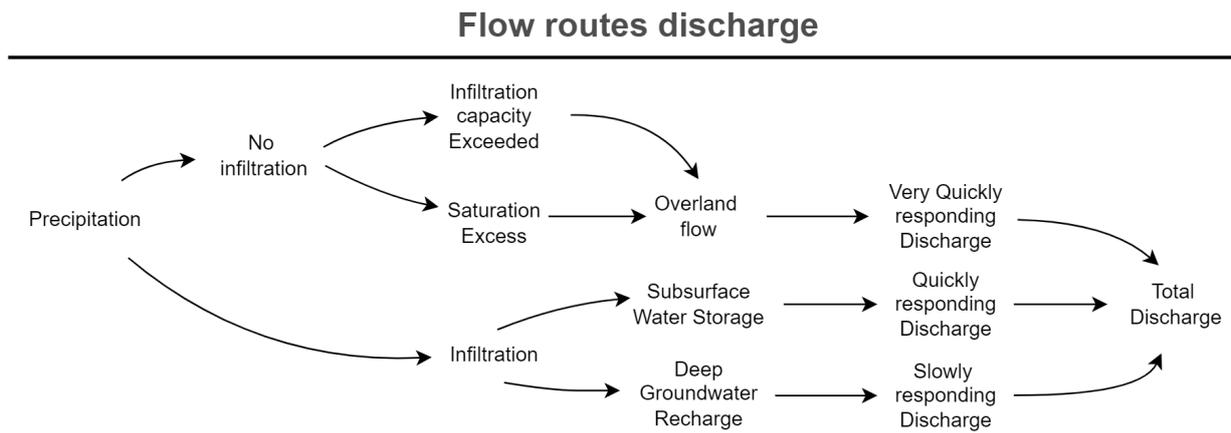


Figure 1.1: Schematic overview of the three flow paths from precipitation to discharge.

1.3 Thesis Outline

This thesis continues with an explanation of the field site and data in Chapter 2. The chapter mainly focuses on the catchment characteristics and the expected dominant discharge processes. In Chapter 3 the methods are explained that are used to create and validate the model, followed by an overview of the results in Chapter 4. Then follow the discussion and conclusion in Chapter 5 and 6, respectively.

2 | Field Site & Data

Discharge data of several catchments with catchment sizes in the same order of magnitude are used. The catchments are situated in Europe (Figure 2.1) and range between Alpine catchments with shallow sub-soils and catchments situated in more lowland areas with deep groundwater tables. Therefore, the flow routes vary and the catchments respond differently to precipitation. The catchments are chosen as such that they comprise a variety of response types, including extremes, therefore being representative for many catchments. Hourly discharge, precipitation and potential evapotranspiration data are collected for all catchments. For catchments experiencing snowfall regularly in winter (mountainous catchments), temperature data are also required to account for snow melt. The years for which the data are available and the data providers are mentioned in Table 2.1, together with an overview of catchment characteristics and the expected dominant discharge process(es).

Noor catchment The first catchment that is investigated is the Noor catchment, situated at the border between the Netherlands and Belgium. The catchment's discharge is assumed to mainly consist of deep groundwater, as the resemblance between the hydrograph and the variation of the groundwater table over time is striking (van Lanen et al., 1995). Additionally, discharge shows little response to precipitation events. This, in combination with the high infiltration capacity of the topsoil, leads to the expectation that little overland flow occurs (Dijkema and van Lanen, 1998).

Hupsel Brook catchment The Hupsel Brook catchment is an intensively investigated lowland catchment located in the east of the Netherlands. The SDS approach is not always performing well for the Hupsel Brook catchment, especially not during dry periods (Brauer et al., 2013). The hydrograph shows that discharge and precipitation are closely linked, as similar patterns can be observed. Additionally, Brauer et al. (2013) found a non-linear dependency of discharge to storage. However, more complex models, among one in which separate flow routes are simulated separately, performed well in predicting the catchment's discharge. This suggests that the catchment might have more than one dominant flow path.

Plynlimon catchment Kirchner (2009) showed very good model performance using the SDS approach as a predictive tool for rainfall-runoff modeling in the Plynlimon catchment in Wales, with Nash-Sutcliffe efficiencies exceeding 0.9. The catchment can be divided into two sub-catchments; Severn and Wye. The two catchments are geologically similar, but are dominated by grassland and conifers, respectively (Brandt et al., 2004). The non-linear dependency of discharge to storage found by Kirchner (2009) and the good performance lead to the believe that the catchment's discharge mainly consists of shallow groundwater based discharge, where overland flow is only a small component of the total discharge.

Rietholzbach catchment The Rietholzbach catchment is a pre-Alpine catchment situated in north-east Switzerland. The discharge of this humid catchment follows precipitation closely. Teuling et al. (2010) showed non-linear dependency of the discharge to subsurface storage and accurate discharge predictions using the SDS approach. However, under dry conditions, overland flow occurs more frequent and the effect of evapotranspiration, which is strongly correlated with the available radiative energy, is more pronounced (Teuling et al., 2010). Therefore, discharge is believed to mainly consist of shallow groundwater flow, occasionally supplemented with with overland flow.

Alptal catchment The pre-Alpine Alptal catchment is situated in the centre of Switzerland (Stähli et al., 2021) and includes three smaller sub-catchments: Erlenbach, Lümpenenbach and Vogelbach. The hydrographs show a fast response to precipitation, in general responding within ten minutes after the event (van Meerveld et al., 2018). Additionally, their soils have a low permeability (Milzow et al., 2006) and shallow groundwater tables that also respond quickly to rainfall (van Meerveld et al., 2018). A correlation was found between groundwater tables and discharge measurements, most pronounced during the growing season. Next, overland flow was observed often during precipitation events, except for the forested areas (van Meerveld et al., 2018). Together, this suggests that discharge will be dominated by the quickest processes; shallow groundwater flow and overland flow. Compared to the other two catchments,



Figure 2.1: Map of Europe showing locations of the studied catchments.

Vogelbach has the highest forested area (65%) (Stähli et al., 2021), which could explain its (slightly) slower discharge response (Staudinger et al., 2017).

Sperbelgraben and Rappengraben catchment The Sperbelgraben and Rappengraben catchments are two parallel catchments with a similar area and channel density, located in the Swiss pre-Alpine Emmental (Stähli et al., 2011). The main difference between the two is the forested area; whereas Sperbelgraben is almost entirely forested, the Rappengraben contains about 50% forest and 50% pasture (Stähli et al., 2011). Discharge responds quickly to rainfall, especially with wet antecedent conditions. The response of discharge to rainfall in the Rappengraben catchment is more pronounced than the response in the Sperbelgraben catchment (Roessel, 1950). As a result of the higher fraction of forest, the fraction of baseflow in the Sperbelgraben catchment is expected to be higher than in the Rappengraben catchment. Yet, the fraction of shallow groundwater based discharge is expected to be larger, based on the quick response to precipitation events.

Atttert catchment The Atttert catchment is located in the west of Luxembourg and can be divided into seven sub-catchments (Martínez-Carreras et al., 2012). Three sub-catchments will be investigated: Huewelerbach, Weierbach, and Wollefsbach, which have similar climatologies (Fencia et al., 2014). During summer, the Huewelerbach catchment shows a very stable base flow, which is assumed to be the result of deep groundwater (Martínez-Carreras et al., 2012; Fencia et al., 2014). During winter, the hydrograph shows more peaky behaviour, directly following precipitation patterns. The Weierbach catchment's hydrograph seems to correlate with precipitation patterns, but with some retention time. Therefore it is assumed that the total discharge consists of mainly deep and shallow sub-surface flow (Fencia et al., 2014). The hydrograph of the Wollefsbach catchment shows most peaky behaviour, quickly responding to precipitation events. It is expected that this is the result of shallow sub-surface flow supplemented with overland flow (Fencia et al., 2014).

Table 2.1: Overview of the catchments explored in this study and the catchment characteristics; Area, discharge (Q), precipitation (P), evapotranspiration (ET) and height difference between the highest and lowest situated point in the catchment. Expected dominant discharge process with GW = deep groundwater dominated, S = shallow groundwater dominated and OF = overland flow dominated. Years for which the data is available and the providers of the data. *Swiss Federal Institute for Forest, Snow and Landscape Research, provided by Stähli (2018) **Discharge data were obtained from the Federal Office for the Environment. ***Luxembourg Institute of Science & Technology. ****ERA5 hourly data on single levels from 1940 to present - reanalysis.

(Sub)Catchment	Area (km^2)	Mean Q (mm/yr)	Mean P (mm/yr)	Mean ET (mm/yr)	Height difference (m)	Expected dominant process(es)	References	Data Availability (<i>years</i>)	Data Provider
Noor	10.56	91-329	799	620	149	GW	van Lanen et al. (1995); Kordík et al. (2002); Dijkma and van Lanen (2001)	1992-2006	Roel Dijkma
Hupsel Brook	6.5	310	790	560	13	S/GW	Brauer et al. (2013); Terink et al. (2018); Brauer et al. (2014)	1976-2022	Brauer et al. (2018)
Severn	8.70	1987	2553	566	419	S	Kirchner (2009)	1992-1996	Kirchner (2009)
Wye	10.55	2111	2599	488	397	S	Kirchner (2009)	1992-1996	Kirchner (2009)
Rietholzbach	3.31	1021	1459	396	286	S/OF	Teuling et al. (2010); Melsen et al. (2014); Seneviratne et al. (2012)	2000-2007	Teuling et al. (2010)
Erlenbach	0.7	1889	2294	512	555	OF/S	Knapp et al. (2020); Staudinger et al. (2017); WSL*	1992-2007	WSL*, ERA5****
Lümpenbach	0.9	1934	2426	425	416	OF/S	Staudinger et al. (2017); WSL*	1992-2007	WSL*, ERA5****
Vogelbach	1.6	1572	2159	558	502	S	Staudinger et al. (2017); WSL*	1992-2007	WSL*, ERA5****
Sperbelgraben	0.544	840	1660	550	292	S	Badoux et al. (2006); Stähli et al. (2011); Rickenmann (1997); Nitsche et al. (2011); Penman (1959)	1958-2023	FOEN**, ERA5****
Rappengraben	0.596	1040	1713	500	260	S	Stähli et al. (2011); Nitsche et al. (2011); Rickenmann (1997); Penman (1959)	1958-2023	FOEN**, ERA5****
Huewelerbach	2.7	189	800	620	120	GW/OF	Fenicia et al. (2014); Martínez-Carreras et al. (2012); Wrede et al. (2015); Onderka et al. (2012)	2001-2022	LIST**, ERA5****
Weierbach	0.42	478	1000	620	90	S/GW	Wrede et al. (2015); Hissler et al. (2021); Fenicia et al. (2014)	2004-2022	LIST**, ERA5****
Wolfsbach	4.5	200	850	620	61	S/OF	Wrede et al. (2015); Kavetski and Fenicia (2011); Fenicia et al. (2014)	2001-2022	LIST**, ERA5****

3 | Methods

To create a model that predicts discharge based on three flow paths, first the data, as described in Chapter 2, were analyzed. For all catchments, the discharge components were identified based on the hydrograph and their relative contribution to the total discharge were evaluated after plotting them in a ternary plot. The next step was to predict the discharge based on the found distribution of the three components. Identifying the components and predicting discharge was done by creating a model in Python. An overview of the model is given in Figure 3.4. To reduce calibration efforts, it was investigated whether the component fractions coincide with already-existing datasets on groundwater recharge. Finally, the new discharge predictions were compared to discharge observations and outputs of two more conventional models, the SDS approach and HBV.

3.1 Data Preparation & Analysis

The collected data were organized in such a way that the data for all catchments are in equal format. The discharge, precipitation and potential evapotranspiration data are measured in units of depth per time, as an average value for the entire catchment. The potential evapotranspiration and precipitation data that were collected from the ERA5 single levels database were transformed from spatial data (horizontal resolution: $0.25^\circ \times 0.25^\circ$) to time series as an average value over the catchment.

3.1.1 Evaporation Correction

First of all, a water balance check was executed for the raw data. Assuming closure of the water balance following a closed system in which conservation of mass is true, precipitation, actual evapotranspiration and discharge fluxes should add up to zero. However, as potential evapotranspiration data were used, an evaporation correction factor, r [-], was determined to close the water balance. The correction fraction was calculated following equation 3.1, where ET_{pot} is the potential evapotranspiration [mm/h] and P precipitation [mm/h]. Using the found evaporation correction factor, the actual evapotranspiration, ET [mm/h] can be calculated as $ET = r * ET_{pot}$.

$$r = \frac{P_{sum} - Q_{total,sum}}{ET_{pot,sum}} \quad (3.1)$$

3.1.2 Discharge Threshold

Additionally, a second evapotranspiration reduction method was applied. Potential evapotranspiration is a measure of actual evapotranspiration, assuming an unlimited water availability (Li et al., 2016). A reduction of the actual evapotranspiration as compared to the potential evapotranspiration occurs as a result of dry soil moisture conditions. In the model, discharge was used as an approach to soil moisture conditions, as storage and discharge are linked monotonically. Following Teuling et al. (2010) and Buitink et al. (2020), the threshold value for discharge was set to 10^{-4} mm/h. When discharge was below this value, actual evapotranspiration was set to a value of zero to prevent modeling negative discharge.

3.1.3 Spectral Analysis

Spectral analysis was executed to gain more insight into the most important time frames at which discharge and precipitation occur. This insight served as a first estimation of the important mechanisms playing part in a catchment. Additionally, power spectra are a useful tool for investigating the differences in precipitation and discharge regimes between catchments. Using Python, the Fast Fourier Transform (FFT) algorithm was applied and FFT frequencies were calculated. The Power Spectral Density (PSD) was computed as the square of the absolute value of FFT, divided by the period.

An example power spectrum for the observed discharge and precipitation of the Rietholzbach and Severn catchment is shown in Figure 3.1. A strong power output indicates a strong response for that certain timescale. It can be seen that for higher frequencies, the power output is larger for precipitation than for discharge for both catchments. This is an indication of water retention by the catchment, as it increases the response time of streamflow to precipitation. A catchment for which the discharge responds very quickly to precipitation due to a high contribution of overland flow would show more similar power outputs for the higher frequencies.

For frequencies larger than 0.2 cycles per day the power output is larger for the Rietholzbach catchment than for the Severn catchment. This indicates that there is more (very) quickly responding discharge in the Rietholzbach catchment than in the Severn catchment.

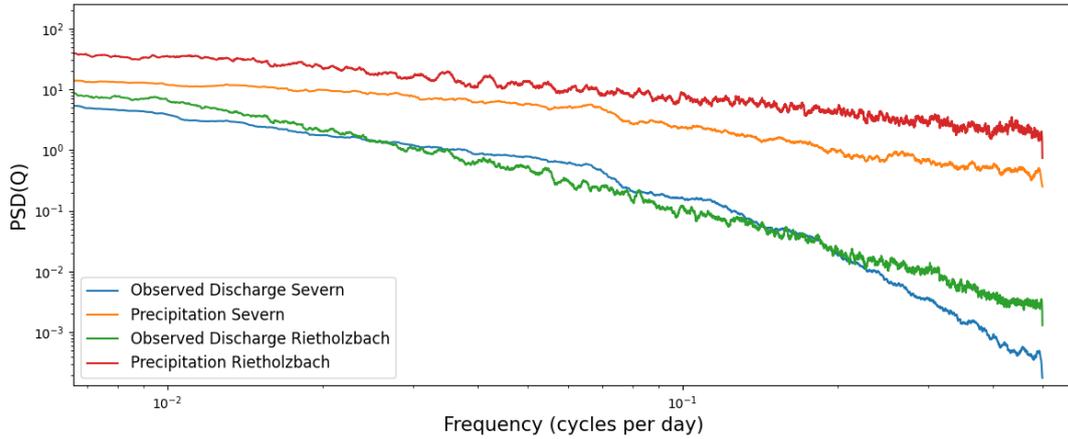


Figure 3.1: Power spectrum for the observed discharge and precipitation of the Severn and Rietholzbach catchment, zoomed in to the higher frequencies. With PSD as the Power Spectral Density of discharge (Q) on the y-axis and frequencies on the x-axis where a frequency of one indicates one cycle per day.

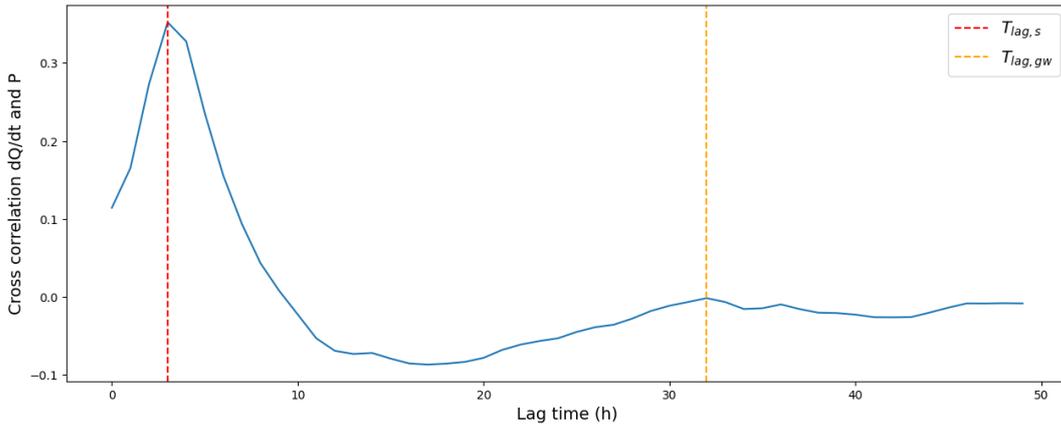


Figure 3.2: Cross correlation between changes in discharge (dQ/dt) and variations in precipitation (P) plotted over time for the Hupsel Brook catchment. Indicating lag times for deep (orange) and shallow (red) groundwater flow.

3.2 Model Development

When building the model, it was assumed that the total discharge, Q_{total} [mm/h], consists of three discharge components (equation 3.2) following the flow paths as explained in the introduction (Figure 1.1): deep groundwater generated discharge, also known as base-flow (Q_{GW}), shallow groundwater generated discharge (Q_S) and overland flow (Q_{OF}). The relative contribution of these components differs between catchments due to differences in catchment characteristics and precipitation regimes.

$$Q_{total} = Q_{GW} + Q_S + Q_{OF} \quad (3.2)$$

3.2.1 Lag Time

Changes in catchment discharge may show a lag with precipitation. Therefore, the lag time was inferred from the peaks in the cross correlation between changes in discharge (dQ/dt) and variations in precipitation. The smallest lag time that was found is used as lag for the shallow groundwater flow, whereas the secondary peaks in lag time were used to approach the lag time of deep groundwater flow. Figure 3.2 shows the determination of the lag time for the Hupsel Brook catchment. The first peak, at a lag time of 3 hours is assigned to the shallow groundwater flow. The second peak is located at a lag time of 32 hours and is assigned to the deep groundwater flow. Lag times were taken into account when modeling discharge in Section 3.2.2 and 3.2.3.

3.2.2 Deep Groundwater Flow

Parameter Identification

Primarily, the discharge generated by deep groundwater flow was identified from the total discharge by executing a baseflow separation. Baseflow separation is a technique that has been widely used to separate the slowly responding runoff as a result of deep groundwater flow (also called the baseflow, Q_b [mm/h]) from the total discharge. Subtracting the baseflow from the total discharge results in the sum of shallow groundwater generated discharge and overland flow (Chapman, 1999). Following Eckhardt (2008), it was assumed that there is a linear response of baseflow to the amount of water stored in the catchment, leading to an exponential baseflow recession (equation 3.3). In this equation, Q_t is discharge at time step t [mm/h], Q_0 discharge at time step zero [mm/h], t the time step [h] and τ represents the recession coefficient [h^{-1}].

$$Q_t = Q_0 * e^{-t/\tau} \quad (3.3)$$

In a frequency spectrum, baseflow gives a higher response to long waves with a low frequency whereas quick flow processes have a higher frequency. Lyne and Hollick (1979) used a filter equation to separate the baseflow from the total discharge. Based on this, Eckhardt (2008) formulated a two parameter Recursive Digital Filter:

$$Q_{b,t} = \frac{(1 - BFI) * z * Q_{b,t-1} + (1 - z) * BFI * Q_t}{1 - z * BFI} \quad (3.4)$$

Where $Q_{b,t}$ is the baseflow at time step t [mm/h], and $Q_{b,t-1}$ the baseflow at the previous time step [mm/h]. The parameter BFI indicates the maximum value of the baseflow index [-], which is the ratio between baseflow and total discharge (Kissel and Schmalz, 2020). Eckhardt (2008) suggested to set $BFI = 0.80$ for perennial streams with porous aquifers, $BFI = 0.50$ for ephemeral streams with porous aquifers, and $BFI = 0.25$ for perennial streams with hard rock aquifers. From the data analysis (Section 3.1.3), a first impression of the difference in precipitation and discharge regimes was gained, which was, together with the first expectations as described in Chapter 2 and the BFI values as suggested by Eckhardt (2008), used as a starting point for the cutoff between deep and shallow groundwater. The model was calibrated on the value of BFI after the deep and shallow groundwater based discharge were determined, to get the most accurate discharge predictions (Section 3.2.4).

The second parameter in equation 3.4, z , is the recession rate, which was determined based on a recession analysis of the measured discharge data. For the recession analysis, only the discharge measurements that are part of a dry period of three days or more were considered. According to Eckhardt (2008), if the recession is long enough that the current day and the next day consist only of baseflow and there is no recharge during this time, discharge at the next time step can be calculated according to equation 3.3, where the recession constant z is as follows:

$$e^{-t/\tau} = z \quad (3.5)$$

Using the polyfit function (Python, using numpy package version 1.26.4), which fits the best line through the points of current discharge plotted against the discharge at the next time step, the value of z was determined as the slope of this line. Consequently, the found value for z could be implemented in equation 3.4 together with the first estimate of BFI to calculate baseflow. The intercept, l , which was found when fitting the regression line will be used in section later to predict deep groundwater flow.

Discharge Prediction

Based on the conditional BFI value, the precipitation and evapotranspiration fraction belonging to the deep groundwater based flow (P_{GW} and ET_{GW} , respectively) are determined by multiplying the total precipitation and evapotranspiration with BFI. Following Kirchner (2009), discharge can be predicted without needing data on storage using equation 3.6. The equation can be solved following a fourth-order Runge-Kutta integration scheme.

Precipitation was measured as the liquid precipitation that falls. Therefore, M_{GW} in equation 3.6 accounts for snow melt [mm/h], as it also contributes to the water availability (Teuling et al., 2010). Snow storage is modelled following the approach of Buitink et al. (2020). In the catchments for which ERA5 data were used, the total water availability was modeled by subtracting snowfall from the total precipitation and adding snow melt. Snow processes were only taken into account in catchments experiencing snowfall regularly in winter.

$$\frac{d(\ln(Q_{GW}))}{dt} = g(Q_{GW}) * \left(\frac{P_{GW} + M_{GW} - ET_{GW}}{Q_{GW}} - 1 \right) \quad (3.6)$$

In equation 3.6, $g(Q_{GW})$ is the sensitivity function of the deep groundwater flow, as defined in equation 3.7. In this function, z is the slope found during the

baseflow separation and k can be calculated as negative one divided by the logarithm of the intercept l .

$$g(Q_{GW}) = k * Q_{GW}^{z+1} \quad (3.7)$$

Lag times were taken into account by shifting the calculated discharge in time ($T_{lag,gw}$) compared to precipitation and evapotranspiration data.

3.2.3 Shallow Groundwater Flow

Parameter Identification

The second component that was identified was the shallow groundwater flow, which was based on the SDS approach by Kirchner (2009). The deep groundwater flow has been determined (Section 3.2.2) and was subtracted from the total observed discharge. The residual discharge, which is the summation of the shallow groundwater flow and overland flow (Chapman, 1999), was used to determine the relative contribution of the shallow groundwater generated discharge to the total discharge. Besides, the residual discharge was used to validate the modeled shallow groundwater flow on. The same was done for precipitation and evapotranspiration; based on the value of BFI, the fraction of precipitation and evapotranspiration used for baseflow generation was subtracted from the total, resulting in P_S and ET_S , respectively.

Following Kirchner (2009), it was assumed that the quickly responding discharge solely depends on the amount of water in storage, S [mm], in the catchment:

$$Q = f(S) \quad (3.8)$$

A storage-discharge relationship was identified from discharge recession and was used later on to predict discharge from precipitation, following the methodology of Kirchner (2009). Storage changes result from the difference between incoming and outgoing water fluxes, as can be seen in equation 3.9.

$$\frac{dS}{dt} = P + M - ET - Q \quad (3.9)$$

From these equations the sensitivity function of shallow groundwater flow, $g(Q_S)$, was derived, which was defined by Kirchner (2009) as the discharge sensitivity to changes in catchment storage. The function follows from the derivative of $f(S)$, leading to equation 3.10.

$$g(Q_S) = \frac{dQ_S}{dS} = \frac{dQ_S/dt}{P_S + M_S - ET_S - Q_S} \quad (3.10)$$

As precipitation, evapotranspiration and snow melt can be measured less accurately than discharge due to their spatial variability, it is best to determine the sensitivity function when discharge is large compared to the other parts of the water balance. Equation 3.10 then simplifies to:

$$g(Q_S) = \frac{dQ_S/dt}{Q_S} \quad (3.11)$$

Data were selected for which the assumption can be made that discharge is larger than the other parts of the water balance. Following Kirchner (2009), data were selected based on the following criterion: discharge at time step i is ten times as large as precipitation and evapotranspiration at the same time step.

After the records have been selected, the recession curve was determined. Non-linearity of the system was assumed (equation 3.12), as multiple studies showed non-linear behaviour of catchments (Kirchner, 2009; Teuling et al., 2010; Rusjan and Mikoš, 2015; Brauer et al., 2013; Adamovic et al., 2015).

$$-\frac{dQ_S}{dt} = a * Q_S^b \quad (3.12)$$

Following the approach of Kirchner (2009), the selected records were used to calculate the discharge recession as the difference between discharge of two consecutive hours ($-dQ_S/dt = Q_{S,i-1} - Q_{S,i}$). Next, the average discharge of two consecutive hours has been determined as follows: $Q_{S,avg} = (Q_{S,i-1} + Q_{S,i})/2$. The two were plotted in a scatter plot in log space. The data were binned into ranges of Q_S , spanning at least 1% of the range of Q_S in log space and fulfilling the criterion that $std.error(-dQ_S/dt) \leq mean(dQ_S/dt)/2$. If the criteria was not fulfilled, binning size increased until the requirements were met. Using the optimization function curve fit (Python, using scipy package version 1.13.0), a quadratic curve was fitted through the binned data to find values for parameters a and β in equation 3.12, which gives the sensitivity function through:

$$g(Q_S) = \frac{dQ_S/dt}{-Q_S} = a * Q_S^{\beta-1} \quad (3.13)$$

Discharge Prediction

As said before, the residual precipitation and evapotranspiration are calculated based on the fraction used for the deep groundwater modelling. In this section, the residuals will be used to approach the precipitation and evapotranspiration fraction belonging to the shallow groundwater based flow.

Similar to the baseflow prediction, shallow groundwater based discharge can be predicted without needing data on storage using equation 3.14, which again can be solved following a fourth-order Runge-Kutta integration scheme. In the equation, M_S accounts for the snow melt [mm/h], which was modeled similar to M_{GW} (section 3.2.2).

$$\frac{d(\ln(Q_S))}{dt} = g(Q_S) * \left(\frac{P_S + M_S - ET_S}{Q_S} - 1 \right) \quad (3.14)$$

The sensitivity function for quickly responding discharge uses the found values for a and β . Lag times were taken into account by shifting the calculated discharge in time ($T_{lag,s}$) compared to precipitation and evapotranspiration data.

3.2.4 Model Calibration

Calibration Methods

The deep and shallow groundwater based discharge were summed, after which the model was calibrated. The BFI value was calibrated by optimizing the Nash-Sutcliffe efficiency (NSE) of the summed discharge predictions and the observed total discharge, using the Nelder–Mead method. The optimization was done based on a time period of one year. NSE values were calculated using the hydroeval package in Python (version 0.1.0), which was developed especially to evaluate the goodness of fit between observed and modeled discharge time series (Hallouin, 2021).

Component Predictions

A test was executed to see whether the relative fractions of discharge can be estimated beforehand based on an already existing dataset: the pan-European high-resolution groundwater recharge map by Martinsen et al. (2022), in which amount of potential recharge [mm/yr] is given. In this dataset, recharge is defined as the water percolating from upper soil layers to the groundwater table and can be calculated as the excess water (precipitation minus actual evapotranspiration, also known as effective precipitation) times the recharge coefficient. The recharge coefficients were compared to the found BFI values and the recharge values were compared to the modeled deep groundwater generated discharge.

3.2.5 Overland Flow

The overland flow component of discharge is the quickest responding part occurring as a result of high intensity rainfall or as a result of soil saturation caused by much antecedent rainfall (de Lima, 1989; Maier and van Meerveld, 2021). Next to this, part of the precipitation will fall directly into the stream, causing discharge without any lag time. Determining when this occurs and in what proportions, was hard to conceptualize. Therefore, the overland flow fraction was determined based on the difference between the total observed discharge and the discharge generated by deep and shallow groundwater flow, as follows:

$$Q_{OF} = Q_{total} - Q_{GW} - Q_S \quad (3.15)$$

The fraction of overland flow was compared to other data, like precipitation, to find a correlation that could be used to predict overland flow. In many catchments, overland flow occurs as a result of paved area. Therefore, the amount of overland flow might be a fraction of the the amount of precipitation that can be linked to the pavement. However, the importance of Hortonian and Dunnian overland flow to the total amount of overland flow was not known. Therefore, the response of the overland flow fraction was also compared to the rainfall intensity and antecedent precipitation. However, none of these showed clear correlations that could be used to make a reasonable estimate of the overland flow fraction.

As a result of soil saturation, the area contributing to overland flow is variable. Therefore, a relationship between the overland flow fraction and antecedent discharge was searched for, to define an equation accounting for a variable source area. This was done by plotting the overland flow part of the discharge divided by the total precipitation against the total discharge, as discharge is an indicator of catchment wetness. The data points were binned into ranges of the total discharge, spanning 1% of the range of the total discharge in the log space. Next, the optimization curve fit function (Python, using scipy package version 1.13.0) was used to fit a quadratic curve through the binned data. The found parameters, c and d , were used to model the fraction of precipitation that forms overland flow (f_{OF}) as a function of the total discharge (equation 3.16). However, when predicting discharge, the total observed discharge is unknown and is therefore approached as the summation of the deep and shallow groundwater based discharge (Q_{total}).

$$f_{OF} = c * Q_{total}^d \quad (3.16)$$

To avoid overestimating the amount of overland flow, the calculated fractions were multiplied with the difference between the total observed precipitation and the modeled deep and shallow groundwater flow:

$$Q_{OF} = f_{OF} * (P - Q_{GW} - Q_S) \quad (3.17)$$

The total modeled discharge could then be calculated as the sum of the three components. The last step was to do another water balance check, in which the modeled and observed discharge are compared.

3.2.6 Ternary Plot

After the steps mentioned above were executed, two recession curves and the fraction of overland flow based on discharge and precipitation were known. The relative fractions were defined as the discharge generated by one mechanism divided by the total modeled discharge. To create an overview of the catchments, the relative fractions were plotted in a ternary plot.

Ternary plots are a useful tool to show the proportions of three variables. Figure 3.3 provides an explanation of the layout of the created ternary plot. The corners of the plot indicate the extremes; when a catchment is located in one of the corners, it means that its total discharge consists of only one (or is strongly dominated by one) discharge component, whereas a catchment located more towards the middle of the plot indicates that the discharge consists of a mixture of all three components.

However, the fraction of overland flow, deep and shallow groundwater based discharge in the total discharge will differ for different climatic/weather conditions. Therefore, the location of a certain catchment on the ternary plot was defined as a yearly average. To show the yearly variability, plots were made showing the relative fractions as a function of a time (over a time span of one year).

3.3 Model Quality Assessment

3.3.1 General Model Performance

NSE was calculated as a measure of the general accuracy of the model, using the hydroeval package in Python (version 0.1.0). NSE values are often used as evaluation criterion for hydrological models. The values for which a model is considered unsatisfactory varies around 0.36 (Eryani et al., 2022) to 0.50 (Moriasi et al., 2015).

Values above 0.50 are considered satisfactory and values above 0.70 are considered good (Eryani et al., 2022; Moriasi et al., 2015).

Additionally, the Kling-Gupta efficiency (KGE) was calculated, also using the hydroeval package (Hallouin, 2021). KGE is an increasingly used metric to indicate model performance in a slightly more balanced way than NSE values do (Knoben et al., 2019). A KGE value is considered good when it is above 0.75 and poor when it is below 0.5 (Rogelis et al., 2016). However, different authors consider different KGE values to be satisfactory, as discussed by Knoben et al. (2019).

A power spectrum analysis was executed to gain more insight into the timescales for which the model works well or less well. For this analysis, the similarity between the power spectrum of the observed discharge and the modeled discharge was compared. It was assumed that the more similar the power spectra are, the more accurate the prediction is.

3.3.2 Comparison to Other Models

The SDS approach by Kirchner (2009) does not take into account the possibility of multiple flow routes in a catchment. This section displays the methodology used to test whether accounting for this possibility improves discharge prediction, or whether it decreases its accuracy, as the number of parameters in the new model does increase significantly. An increase in the number of model parameters could be at hand with simulating additional processes, but could also cause overfitting when the available data is limited, which can introduce a noise increasing model uncertainty.

HBV does account for different response times, but has a different model concept using an upper and lower groundwater box. The discharge predictions of the HBV model will also be compared to the output of the new model. Discharge predictions based on the SDS approach were made following the methodology of Kirchner (2009). Additionally, the model by Kirchner (2009) was calibrated based on the Levenberg-Marquardt algorithm, to optimize the model prediction. The other discharge predictions are made using HBV light, using a Matlab version by McGuire (2012), which was rewritten into Python. The HBV light model uses a Monte-Carlo calibration, which finds the most suitable parameter set resulting in the highest NSE values, based on 1000 model runs.

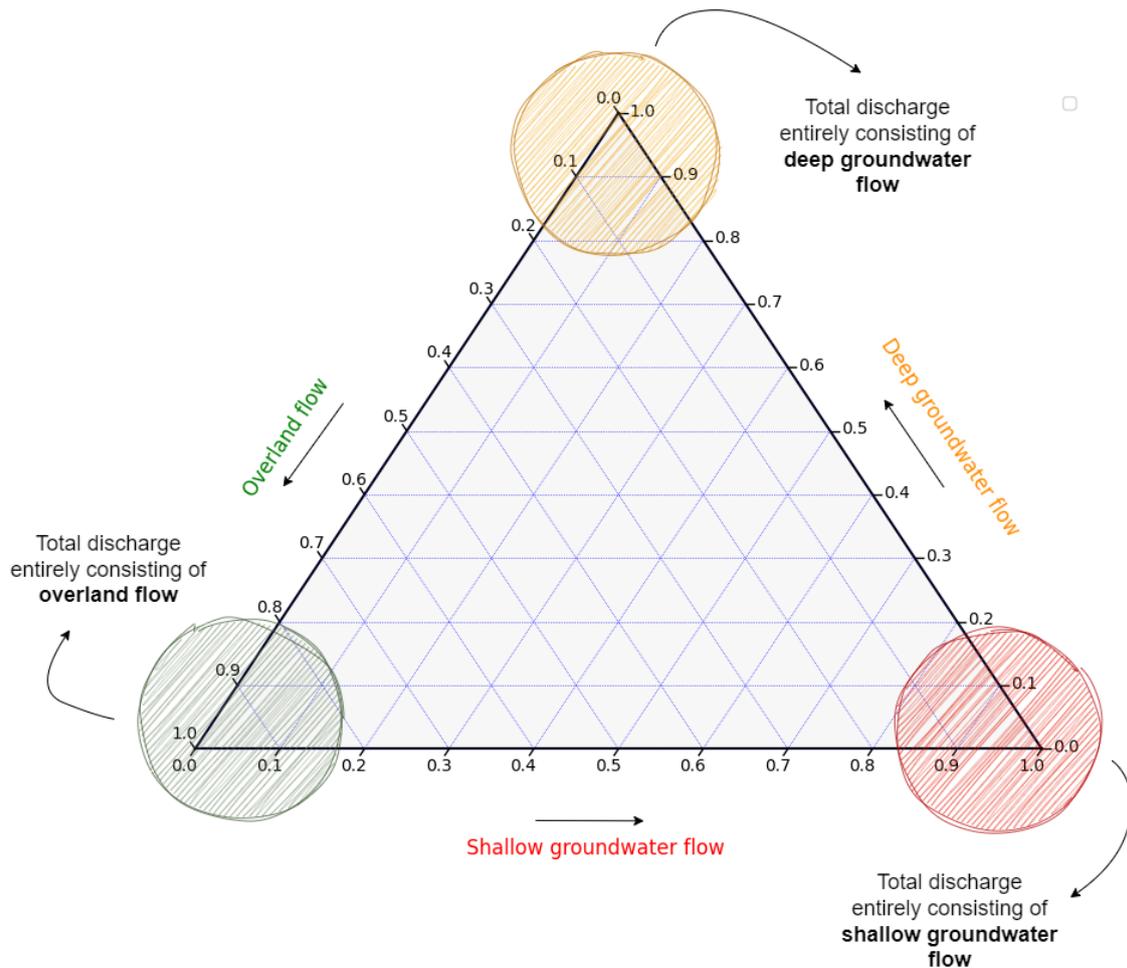


Figure 3.3: Explanation of the ternary plot that was created to show the relative distribution of the discharge components for the different catchments.

First, the model performance was compared based on the NSE and KGE. Next to these statistical measures, the models were evaluated based on other hydrological signatures. In his paper, McMillan (2021) shows the relevance of hydrological signatures to evaluate the similarity of hydrological time series, which could also be used for model evaluation. Two of those signatures, the rising limb density (RLD) and the flow duration curve (FDC) were used to assess how similar the modeled discharge is to the observed discharge.

RLD is a measure of the time it takes for the catchment to reach the peak discharge after a rainfall event, influenced by e.g. slope, moisture conditions and soil type (Mathai and Mujumdar, 2022). RLD can therefore be used to investigate the similarity in runoff mechanisms between catchments (Mathai and Mujumdar, 2022), but in this case it was used to compare the correctness of the modeled storage-discharge relationship and water retention times.

FDC is a curve that shows the percentage of time during which a certain discharge is exceeded (Searcy, 1959). It is a statistical measure of the flow distribution (McMillan, 2021), which is unique per catchment. It can be used to evaluate similarities between catchments, but is also used to calibrate hydrological models on (Westerberg et al., 2011). In this case, FDC was used to evaluate the similarity between the modeled and observed hydrographs.

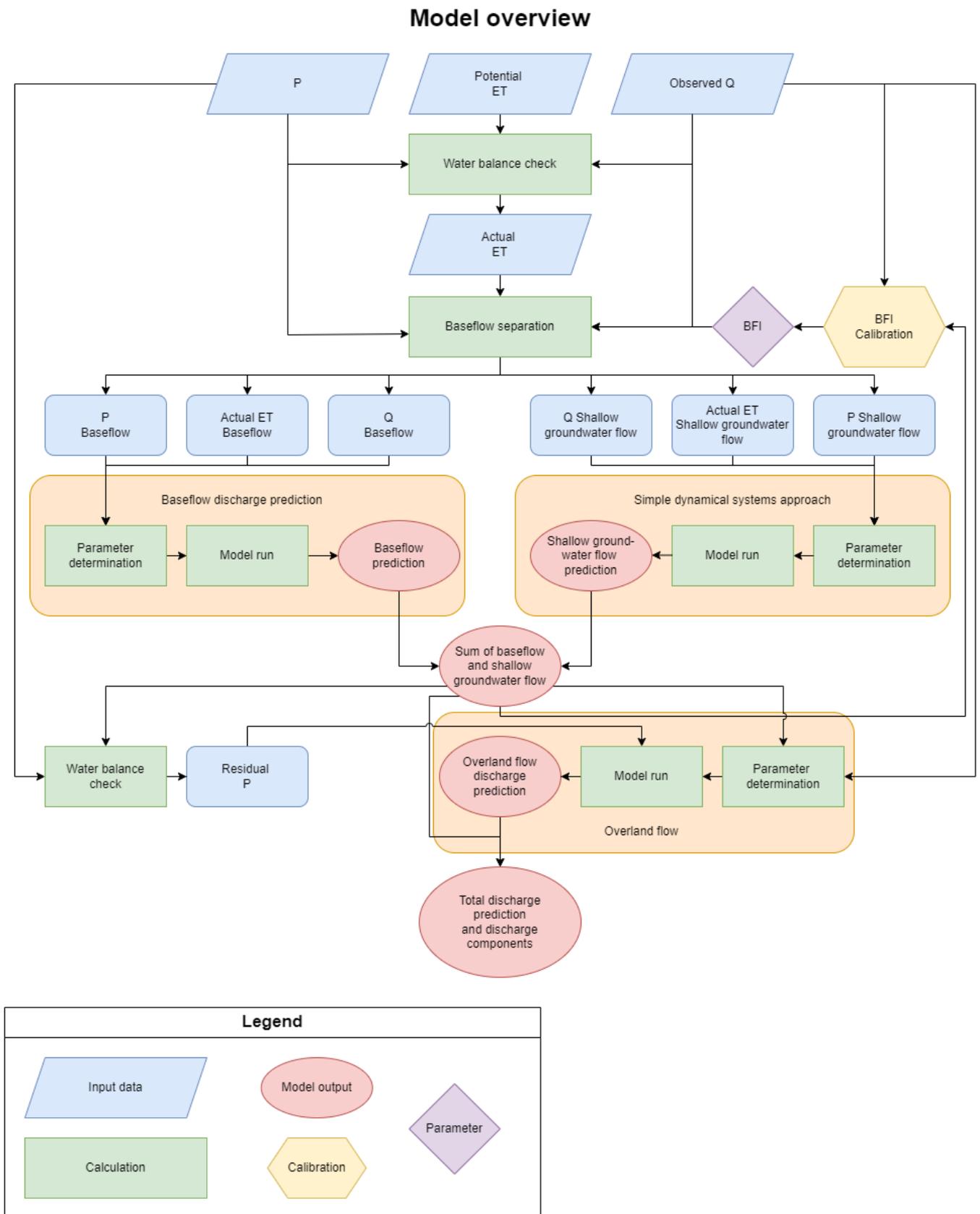


Figure 3.4: Graphical overview of the methods with P = precipitation, ET = evapotranspiration and Q = discharge.

4 | Results & Interpretation

This chapter provides an overview of the found results. Figures for parameter identification, discharge prediction (Section 4.2) and model comparison (Section 4.3.2) are only shown for the Severn catchment, but the same steps were taken for the other catchments. Severn was chosen as a representative catchment as high quality and complete data was available.

4.1 Data Preparation & Analysis

4.1.1 Evaporation Correction

First of all, the raw data was analyzed to investigate the differences between the catchments. In Table 4.1, the evapotranspiration correction factors (r) are displayed. These values are used to approach the actual evapotranspiration (ET) from potential evapotranspiration (PET) in order to close the water balance. What is striking about the found values of r is that they tend to differ quite a lot between catchments close together, such as Erlenbach and Lümpenenbach.

The determination of r is done based on the assumption that the multi-year water balance is closed. This means that there is no change in storage over the years, implying that precipitation (P), actual ET and discharge (Q) add up to zero. A small actual ET as compared to the PET is caused by a water shortage. Having a stable baseflow could be an indication of a steady water availability throughout the year, resulting in an energy limited system instead of water limited, for which PET is a good approximation of actual ET. This could serve as an explanation for the differences in the found r -values. However, no clear evidence of this relation was found for the studied catchments.

For some catchments, the found correction factor is above 1, which is counter-intuitive considering that the input data is PET. However, the catchments for which this was found are catchments that use ERA5 data, which have a larger spatial coverage than the size of the catchment. Therefore, the PET data could be an underestimation of ET occurring in the catchment as it represents an average for a larger region.

4.1.2 BFI Estimation

The expected BFI values (Table 4.1) are based on knowledge of the catchment's discharge characteristics (Chap-

ter 2) and analysis of the Power Spectral Density (PSD) of P and Q . The highest expected BFI was found for the Noor catchment, as it is known that this catchment's discharge mainly consists of deep groundwater flow, showing little response to rainfall events. Besides, the power output for discharge at higher frequencies was very low, showing a large difference with the output of rainfall data, which supports the high expected BFI.

Based on the hydrograph, BFI for the Hupsel Brook catchment is expected to be slightly larger than BFI for the Severn and Wye catchment. This is supported by the PSD of Q , which showed a slightly smaller power output at the higher frequencies for the Hupsel Brook catchment than for the Severn and Wye catchment. Rietholzbach is a pre-Alpine catchment for which streamflow is mainly a function of subsurface water storage. The power spectrum showed a strong output for P and for the higher frequencies the difference between P and Q was remarkable. However, the power output for Q was still higher than for the other four catchments analyzed before, therefore leading to a smaller expected BFI.

Erlenbach, Lümpenenbach and Vogelbach showed very similar power spectra, with a distinct peak at a frequency of 0.035 per day, equal to about a month. The discharge is expected to mainly consist of shallow groundwater based discharge and the power spectra showed similar results to Rietholzbach. Therefore, the BFI is expected to be low. The power spectra of Sperbelgraben and Rappengraben are again very similar, only the output being slightly higher for Rappengraben than Sperbelgraben. The power spectra are comparable to that of the Rietholzbach, except for the higher frequencies, where the power output is larger than for Rietholzbach, leading to a smaller expected BFI.

The sub-catchments of the Attert catchment showed differences mainly for the higher frequencies. This is due to the lack of overland flow in the Weierbach catchment, whereas this is present in the Huewelerbach catchment. Therefore, Weierbach is expected to have the highest BFI and was chosen similar to that of the Hupsel Brook catchment, as their power spectra were comparable. The power spectrum of Wollefsbach was comparable to that of Severn and therefore, the BFI value was expected to be similar.

Table 4.1: Overview of the studied catchments, the found evaporation correction factors (r) and expected baseflow indices (BFI).

Catchment	r	Expected BFI
Noor	1.19	0.75
Hupsel Brook	0.87	0.40
Severn	0.86	0.35
Wye	1.00	0.35
Rietholzbach	0.79	0.10
Erlenbach	0.98	0.10
Lümpenenbach	0.65	0.10
Vogelbach	1.11	0.10
Sperbelgraben	0.77	0.08
Rappengraben	0.80	0.08
Huwelerbach	1.19	0.30
Weierbach	0.79	0.40
Wollefsbach	0.96	0.35

4.2 Model Development

4.2.1 Lag Time

Lag times were calculated for all catchments and are displayed in Table 4.2. Some catchments show a clear distinction between lag time for the shallow groundwater flow and deep groundwater flow, while this distinction is less clear for other catchments. The found lag times range between 0 to 3 hours for the shallow groundwater flow. The largest differences in lag times can be found in the deep groundwater based discharge. These lag times range from 5 hours for catchments with shallow sub-soils or steep slopes (Alpine catchments) and lag times larger than one hundred hours in the Noor catchment, which is characterized by its high infiltration capacity and deep groundwater tables.

These first findings serve as a good insight into the processes taking place inside the catchments. For the catchments showing a long retention time for the deep groundwater, the importance of the memory of the discharge to precipitation is highlighted. For the catchments showing only a small difference in retention time between the two, one may question the effectiveness of explicitly modelling deep groundwater separately and the necessity of using the new model.

4.2.2 Deep Groundwater Flow

Parameter Identification

The first step in predicting the deep groundwater flow, is executing a baseflow separation. For this only those time

steps are selected that are part of a discharge recession of three days or longer (72 consecutive hours). Figure 4.1a shows an hourly time series of total discharge in the Severn catchment. It can be seen that the discharge peaks follow the timing of precipitation closely, whereas this behaviour was less clear in for example the Noor catchment. During some instances the discharge seems to recede for a longer time period, however, no time steps were selected. In these cases, the hourly discharge did not recede continuously for 72 hours, even though the average discharge trend is decreasing.

The selected discharge values are plotted in Figure 4.1b, against the associated discharge at the next time step. It can be seen that the data points are situated in a fairly neat line, through which a least-squares linear regression line is fitted. The R^2 value of the fit is 0.999, associated with small residuals (Figure 4.1c). However, the residual plot does indicate some heteroscedasticity, which could be an indicator of unequal variance over the measurement data. The found intercept and slope of the linear regression line are used later on to perform the baseflow separation. The same procedure is executed for all catchments. Baseflow was assumed to depend (approximately) linearly on storage (section 3.2.2), which is true for $z = 1$. This is in accordance with the found slopes, which vary between 0.948 (Wollefsbach) and 0.989 (Weierbach). Similar to the results in Figure 4.1, the R^2 values found are high, ranging from 0.993 (Lümpenenbach) to 0.999 (Severn). The found parameters are displayed in 4.2 for all catchments.

In literature, the point selection was often done based on a coarser time scale (e.g. daily or multiple hours). As a result of the requirement that the discharge needs to recede continuously for 72 time steps, not that much data points were selected. This raises the question whether the limited amount of selected data points represents the catchments low-flow behaviour properly, or whether the data pool is too small to represent all data.

Discharge Prediction

Figure 4.2a shows the precipitation associated with deep groundwater flow, calculated as $BFI * P$. This precipitation was used to model baseflow, which can be seen in Figure 4.2b. The smoothed baseflow prediction shows a quite accurate resemblance with the observed baseflow (the results from baseflow separation are assumed to be the 'true' deep groundwater flow), whereas the non-smoothed prediction shows too peaky behaviour.

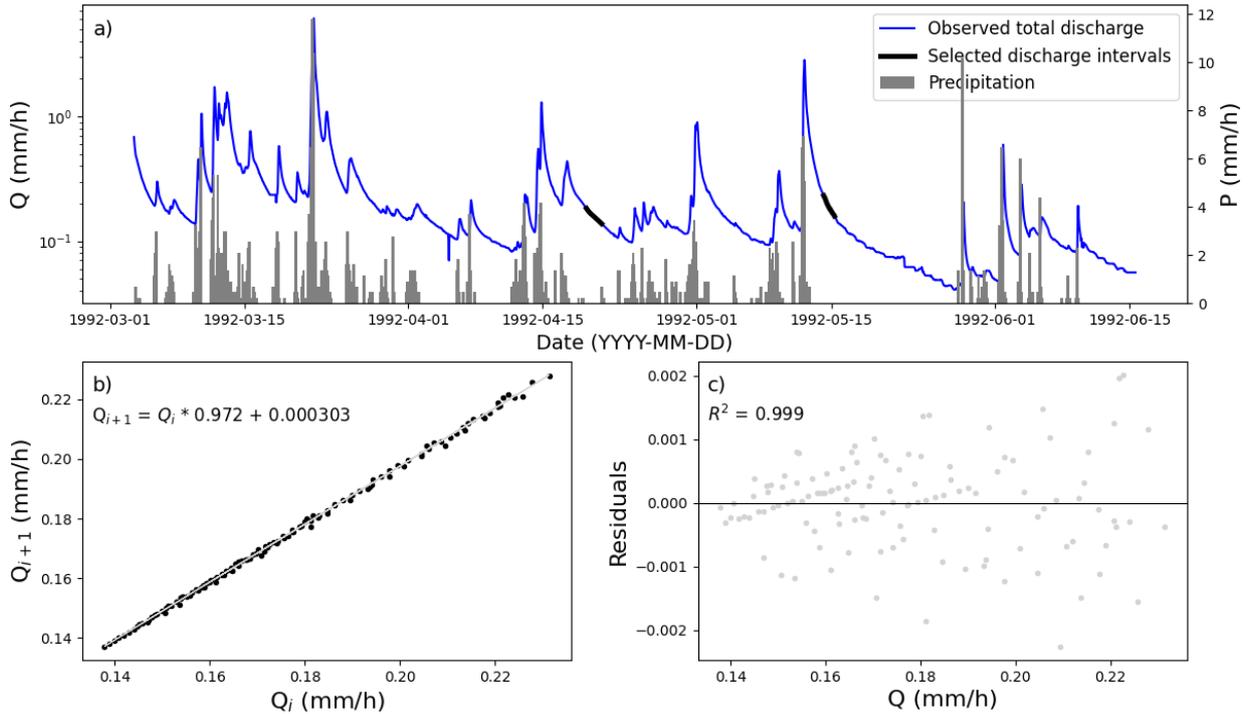


Figure 4.1: Plots showing baseflow separation steps and parameter determination for the Severn catchment. a) Semi-log plot of total discharge, indicating the selected intervals for which a discharge recession of a minimum of three days is true. On the right axis total precipitation. b) Selected discharge time steps with time step i on x-axis and $i+1$ on the y-axis (black), and the best fit line found using least-squares linear regression (grey). c) Residuals from the best fit line (grey), plotted against discharge.

Observed baseflow was obtained from the baseflow separation, and is assumed to be the true value of baseflow. The accurate resemblance is also reflected in the NSE, which is 0.96 (Table 4.2). In log-space (Figure 4.2c), which is a useful tool to investigate the prediction quality during low flows, it can be seen that the prediction sometimes over- or underestimates the peak and low flows. The timing of precipitation and discharge peaks only show a small deviation, indicating a small time lag. This time lag is implemented in the modeled baseflow, which therefore shows discharge peaks at the same time as the observed discharge.

The power output for the higher frequencies (larger than 0.1 cycles per day) overestimates the output for the observed baseflow. This also shows in the discharge time series, where the modeled baseflow is showing more short term variation than the observed baseflow. The quicker response of modeled baseflow as compared to the observed baseflow also shows in the rising limb density (RLD). RLD is a value that reflects on the travel times in a catchment by comparing the number of discharge peaks and the average time it takes to reach such a peak. For the Severn catchment, the RLD of the observed baseflow is 0.0620, whereas that of the modeled

baseflow is 0.285, which is more than a factor four difference. The smoothed modeled baseflow, on the other hand, has a RLD value much closer to the observed baseflow, showing the importance of smoothing.

For the other catchments, the model performance (looking at NSE) is less accurate than it is for the Severn catchment. However, for most catchments the model performance can be considered satisfactory or even good, except for Rappengraben and Huewelerbach, which both have a NSE below 0.36, indicating unsatisfactory model performance (Table 4.2).

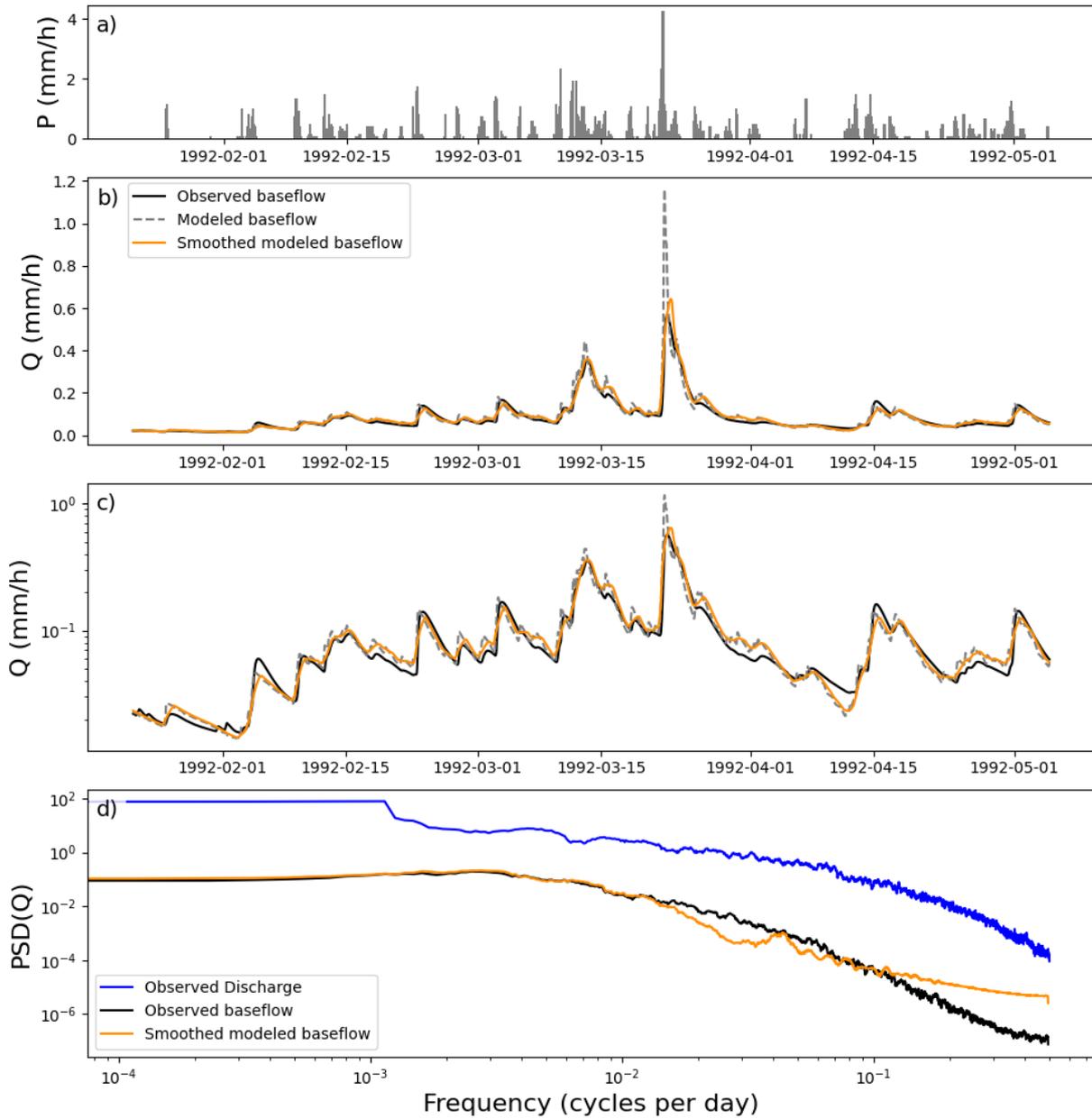


Figure 4.2: Baseflow prediction for the Severn catchment for the year 1992. a) Precipitation used for baseflow prediction over time (with the date given as (YYYY-MM-DD)). b) Observed baseflow and (smoothed) modeled baseflow plotted over time. c) Observed and (smoothed) modeled baseflow plotted over time on semi-log axis. d) Power spectra of the observed discharge, observed baseflow and smoothed modeled baseflow.

4.2.3 Shallow Groundwater Flow

Parameter Identification

After the parameters for the deep groundwater flow have been determined and the baseflow is predicted, the residual P, ET and Q data are calculated as the difference between the total data and the data used for deep groundwater prediction. The residual discharge is assumed to be the true value of shallow groundwater flow, and will be used to evaluate model predictions. Firstly, discharge intervals are selected for which the discharge is ten times larger than both evapotranspiration and precipitation. These intervals are highlighted in Figure 4.3a, which displays a time series of total discharge and precipitation.

For the selected intervals flow recession rates are calculated and plotted as a function of Q (Figure 4.3b). Similar to the analysis by Kirchner (2009), a significant variation of the scatter can be seen for smaller values of Q. This can be caused by multiple factors, such as a measurement noise. As the variation of data points is larger for smaller values of Q, it is also an indication of heteroscedasticity, which could indicate underlying problems in the data causing the unequal variance over Q.

Negative or zero values for $-dQ/dt$ cannot be plotted on log-axis. Even though these values do not correspond with a discharge recession, the values could still reflect variations in Q around the average recession trends, such as daily fluctuations. It can therefore not be justified to excluded all negatives and zeros from the determination of $g(Q)$, as the positive variations around the recession trend are not excluded, which could give a biased estimation of $g(Q)$. Therefore, the data binning has been done based on all data points, so the negative $-dQ/dt$ are not neglected. It is therefore assumed that the estimation of $g(Q)$ is not biased towards the positive variations.

The binned data are shown in Figure 4.3b, c (logarithm of the binned values) and d (regular axis). Smooth curves are fitted through the binned data using least squares non-linear regression. In Figure 4.3c a second order polynomial is fitted and in Figure 4.3d a power-law is fitted. The residuals of the fitted power-law are plotted in Figure 4.3e. The residuals show somewhat sinusoidal behaviour, especially for lower Q, indicating that the found regression line might not adequately reflect the relationship between $-dQ/dt$ and Q. There can be many factors causing this behaviour, among which an imperfect measurement method or scaling errors for laboratory derived Q-h relationships. Additionally, the ob-

served sinusoidal behaviour could be the result of daily cycles, for example that of evapotranspiration. As the residuals only deviate a few percent from the discharge values and the found R^2 value is high (0.958), the observed behaviour was assumed not to have a significant impact on the results. The same sinusoidal tendency was observed for other catchments, and R^2 values varied between 0.611 (Rappengraben) and 0.911 (Huwelerbach).

The found parameters are displayed in Table 4.2. The table shows β values ranging from 1.75 (Weierbach) to 3.25 (Sperbelgraben), indicating non-linear and storage dependent behaviour, as was expected based on literature (Kirchner, 2009; Teuling et al., 2010; Brauer et al., 2013; Buitink et al., 2020).

Discharge Prediction

Shallow groundwater based discharge was predicted using the residual precipitation (shown as a time series in Figure 4.4a) and residual evapotranspiration. Figure 4.4b shows the modeled shallow groundwater based discharge together with the observed shallow groundwater based discharge as a function of time. It can be noticed that the discharge peaks follow precipitation. The timing of the modeled discharge and the observed discharge are in accordance, however, discharge peaks are often underestimated in the model prediction. This also shows in Figure 4.4c, which again shows a time series of modeled and observed shallow groundwater based discharge, but now on a logarithmic axis. Besides, in this plot it can also be seen that the low flows are underestimated, especially more towards spring times (right side of the plot). Still, the modeled discharge is a good approximation of the observed discharge, as shows from the found NSE value of 0.94.

The power spectrum of the modeled and observed discharge (Figure 4.4d) show very similar patterns, with a slight underestimation of the power output for frequencies between 0.01 and 0.4 cycles per day. The summation of deep and shallow groundwater based discharge gives an approximation of the total discharge. For the Severn catchment this approximation is close to the total discharge, as shows in the found NSE value of the summed discharge, which is 0.94 (Table 4.2).

The discharge prediction is not as accurate for all catchments. Seven out of thirteen catchments show a NSE lower than 0.6 for the modeled shallow groundwater based discharge, of which five catchments even show unsatisfactory model performance. However, there are

Table 4.2: Found BFI values and parameters after calibrating the model, lag times of both deep and shallow groundwater based discharge ($T_{lag,gw}$ and $T_{lag,s}$, respectively) in hours. NSE values calculated separately for baseflow (NSE_{bf}) and shallow groundwater based discharge (NSE_s), and for the summation of the two (NSE_{sum}).

Catchment	BFI	$T_{lag,gw}$ (h)	z	k	NSE_{bf}	$T_{lag,s}$ (h)	a	β	NSE_s	NSE_{sum}
Noor	0.91	900	0.988	0.0683	0.637	1	1.23	2.33	0.611	0.691
Hupsel Brook	0.30	32	0.984	0.165	0.882	3	0.120	2.07	0.780	0.792
Severn	0.36	23	0.972	0.115	0.958	2	0.151	1.93	0.937	0.944
Wye	0.30	23	0.977	0.0775	0.911	2	0.194	2.07	0.943	0.954
Rietholzbach	0.05	28	0.978	0.620	0.741	1	0.0921	1.92	0.661	0.681
Erlenbach	0.11	5	0.982	0.274	0.765	1	0.217	2.15	0.728	0.738
Lümpenbach	0.18	5	0.951	0.185	0.732	0	0.224	2.27	0.506	0.512
Vogelbach	0.14	5	0.955	0.212	0.570	0	0.0909	1.81	0.507	0.551
Sperbelgraben	0.07	12	0.993	0.304	0.535	0	0.461	3.25	0.275	0.319
Rappengraben	0.09	12	0.987	0.119	0.305	0	0.103	2.19	0.216	0.243
Huwelerbach	0.25	28	0.977	0.348	-9.01	0	1.22	3.10	-0.312	-2.53
Weierbach	0.38	192	0.989	0.114	0.619	0	0.0214	1.75	0.326	0.460
Wollefsbach	0.25	12	0.948	0.482	0.643	0	0.315	1.97	0.553	0.588

also some catchments for which the model performance is better, some even showing a good model performance. The NSE values are in general higher for the modeled deep groundwater based discharge. Therefore, the NSE of the sum of both predictions, as compared to the total observed discharge, is slightly higher than the NSE values for the shallow groundwater based discharge. However, still many catchments show NSE values below 0.6, and Huwelerbach even shows a negative NSE.

4.2.4 Component Prediction

After the BFI values were determined, a test was executed to check whether the amount of deep groundwater generated flow or the BFI value can be predicted based on a groundwater recharge dataset by Martinsen et al. (2022), shown in Figure B.1. In the used dataset the groundwater recharge coefficient is defined as the difference between the potential groundwater recharge and the effective precipitation. The values were read from the potential recharge groundwater map, which is shown in Figure B.2, together with the locations of the researched catchments. These fractions, however, showed no agreement with the found BFI values.

Next, the yearly potential groundwater recharge per catchment was read from Figure B.1. The found values were plotted against the modeled groundwater based discharge, as can be seen in Figure B.4. The potential recharge is often higher than the modeled baseflow, as most catchments are located below the 1:1 line. However, the Wye catchment is located almost perfectly on the line and the Noor catchment is close.

As the potential recharge is often higher than the modeled baseflow, the potential recharge was reduced based on the actual evapotranspiration (which was based on the same dataset, Figure B.3). Figure B.5 shows the modeled baseflow plotted against the potential recharge minus the actual ET times BFI, to simulate the effect of water loss by evapotranspiration on baseflow. It shows that this water reduction is way too large for the Noor catchment, and also the Severn and Wye catchment are further away from the 1:1 line than before. The Hupsel Brook and Huwelerbach catchment, on the other hand, are located on the 1:1 line. For these catchments the reduction is effective.

The figure also shows that the reduction is too little for the catchments that are still located on the right side of the line (Rietholzbach, Sperbelgraben, Rappengraben, Erlenbach and Vogelbach). Striking is that these are the catchments with the smallest found BFI values. The catchments with larger found BFI values are located more towards the left side of (or on) the 1:1 line. The outcome of the reduction raises the question whether the amount of potential recharge that is lost depends on BFI, which indirectly links to the depth of the aquifer (Hare et al., 2021). However, the test shows little resemblance between potential groundwater recharge and deep groundwater generated discharge, even after reducing with evapotranspiration, therefore not being useful in predicting the deep groundwater fraction.

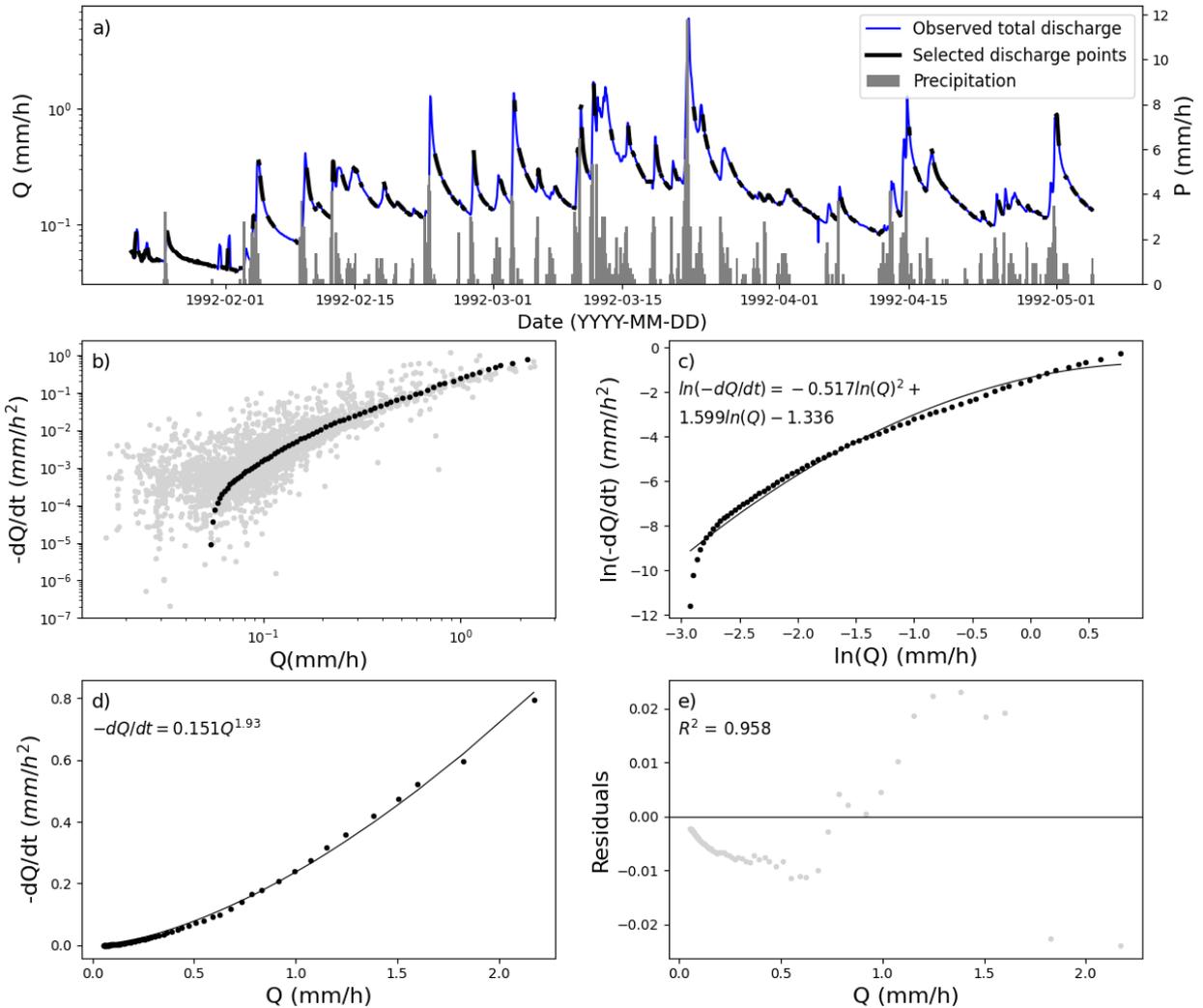


Figure 4.3: Plots showing steps taken in parameter determination for shallow groundwater flow in the Severn catchment. a) semi-log plot of total discharge, indicating the selected discharge intervals for which $Q > 10 \cdot P$ and $Q > 10 \cdot ET$. With on the right axis precipitation on a regular scale. b) Flow recession rates ($-dQ/dt$) plotted over Q for the selected discharge points (grey), binned as described in section 3.2.3 (black). c) Natural logarithm of the binned averages (black) and the associated best fit line (grey) calculated using least-squares regression. d) Binned averages (black) and the associated best fit line (grey) calculated using least-squares regression. e) Residuals from the best fit line (d), plotted over discharge Q .

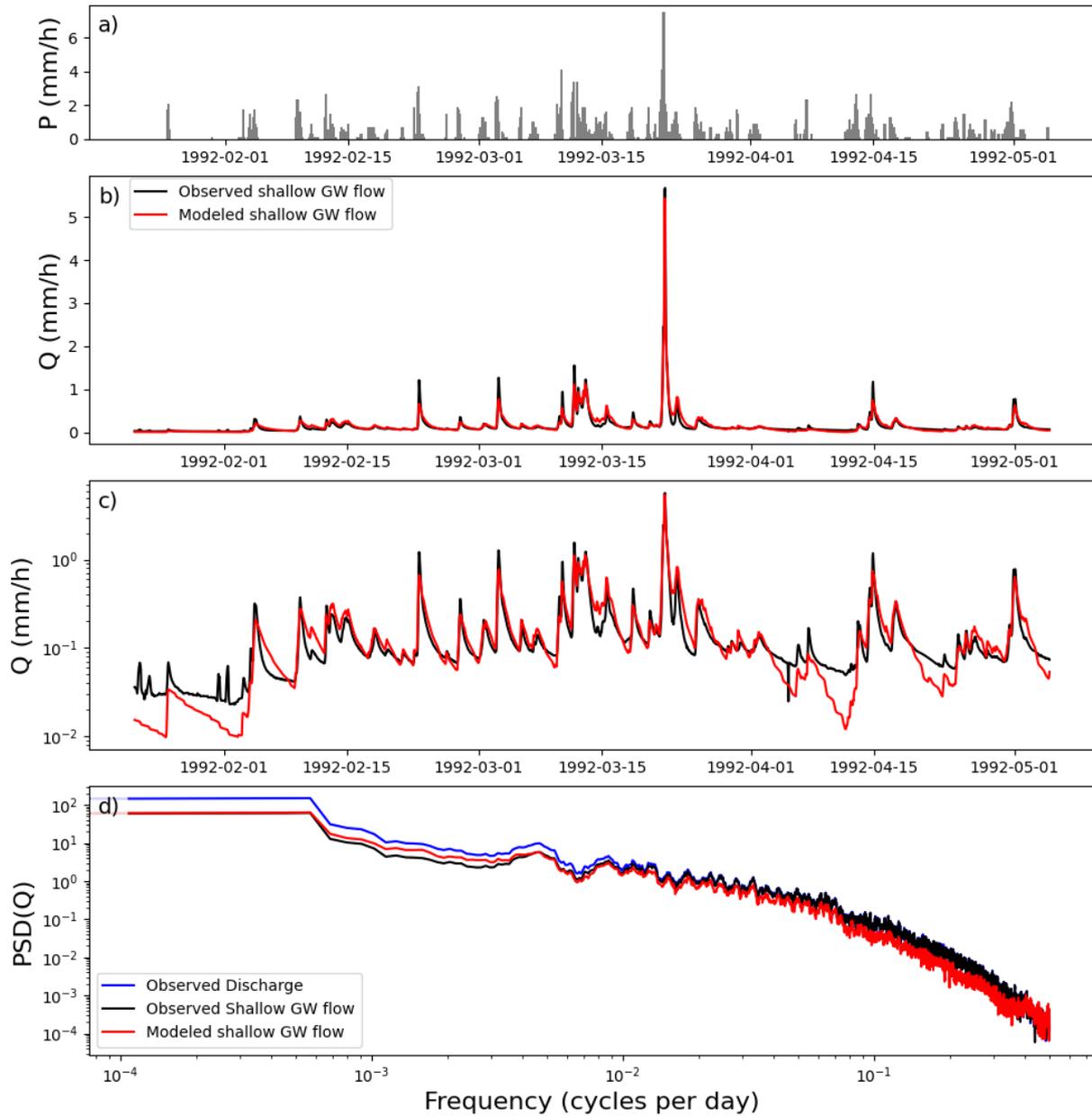


Figure 4.4: Shallow groundwater flow prediction for the Severn catchment for the year 1992. a) Precipitation used for shallow groundwater flow prediction over time. b) Observed shallow groundwater flow, modeled shallow groundwater flow plotted over time. c) Observed, modeled shallow groundwater flow plotted over time on semi-log axis. d) Power spectra of the Observed discharge, observed shallow groundwater flow and modeled shallow groundwater flow.

4.2.5 Overland Flow

Parameter Identification

The last step in the development of the model is to predict overland flow occurrence. The summation of the modeled deep and shallow groundwater based discharge is shown as a time series in Figure 4.5a, together with the total observed discharge and total precipitation. It can be seen that the timing of the discharge peaks is in accordance, although sometimes underestimated by the model. Therefore, parameters are determined to model overland flow. Figure 4.5b shows the difference between observed and modeled discharge normalized by the precipitation as a function of discharge.

Significant scatter can be observed over the whole range of Q , and therefore the data was binned into ranges of Q . The natural logarithms of the binned data points are shown in Figure 4.5c, as a function of the natural logarithm of Q . Through the binned data, a line is fitted as determined using least-squares regression analysis. The residuals of the line in Figure 4.5d are shown in Figure 4.5e, again showing a somewhat sinusoidal pattern, but with a R^2 value of 0.874. The sinusoidal pattern is expected to result from the same reasons as mentioned in Section 4.2.3. Similar, the residuals are small, deviating only a few percent from the the difference between observed and modeled Q divided by P . This, together with the high R^2 , leads to the assumption that the sinusoidal behaviour does not have a significant impact on the results.

Discharge Prediction

Consequently, overland flow is modeled as a fraction of the precipitation, where the fraction varies as a result of a variable source area, entailing that saturated areas are the main contributors to overland flow. Therefore, the variable source area is represented by the catchment wetness, which is again determined based on discharge. This means that for a higher discharge, the fraction of precipitation that becomes overland flow is larger. This could also be seen in Figure 4.5d, in which the data showed an upward trend for increasing Q .

Figure 4.6a shows the residual precipitation, calculated as the difference between observed precipitation and modeled discharge. In Figure 4.6b the modeled overland flow is shown over time, calculated as the residual precipitation times the fraction (f_{OF}), which was calculated per time step based on discharge. Remarkable is

that the values are low, which results from low values for the determined fractions. For other catchments, which showed a larger difference between the modeled and observed discharge, there is more overland flow modeled.

The total observed discharge is shown in Figure 4.6c, together with the total modeled discharge (summation of the three components). It shows that the peak in March 1992 is approached more closely than it was before, due to the addition of the overland flow. Finally, the power spectrum of the two are shown (Figure 4.6d), showing a similar pattern as was seen for the shallow groundwater based flow. For the Severn catchment, no significant improvement could be seen after adding overland flow, mainly due to the small values of overland flow. For catchments in which the overland flow fraction played a more important role (Sperbelgraben, Rappengraben and Huewelerbach), the power spectrum did show an improved resemblance for the higher frequencies.

As a result of the added overland flow, the modeled discharge shows a more peaky pattern than was observed, which decreases the NSE values as compared to the NSE values found for the summation of the deep and shallow groundwater based discharge (for most catchments). On the other hand, the summation of the modeled discharge did approach the summation of the observed total discharge more closely, thus decreasing the bias. The found NSE values and model performance are discussed more elaborately in section 4.3.1.

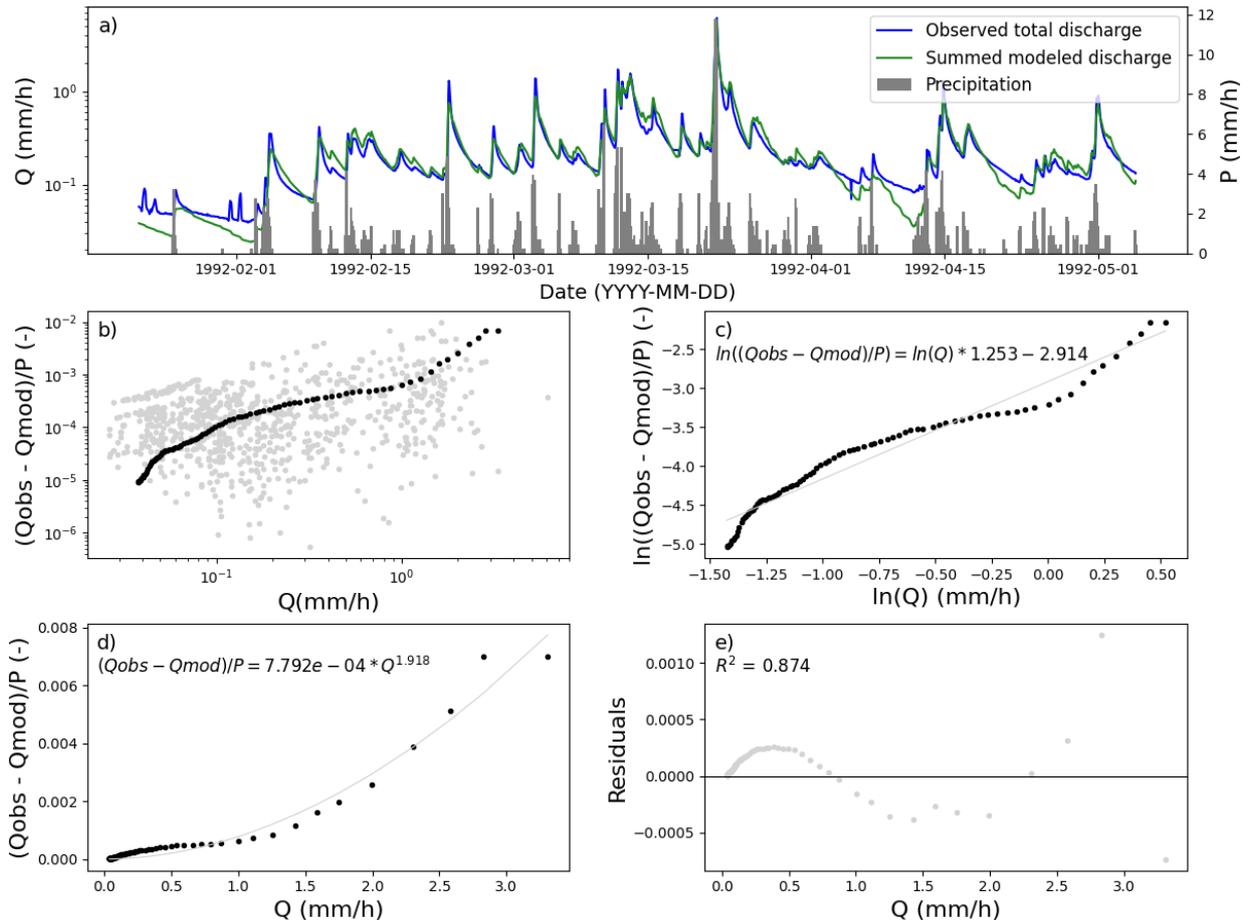


Figure 4.5: Plots showing steps taken in parameter determination for overland flow in the Severn catchment. a) Plot of total discharge and the modeled discharge (sum of deep and shallow groundwater based discharge). b) Difference in observed and modeled discharge, normalized by precipitation plotted over Q (grey), binned as described in section 3.17 (black). c) Natural logarithm of the binned averages (black) and the associated best fit line (grey) calculated using least-squares regression. d) Binned averages (black) and the associated best fit line (grey) calculated using least-squares regression. e) Residuals from the best fit lines, plotted over discharge Q .

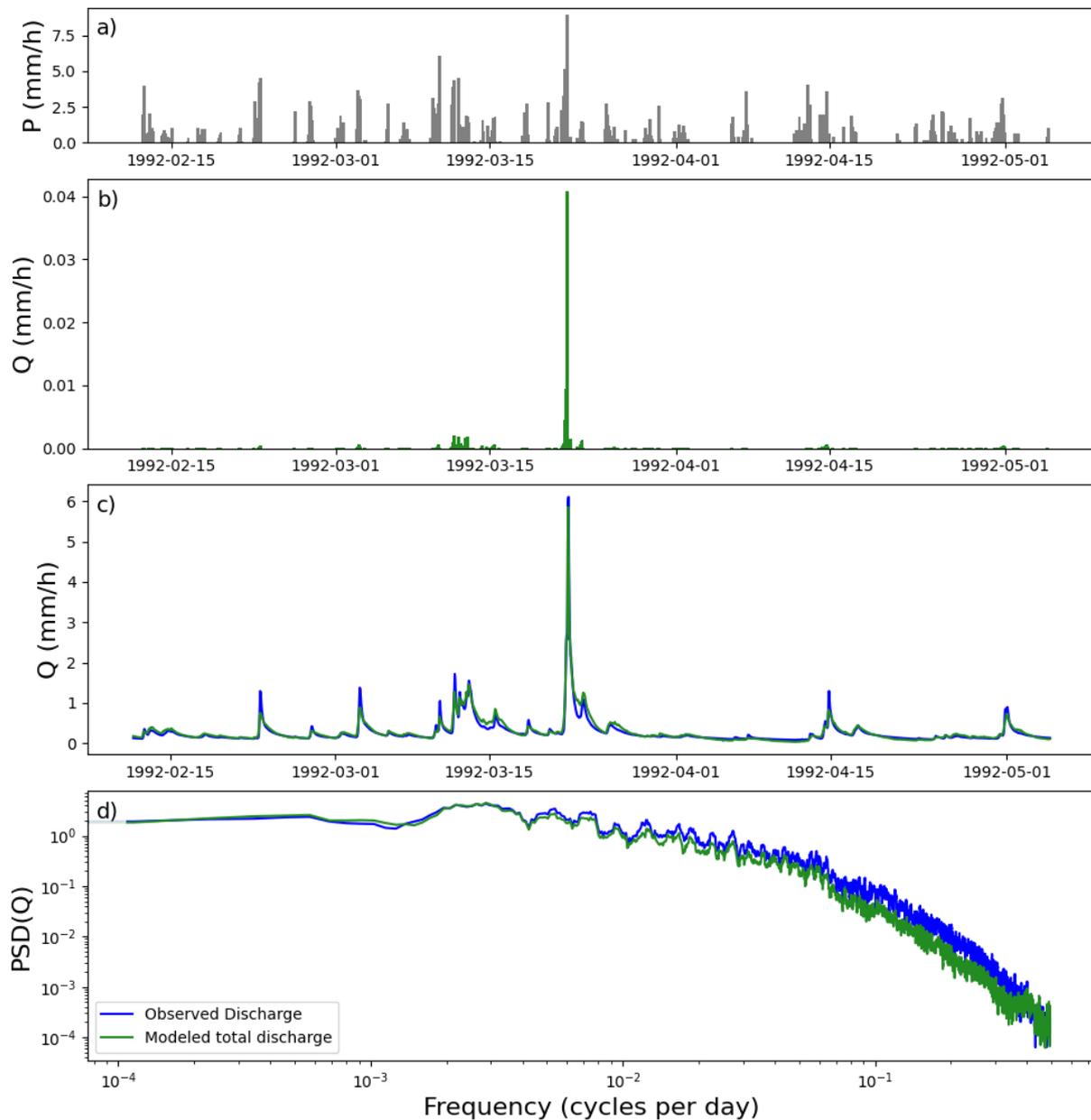


Figure 4.6: Overland flow and total discharge prediction for the Severn catchment for the year 1992. a) Precipitation used for overland flow prediction over time. b) Modeled overland flow over time. c) Observed and modeled discharge over time. d) Power spectra of the observed discharge (blue) and modeled total discharge (green).

4.2.6 Ternary Plot

Yearly Average

Figure B.6 in the Appendix shows a ternary plot of the yearly average distribution of the discharge components for the different catchments. Remarkable is that all catchments are situated towards the right side of the plot, indicating a small amount of overland flow. This is because the average was calculated over all time steps in a year. As overland flow occurs only when precipitation occurs, the average contribution of overland flow will be low. Figure B.7 shows the distribution of the discharge components when precipitation is larger than zero. The catchments are still predominantly situated towards the right side. As overland flow occurs in different situations, plots were made for different precipitation amounts ($P > \text{average } P$, $P > 2 \text{ times average } P$ and $P > 2 \text{ mm/h}$), as shown in the Appendix.

The plot for precipitation larger than 2 mm/h is displayed in Figure 4.7. It shows that Huewelerbach, Sperbelgraben and Rappengraben have the highest overland flow fraction, whereas Lümpenenbach shows the smallest fraction even for a precipitation amount above 2 mm/h. In the plot it can also be seen that the Noor catchment is the only catchment dominated by groundwater. This is according to the expectations, as little response to rainfall was observed in its hydrograph.

The discharge compositions of the Rietholzbach, Vogelbach and Wollefsbach are similar to the situation where precipitation is larger than 2 mm/h, whereas they are located further away from each other when the precipitation amount is smaller, resulting from their difference in BFI. The fraction of deep groundwater based discharge has decreased significantly for the Vogelbach and Wollefsbach catchment with increasing precipitation.

Additionally, the overland flow fraction for the Erlenbach and Lümpenenbach catchments remains below 0.15, even for precipitation amounts larger than 2 mm/h. Contradicting, in the Alptal catchments overland flow was said to be observed often, except for the forested areas (van Meerveld et al., 2018). The difference could be explained by an underestimation of the forest fraction, or maybe observed overland flow only occurred in small quantities, therefore contributing little to the total discharge. The overland flow fraction for the Huewelerbach catchment is a bit larger. However, as a result of the low model performance, the correctness of this distribution is questionable.

Sperbelgraben and Rappengraben showed the largest overland flow fractions in the ternary plot. The discharge was expected to respond quickly to rainfall, which is according to the results; little contribution by deep groundwater. The discharge in the Sperbelgraben catchment was expected to respond slower with a larger contribution of baseflow, resulting from the higher forested area (Stähli et al., 2011). Baseflow in the Sperbelgraben catchment was not found to be higher, but the contribution by overland flow is lower, which is in line with expectations based on literature (Roessel, 1950).

In the ternary plots with all time steps and all time steps for which $P > 0$, the catchments Erlenbach, Sperbelgraben and Rappengraben are located close together. In Figure 4.7, however, the catchments do not show the same composition, as the fraction of shallow groundwater based discharge and overland flow vary. This shows the difference in sensitivity to overland flow formation as a result of rainfall per catchment.

Taking into account all time steps, the composition of discharge fractions for Severn and Weierbach and for Wye and Hupsel Brook are nearly identical, whereas this is not the case when taking into account only those time steps where $P > 0$. The Severn and Wye catchment only show a slight increase in the overland flow fraction as a result of precipitation. The Hupsel Brook catchment shows a little more overland flow, but not as much as the Weierbach.

Yearly Variation

The distribution of the discharge components are plotted as a function of time (hourly time step, within one year) per catchment. This is done for all time steps and for the time steps for which precipitation is not zero. In the Appendix all plots can be found. The figures showing only those time steps for which precipitation is not zero, display fewer time steps, but to keep the figures comparable, the time range is similar. It was noticed that the plots with all time steps show very similar patterns as the plots for $P > 0$.

In this chapter, some plots are discussed more elaborately. Figure 4.8 shows the component distribution of the Noor catchment. Compared to the other catchments, the compositions of discharge fractions for the Noor are located more towards the top corner. This is due to the high contribution of baseflow to the total flow. The plot also shows that the fraction of deep groundwater based discharge is smallest at the beginning of the

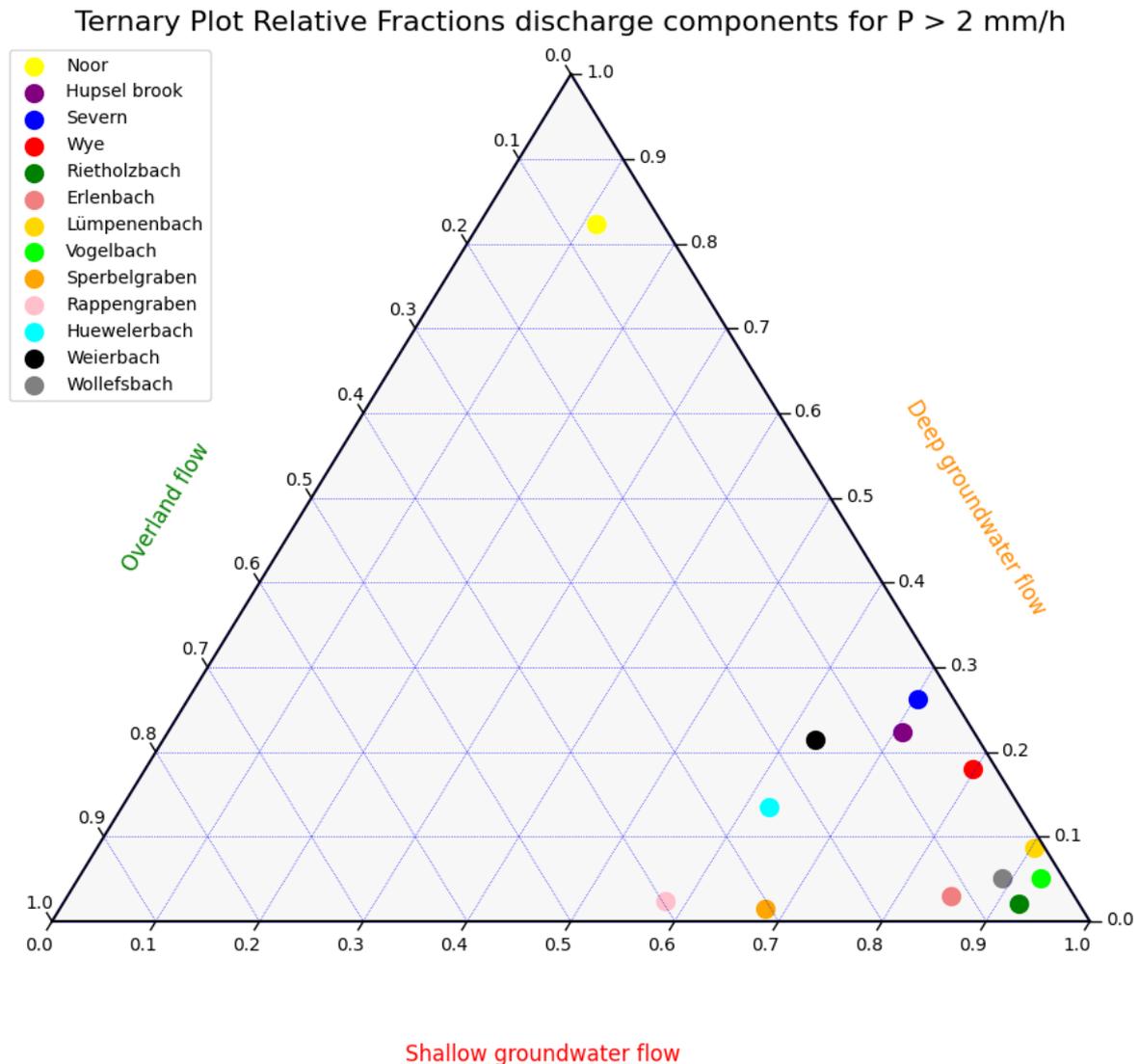


Figure 4.7: Ternary plot of the yearly average discharge components of the studied catchments, calculated based on all time steps where precipitation is larger than 2 mm/h.

year (January) and increases towards the end of the year.

The Lümpenenbach catchment shows a larger spread of compositions than most catchments (Figure 4.9). The fractions of the beginning and end of the year are located in the same corner, approximately, with a large amount of shallow groundwater based discharge. Halfway through the year, the compositions move a bit more towards the top of the plot, indicating more deep groundwater based discharge, but also overland flow occurs more.

The Huewelerbach catchment shows a remarkable pattern; the discharge fractions show a deep groundwater based discharge fraction of (almost) zero throughout most of the year, with the exception of the first 2000 hours (about 2.5 months) of the year (Figure 4.10). Dur-

ing this time of year the deep groundwater fraction is higher, reaching up to almost 0.8.

Lastly, the yearly variation for the Hupsel Brook catchment is shown in Figure 4.11, which demonstrates very little variation throughout the year, with a overland flow fraction that remains low. The plot shows some variation in the amount of deep groundwater based discharge with some time steps showing more overland flow. However, no clear seasonal pattern can be distinguished. The catchment is known to have an intense drainage network (Brauer et al., 2013), which might reduce the amount of overland flow due to the availability of fast flow routes.

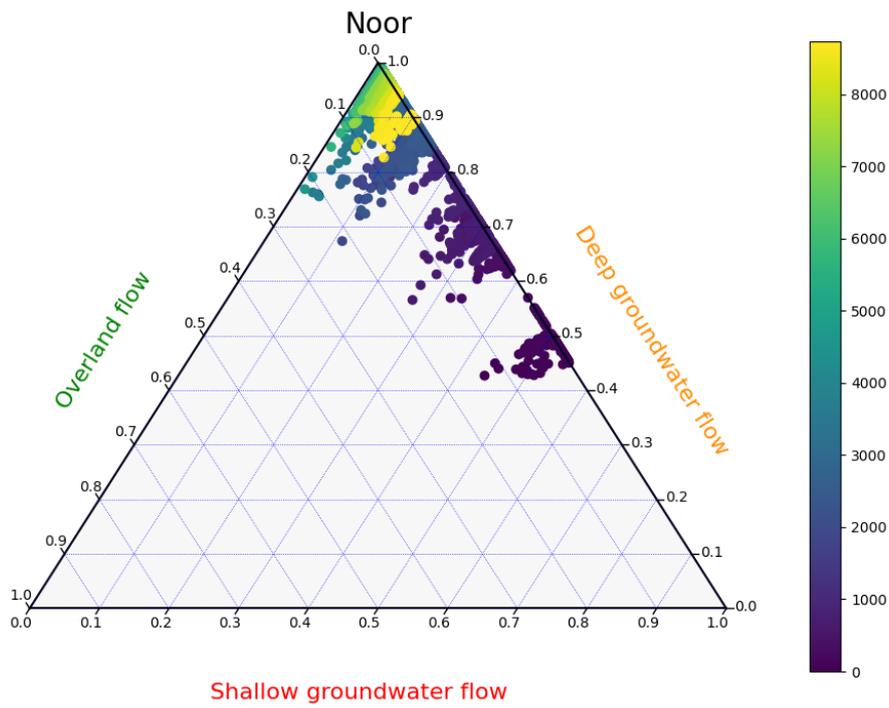


Figure 4.8: Ternary plot of the yearly variation in discharge components of the Noor catchment for the year 1994, calculated per time step where $P > 0$ mm/h. In the legend, 0 indicates the first hour of the year (January first, 01:00).

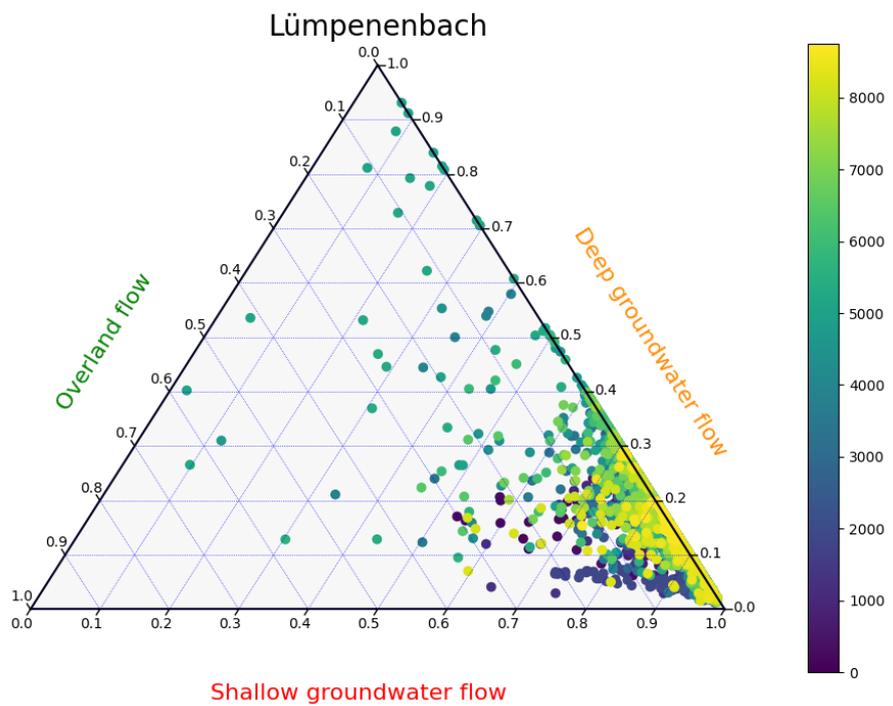


Figure 4.9: Ternary plot of the yearly variation in discharge components of the Lümpenenbach catchment for the year 1993, calculated per time step where $P > 0$ mm/h. In the legend, 0 indicates the first hour of the year (January first, 01:00).

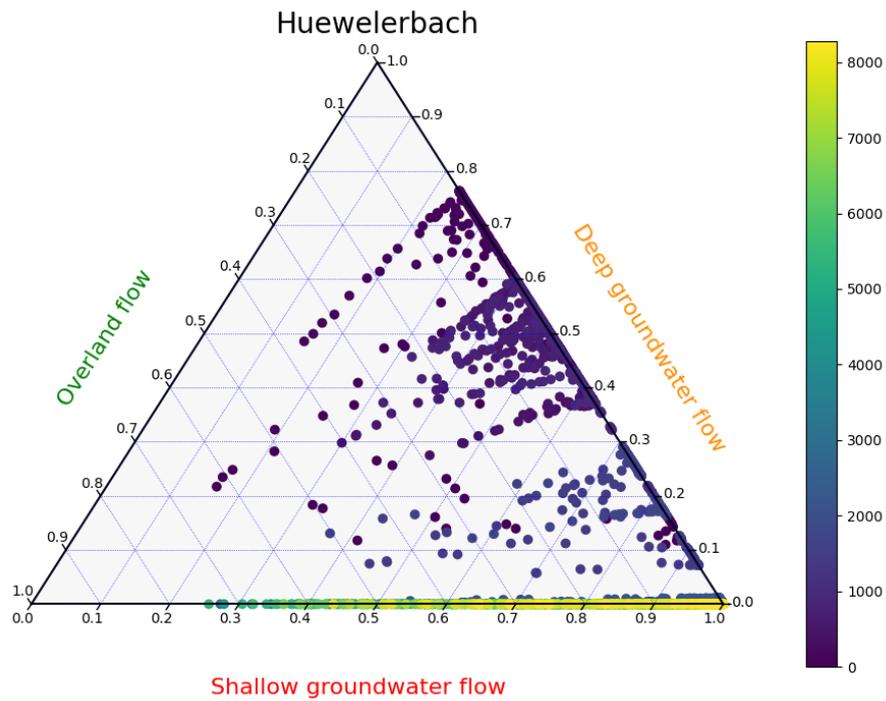


Figure 4.10: Ternary plot of the yearly variation in discharge components of the Huewelerbach catchment for the year 2003, calculated per time step where $P > 0$ mm/h. In the legend, 0 indicates the first hour of the year (January first, 01:00).

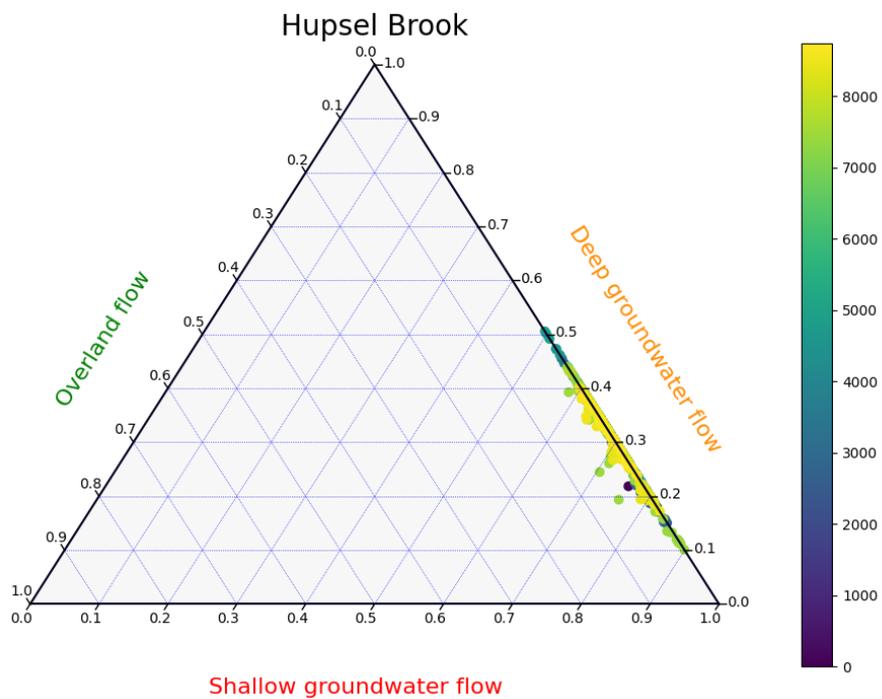


Figure 4.11: Ternary plot of the yearly variation in discharge components of the Hupsel Brook catchment for the year 1998, calculated per time step where $P > 0$ mm/h. In the legend, 0 indicates the first hour of the year (January first, 01:00).

4.3 Model Quality Assessment

4.3.1 General Model Performance

The general model quality is assessed by looking at the NSE and KGE of the model, and additionally by looking at the similarity of the RLD of the observed and modeled discharge. Table 4.3 shows these values for all catchments. For eight catchments, the NSE values are lower for the total discharge than for the summation of the deep and shallow groundwater discharge, with the exception of the Erlenbach, Vogelbach, Rappengraben, Huewelerbach and Wollefsbach catchment. For most catchments, the found KGE value can be considered satisfactory or even good. However, the model performance of Sperbelgraben, Huewelerbach and Weierbach is poor, as the KGE values remain below 0.4.

Looking at the rising limb density (RLD) of the model, it can be seen that for some catchments the modeled discharge simulates this discharge metric accurately. This can be seen from RLD that are similar to the RLD of the observed discharge. For example for the Erlenbach catchment, the RLD of the observed and modeled discharge are similar. However, for the Noor catchment the RLD is overestimated, indicating a modeled discharge which has a shorter rising limb than the observed discharge. This results from the model which shows a quicker change in discharge after a precipitation event than is observed. For the Wollefsbach catchment, the opposite is the case.

4.3.2 Comparison to Other Models

Discharge is predicted for all three models (the new model, SDS approach and HBV). Figure 4.12 shows the discharge prediction and the observed discharge on both normal (Figure 4.12b) and log scale (Figure 4.12c) based on the total observed precipitation, which is shown in Figure 4.12a. It can be noted that the discharge peaks are often underestimated for all three models. The HBV model shows most peaky behaviour, also when there are no discharge peaks in the observed discharge data. In the semi-log plot it can be seen that the HBV model often underestimates the discharge whereas this happens less often (and less extreme) for the other two models. The peaky behaviour is also reflected in the RLD (Table 4.3), which is highest for the HBV model.

The power spectra of the different models are shown in Figure 4.12d. It can be seen that the output for the HBV model is larger than the other two models and the

observed discharge, resulting from the quickly changing discharge behaviour. The new model and the SDS approach show very similar patterns, except for the lower frequencies, where the SDS approach gives a higher output, which is a closer resemblance to the observed discharge.

In Table 4.3 the RLD values are shown for the observed and modeled discharge. The RLD showing the closest resemblance to the observed discharge is marked in green, whereas the value which is farthest away from the observations is marked red. Looking at these colors, it can be seen that the RLD of the HBV models are most often marked red, usually due to an overestimation, indicating a rising limb which responds more quickly to precipitation than that the observed data does. The new model and SDS approach both have some green marked values and some red ones. The red values differ between over and underestimating the RLD.

The table also shows the found NSE and KGE values, again marking the best values in green and the worst in red. It can be noted that the new model shows most green values for the first five models, whereas the HBV model performs least well for these catchments. For the following four catchments however, the HBV model outperforms both the new model and the SDS approach. Striking is that both the SDS approach and HBV model have negative NSE and KGE values for the Noor catchment, but that the model performance for the new model is reasonable. For the Huewelerbach, the new and HBV model show bad model performance. The model performance for the SDS approach is better, but still not satisfactory.

Boxplots of the NSE and KGE values are shown in Figure 4.13, to display the general model performance. In the top row boxplots are shown for the NSE values. The left panels in Figure 4.13 provide an overview of the general model performance, taking into account all catchments. The quality of the forcing data is expected to affect the model performance. Therefore, separate plots were made for catchments forced with local data (middle panel) and catchments forced with ERA5 precipitation data (right panel).

In the NSE and KGE values for all catchments, little differences are seen between the median values. The median of the HBV model is slightly higher than it is for the other two models. The median value for the new model is lowest, but the first quartile is higher than it is for the other two models, resulting from the absence of very low model performance. Using only the catchments that are not forced with ERA5

precipitation data (middle panel), it can be seen that the median values are higher than before. The new model now shows the best model performance, however only slightly higher than the other models. Both NSE and KGE show lower median values for the ERA5-forced simulations. The performance as compared to all catchments is decreased least for the SDS approach.

Figure 4.14a shows a scatter plot of the absolute error of the different models, calculated as the modeled discharge minus the observed discharge. From this figure it can be seen that for low discharge HBV tends to both over- and underestimate the discharge and the new model mainly underestimates the discharge, which is in agreement with Figure 4.12.

Additionally, in Figure 4.14b the flow duration curves (FDC) of the observed and modeled discharges are seen. As the curves were very similar, the shown plot is zoomed in to better visualize the differences. The original FDC can be seen in Figure B.16 in the Appendix. Figure 4.14b shows that the SDS approach and HBV model almost perfectly follow the curve of the observed discharge, whereas the new model shows a higher exceedance probability than the observed discharge. This indicates that for the new model, higher flow rates have a (slightly) higher probability of occurring.

Figures 4.14c-g show scatter plots of the discharge per time step for the different models and observations, with in black a 1:1 line. In the plots it can be seen that the SDS approach and new model have the most similar modeled discharges, as they are close to the line (Figure 4.14f) and show similar patterns when plotted over the observed discharge (Figure 4.14c and d). The HBV model shows most variation around the 1:1 line (Figure 4.14e) and is less comparable to the new model (Figure 4.14g).

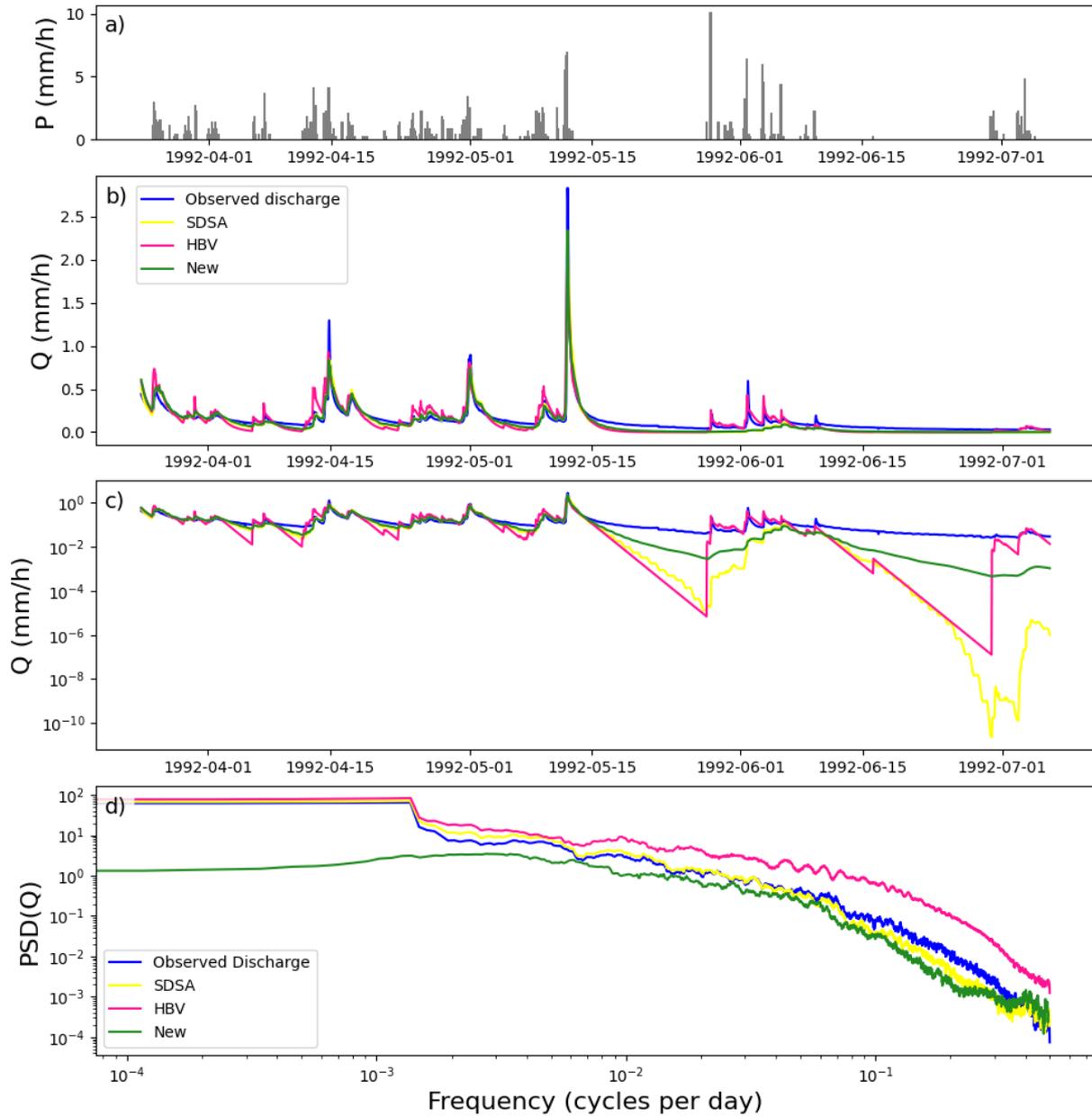


Figure 4.12: Comparison of discharge prediction for the Severn catchment using three different models: the new model, simple dynamical systems approach (sdsa) and HBV with a) total precipitation as a function of time. b) Time series of observed and modeled discharge. c) Time series of observed and modeled discharge on semi-log axis. d) Power spectra of the observed and modeled total discharge.

Table 4.3: Rising limb density (RLD) and model performance statistics (Nash-Sutcliffe efficiency; NSE and Kling-Gupta efficiency; KGE) over all researched catchments and three models, with the best (worst) performing model per RLD or statistic marked in green (red).

Catchment	Obs RLD	New			SDSA			HBV		
		RLD	NSE	KGE	RLD	NSE	KGE	RLD	NSE	KGE
Noor	0.0486	0.247	0.685	0.786	0.176	-8.23	-0.468	0.202	-2.30	-0.930
Hupsel Brook	0.0801	0.203	0.785	0.845	0.381	0.792	0.833	0.397	0.778	0.875
Severn	0.242	0.213	0.931	0.961	0.274	0.927	0.905	0.350	0.851	0.925
Wye	0.288	0.265	0.932	0.962	0.310	0.938	0.855	0.372	0.824	0.908
Rietholzbach	0.236	0.287	0.669	0.836	0.302	0.656	0.828	0.363	0.679	0.815
Erlenbach	0.256	0.276	0.749	0.844	0.289	0.789	0.882	0.326	0.764	0.822
Lümpenenbach	0.272	0.334	0.506	0.724	0.277	0.660	0.810	0.320	0.707	0.798
Vogelbach	0.290	0.261	0.585	0.755	0.2741	0.586	0.795	0.322	0.668	0.795
Sperbelgraben	0.157	0.149	0.313	0.363	0.153	0.275	0.0501	0.173	0.379	0.452
Rappengraben	0.170	0.173	0.251	0.500	0.153	0.180	-0.0535	0.179	0.333	0.524
Huewelerbach	0.386	0.450	-1.44	0.116	0.190	0.169	0.538	0.237	-26.7	-3.55
Weierbach	0.100	0.278	0.446	0.389	0.205	0.372	0.310	0.211	0.316	0.304
Wollefsbach	0.312	0.0834	0.592	0.707	0.189	0.614	0.741	0.249	0.400	0.539

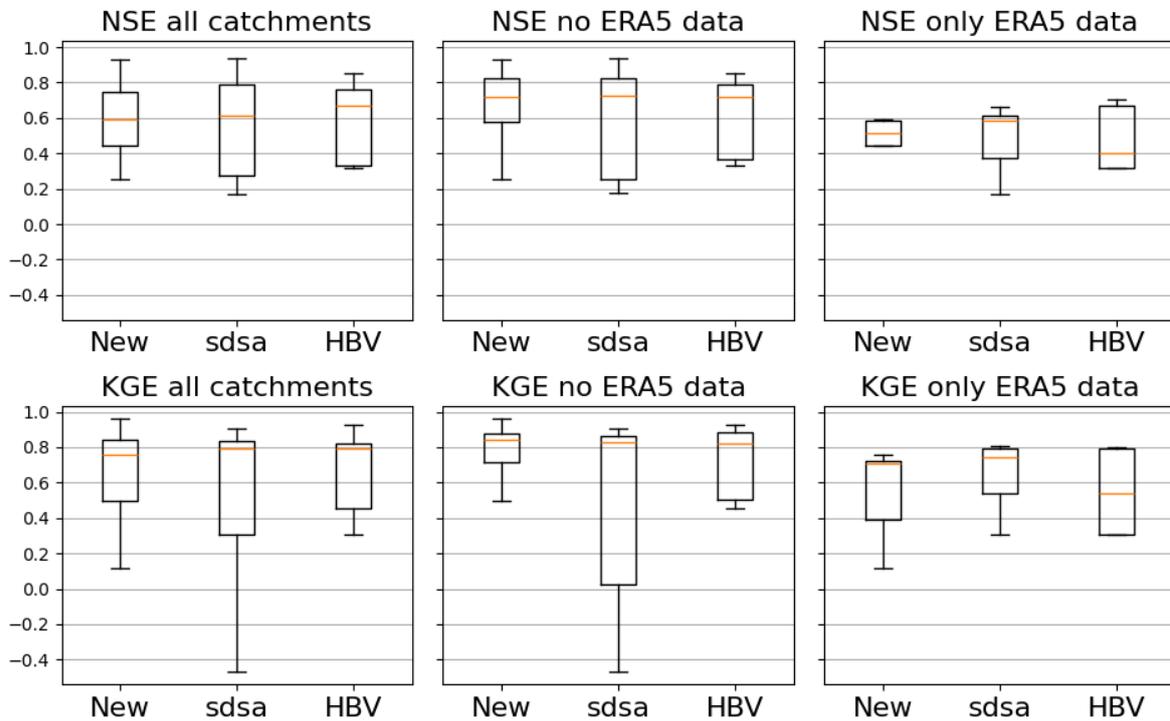


Figure 4.13: Boxplot to display general model performance (Nash-Sutcliffe efficiency; NSE in the top and Kling-Gupta efficiency; KGE in the bottom), taking into account all catchments (left panels), only the catchments with local precipitation data, i.e. excluding ERA5-forced catchments (middle panels) and only the catchments for which ERA5 precipitation data is used (right panels). General model performance is shown for the newly created model (New), the simple dynamical systems approach (sdsa) and for the HBV model.

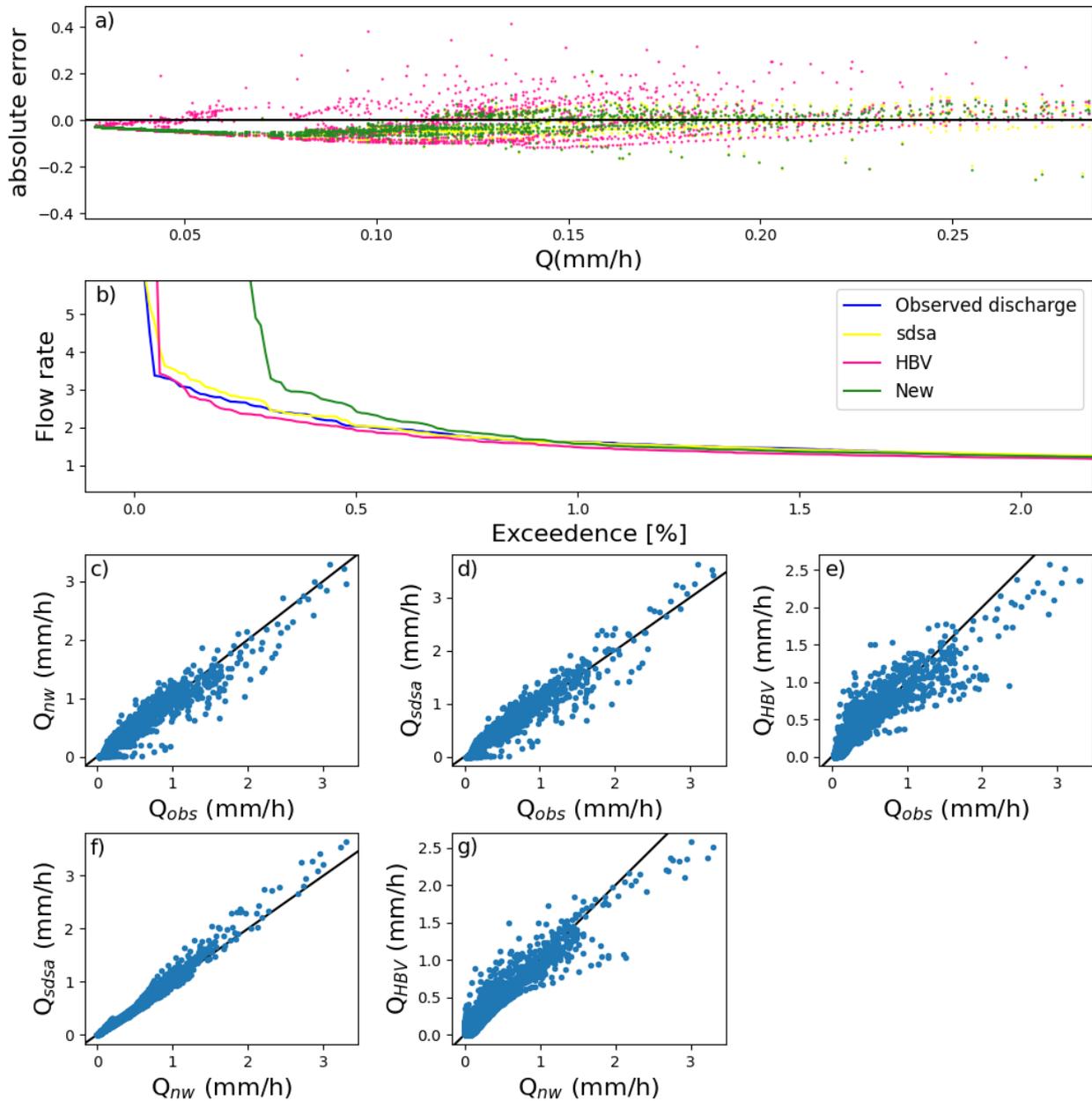


Figure 4.14: Comparison of discharge predictions for the Severn catchment using three different models: the new model, simple dynamical systems approach (sdsa) and HBV with a) absolute errors per model, calculated as the difference between modeled and observed discharge, plotted over discharge. b) Flow duration curves of the observed and modeled discharge, zoomed in to the low exceedance probabilities. c) New modeled discharge over observed discharge (blue) with 1:1 line in black. d) Sdsa modeled discharge over observed discharge (blue) with 1:1 line in black. e) HBV modeled discharge over observed discharge (blue) with 1:1 line in black. f) Sdsa modeled discharge over new modeled discharge (blue) with 1:1 line in black. g) HBV modeled discharge over new modeled discharge (blue) with 1:1 line in black.

5 | Discussion

In this study, a model was build that predicts discharge based on precipitation and evapotranspiration data using knowledge on the relative distribution of three discharge components. The aim was to develop an approach that gives high quality discharge predictions by simulating the the different flow routes from precipitation to discharge, while keeping the model simple with a limited number of parameters. As a result of its simplicity, the model has limitations, which can be linked to the original limitations of the SDS approach, limitations of a baseflow separation and to the implementation of overland flow to the model. Besides, data availability and quality can also act as a limitation factor for a model to perform well. This chapter provides an overview of these limitations and their consequences.

5.1 Model Structure

Potential evaporation data is used as input for the model, which is consequently converted using an evaporation correction factor. Additional to the evaporation correction factor, the model includes an evaporation reduction switch, that is used when discharge reaches a set minimum. This is done to prevent modeling negative discharge as a result of an overestimation of evapotranspiration. When the discharge reaches the threshold value, evaporation is set to zero, which is not according to real world processes, in which evaporation reduction depends on many factors, such as soil moisture, and does not simply work as an on/off switch.

The next step in the development of the model was to model the discharge which results from deep groundwater flow, for which a baseflow separation was executed. Consequently, the derived baseflow was used as the 'true' amount of deep groundwater flow in the catchment. The found baseflow was used to identify parameters used in the model and was used to assess the quality of the model output. However, in reality, a baseflow separation is not without uncertainties, especially when using a methodology based on Recursive Digital Filter, instead of using tracer measurements. However, Eckhardt (2008) showed that using tracer measurements and equation 3.4, which was used to determine baseflow, showed similar results when a correct filter parameter was chosen for BFI. Besides, this approach returned hydrological more plausible results than the methods HY-

SEP1, HYSEP2, HYSEP3, PART, BFLOW and UKIH (Eckhardt, 2008).

Adding to that, Nathan and McMahon (1990) proposed typical recession constants ranging from 0.93 to 0.995, which is in agreement with the parameters found during the baseflow separation (Duncan, 2019). Therefore, it is assumed that the baseflow separation executed provides reasonable estimates for the amount of baseflow in a catchment, considering that the BFI values are correct. However, as the true values for BFI are unknown, the correctness of the baseflow estimation remains unknown, even after BFI was calibrated. Tracer measurements can be at help with determining the most correct value for BFI.

A limitation of the simple dynamical systems approach, and consequently also of the newly presented model, is that the same storage-discharge relationship is assumed for both the rising and falling limb, which cannot be justified for catchments showing hysteresis. However, as the research includes the parameterization of the sensitivity function based on discharge observations, it is assumed that the sensitivity function reflects the average behaviour of the catchment. Previous studies using the same methodology showed that this did not limit the model performance (Kirchner, 2009; Teuling et al., 2010).

Multiple mechanisms exist that can cause overland flow. In the new model a fraction was determined that represents the fraction of precipitation that becomes overland flow, f_{OF} . Following the concept of a variable source area, in which the area contributing to saturation overland flow changes as a result of antecedent catchment wetness, f_{OF} was determined based on discharge, as it is assumed that discharge and storage are linked monotonically. Therefore, Dunnian overland flow is represented well by the model. Hortonian overland flow, occurring as a result of exceedance of the infiltration capacity of the soil is not explicitly implemented into the model, as it was unknown when the infiltration capacity is exceeded. Additionally, the fraction of precipitation falling directly into the stream was not added to the model. These two mechanisms could be a valuable addition to the model to represent more real world processes.

Another limitation of the overland flow modelling is that f_{OF} is determined based on the modeled discharge,

to be able to predict discharge even when observations are not available. The modeled discharge, which is the summation of deep and shallow groundwater based discharge, is assumed to be an approximation of the actual discharge. However, when the model underestimates the actual discharge, the determined fraction will also be small, whereas an overestimation will lead to a large fraction, leading to more extreme over-/underestimations.

The model performance was compared to two other models, using different calibration methods. HBV was calibrated using a Monte Carlo algorithm which finds the set of parameters resulting in the highest NSE value. The SDS approach was calibrated using the Levenberg-Marquardt method to optimize the model performance. As these are two different calibration methods, the reliability of comparing the model performance based on NSE and KGE values is questionable, especially since the NSE and KGE values are similar. Additionally, the purpose of using a model is very important in model choice and might be more important than general model performance. Therefore, no unambiguous conclusion can be drawn on which model performs best.

As Kirchner (2009) mentioned in his paper, *"any analysis is only as good as the data it is based on"*. This is also reflected by the model evaluation. The catchments which were forced with ERA5 precipitation data showed a considerably lower prediction quality (looking at NSE and KGE). This can be reflected back to the horizontal resolution of the ERA5 precipitation data, which is $0.25^\circ \times 0.25^\circ$, covering an area larger than the analyzed catchments. Due to this, the precipitation data used was equal for catchments located close to each other, while in reality those areas will show heterogeneous precipitation patterns. Therefore, the model output quality is reduced, which can also be inferred from the fact that the NSE and KGE values for all three models are considerably smaller than average.

5.2 Broader Model Discussion

The newly presented model is more complex than the SDS approach by Kirchner (2009). It includes multiple flow routes that are modeled separately and therefore has more parameters which were inferred from observations. Studies showed that the optimal number of free parameters is between three and five (Peel and McMahon, 2020). Perrin et al. (2001) demonstrated that more complex models tend to outperform the simpler ones after calibration, however, lacking stability and

therefore having a lower verification performance. Over-parameterization as a result of an inadequate model complexity compared to the available data can lead to parameter uncertainty. This is a risk for the model accuracy and prevents a model from reaching its optimal performance level. Corresponding, Grayson and Blöschl (2001) discussed the relationship between data availability and model complexity and the effect on the model's predictive performance. It was shown that to reach an optimal performance, the available data and model complexity should be adequate, and that an increased data availability does not always lead to a better performance, e.g. when the model too simple to exploit all available data. Vice versa, when the model is too complex compared to the available data, this can lead to improper parameter identification for which the correctness of the model output is questionable.

Perrin et al. (2001) stated that simple model structures often perform as well as complex ones for many purposes. This corresponds with the result that were found in this thesis, which showed similar model performance (based on NSE and KGE) for the SDS approach, HBV and the newly presented model. For the catchments forced with ERA5 data, the median NSE values found were highest for the SDS approach, which could result from the optimum between model complexity and data availability. The lower data availability has shown to stronger affect the other models, which could be due to their higher number of parameters.

The modeled overland flow fraction is really small in some catchments, resulting from a modeled discharge (sum of deep and shallow groundwater based discharge) that is already close to the total observed discharge. Additionally, some catchments showed to be dominated strongly by one process, such as the Vogelbach catchment. This leads to the higher question what the use of the newly presented model is and whether distinguishing three flow routes is necessary. The answer to this might lie in the purpose of the modelling study. As compared to the SDS approach and HBV, the new model has the advantage that it provides information on the distribution of components. When choosing a model, one should consider the *fit for purpose* concept, which entails choosing a model based on the task (required output), but also on the available input data (Peel and McMahon, 2020). Based on this, one chooses the most appropriate model. Therefore, the use of the new model lies in studies interested in the relative distribution of components throughout the year. For studies interested

in discharge predictions, one could question the necessity of using the new model for catchments dominated by one flow route.

An advantage of the SDS approach as compared to other models is that the simplicity of the sensitivity function allows for the equation to be invertible, which means that it can be used to estimate precipitation and evapotranspiration from discharge; *doing hydrology backwards* (Kirchner, 2009). Adding to that, Kirchner (2009) mentions that the approach is simple and general as no distinction is made between base and quick flow and therefore no baseflow separation is needed. Expanding the model by distinguishing multiple flow contributors reduces the simplicity of the model and thereby also complicates the possibility to infer precipitation from discharge. Additionally, the the new model requires more steps reducing its simplicity and ease of use. Moreover, the increased number of parameters leads to an increase in risk of over-parameterization which could reduce the model's accuracy.

However, a limitation of the SDS approach as mentioned by Kirchner (2009) is its simplicity, reducing the generality of the model to be used in different catchment types. This limitation is reduced by expanding the model with two discharge generating mechanisms, making the model more general in usage, for example in catchments characterized by much overland flow, that were not represented well by the SDS approach.

Additionally, Buitink et al. (2020) mentioned the lack of precipitation dispersion as a result of water retention in the SDS approach. In the paper, this was justified by saying that for small catchments, dispersion does not have a significant impact on the hydrograph. Dispersion is also not implemented explicitly in the new model, however, accounting for two separate lag times ($T_{lag,gw}$ and $T_{lag,s}$) does somewhat represent dispersion as the precipitation peak is split into two and will reach the stream at different times.

The HBV model, on the other hand, is not as simple as the SDS approach, including nine parameters, representing snow processes and an upper and lower groundwater box. This increased number of parameters provides for the opportunity to model more complex or additional processes. However, as explained before, the risk of a reduced model accuracy due to over-parameterization will be higher compared to less complex models. Again, the new model has the advantage that it explicitly provides results on the distribution of discharge components in the catchment.

5.3 Model Prospect

The catchment selection was done as such to include a variety of flow processes and including extremes, such as the Noor catchment which is unique due to its very deep groundwater tables and little response to precipitation events. However, the question remains whether the set of catchments spans a range wide enough to be representative for the whole world. Therefore, in future research, the model should be tested on catchments further away, with different characteristics such as soil type and different climatic zones. Additionally, the model should be tested using data with similar quality, to be able to make a statement on the applicability of the model to catchment types, whereas this is complicated now due to the differences in data quality and resolution.

Additionally, to test the accuracy of the found distribution of discharge components, the model outputs should be compared with tracer experiments that can provide more information on the response times of a catchment. Tracer experiments can be used to determine the value of BFI beforehand (since BFI cannot be estimated based on the groundwater recharge dataset), reducing both the calibration time needed and the structural uncertainty which exists due to the lack of confirmation on the correctness of the found BFI values.

6 | Conclusion

The model presented in this thesis is a rainfall-runoff model that takes into account three separate discharge components (deep groundwater, shallow groundwater and overland flow) to model total discharge based on precipitation and evapotranspiration data. The total discharge is calculated by summing the modeled components, where deep groundwater based discharge is modeled following a linear relationship with precipitation, shallow groundwater based discharge is modeled according to the SDS approach by Kirchner (2009), and lastly overland flow was modeled as a function of catchment wetness, represented by discharge, and precipitation.

Applied to thirteen catchments in Europe with areas in the same order of magnitude, ranging from lowland to Alpine catchments with different climates, the model has proven to be successful in determining the relative contribution of the three discharge components based on the hydrograph. Starting off with a baseflow separation, deep and shallow groundwater based discharge were modeled after which the model was calibrated on the most suitable cutoff between the two (BFI). Afterwards, the overland flow component was added to the discharge prediction to improve discharge peak modeling. After all three components had been modeled, the relative contribution of the components was calculated both as yearly average fraction (for various values of P) and as an fraction per time step (hourly), showing the course of the component distribution throughout the year.

While this new approach manages to produce reasonable discharge predictions, the model did not prove to be superior to two more conventional models (SDS approach and HBV). The general model performance (as measured by the Nash-Sutcliffe efficiency and Kling-Gupta efficiency) showed to be relatively similar to, but not better than, the other two. However, the lower quartile of the new model shows a better model performance than the lower quartile of the other two models. This is due to the absence of extreme low model performance per catchment for the new model, whereas these are present for the other two models. Nevertheless, due to the different calibration methods and the similar performance, no significant conclusion can be drawn that states which model is best. Additionally, model choice also highly depends on research purpose.

The new model has two unique strengths compared

to the SDS approach by Kirchner (2009): (1) it can both use knowledge on the distribution of discharge components to predict discharge and the model can be used to determine the distribution of discharge components based on the hydrograph; (2) as total discharge is computed as the sum of three discharge components, it provides the opportunity to be used more generally for different types of catchments. On the other hand, due to the elaboration of the model as compared to the SDS approach, model usage is less straightforward and the risk of decreased model accuracy resulting from model-uncertainty increases, due to the increased number of parameters which could cause over-parameterization when using inadequate data.

Lastly, the possibility to predict discharge components based on already existing recharge datasets to reduce calibration efforts has not proven to be successful. The found recharge fractions were not comparable to the calibrated BFI values and the modeled yearly baseflow did not show significant correspondence with the yearly potential recharge.

Overall, the newly presented model is a valuable addition to the conceptual rainfall-runoff models, as it provides the opportunity to determine the distribution of discharge components in a relatively simple and understandable manner. However, as a predictive tool, no evidence was found that the model outperforms the conceptual rainfall-runoff model HBV or the SDS approach.

Acknowledgements

First of all, I would like to thank my supervisors Ryan Teuling and Lieke Melsen, as they helped me throughout the period that I worked on my thesis. They guided me towards conceptualizing the model idea and provided feedback during our meetings. Besides, I would like to thank Claudia Brauer for being my examiner.

Without available data, I would not have been able to write this thesis. Therefore, I would like to thank Roel Dijkema for providing discharge data for the Noor catchment and for taking the time to explain me the most important findings about this catchment. Next, I would like to thank Laurent Pfister and the Luxembourg Institute of Science and Technology (LIST), Hydro-Climatological Observation network (HOST) for giving me discharge data of the Attert catchment. Thirdly, I would like to thank Manfred Stähli (WSL) for providing me the Alptal discharge data. Adding to the data availability, many regards to Jelle ten Harkel for helping me download the ERA5 data and changing it to time series.

Bibliography

- Adamovic, M., Braud, I., Branger, F., and Kirchner, J. W. (2015). Assessing the simple dynamical systems approach in a mediterranean context: application to the ardèche catchment (france). *Hydrology and Earth System Sciences*, 19(5):2427–2449.
- Arnell, N. and Gosling, S. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, 134:387–401.
- Badoux, A., Jeisy, M., Kienholz, H., Lüscher, P., Weingartner, R., Witzig, J., and Hegg, C. (2006). Influence of storm damage on the runoff generation in two sub-catchments of the sperbelgraben, swiss emmental. *European Journal of Forest Research*, 125:27–41.
- Batelaan, O., De Smedt, F., and Triest, L. (2003). Regional groundwater discharge: phreatophyte mapping, groundwater modelling and impact analysis of land-use change. *Journal of Hydrology*, 275(1):86–108.
- Braithwaite, J. J. (2007). Occam's razor: The principle of parsimony. *Academia. edu*.
- Brandt, C., Robinson, M., and Finch, J. W. (2004). Anatomy of a catchment: the relation of physical attributes of the plynlimon catchments to variations in hydrology and water status. *Hydrology and Earth System Sciences*, 8(3):345–354.
- Brauer, C., van der Velde, Y., Teuling, A., and Uijlenhoet, R. (2018). The hupsel brook catchment: Insights from five decades of lowland observations. *Vadose Zone Journal*, 17(1):180056.
- Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., and Uijlenhoet, R. (2013). Investigating storage-discharge relations in a lowland catchment using hydrograph fitting, recession analysis, and soil moisture data. *Water Resources Research*, 49(7):4257–4264.
- Brauer, C. C., Torfs, P. J. J. F., Teuling, A. J., and Uijlenhoet, R. (2014). The wageningen lowland runoff simulator (walrus): application to the hupsel brook catchment and the cabauw polder. *Hydrology and Earth System Sciences*, 18(10):4007–4028.
- Buitink, J., Melsen, L. A., Kirchner, J. W., and Teuling, A. J. (2020). A distributed simple dynamical systems approach (ds2 v1.0) for computationally efficient hydrological modelling at high spatio-temporal resolution. *Geoscientific Model Development*, 13(12):6093–6110.
- Chapman, T. (1999). A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, 13(5):701–714.
- de Lima, J. (1989). Overland flow under rainfall : some aspects related to modelling and conditioning factors. *Landbouwniversiteit Wageningen*.
- Dijksma, R. and van Lanen, H. (1998). Monitoring and modelling of springflow in the noor catchment (the netherlands). In Bucek, J., Sir, M., and Tesar, M., editors, *Catchment hydrological and biochemical processes in changing environment : conference proceedings Liblice, Czech Republic, September 22-24, 1998*, pages 21–24.
- Dijksma, R. and van Lanen, H. (2001). De afvoer van de noor (zuid-limburg): periode 1992-2000. Technical report, Landbouwniversiteit Wageningen.
- Driessen, T. L. A., Hurkmans, R. T. W. L., Terink, W., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R. (2010). The hydrological response of the ourthe catchment to climate change as modelled by the hbv model. *Hydrology and Earth System Sciences*, 14(4):651–665.
- Duncan, H. P. (2019). Baseflow separation – a practical approach. *Journal of Hydrology*, 575:308–313.
- Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, 352(1):168–173.
- Eryani, I. G. A., Jayantari, M. W., and Wijaya, I. K. (2022). Sensitivity analysis in parameter calibration of the weap model for integrated water resources management in unda watershed. *Civil Engineering and Architecture*, 10:455–469.
- Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., and Freer, J. (2014). Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrological Processes*, 28(4):2451–2467.

- Gore, J. A. and Banning, J. (2017). *Chapter 3 - Discharge Measurements and Streamflow Analysis*, pages 49–70. Academic Press, Boston, third edition edition.
- Grayson, R. and Blöschl, G. (2001). Spatial modelling of catchment dynamics. *Spatial patterns in catchment hydrology: observations and modelling*, pages 51–81.
- Hallouin, T. (2021). hydroeval: an evaluator for streamflow time series in Python.
- Hare, D. K., Helton, A. M., Johnson, Z. C., Lane, J. W., and Briggs, M. A. (2021). Continental-scale analysis of shallow and deep groundwater contributions to streams. *Nature Communications*, 12(1450).
- Hattermann, F., Krysanova, V., Wechsung, F., and Wattenbach, M. (2004). Integrating groundwater dynamics in regional hydrological modelling. *Environmental Modelling Software*, 19(11):1039–1051.
- Hissler, C., Martínez-Carreras, N., Barnich, F., Gourdol, L., Iffly, J. F., Juilleret, J., Klaus, J., and Pfister, L. (2021). The weierbach experimental catchment in luxembourg: A decade of critical zone monitoring in a temperate forest - from hydrological investigations to ecohydrological perspectives. *Hydrological Processes*, 35(5):e14140.
- Holländer, H. M., Blume, T., Bormann, H., Buytaert, W., Chirico, G. B., Exbrayat, J.-F., Gustafsson, D., Hölzel, H., Kraft, P., Stamm, C., Stoll, S., Blöschl, G., and Flühler, H. (2009). Comparative predictions of discharge from an artificial catchment (chicken creek) using sparse data. *Hydrology and Earth System Sciences*, 13(11):2069–2094.
- Kalantari, Z., Lyon, S. W., Jansson, P.-E., Stolte, J., French, H. K., Folkesson, L., and Sassner, M. (2015). Modeller subjectivity and calibration impacts on hydrological model applications: An event-based comparison for a road-adjacent catchment in south-east norway. *Science of The Total Environment*, 502:315–329.
- Kavetski, D. and Fenicia, F. (2011). Elements of a flexible approach for conceptual hydrological modeling: 2. application and experimental insights. *Water Resources Research*, 47(11).
- Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45(2).
- Kissel, M. and Schmalz, B. (2020). Comparison of baseflow separation methods in the german low mountain range. *Water*, 12(6).
- Knapp, J. L. A., von Freyberg, J., Studer, B., Kiewiet, L., and Kirchner, J. W. (2020). Concentration–discharge relationships vary among hydrological events, reflecting differences in event characteristics. *Hydrology and Earth System Sciences*, 24(5):2561–2576.
- Knoben, W. J. M., Freer, J. E., and 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.
- Kordík, J., Dijkma, R., and Lanen, H. A. v. (2002). Monitoring and the flonet/trans model as tools to characterize the nitrate distribution and transport in the noor catchment (the netherlands). *Slovak Geological Magazine*, 8(2):147–158.
- Li, H., Beldring, S., and Xu, C.-Y. (2013). Implementation and testing of routing algorithms in the distributed hydrologiska byråns vattenbalansavdelning model for mountainous catchments. *Hydrology Research*, 45(3):322–333.
- Li, S., Kang, S., Zhang, L., Zhang, J., Du, T., Tong, L., and Ding, R. (2016). Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions. *Journal of Hydrology*, 543:450–461.
- Lyne, V. and Hollick, M. (1979). *Stochastic time-variable rainfall-runoff modelling*, volume 79, pages 89–93.
- Maier, F. and van Meerveld, I. (2021). Long-term changes in runoff generation mechanisms for two proglacial areas in the swiss alps i: Overland flow. *Water Resources Research*, 57(12):e2021WR030221.
- Martínez-Carreras, N., Krein, A., Gallart, F., Iffly, J.-F., Hissler, C., Pfister, L., Hoffmann, L., and Owens, P. N. (2012). The influence of sediment sources and hydrologic events on the nutrient and metal content of fine-grained sediments (attert river basin, luxembourg). *Water, Air, & Soil Pollution*, 223:5685–5705.
- Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A. J., Mansour, M., Sallasmaa, O.,

- Pulido-Velazquez, D., Williams, N. H., Zaadnoordijk, W. J., and Stisen, S. (2022). Developing a pan-european high-resolution groundwater recharge map – combining satellite data and national survey data using machine learning. *Science of The Total Environment*, 822:153464.
- Mathai, J. and Mujumdar, P. P. (2022). Use of stream-flow indices to identify the catchment drivers of hydrographs. *Hydrology and Earth System Sciences*, 26(8):2019–2033.
- McMillan, H. K. (2021). A review of hydrologic signatures and their applications. *WIREs Water*, 8(1):e1499.
- Melsen, L. A., Teuling, A. J., van Berkum, S. W., Torfs, P. J. J. F., and Uijlenhoet, R. (2014). Catchments as simple dynamical systems: A case study on methods and data requirements for parameter identification. *Water Resources Research*, 50(7):5577–5596.
- Milzow, C., Molnar, P., McArdeell, B. W., and Burlando, P. (2006). Spatial organization in the step-pool structure of a steep mountain stream (vogelbach, switzerland). *Water Resources Research*, 42(4).
- Moriasi, D., Gitau, M., Pai, N., and Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*, 58:1763–1785.
- Nathan, R. J. and McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, 26(7):1465–1473.
- Nitsche, M., Rickenmann, D., Turowski, J., Badoux, A., and Kirchner, J. (2011). Evaluation of bedload transport predictions using flow resistance equations to account for macro-roughness in steep mountain streams. *Water Resources Research*, 47(8).
- Nonki, R. M., Lenouo, A., Tshimanga, R. M., Donfack, F. C., and Tchawoua, C. (2021). Performance assessment and uncertainty prediction of a daily time-step hbv-light rainfall-runoff model for the upper benue river basin, northern cameroon. *Journal of Hydrology: Regional Studies*, 36:100849.
- Onderka, M., Krein, A., Wrede, S., Martínez-Carreras, N., and Hoffmann, L. (2012). Dynamics of storm-driven suspended sediments in a headwater catchment described by multivariable modeling. *Journal of Soils and Sediments*, 12:620–635.
- Peel, M. C. and McMahon, T. A. (2020). Historical development of rainfall-runoff modeling. *WIREs Water*, 7(5):e1471.
- Penman, H. (1959). Notes on the water balance of the sperbelgraben and rappengraben. *Mitt. Schweiz. Anst. Forst. Versuchswes*, 35:99–109.
- Perrin, C., Michel, C., and Andréassian, V. (2001). Does a large number of parameters enhance model performance? comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242(3):275–301.
- Puy, A., Beneventano, P., Levin, S. A., Piano, S. L., Portaluri, T., and Saltelli, A. (2022). Models with higher effective dimensions tend to produce more uncertain estimates. *Science Advances*, 8(42):eabn9450.
- Rickenmann, D. (1997). Sediment transport in swiss torrents. *Earth Surface Processes and Landforms*, 22(10):937–951.
- Roessel, B. W. P. (1950). Hydrologic problems concerning the runoff in headwater regions. *Eos, Transactions American Geophysical Union*, 31(3):431–442.
- Rogelis, M. C., Werner, M., Obregón, N., and Wright, N. (2016). Hydrological model assessment for flood early warning in a tropical high mountain basin. *Hydrology and Earth System Sciences Discussions*, 2016:1–36.
- Rosbjerg, D. and Rodda, J. (2019). lahs: a brief history of hydrology. *History of Geo- and Space Sciences*, 10(1):109–118.
- Rusjan, S. and Mikoš, M. (2015). A catchment as a simple dynamical system: Characterization by the streamflow component approach. *Journal of Hydrology*, 527:794–808.
- Searcy, J. K. (1959). *Flow-duration curves*. Number 1542. US Government Printing Office.
- Seibert, J. (1996). Hbv light. *User's manual, Uppsala University, Institute of Earth Science, Department of Hydrology, Uppsala*.
- Seibert, J. and Bergström, S. (2022). A retrospective on hydrological catchment modelling based on half a century with the hbv model. *Hydrology and Earth System Sciences*, 26(5):1371–1388.

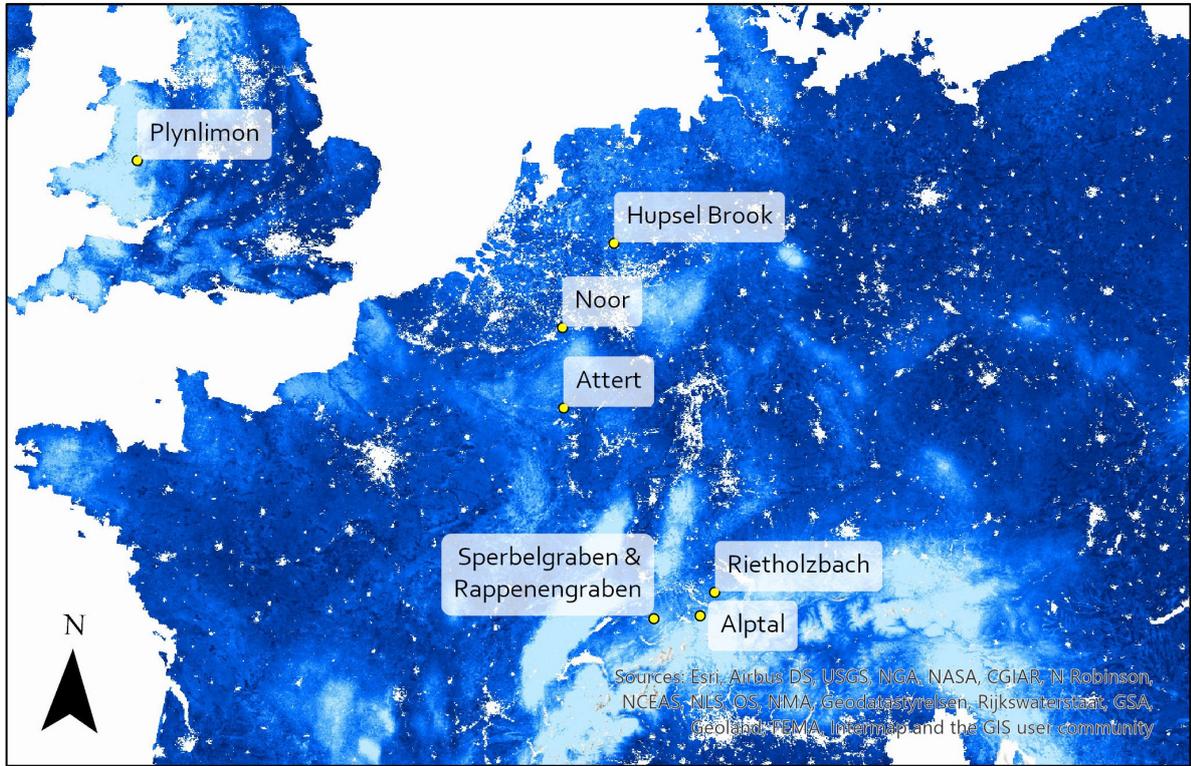
- Seneviratne, S., Lehner, I., Gurtz, J., Teuling, A., Lang, H., Moser, U., Grebner, D., Menzel, L., Schroff, K., Vitvar, T., and Zappa, M. (2012). Swiss prealpine rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event. *Water Resources Research*, 48.
- Snowling, S. and Kramer, J. (2001). Evaluating modelling uncertainty for model selection. *Ecological Modelling*, 138(1):17–30.
- Speir, S. L., Rose, L. A., Blaszcak, J. R., Kincaid, D. W., Fazekas, H. M., Webster, A. J., Wolford, M. A., Shogren, A. J., and Wymore, A. S. (2023). Catchment concentration–discharge relationships across temporal scales: A review. *WIREs Water*, page e1702.
- Staudinger, M., Stoelzle, M., Seeger, S., Seibert, J., Weiler, M., and Stahl, K. (2017). Catchment water storage variation with elevation. *Hydrological Processes*, 31(11):2000–2015.
- Stähli, M. (2018). Longterm hydrological observatory alptal (central switzerland).
- Stähli, M., Badoux, A., Ludwig, A., Steiner, K., Zappa, M., and Hegg, C. (2011). One century of hydrological monitoring in two small catchments with different forest coverage. *Environmental Monitoring and Assessment*, 174(1):91–106.
- Stähli, M., Seibert, J., Kirchner, J. W., von Freyberg, J., and van Meerveld, I. (2021). Hydrological trends and the evolution of catchment research in the alptal valley, central switzerland. *Hydrological Processes*, 35(4):e14113.
- Terink, W., Leijnse, H., Eertwegh, G., and Uijlenhoet, R. (2018). Spatial resolutions in areal rainfall estimation and their impact on hydrological simulations of a lowland catchment. *Journal of Hydrology*, 563.
- Teuling, A. J., Lehner, I., Kirchner, J. W., and Seneviratne, S. I. (2010). Catchments as simple dynamical systems: Experience from a swiss prealpine catchment. *Water Resources Research*, 46(10).
- Todini, E. (2011). History and perspectives of hydrological catchment modelling. *Hydrology Research*, 42(2-3):73–85.
- Uhlenbrook, S., Seibert, J., Leibundgut, C., and Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, 44(5):779–797.
- van der Velde, Y., Rozemeijer, J. C., de Rooij, G. H., van Geer, F. C., and Broers, H. P. (2010). Field-Scale Measurements for Separation of Catchment Discharge into Flow Route Contributions. *Vadose Zone Journal*, 9(1):25–35.
- van Lanen, H., van de Weerd, B., Dijkma, R., ten Dam, H., and Bier, G. (1995). *Hydrogeologie van het stroomgebied van de Noor en de effecten van grondwateronttrekkingen aan de westrand van het Plateau van Margraten*. Number 57 in Rapport / Landbouwniversiteit, Vakgroep Waterhuishouding. Landbouwniversiteit Wageningen, Netherlands. i.o.v. NV Waterleiding Maatschappij Limburg.
- van Meerveld, H., Fischer, B., Rinderer, M., Stähli, M., and Seibert, J. (2018). Runoff generation in a prealpine catchment: A discussion between a tracer and a shallow groundwater hydrologist. *Cuadernos de Investigación Geográfica*, 44(2):429–452.
- Wagener, T., Sivapalan, M., Troch, P., and Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, 1(4):901–931.
- Wainwright, J. and Mulligan, M. (2005). *Environmental Modelling: Finding Simplicity in Complexity*. Wiley.
- Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C.-Y. (2011). Calibration of hydrological models using flow-duration curves. *Hydrology and Earth System Sciences*, 15(7):2205–2227.
- Wrede, S., Fenicia, F., Martínez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H. H. G., Uhlenbrook, S., Kavetski, D., and Pfister, L. (2015). Towards more systematic perceptual model development: a case study using 3 luxembourgish catchments. *Hydrological Processes*, 29(12):2731–2750.

A | List of Variables

Table A.1: List of variables

Variable name	Description	Unit
r	Evaporation reduction factor	-
ET_{pot}	Potential evapotranspiration	mm/h
P	Precipitation	mm/h
ET	Actual evapotranspiration, calculated as $r * ET_{pot}$	mm/h
FFT	Fast Fourier Transform frequency	s^{-1}
PSD	Power Spectral Density	-
Q_{total}	Total discharge, summation of the three discharge components	mm/h
Q_{GW}	Deep groundwater generated discharge	mm/h
Q_S	Quickly responding discharge	mm/h
Q_{OF}	Very quickly responding discharge, overland flow generated	mm/h
Q_b	Baseflow, where Q_b is equal to Q_{GW}	mm/h
Q_t	Discharge at time step t	mm/h
Q_0	Discharge at time step 0	mm/h
τ	Recession coefficient	h^{-1}
$Q_{b,t}$	Baseflow at time step t	mm/h
$Q_{b,t-1}$	Baseflow at time step t - 1	mm/h
BFI	Maximum value baseflow index; ratio between baseflow and total discharge	-
z	Recession rate, defined as $e^{-t/\tau}$	-
S	Water storage	mm
M	Snow melt	mm/h
$g(Q)$	Sensitivity function, defined as the discharge sensitivity to changes in catchment storage	-
a	$g(Q)$ parameter - intersect	-
b	$g(Q)$ parameter - slope	-
β	$g(Q)$ parameter - slope	-
$T_{lag,GW}$	Lag time of deep groundwater based discharge	h
$T_{lag,s}$	Lag time of shallow groundwater based discharge	h
f_{OF}	Fraction of precipitation which becomes overland flow	-

B | Additional Figures



● Catchment locations

0 90 180 360 Kilometers

Potential Groundwater Recharge (mm/yr)

Value



Figure B.1: Potential Groundwater Recharge map (Martinsen et al., 2022) in mm/yr with the researched catchments marked in yellow.

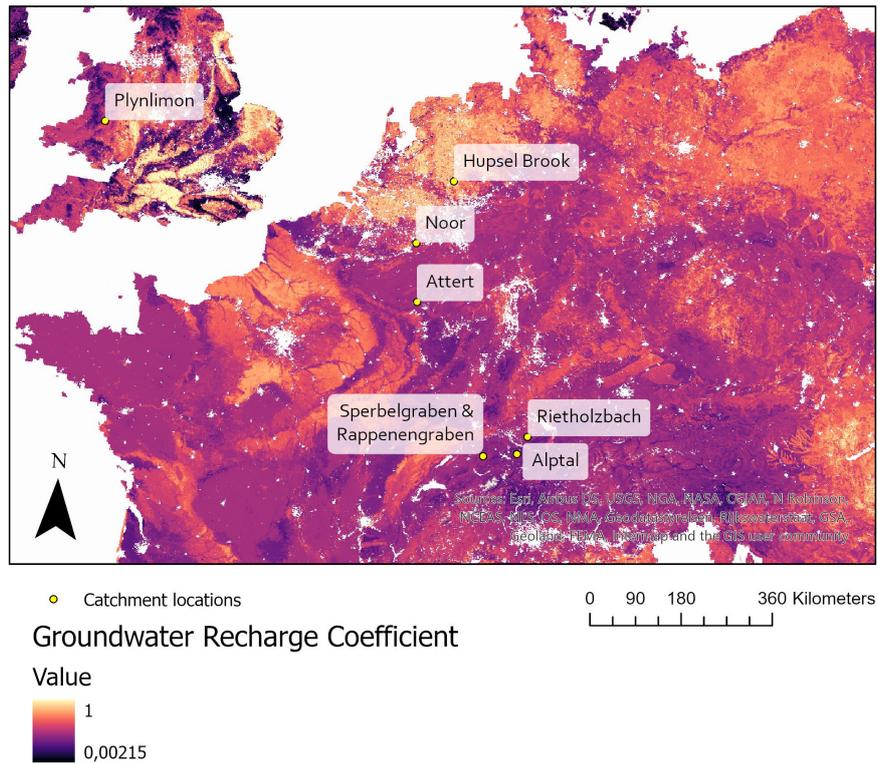


Figure B.2: Groundwater Recharge coefficient map (Martinsen et al., 2022) with the researched catchments marked in yellow.

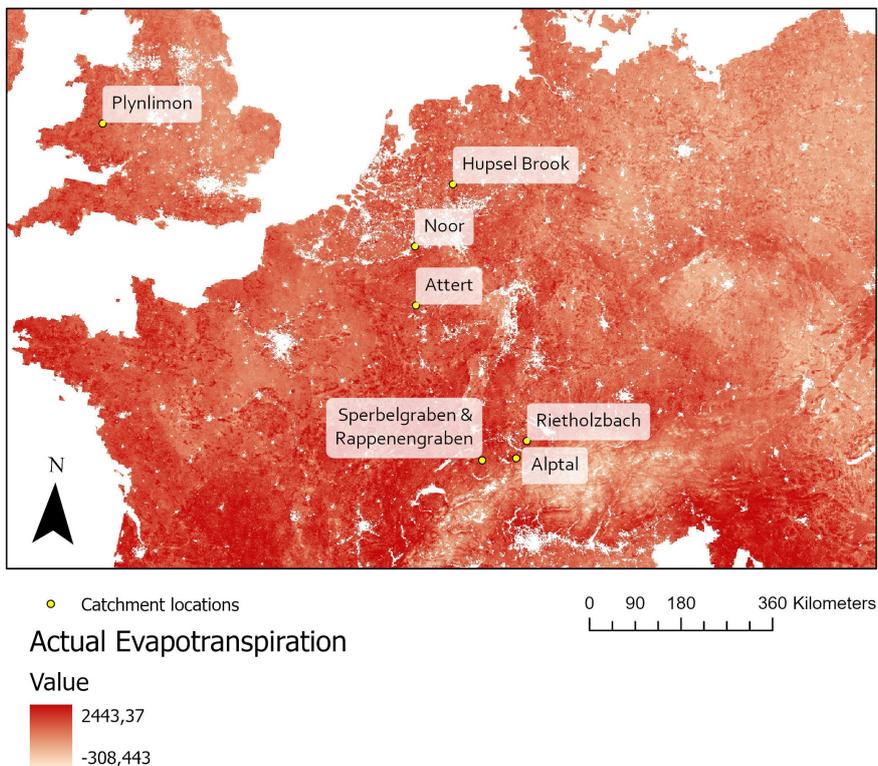


Figure B.3: Actual evapotranspiration map (Martinsen et al., 2022) in mm/yr with the researched catchments marked in yellow.

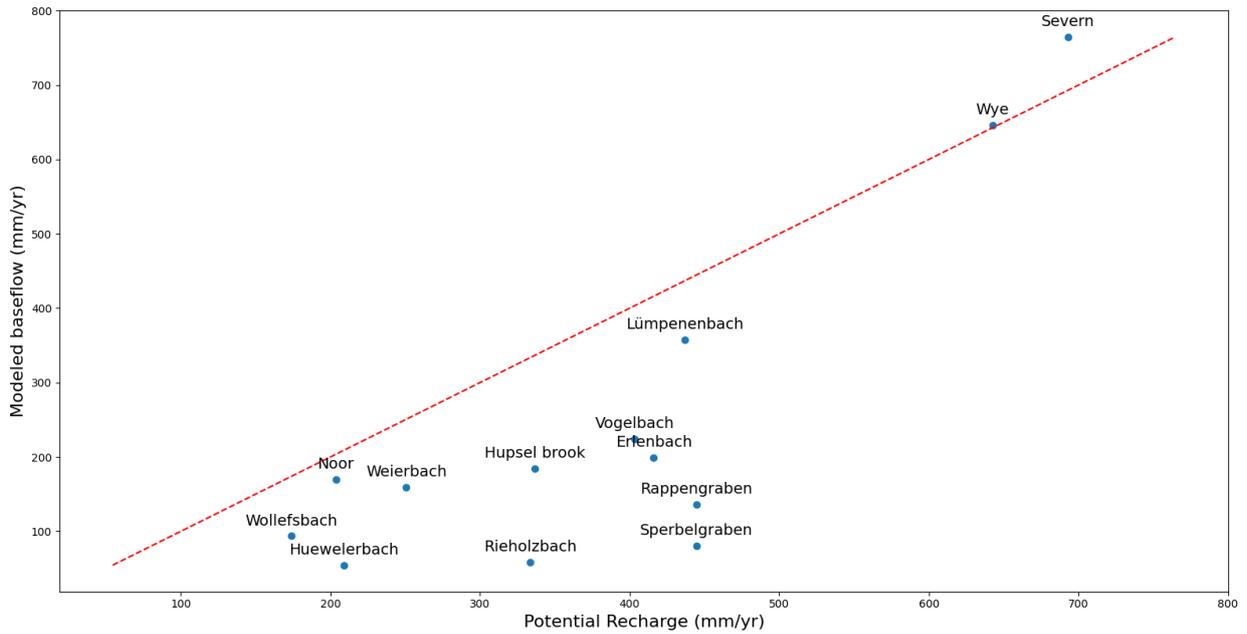


Figure B.4: Modeled total baseflow plotted against potential recharge (Martinsen et al., 2022) per researched catchment with 1:1 line in red.

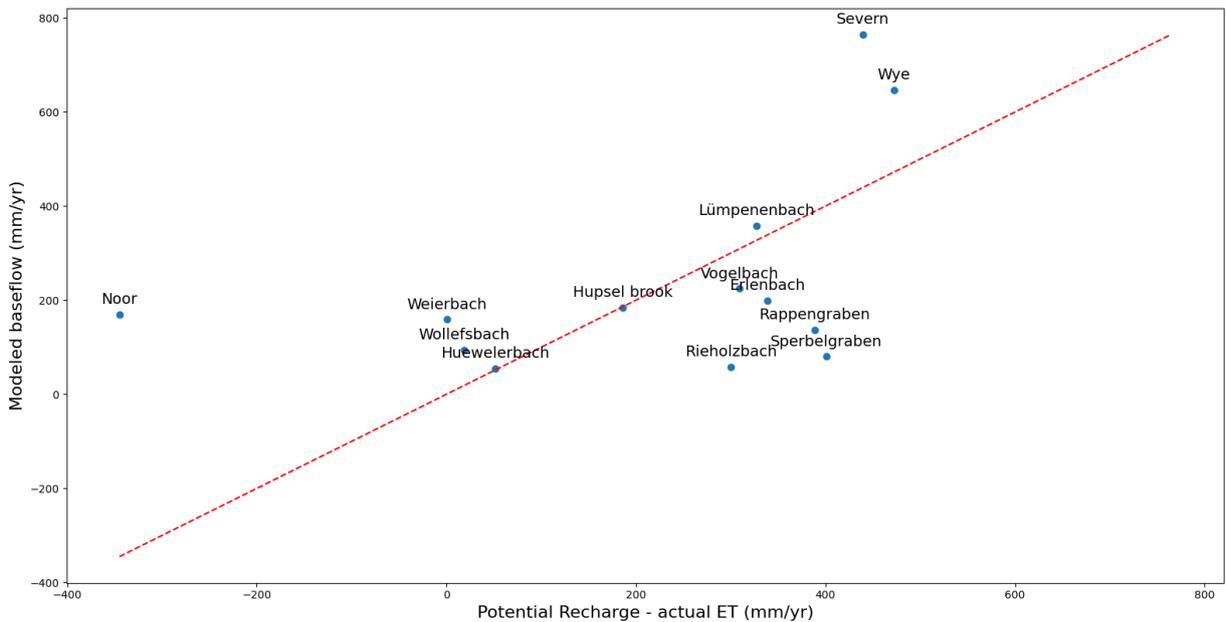


Figure B.5: Modeled total baseflow plotted against potential recharge (Martinsen et al., 2022) minus actual evapotranspiration*BFI per researched catchment with 1:1 line in red.

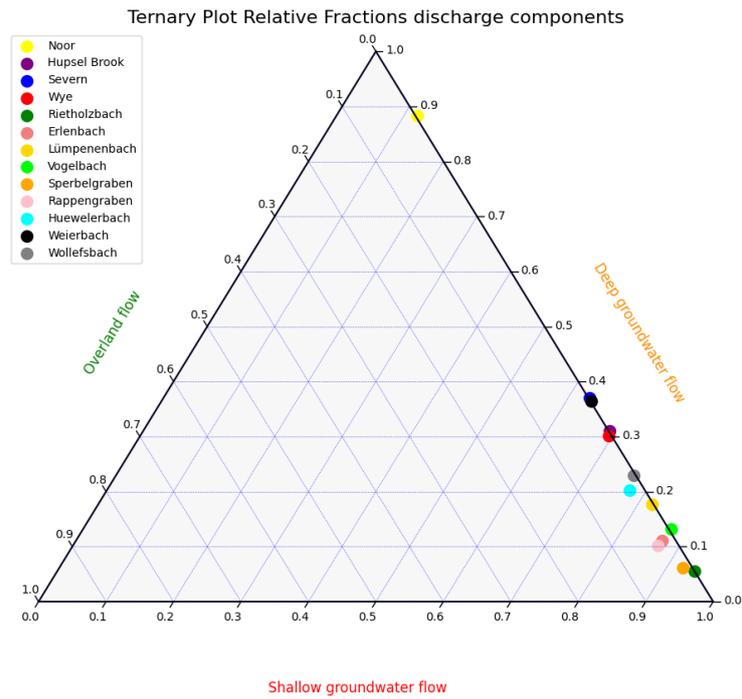


Figure B.6: Ternary plot of the yearly average discharge components of the researched catchments, calculated based on all time steps.

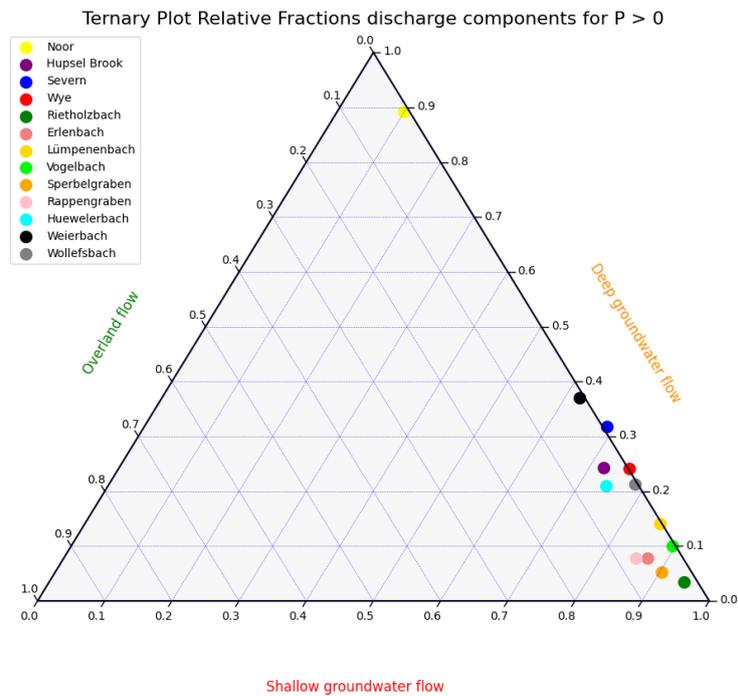


Figure B.7: Ternary plot of the yearly average discharge components of the researched catchments, calculated based on all time steps where precipitation is larger than zero.

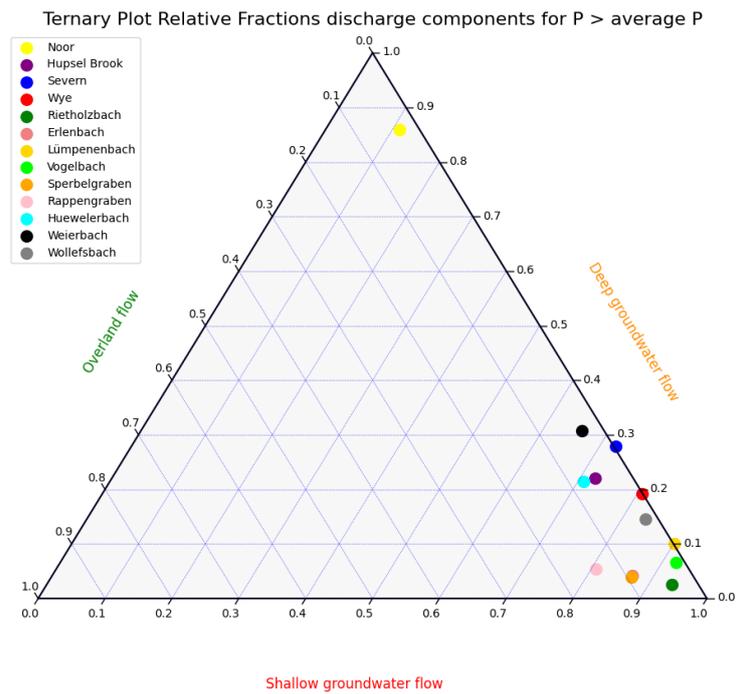


Figure B.8: Ternary plot of the yearly average discharge components of the researched catchments, calculated based on all time steps where precipitation is larger than the average precipitation (average of all time steps with precipitation).

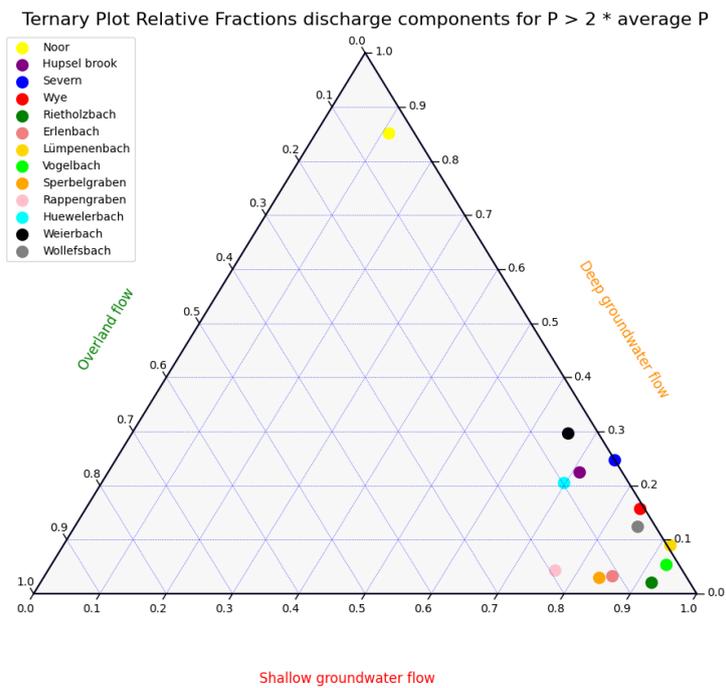


Figure B.9: Ternary plot of the yearly average discharge components of the researched catchments, calculated based on all time steps where precipitation is larger than two times the average precipitation (average of all time steps with precipitation).

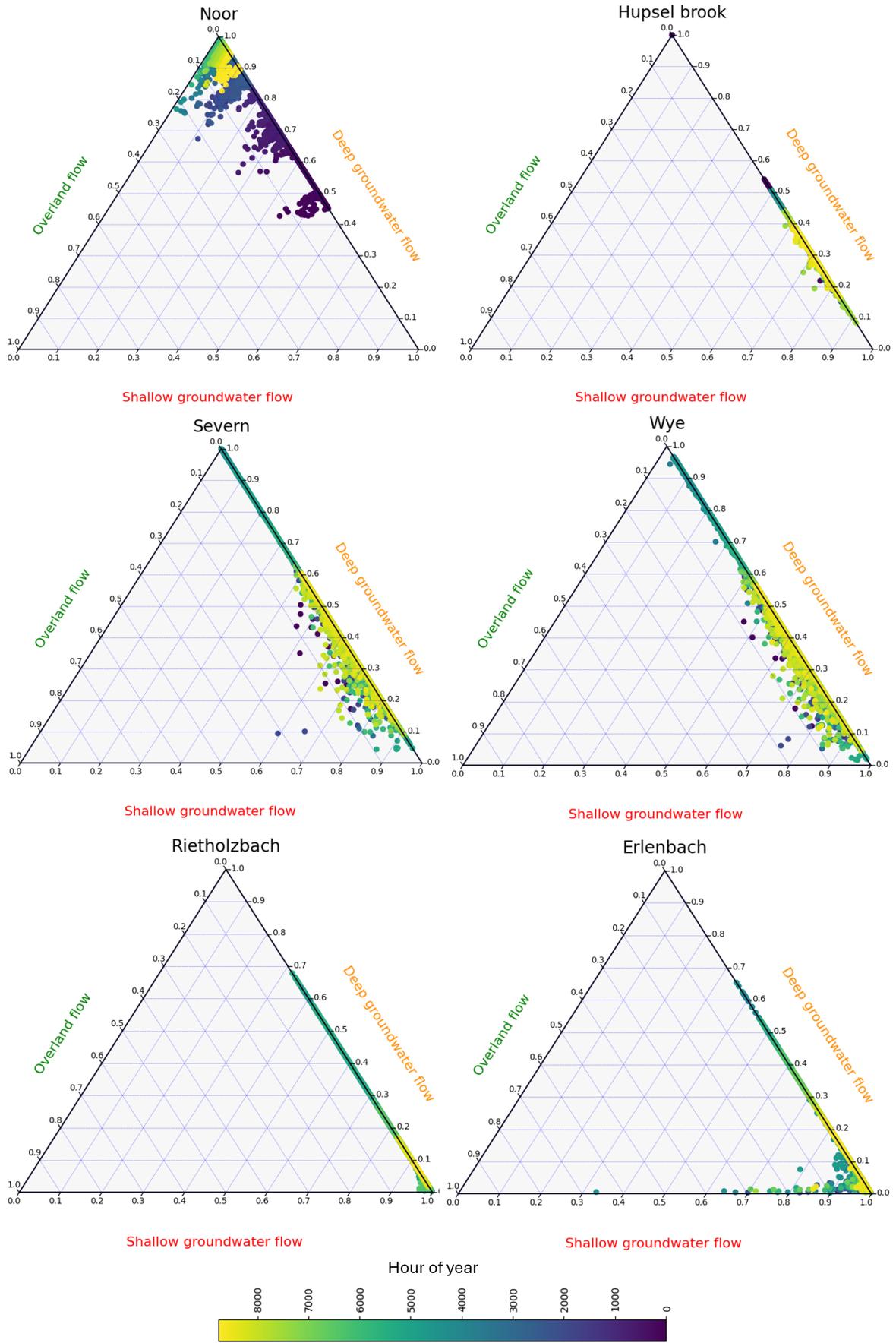


Figure B.10: Ternary plot of the yearly variation in discharge components of the researched catchments, shown over time as hour of year.

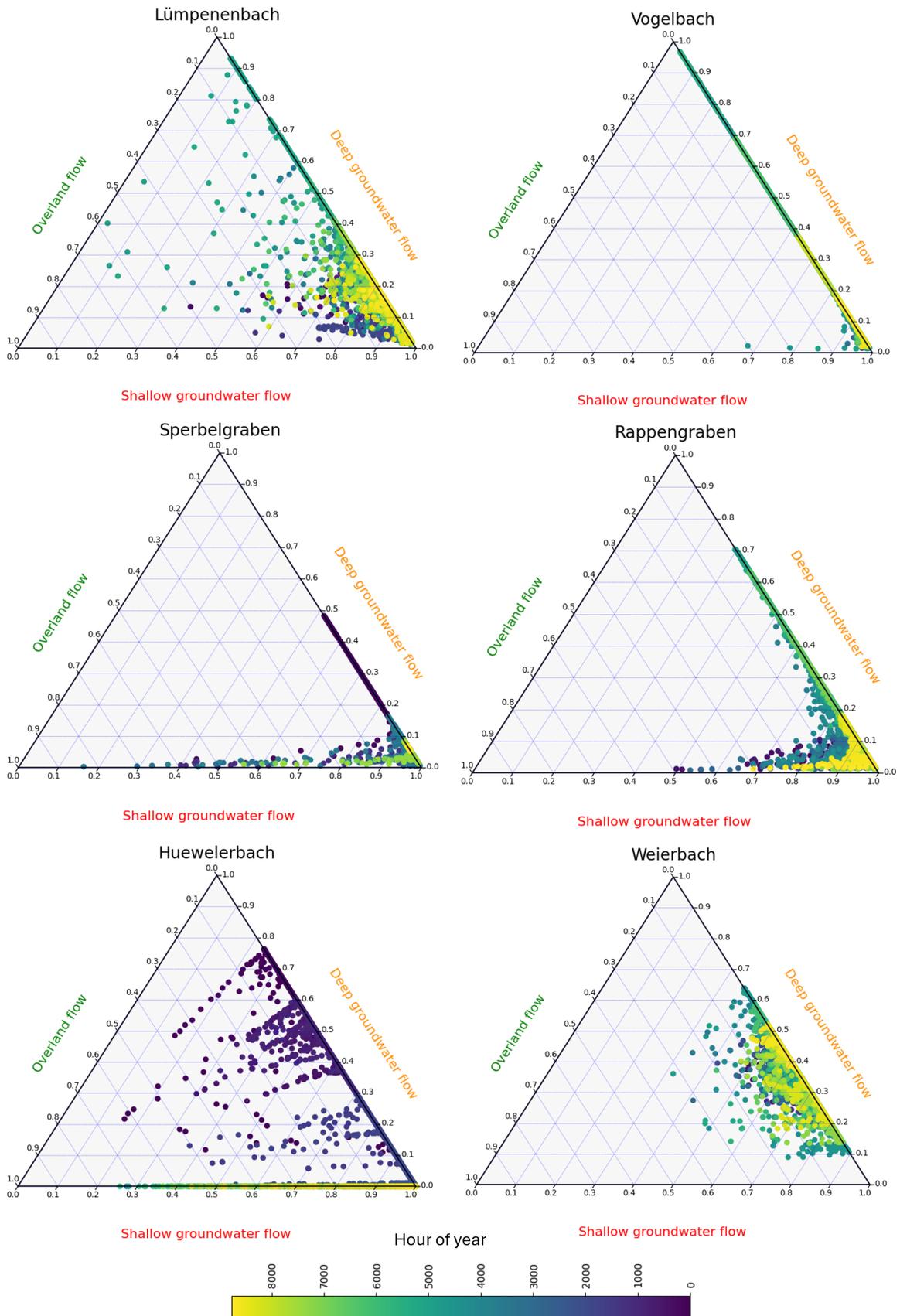


Figure B.11: Ternary plot of the yearly variation in discharge components of the researched catchments, shown over time as hour of year.

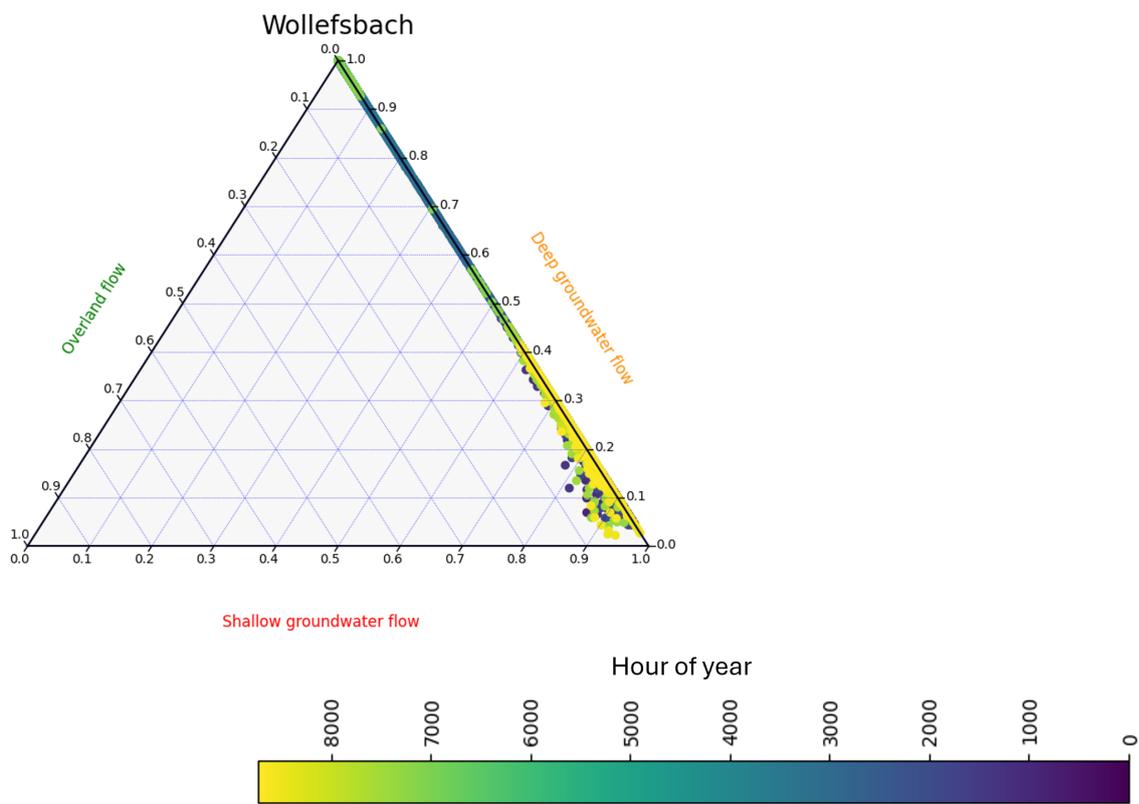


Figure B.12: Ternary plot of the yearly variation in discharge components of the researched catchments, shown over time as hour of year.

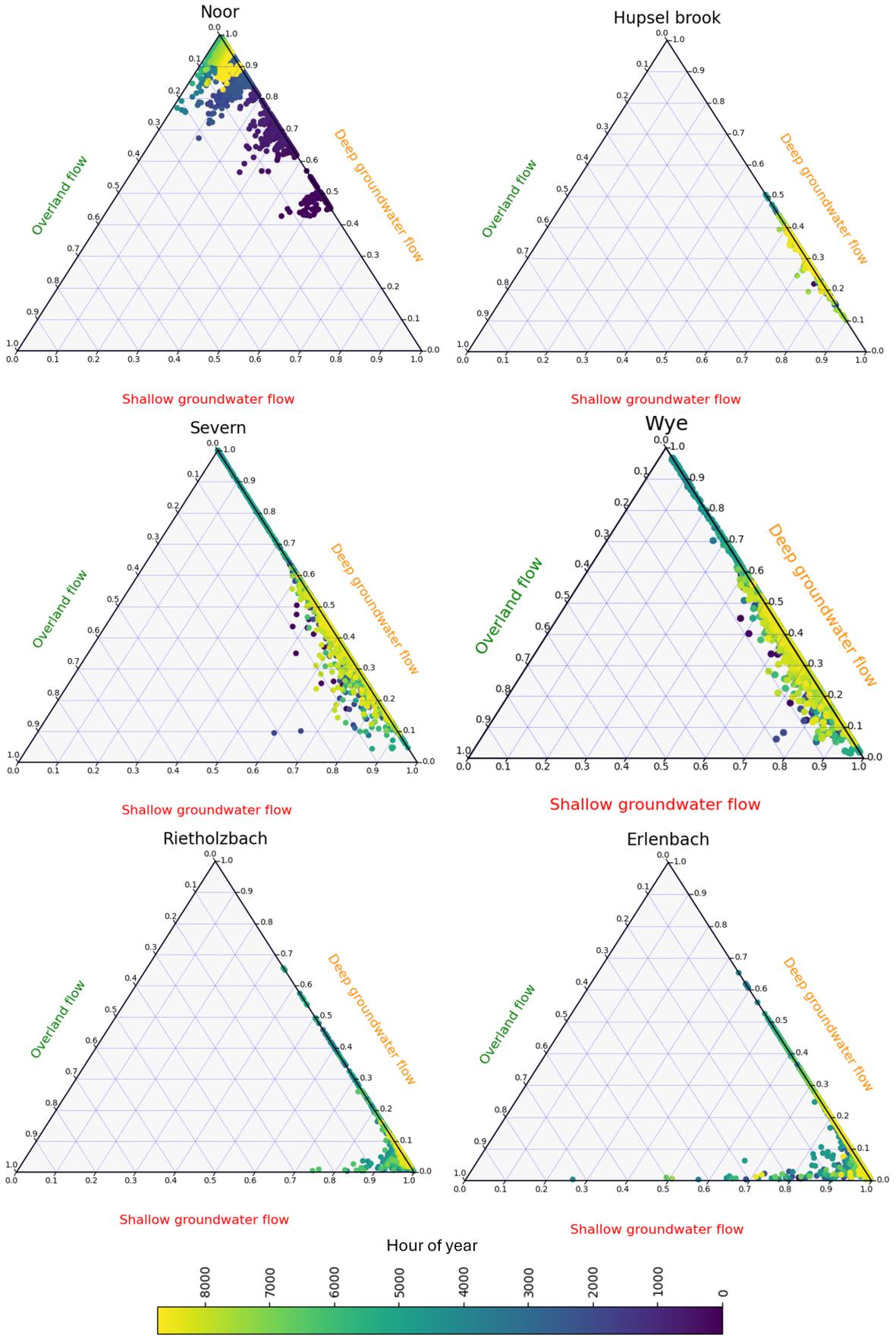


Figure B.13: Ternary plot of the yearly variation in discharge components of the researched catchments for the time steps where precipitation is larger than zero, shown over time as hour of year.

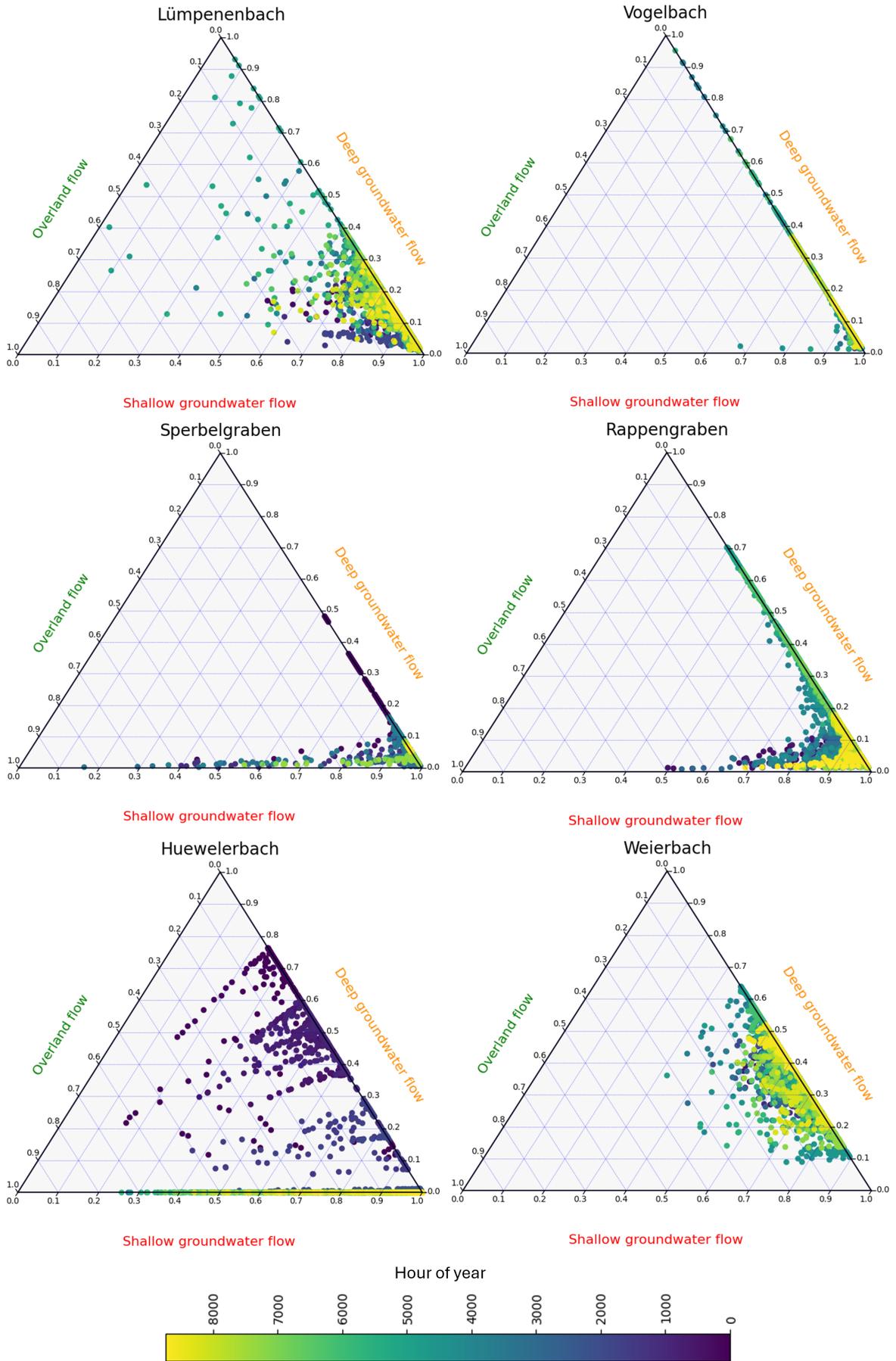


Figure B.14: Ternary plot of the yearly variation in discharge components of the researched catchments for the time steps where precipitation is larger than zero, shown over time as hour of year.

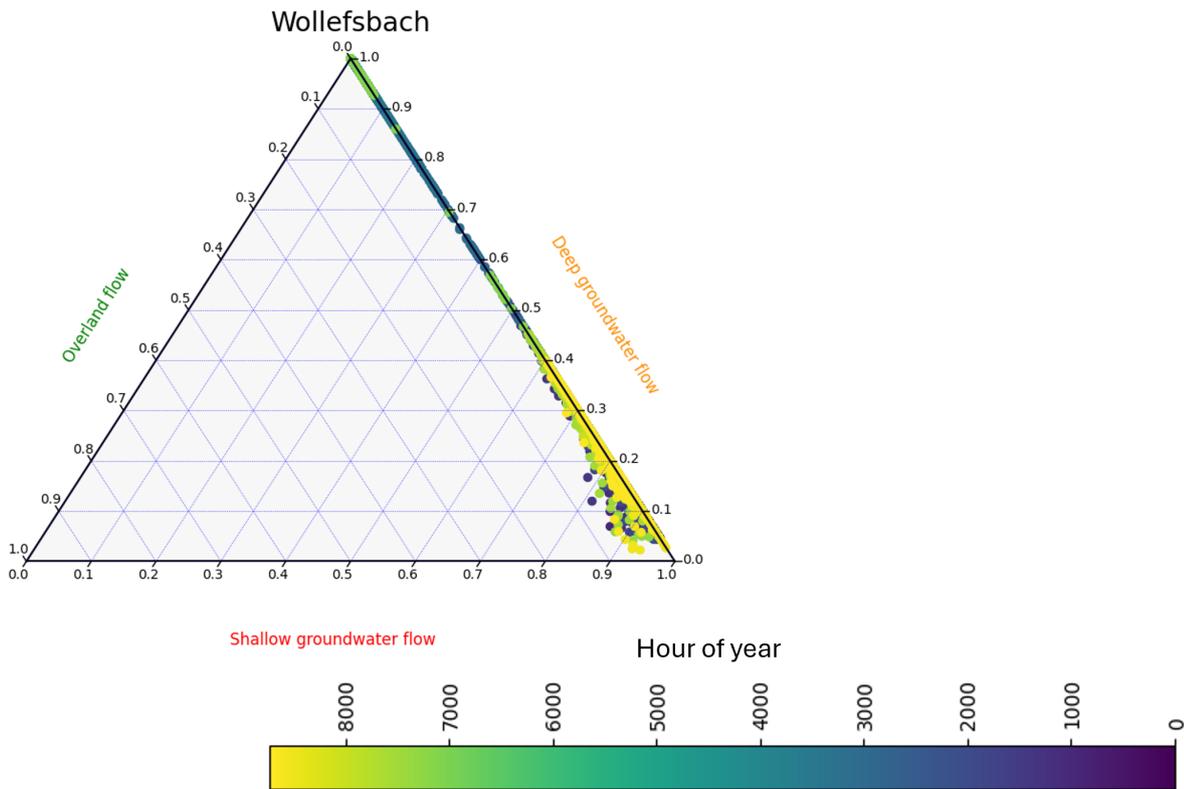


Figure B.15: Ternary plot of the yearly variation in discharge components of the researched catchments for the time steps where precipitation is larger than zero, shown over time as hour of year.

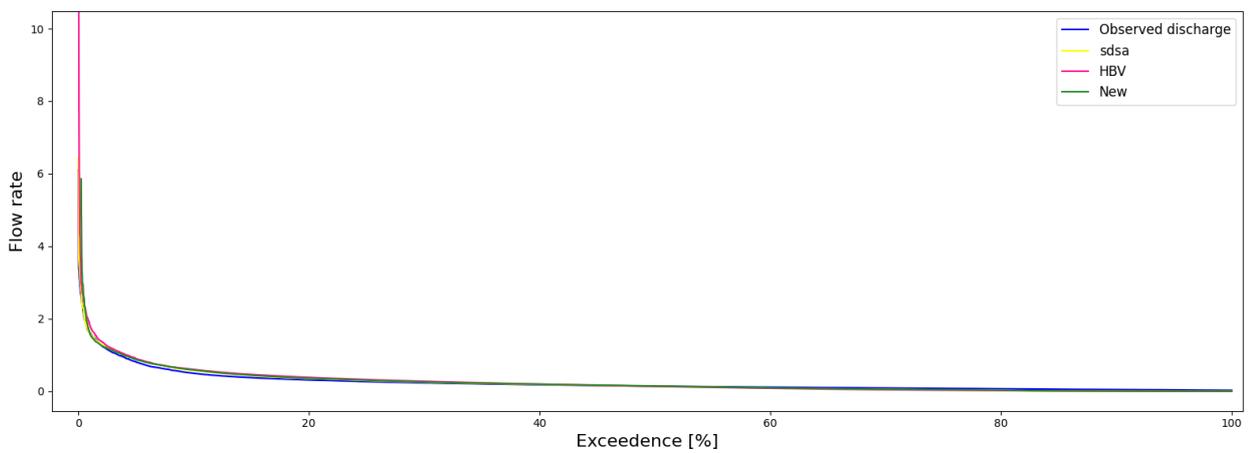


Figure B.16: Flow duration curves of the observed and modeled discharge, plotted against discharge.