A comparison of observed and modelled precipitation-streamflow relationships in two Dutch catchments using Ensemble Rainfall-Runoff Analysis (ERRA)

Max van Gorkum, 2647725
Supervisor: Wouter Berghuijs

VRIJE UNIVERSITEIT AMSTERDAM

Abstract

The rainfall-runoff relationship is one of the most fundamental aspects of catchment functioning. The rainfall-runoff conversion can drastically shift when precipitation intensity and catchment wetness change. Several existing methods quantify the rainfall-runoff relationship, but most of these methods have fundamental limitations. For example, conventional unit hydrographs often disregard the nonstationary and nonlinear behaviour of the rainfall-runoff relationship. Hydrological models often parameterize for nonlinearity and non-stationarity, but these models tend to suffer from overparameterization and equifinality. Here, we use multiple years of hourly observational data and the Ensemble Rainfall-Runoff Analysis (ERRA) to study the response functions of two Dutch catchments (the Hupselse Beek and the Ramsbeek) to quantify the rainfall-runoff response of these catchments. Results show that both catchments react quickly to rainfall, with the Hupselse Beek response peaking 8.5 hours and the Ramsbeek peaking 11 hours after rainfall. Initial runoff contributions are dominated by unsaturated zone processes, with groundwater contributions dominating later on. The groundwater contributions are largest at the Hupselse Beek, where they become the dominant contribution after 15 hours, and eventually make up the vast majority of the runoff response. In the Ramsbeek the groundwater contributions are less dominant. Both catchments' runoff responses are strongly nonlinear and nonstationary with (per unit rainfall) runoff responses growing with higher precipitation intensities and higher antecedent groundwater levels. An existing hydrological model widely applied to the Hupselse Beek (the WALRUS model) captures part of the catchments non-linearity and non-stationarity, but peak responses during high rain intensities and high antecedent groundwater levels tend to be underestimated. This approach highlights a new approach to evaluating hydrological models and could be more broadly applied in hydrology to better understand if models capture important climate sensitivities of catchments.

Table of Contents

1. Introduction

The conversion of rainfall into streamflow is one of the most fundamental properties of catchment functioning. This conversion could be described as "how much streamflow does precipitation generate and how is this distributed over time". This conversion determines rates and dynamics of streamflow (Berghuijs et al., 2020; Botter et al., 2013), and the nature of hydrological extremes such as floods and droughts (Merz et al., 2023; Van Loon, 2015). Without a quantitative understanding of this conversion, streamflow cannot be understood and predicted.

Given its central importance in hydrology, the precipitation-streamflow relationship has been quantified before with various methods across many different catchments. Main methods to assess this relationship are (i) (conceptual, physically based, and statistical) hydrological models (Beven, 2012) and (ii) conventional unit hydrographs (Dooge, 1959). The unit hydrograph method is a data-driven method to estimate the response of a catchment to a unit of rainfall but these unit hydrographs approaches generally assume a linear (streamflow response is proportional to effective precipitation) and stationary (streamflow response to a given unit of rainfall is identical, regardless of when it falls) response (Dooge, 1959; James and Johanson, 1999; Gao et al., 2025). However, the conversion of precipitation to streamflow can act differently according to precipitation intensity (Chang et al., 2023) and the water that is already stored in the catchment before the rainfall event (Nippgen et al., 2016). In reality, most catchments do not behave linear and stationary (Kirchner et al., 2023; Kirchner, 2024; Knapp et al., 2024; Mathias et al., 2016) and similar behaviour is also observed at other scales such as hillslopes (McDonnell et al., 2021; Tromp-van Meerveld et al., 2007).

A second method to quantify the rainfall-runoff relationship is with the use of (conceptual, physically based, and statistical) hydrological models. Hydrological models are a key tool in scientific and operational hydrology, as they can be used to, for example, forecast streamflow (Dion et al., 2021), assess long-term hydrological changes under climate change (Haddeland et al., 2014), and test hypotheses (Beven, 2018; Perrini et al., 2025). However, almost every hydrological model is overparameterized and suffers from equifinality (Beven, 2006; Khatami et al., 2019; Kirchner, 2009) as streamflow timeseries (which are typically used for model testing) can only constrain very few model parameters (Jakeman & Hornberger, 1993). In addition, hydrological models make pre-defined assumptions about the rainfall-runoff relationship (i.e. the model structure is typically fixed), and only implicitly model this relationship but do not show the relationship (and its nonstationarity and nonlinearity) explicitly. Consequently, a catchment's streamflow response to rainfall typically remains largely unconstrained and is not exposed explicitly, even after a model is developed for the catchment. This applies both to the overall response of the catchment, but also the sources of water that contribute to this streamflow generation.

In this study, we use Ensemble Rainfall-Runoff Analysis (ERRA) (Kirchner, 2024) to quantify the rainfall-runoff relationship of catchments using hourly observational data. Ensemble Rainfall–Runoff Analysis (ERRA) is a data-driven, model-independent approach used to quantify nonlinear, nonstationary, and spatially heterogeneous rainfall–runoff relationships directly from observed data. It operates across a range of time lags and uses nonlinear deconvolution to extract catchment-specific runoff responses to varying precipitation intensities. We analyse two Dutch catchments using ERRA, namely the Hupselse Beek and the Ramsbeek. For these catchments we will (*i*) quantify typical responses to precipitation, (*ii*) estimate the groundwater contribution to these responses, (*iii*) quantify the nonlinearity and nonstationarity of the runoff responses and (*iv*) test if an existing hydrological model (WALRUS (Brauer et al., 2014)) captures the same behaviour as the ERRA method.

2. Methods

2.1 Study sites and data

This study focusses on catchments, the Hupselse Beek and the Ramsbeek (Fig. 1). The Hupselse Beek catchment has an area of approximately 6.5 km^2 and the Ramsbeek catchment has an area of approximately 41 km^2. The average yearly precipitation (1969-2017) in this region is approximately 772 mm per year. These catchments were chosen because they have been extensively studied in the past. The Hupselse Beek has been used since 1964 as an example rural lowland area for research. The Hupselse Beek was selected as part of a larger project attempting to combat drought damages. It has been used for understanding for example rainfall-runoff relationships and solute transport and how they are affected by land use change (Brauer et al., 2018). For this reason, hydrological and meteorological measurements have been taken nearly continuously in the catchment since 1964. The Ramsbeek has also been the subject of studies, although it has not been studied as extensively as the Hupselse Beek (Geertsema et al., 2020).

Both lowland catchments are very flat, with the elevation of the Hupselse Beek ranging between 22 and 35 meters above sea level. The soil underlaying both catchments is a loamy sand layer with a maximum thickness of 10 meters. This is underlain by an impermeable clay layer, resulting in a single aquifer draining to the streams (Brauer et al., 2018). Since no specific yield measurements are available, a specific yield of 0.1 was used for this study.

Hourly discharge measurements were obtained for both these catchments. For the Hupselse Beek these hourly measurements were obtained for a period of 18 January 2012 until 6 November 2024. Hourly discharge measurements for the Ramsbeek were obtained for a period between 11 December 2012 and 11 October 2018 (Waterschap Rijn en IJssel, 2024). Hourly precipitation was obtained from the KNMI weather station located in the north of the Hupselse Beek catchment (Fig. 1) for the same period as the discharge timeseries. Since the Ramsbeek catchment is located within 4 km of this weather station, the same precipitation time series was used for both catchments (KNMI, 2024).

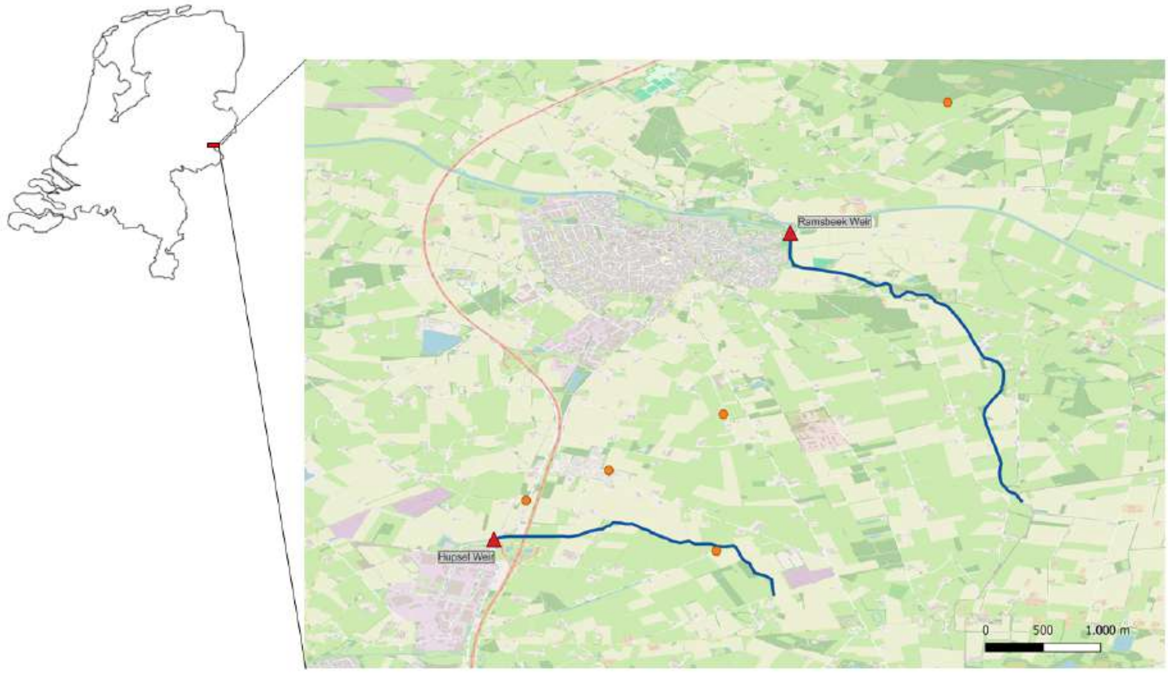


Figure 1: Map of the Hupselse Beek and the Ramsbeek. The map shows the location of both weirs, as well as the groundwater wells used in this study. The northern stream is the Ramsbeek and the southern stream is the Hupselse Beek. The three southern groundwater wells are used in the Hupselse Beek, the thee northern ones for the Ramsbeek.

For both catchments, hourly measurements of the groundwater level were obtained in m+ NAP (Waterschap Rijn en IJssel, 2024). Three groundwater wells were used per catchment, and an average groundwater level was calculated. This was done to filter out noise in the groundwater data. The location of these groundwater wells are shown in figure 1. The Ramsbeek catchment only contained two useable groundwater wells with sufficient hourly data, therefore the closest groundwater well in the Hupselse Beek catchment was also used for the Ramsbeek. With this groundwater data, the groundwater recharge was calculated using the water table fluctuation method (Healy & Cook, 2002). This means that recharge was calculated by multiplying the increase in groundwater level with the estimated specific yield of 0.1. When the groundwater level decreased, it was assumed that no groundwater recharge occurred and the recharge was set to zero.

2.2 Ensemble Rainfall-Runoff Analysis (ERRA)

This study used the ensemble rainfall-runoff analysis (ERRA)(Kirchner, 2024) to quantify the relationship between precipitation and streamflow. ERRA is a data driven method based on mathematical methods of estimating impulse response functions (Kirchner, 2022). ERRA is based on conventional unit hydrograph approaches, but with classic hydrograph approaches, the runoff is split into baseflow and stormflow and the precipitation is split into effective precipitation and evaporative losses. These divisions require assumptions, which introduce uncertainty in the resulting hydrographs. ERRA can instead analyse full timeseries of precipitation and discharge, removing these assumptions and therefore removing the uncertainties (Kirchner, 2024). Unlike most conventional unit hydrograph approaches, ERRA does not assume linearity or stationarity. ERRA can quantify streamflow response to precipitation and see how this differs under changing precipitation intensities and changing catchment wetness.

In this study, two main outputs of the ERRA method are shown. Firstly the runoff response distribution (RRD). The RRD can quantify the response of a system to an input of precipitation or groundwater recharge over time. The RRD does not take nonlinearity or nonstationarity into account, and can thus be seen as the average response of a system to a unit input of precipitation. ERRA calculates the RRD using equation 1.

$$Q(t) = \int_{\tau=0}^{\infty} \mathrm{RRD}(\tau) P(t-\tau) \mathrm{d}\tau \quad (1)$$

In this equation Q is the streamflow and P is the precipitation. The RRD in this equation quantifies the coupling of the precipitation and the streamflow at timestep τ. This equation assumes a continuous time series. In this study we deal with discrete time steps of 1 hour. Therefore the equation used in our study to estimate the runoff response distribution becomes

$$Q_j = \sum_{k=0}^{m} \mathrm{RRD}_{\{k\}} P_{\{j-k\}} \Delta t \quad (2)$$

Where Q_j is the streamflow at timestep j, P_{j-k} is the precipitation occurring k timesteps earlier, RRD_k is the impulse response of streamflow to precipitation at lag k and m is the total number of lag steps considered in the ERRA simulation (Kirchner, 2024).
For the RRD's shown in this study, the total number of lag steps considered was always set to 500, to include the tail of the responses.

The second main output of ERRA is the nonlinear response function (NRF). The non-linear response function can quantify the response of a system to precipitation of a certain intensity. Therefore, the NRF can show the nonlinearity of a system. In ERRA this is done by making the RRD in equation 2 a function of precipitation intensity. Equation 2 then becomes:

$$Q_j = \sum_{k=0}^{m} P_{\{j-k\}} \mathrm{RRD}_{\{k\}}(P_{\{j-k\}}) \Delta t \quad (3)$$

In this equation $RRD_{\{k\}}(P_{\{j-k\}})$signifies the impulse response of the streamflow at a certain precipitation intensity over a certain timestep.
The NRF is then defined as the product of the RRD at a certain precipitation intensity $RRD_{\{k\}}(P_{\{j-k\}})$ and the precipitation rate (P):

$$\mathrm{NRF}_{\{k\}}\left(P_{\{j-k\}}\right) = P_{\{j-k\}} \mathrm{RRD}_{\{k\}}(P_{\{j-k\}}) \quad (4)$$

The complete NRF function is then made by combining equation 3 and 4:

$$Q_j = \sum_{k=0}^{m} \mathrm{NRF}_{\{k\}}(P_{\{j-k\}})\Delta t \tag{5}$$

This equation gives the sum of the nonlinear response functions at the chosen precipitation intensity ranges. Q_j is the streamflow at timestep j, P_{j-k} is the precipitation that falls k time steps earlier, NRF_k is the nonlinear response function to precipitation at rate P_{j-k} and lasts for timestep Δt and m is the total lag steps considered in the ERRA simulation (Kirchner, 2024).
We work exclusively with hourly data, so Δt is equal to 1, and the precipitation intensity is in mm/h over a time step of 1 hour. Because of this, the resulting NRF has a unit of mm/h.

	Hupselse Beek	**Ramsbeek**
m	168	168
xknot type	values	values
Precipitation intensity range 1	<0.9	<0.9
Precipitation intensity range 2	0.9 – 1.8	0.9 – 1.8
Precipitation intensity range 3	1.8 – 2.7	1.8 – 2.7
Precipitation intensity range 4	2.7 – 3.5	2.7 – 3.5
Antecedent groundwater level range 1	27.2 – 27.8	25.3 – 25.8
Antecedent groundwater level range 2	27.8 – 27.9	25.8 – 25.9
Antecedent groundwater level range 3	27.9 – 28.1	25.9 – 26.1
Antecedent groundwater level range 4	28.1 – 28.4	26.1 – 26.3
Antecedent groundwater level range 5	28.4 – 29.1	26.3 – 26.6

Table 1: Settings used in ERRA for this study. The precipitation ranges are the precipitation classes used in the NRF's in mm/h. The antecedent groundwater level ranges are the catchment average groundwater levels as calculated with 3 groundwater weirs per catchment in m+ NAP. The number of time steps used for the NRF estimation is given by m.

Table 1 shows the precipitation intensity ranges and antecedent groundwater level ranges in this study as well as the number of timesteps used in ERRA (m) for the NRF estimation of both catchments.

One objective of this study was to quantify the groundwater contribution over time for both streams. Because the precipitation, discharge and groundwater recharge are all in mm/h, ERRA can be used to estimate the effect of precipitation on the groundwater recharge and subsequently the effect of groundwater recharge on the streamflow. In order to estimate the effect of precipitation on groundwater recharge, ERRA can be run with precipitation as input and groundwater recharge as output (instead of discharge). This will result in the groundwater response distribution to one unit of precipitation (GRRDp). This distribution describes the response of groundwater recharge over time to one unit of precipitation input. If we take the groundwater recharge as input and the discharge as output, we can look at the effect of groundwater recharge on the streamflow. This results in a runoff response distribution driven by groundwater recharge (RRDgr). This signifies the effect of groundwater recharge on the streamflow over time. By convolving the GRRDp and the RRDgr, we combine these two stages (rainfall to recharge and recharge to streamflow) into a single response function. The resulting distribution represents the runoff response due to groundwater pathways, meaning how much of the precipitation eventually enters the stream via the groundwater system (Gao et al., 2024). When we then subtract this distribution from the RRDp, we get the runoff response distribution of water released into the stream from quickflow.

To estimate the fraction of precipitation that is converted into streamflow or groundwater recharge, we use the runoff coefficient, which is defined as the integral of the runoff response distribution (RRD) over time. The runoff coefficient represents the fraction of a unit precipitation input that is converted to streamflow (or groundwater recharge) over the entire response period. With this, we can quantify how much precipitation ends up as streamflow, and we can quantify how much of this streamflow consists of groundwater contribution and how much consist of quickflow.

Since an NRF is the response function with a specific precipitation intensity, the effect of precipitation on groundwater recharge and the effect of groundwater recharge can also be assessed for NRFs. By making the same convolution as done for the RRD (convolving the nonlinear groundwater recharge response with the nonlinear streamflow response to recharge), we can investigate how precipitation intensities and antecedent wetness levels change the streamflow composition over time.

2.3 Model evaluation of WALRUS

The model we chose to compare our ERRA results with is the Wageningen Lowland Runoff Simulator (WALRUS) model, developed at the university of Wageningen. The WALRUS model is made up of three reservoirs, a coupled groundwater–vadose zone reservoir, a quickflow reservoir and a surface water reservoir (Brauer et al., 2014). Because the WALRUS model is mainly used in the Netherlands, it is optimized for hydrological processes that are important in the lowland areas of the Netherlands. (Yan et al., 2016). This is important, because many hydrological models are developed for sloping catchments, which may present errors when applied to lowland catchments (Bormann and Elfert, 2010).
The full overview of the model structure is shown in figure 2. The model requires a precipitation time series as input, as well as a potential evapotranspiration timeseries. For the precipitation timeseries we used the same precipitation timeseries that we used when running ERRA. A potential evapotranspiration timeseries is not openly available, and was provided by Claudia Brauer of the Wageningen university. Both these timeseries also consisted of hourly data.

The WALRUS model was first run with these timeseries to obtain a modelled hourly discharge timeseries. This discharge was then compared against the observed discharge in order to evaluate the model accuracy. To assess whether the WALRUS model can reproduce the same nonlinear and nonstationary behaviour observed in real-world data, the ERRA method was applied using modelled discharge from WALRUS instead of the observed discharge.

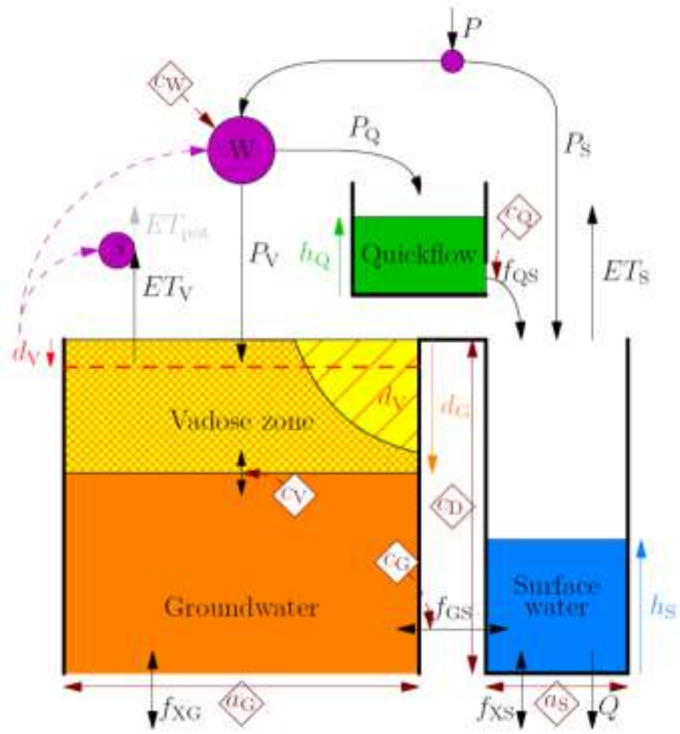


Figure 2: Overview of the model structure of the WALRUS model

As mentioned before, the WALRUS model is split up into three reservoirs with fluxes between these reservoirs. Because the WALRUS model has a coupled groundwater-vadose zone reservoir, feedbacks between the groundwater and vadose zone can be modelled, unlike other rainfall-runoff models, which often treat these zones as two different reservoirs (Brauer et al., 2014).

The model works by calculating the fluxes between these reservoirs. First, the precipitation is split between the three reservoirs. By multiplying the precipitation by the surface water area fraction the amount of precipitation that directly falls in the stream is calculated. Next, the precipitation is split further in the other two reservoirs:

$$Pv = P \cdot (1 - W) \cdot aG \tag{6}$$

$$Pq = P \cdot W \cdot aG \tag{7}$$

Equation 6 shows the amount of precipitation that is transported to the vadose zone in mm/h, and equation 7 shows the amount of precipitation that is transported into the quickflow reservoir in mm/h. In these equations P is the precipitation in mm/h, W is the wetness index and aG is the groundwater reservoir area fraction (defined as all surface area which is not covered by the stream).

The WALRUS model also accounts for evapotranspiration, both from the vadose zone (equation 8) and from the surface water (equation 9).

$$ETv = ETpot \cdot \beta \cdot aG \tag{8}$$

$$ETs = ETpot \cdot aS \tag{9}$$

In these equations, the ETv is the actual evapotranspiration from the vadose zone in mm/h, ETpot is the potential evapotranspiration in mm/h, β is the evapotranspiration reduction factor, ETs is the actual evapotranspiration from the surface water in mm/h and aS is the surface water area fraction.

Furthermore, the WALRUS model computes the flux between the groundwater and the surface water:

$$fGS = \frac{(cD - dG - hS) \cdot MAX((cD - dG), hS)}{cG} \cdot aG \tag{10}$$

In which fGS is the flux between the groundwater and the surface water in mm/h, cD is the channel depth in mm , dG is the groundwater depth in mm, hS is the surface water level in mm and cG is the groundwater reservoir constant in mm h.

The final internal flux computed in the WALRUS model is the flux from quickflow to surface water (equation 11):

$$fQS = \frac{hQ}{cQ} \cdot aG \tag{11}$$

In which fQS is the flux from quickflow to surface water in mm/h, hQ is the level of the quickflow reservoir in mm and cQ is the quickflow reservoir constant in h.

The parameters used in the WALRUS model were obtained from the WALRUS tutorial on the Hupselse Beek and are listed in appendix table 1.

A more detailed description of the WALRUS model and its parameters can be found at Brauer et al., (2014).

3. Results

3.1 Typical catchment responses

3.1.1 Hupselse Beek

As a first step, we illustrate the typical response of the Hupselse Beek catchment to incoming rainfall (Fig. 3). This response consists of the streamflow response to incoming rainfall (Fig. 3a), the recharge response to incoming rainfall (Fig. 3b), and the response of streamflow to recharge (Fig. 3c), and their combined responses that also highlights the relative contribution of groundwater (and unsaturated zone processes) to streamflow (Fig. 3d). All these functions show the average response that is generated by one unit of precipitation.

Streamflow of the Hupselse Beek tend to respond relatively fast to incoming rainfall but also sustains a response long after the rain has fallen. The runoff response distribution of the Hupselse Beek to one unit of precipitation over time (RRDp, Figure 3a) shows that the runoff response peaks after approximately 8.5 hours, and the height of the peak is $0.0045 \pm 4.35 \times 10^{-5}$ indicating that during the peak hour typically only ~0.45% of incoming precipitation is drained through the stream. While this peak is fast, streamflow responses are sustained for a long time, with still measurable (but very small) responses 480 hours (20 days) after the rain fell. The runoff coefficient of the RRDp is approximately 0.42, 42% of precipitation is turned into streamflow over the full response period.

The response of (groundwater) recharge to incoming precipitation is faster and lasts much shorter compared to streamflow (Fig. 3b). This figure shows the average response of the groundwater recharge to one unit of precipitation over time under (GRRDp). Groundwater recharge response to precipitation is almost instantaneous, with groundwater recharge peaking after approximately 40 minutes and reaching a peak height of $0.38 \pm 2.6 \times 10^{-3}$. After approximately 10 hours, the groundwater recharge response to precipitation stops (or becomes unmeasurably small). Thus timescales of recharge are more than an order of magnitude faster than those of streamflow within the Hupselse Beek. The runoff coefficient of the GRRDp is approximately 0.56, meaning that one unit of precipitation is turned into 0.56 units of groundwater recharge.

The response of streamflow to groundwater recharge is again slower and more sustained (Fig. 3c). The peak of the response is after approximately 11 hours, and the height of the peak is approximately $0.0048 \pm 8.2 \times 10^{-5}$, very similar to the RRDp. The runoff coefficient of the RRDgr is approximately 0.53. This means that when the groundwater recharge increases with one unit, the streamflow increases with 0.53 units.

Volumetrically, the total unit streamflow volume per unit precipitation is $0.42 \pm 2.9 \times 10^{-3}$ and the GRRDp has a runoff coefficient of approximately $0.56 \pm 8.6 \times 10^{-3}$. This shows that one unit of precipitation under the average catchment wetness conditions generates more groundwater recharge then streamflow, with approximately 42% of the precipitation turning into streamflow and 56% turning into groundwater recharge. It is important to note that the 42% of precipitation that is turned into streamflow is not just quickflow, but it also includes the contribution of groundwater into the stream as a result of the groundwater recharge.
We can also calculate how much of the streamflow consists of water originating from quickflow processes and how much of the streamflow enters the stream via the groundwater pathway. To do this we have to multiply the runoff coefficient of the GRRDp with the RRDgr. This will give us the fraction of the streamflow that consists of water that entered the stream via the

groundwater after one unit of precipitation. For the Hupselse Beek this turns out to be $0.56 \times 0.53 \approx 0.3$. This means that volumetrically approximately 30% of precipitation that falls in the Hupselse Beek catchment turns into streamflow via the groundwater pathway. If we keep in mind that in total approximately 42% of precipitation is converted to streamflow, this means that approximately 71% of streamflow consists of water originating from groundwater, and 29% of streamflow consists of quickflow.

In summary, one unit of precipitation under average wetness conditions in the Hupselse Beek leads to 0.56 units of groundwater recharge and 0.42 units of streamflow, 0.3 of which can be attributed to groundwater contribution, while the remaining 0.12 units can be attributed to quickflow. This means that the streamflow response of the Hupselse Beek consists for approximately 71% of groundwater and 29% of quickflow.

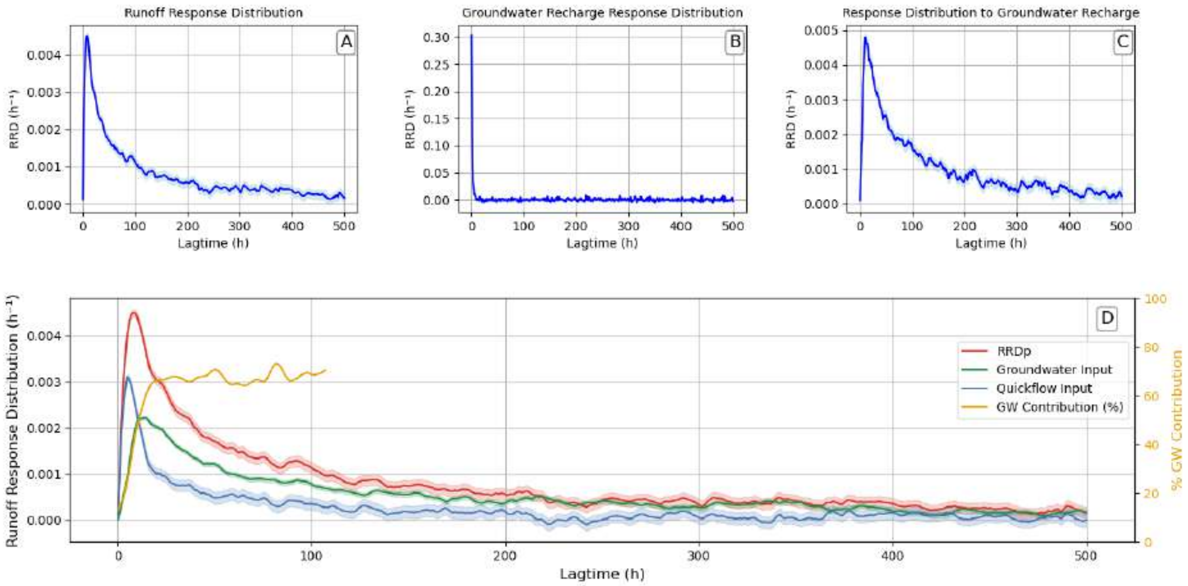


Figure 3: Response distributions at the Hupselse Beek. (a) Runoff response distribution to a unit of precipitation (RRDp). (b) Groundwater recharge response distribution caused by one unit of precipitation (GRRDp). (c) Runoff response distribution generated by one unit of groundwater recharge. (d) Composition of the runoff in the Hupselse Beek. The red line shows the RRDp, which is the average runoff generated by one unit of precipitation across all the data. The green line shows the amount of groundwater that is turned into streamflow over time after one unit of precipitation. The blue line signifies the non-groundwater input to the Hupselse Beek over time.

These individual response functions can be synthesized into the overall catchment response that summarizes the overall catchment response, including the relative contribution of groundwater (and unsaturated zone processes) to streamflow (Fig. 3d). The red line shows the total streamflow response (RRDp), while the green line represents the contribution from groundwater (calculated as the convolution of GRRDp and RRDgr), and the blue line indicates the quickflow component (the residual of RRDp after subtracting the groundwater contribution). During the first hours after the precipitation the streamflow mostly consists of quickflow processes, peaking after approximately 6 hours. The peak of the groundwater contribution to the stream is reached after approximately 17 hours, 7 hours later than the peak of the RRDp. Based on both these responses we can quantify the percentage of the streamflow that consists of groundwater contribution after one unit of precipitation. This is visualized by the yellow line. After approximately 18 hours, groundwater contribution takes over as the dominant input into the stream. After 100 hours, the groundwater input approaches 75%.

3.1.2 Ramsbeek

Next we will show the typical response of the Ramsbeek to rainfall (Fig. 4).
The streamflow response distribution to one unit of precipitation (RRDp, figure 4a) peaks after approximately 11 hours, and reaches a height of $0.0032 \pm 3.6 \times 10^{-5}$. The runoff coefficient during this runtime is approximately 0.22. This means that under average catchment wetness conditions approximately 22% of the precipitation is converted into streamflow in the Ramsbeek. When comparing the streamflow response of the Ramsbeek to the streamflow response of the Hupselse Beek, the results show that the Hupselse Beek has a quicker response to one unit of precipitation, with the peak response occurring after 8.5 hours, instead of the 11 hours seen at the Ramsbeek. Also, the peak height is larger at the Hupselse Beek, $0.0045 \pm 4.35 \times 10^{-5}$ at the Hupselse Beek and $0.0032 \pm 3.6 \times 10^{-5}$ at the Ramsbeek. Furthermore, in the Hupselse Beek approximately 42% of the precipitation ins converted into streamflow compared to the 22% in the Ramsbeek.

The groundwater response to precipitation (GRRDp, Fig. 4b) is also almost instant for the Ramsbeek, with a peak after approximately 35 minutes and reaching a height of $0.19 \pm 4.3 \times 10^{-3}$. The runoff coefficient of the GRRDp of the Ramsbeek is approximately 0.28, meaning that 28% of the precipitation becomes groundwater recharge. Compared to the Hupselse Beek, the groundwater recharge of the Ramsbeek reacts slightly faster (peaking after 35 minutes vs peaking after 40 minutes respectively), however, groundwater recharge responds less to precipitation in the Ramsbeek, with 28% of the precipitation turning into recharge compared to the 56% in the Hupselse Beek.

The response distribution of streamflow to groundwater recharge (RRDgr, Fig. 4c) shows that the peak response occurs approximately 17 hours after the precipitation, and reaches a peak height of $0.0034 \pm 6.5 \times 10^{-5}$. The runoff coefficient of the RRDgr is approximately 0.31, meaning that 31% of the groundwater recharge turns into streamflow during the duration of this run.

To estimate the total contribution of precipitation to streamflow via the groundwater pathway, we multiply the runoff coefficient of the GRRDp (0.28) with the runoff coefficient of the RRDgr (0.31). This gives $0.28 \times 0.31 \approx 0.087$, meaning that volumetrically 8.7% of the precipitation in the Ramsbeek enters the stream via the groundwater pathway. Given that the total runoff coefficient (from RRDp) is approximately 0.22, this implies that ~40% of the streamflow over the event is supplied by groundwater, while the remaining 60% originates from quickflow processes. Compared to the Hupselse Beek, the runoff response of the Ramsbeek is less groundwater dominated, with the total response after 500 hours consisting for 40% of groundwater compared to the 71% observed in the Hupselse Beek.

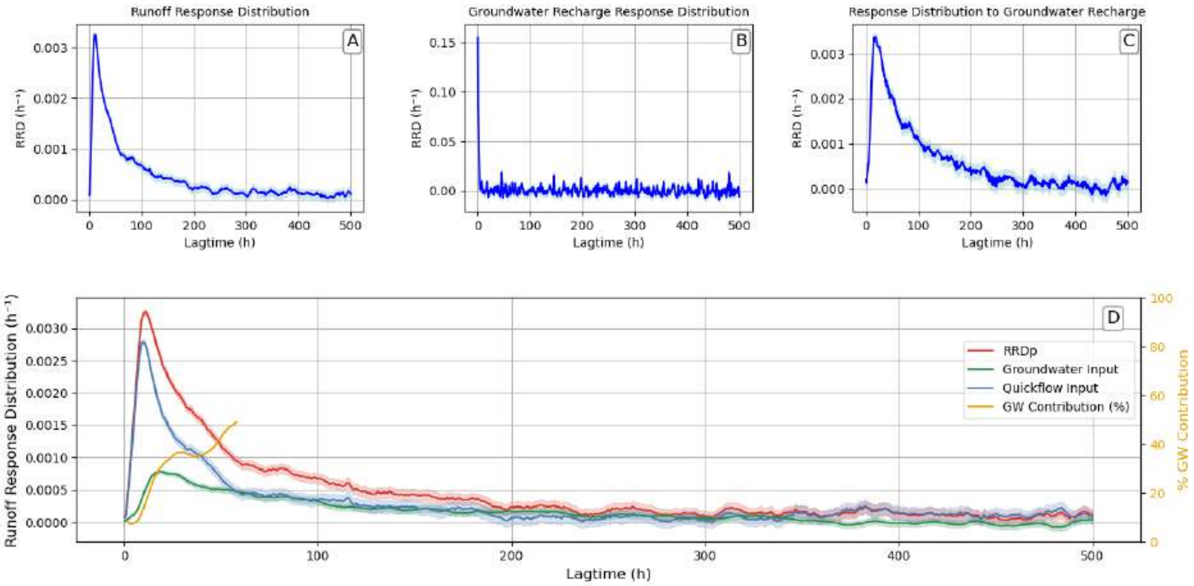


Figure 4: Response distributions at the Ramsbeek. (a) Runoff response distribution to a unit of precipitation (RRDp). (b) Groundwater recharge response distribution caused by one unit of precipitation (GRRDp). (c) Runoff response distribution generated by one unit of groundwater recharge. (d) Composition of the runoff in the Ramsbeek. The red line shows the RRDp, which is the average runoff generated by one unit of precipitation across all the data. The green line shows the amount of groundwater that is turned into streamflow over time after one unit of precipitation. The blue line signifies the quickflow contribution to the Ramsbeek over time.

When we look at the composition of the Ramsbeek over time (Fig. 4d), it becomes visually evident that the Ramsbeek is a surface runoff dominated stream. The quickflow input is larger than the groundwater contribution during most of the runoff response peak. After 80 hours, the input from both the groundwater source and the quickflow source is approximately equal, however at this time, the response is already very weak. The groundwater contribution peaks after 16 hours, but even here, the quickflow contribution is more than 2 times larger. The quickflow contribution

In summary, when comparing the responses of the Hupselse Beek and the Ramsbeek, the Hupselse Beek exhibits a slightly faster and stronger response to precipitation. A larger fraction of precipitation is converted into streamflow in the Hupselse Beek, and a greater proportion of this streamflow originates from groundwater contributions, indicating that the Hupselse Beek is more groundwater-dominated than the Ramsbeek.

3.2 Nonlinearity and nonstationary of catchment response

In the previous section, we have characterized the mean response of catchments, but disregarded any potential non-linearity and non-stationarity present in the real systems. In the real world, streams tend to react differently to precipitation of different intensities. Furthermore, streams also tend to behave differently according to their catchment wetness. To quantify this non-linearity and non-stationarity directly from data, ERRA can be used to split the data in different precipitation intensity classes, as well as different catchment wetness conditions. In this study, the groundwater level is used as a signal for catchment wetness. Here, the data are split into four precipitation intensity classes and five antecedent groundwater levels, and then analysed with ERRA.

Streamflow response of the Hupselse Beek is strongly dependent both on the antecedent groundwater levels and on precipitation intensity (Fig. 5). From left to right, the figure shows an response function for increasing in precipitation intensity (ranges shown on top in mm/h), and from top to bottom, as a function of increasing antecedent groundwater level (ranges shown on the right in m + NAP). The red line shows the streamflow response to precipitation of a given intensity at a certain groundwater level, the green line signifies the groundwater contribution to the streamflow and the blue line shows the quickflow contribution to the streamflow.

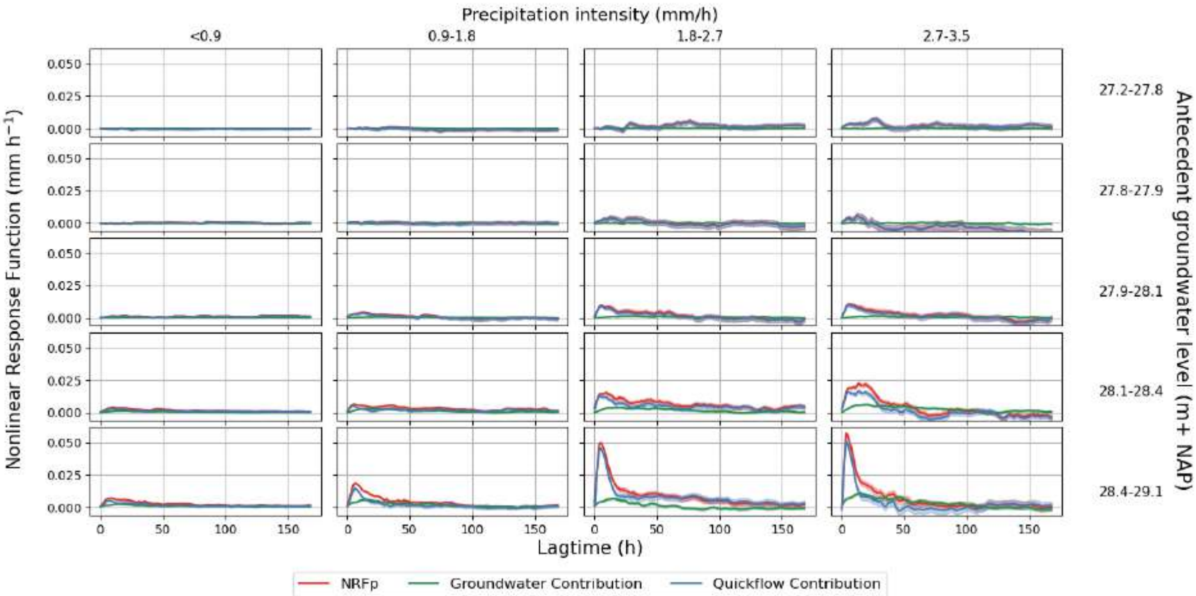


Figure 5: Nonlinear response functions (NRF) for the Hupselse Beek. From top to bottom an increase in antecedent groundwater level is shown. From left to right, an increase in precipitation intensity is shown. The red line shows the NRFp, so the total streamflow response to one hour of a certain precipitation intensity. The green line shows the groundwater input to the stream and the blue line shows the non-groundwater input to the stream.

Streamflow responses grow strongly with precipitation intensity. When the precipitation intensity is lower than 0.9 mm/h, there is almost no streamflow response, even with high antecedent groundwater levels (top row). When the precipitation intensity increases, between 0.9 and 1.8 mm/h, still no streamflow response is visible when the antecedent groundwater level is lower than 28.1 m+ NAP. A slight streamflow response can be seen when the antecedent groundwater level becomes higher. When the antecedent wetness exceeds 28.4 m+ NAP, a peak streamflow response of approximately 0.02 mm/h is generated. If the precipitation intensity increases further, the streamflow responses become larger. It is interesting to note that even with the highest precipitation intensities, almost no streamflow response is observed with the lowest antecedent groundwater levels. With the highest antecedent groundwater level (>28.4 m+ NAP) and the highest precipitation intensity (>2.7 mm/h), also the highest NRF peak is observed at approximately 0.055 mm/h. For this scenario, the total runoff coefficient of the streamflow response to this precipitation intensity is approximately 0.65, meaning that 65% of the precipitation converted into streamflow. This is significantly larger than the 42% under catchment average conditions as estimated with the RRD (figure 3a). This shows that the Hupselse Beek shows a nonlinear streamflow response with increasing precipitation intensity.

The composition of the discharge also changes as a function of precipitation intensities and antecedent groundwater levels. When antecedent groundwater levels are high (>28.1 m+ NAP), and the precipitation intensity is low (<1.8 mm/h), most streamflow response consists of groundwater. When the antecedent groundwater levels are low (<28.1 m+ NAP) and the precipitation intensity is high (>1.8 mm/h), almost 100% of the streamflow response consists of quickflow over all time lags. If we look at the highest precipitation intensity class (2.7 – 3.5 mm/h) and the highest antecedent groundwater level class (28.4 – 29.1 m+ NAP), we can see that most of the peak response consists of quickflow. Similar to the RRD approach, we can estimate the total contribution of groundwater and quickflow over these lag times. To do this, we multiply the runoff coefficient of the recharge response to precipitation (0.24) by the streamflow response to groundwater recharge (0.78), resulting in 0.24 × 0.78 = 0.19. This means that approximately 19% of the precipitation is converted to streamflow via the groundwater pathway in this scenario.

When expressed as a proportion of the total streamflow (runoff coefficient = 0.65), this 0.19 represents roughly 29% of the total streamflow, with the remaining 71% originating from quickflow processes. This is the exact opposite of what is found with the RRD under average catchment conditions, and further illustrates that with increasing precipitation intensity and antecedent groundwater levels, streamflow in the Hupselse Beek becomes increasingly more quickflow dominated.

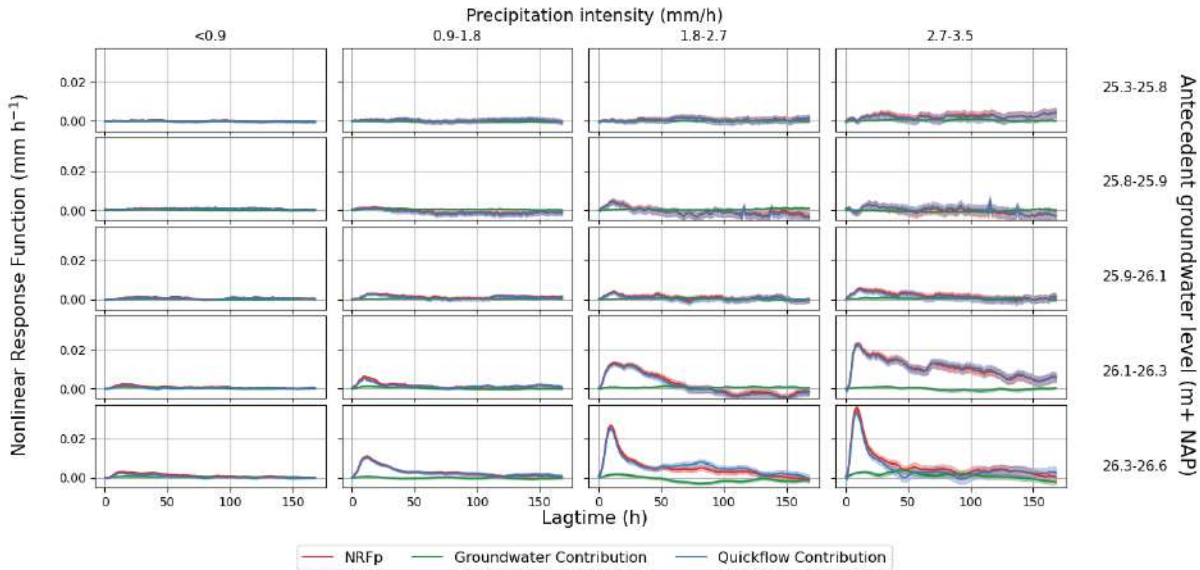


Figure 6: Nonlinear response functions (NRF) for the Ramsbeek. From top to bottom an increase in antecedent groundwater level is shown. From left to right, an increase in precipitation intensity is shown. The red line shows the NRFp, so the total streamflow response to one hour of a certain precipitation intensity. The green line shows the groundwater input to the stream and the blue line shows the non-groundwater input to the stream.

Comparable patterns of nonlinearity and nonstationarity are observed in the Ramsbeek (Fig. 6), using the same precipitation intensity classes as for the Hupselse Beek. A similar trend emerges: at low precipitation intensities (<0.9 mm/h), no streamflow response is observed, regardless of antecedent groundwater level. However, when groundwater levels exceed 26.1 m+ NAP, a streamflow response becomes apparent once precipitation intensity surpasses 0.9 mm/h. The strongest streamflow response occurs when groundwater levels are above 26.3 m+ NAP and precipitation intensity exceeds 2.7 mm/h. Here, a peak streamflow response is observed of approximately 0.036 after 9 hours. The runoff coefficient for this scenario is approximately 0.35, meaning that volumetrically 35% of precipitation is converted to streamflow.

This response is almost entirely quickflow dominated. By multiplying the runoff coefficient of the groundwater recharge response to precipitation (0.035) with the runoff coefficient of the streamflow response to groundwater recharge (0.75), we find that approximately 0.035 × 0.75 = 0.026, meaning 2.6% of the precipitation is converted into streamflow via the groundwater pathway. When expressed as a proportion of the total streamflow (runoff coefficient = 0.35), this 0.026 represents roughly 7.4% of the total streamflow, with the remaining 92.6% originating from quickflow processes. This demonstrates that the Ramsbeek also exhibits strong nonlinear behaviour in response to increasing precipitation intensity and antecedent groundwater levels. Additionally, the streamflow response becomes increasingly dominated by quickflow under these conditions.

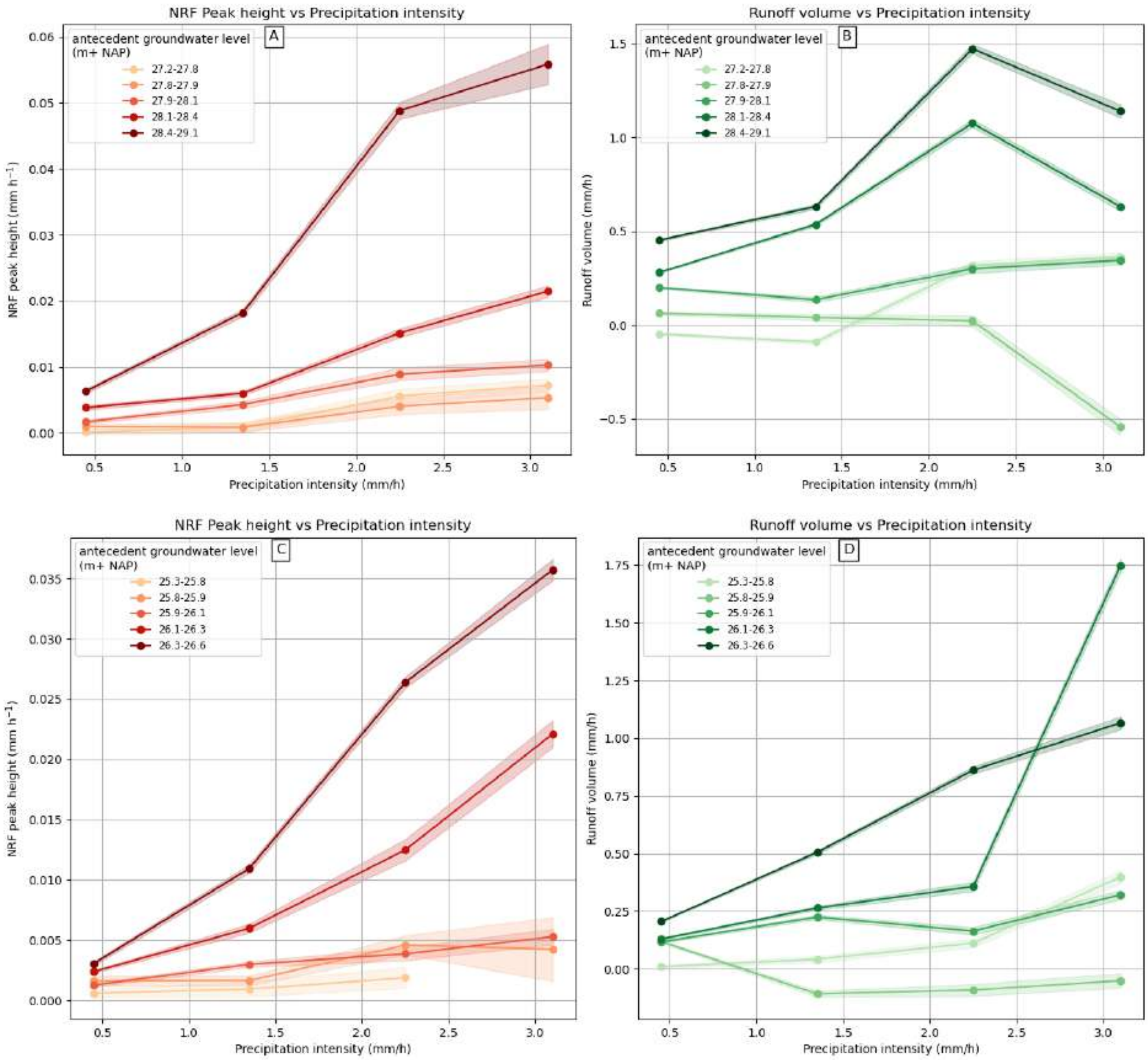


Figure 7: Peak stats for the Hupselse Beek and Ramsbeek. (a) Peak height of the nonlinear response function of the Hupselse Beek. (b) runoff volume response of the at different precipitation intensities of the 5 antecedent groundwater level classes of the Hupselse Beek. (c) Peak height of the nonlinear response function of the Ramsbeek. (d) runoff volume response of the at different precipitation intensities of the 5 antecedent groundwater level classes of the Ramsbeek.

Figure 7 illustrates how the peak response and total runoff volume of catchment response changes with antecedent groundwater level and precipitation intensity. The peak height of the streamflow response of the Hupselse Beek of the 5 antecedent groundwater level classes is plotted against the precipitation intensity (Fig. 7a). The results shows that when the catchment

wetness increases, the peak height of the response also increases. This is especially true for higher precipitation intensities. Comparing the NRF peak height of the two classes with the highest antecedent groundwater levels (28.1–28.4 m+ NAP and 28.4–29.1 m+ NAP), we see that at a precipitation intensity of 1.5 mm/h, the NRF peak height is approximately 0.08 mm/h for the 28.1–28.4 m+ NAP class, and about 0.22 mm/h for the 28.4–29.1 m+ NAP class, more than 2.5 times higher. When the precipitation intensity increases to 3 mm/h, the peak height for the 28.1–28.4 m+ NAP class rises to around 0.20 mm/h — an increase of 0.12 mm/h. In contrast, for the wettest class (28.4–29.1 m+ NAP), the peak height increases to approximately 0.56 mm/h — a rise of 0.34 mm/h. In contrast, for lower groundwater levels (below 27.9 m+ NAP), almost no peak in streamflow is observed following precipitation, even at high precipitation intensities.
These results shows a strong nonlinear behaviour of the Hupselse Beek, under wetter conditions and with higher precipitation intensities, the stream reacts stronger to precipitation.

The same trend is visible in the Ramsbeek (Fig. 7c). With low antecedent groundwater levels (below 26.1 m+ NAP) barely any peak in streamflow is estimated. When the antecedent groundwater level increases, the peak height increases with an increasing precipitation intensity. This peak height is highest for the highest antecedent groundwater level (26.3-26.6 m+ NAP) at approximately 0.035 mm/h at a precipitation intensity of 3 mm/h. It is important to note that the peak height of the Ramsbeek is lower than at the Hupselse Beek across all antecedent wetness ranges and precipitation intensities.

Figure 7b and 7d show the runoff volume generated by one hour of precipitation at a certain precipitation intensity for the 5 antecedent wetness classes for the Hupselse Beek and Ramsbeek respectively. The results show that for both the Hupselse Beek and the Ramsbeek runoff is mainly generated during periods of high groundwater level. The Hupselse Beek barely generates any runoff when the groundwater level is lower than 27.9 m+ NAP, regardless of the precipitation intensity. As groundwater levels increase slightly (27.9–28.1 m+ NAP), some runoff is produced, but runoff volume increases only marginally with higher precipitation intensities. When the groundwater level exceeds 28.1 m+ NAP, a significant rise in runoff volume generated is observed. The amount of runoff generated also increases with precipitation intensity.

A similar trend is visible for the Ramsbeek (Figure 7d), with very little runoff generated at low antecedent groundwater levels (below 26.1 m+ NAP), regardless of the precipitation intensity. When the antecedent groundwater level increases, runoff is generated, increasing with precipitation intensity.

3.3 WALRUS model comparison

The WALRUS model has been specifically developed to accurately simulate catchment behaviour in lowland areas with flat topography. (Brauer et al., 2014; Yan et al., 2016).
One of the catchments that has been studied extensively with the use of the WALRUS model is the Hupselse Beek. It is therefore important to understand if this model can also reproduce the nonlinearity and nonstationarity in the Hupselse Beek as shown by ERRA in our results. To investigate this, the WALRUS model is first run with the groundwater and precipitation timeseries gathered for this study. The model output is shown in appendix figure 1. The WALRUS model performs reasonably well, with a Nash-Sutcliffe efficiency of 0.65, however some of the discharge peaks are not present in the modelled discharge. This modelled discharge is then used in ERRA instead of the previously used observed discharge.

First, we illustrate the runoff response of the Hupselse Beek without accounting for nonlinearity or nonstationarity, using the modelled discharge from the WALRUS model (Fig. 8).

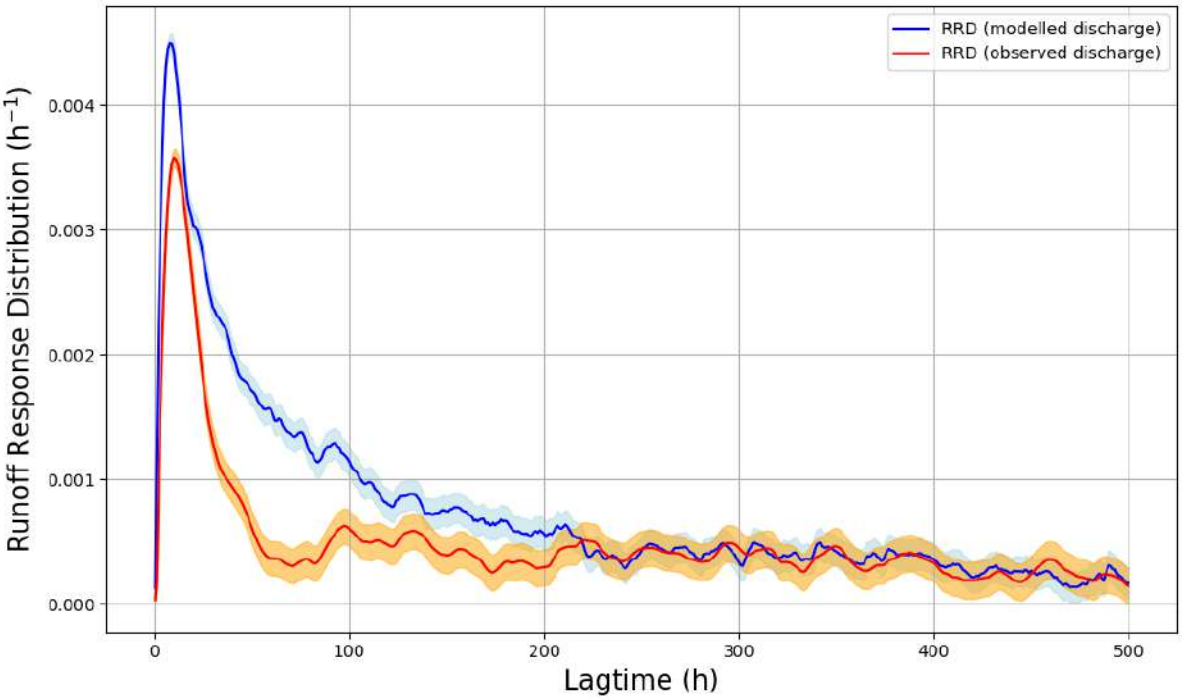


Figure 8: Runoff response distribution of the Hupselse Beek streamflow to one unit of precipitation. In blue the RRDp of ERRA with the observed discharge is shown. In red the RRDp of the modelled discharge obtained with the WALRUS model is shown.

In figure 8, the runoff response distribution of the Hupselse Beek based on observed discharge (blue) is compared to the distribution based on modelled discharge from the WALRUS model (red). The figure shows that when using the modelled discharge from WALRUS, the runoff response to a unit of precipitation is lower compared to that based on observed discharge. The peak height with the modelled discharge is approximately 0.0036 and the peak height with the observed discharge is approximately 0.0044. This means that the WALRUS model underestimates the peak response by approximately 18%.

The runoff coefficient of the runoff response distribution estimated from the modelled discharge is approximately 0.26, indicating that about 26% of precipitation is converted to streamflow when using the discharge simulated by the WALRUS model. In Figure 3, we established that 42% of precipitation was converted to streamflow when using the observed discharge. This represents an underestimation of approximately 38% relative to the streamflow estimated from the observed discharge.

Furthermore, looking at figure 8, the WALRUS model also estimates the peak runoff response to be slightly later. As mentioned before, ERRA ran with observed discharge estimates the peak runoff response to occur after approximately 8.5 hours, however with the modelled discharge, the peak is estimated to occur after approximately 11 hours.

We then divide the data by precipitation classes and antecedent groundwater level again in order to see if the WALRUS model can capture the non-linearity and non-stationarity present in the real world (Fig. 9). To allow for direct comparison, the data is divided in the same classes used for the Hupselse Beek and shown in figure 6.

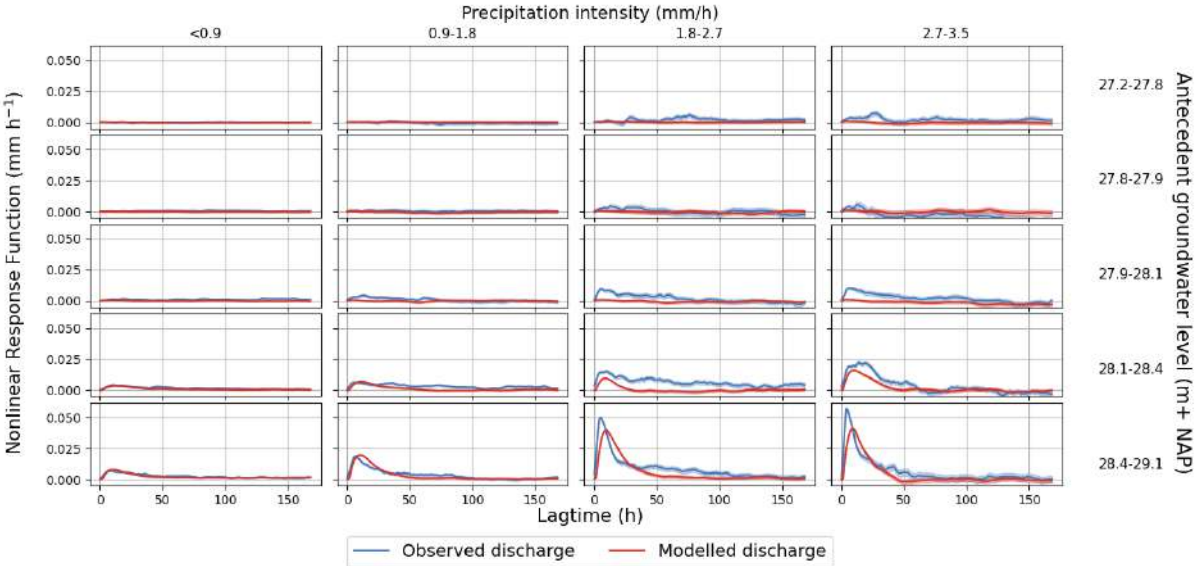


Figure 9: non-linear response function of the Hupselse Beek. In blue, the NRF is shown using the observed discharge. In red the NRF is shown using the modelled discharge obtained from the WALRUS model.

Figure 9 shows a comparison between the NRF obtained with observed discharge (blue) versus the NRF obtained with the modelled discharge (red). We see that the WALRUS model can model non-linearity and non-stationarity. At low precipitation intensity (<1.8 mm/h) and a low antecedent groundwater level (below 27.9 m+ NAP), no streamflow response is visible. When the antecedent groundwater level is increased (>28.1 m+ NAP), some streamflow is generated. Also with the WALRUS model, the largest streamflow response is observed with a high precipitation intensity (2.7-3.5 mm/h) and a high antecedent groundwater level (28.4-29.1 m+ NAP).

It is important to note that the underestimation of the streamflow response by the WALRUS model occurs mainly at higher precipitation intensities (>1.8 mm/h). Figure 8 shows good agreement between the modelled and observed nonlinear response functions for precipitation intensities below 1.8 mm/h across all antecedent groundwater levels. However, as precipitation intensity increases, the WALRUS model increasingly underestimates the runoff response.
This can be quantified using the most extreme scenario, the scenario with the highest precipitation intensity (2.7-3.5 mm/h) and highest groundwater level (28.4-29.1 m+ NAP). In this case, the runoff coefficient derived from the observed NRF is approximately 0.65, whereas the coefficient derived from the modelled NRF is approximately 0.25. This means that for this scenario, the WALRUS model estimates that 25% of the precipitation is converted to streamflow, while ERRA using observed discharge estimates 65% of the precipitation to be converted into streamflow. This is an relative underestimation of approximately 62% compared to the observed runoff response.

These results indicate that the WALRUS model is capable of capturing the nonlinearity and nonstationarity in the Hupselse Beek, but it underestimates the magnitude of these dynamics compared to those observed in real-world data.

4. Discussion

4.1 Average catchment responses

This study shows that the streamflow response for both the Hupselse Beek and the Ramsbeek is quick and flashy. Both catchments show a peak response within the first 12 hours after precipitation under average catchment wetness conditions and average precipitation intensities. The runoff response distributions for both catchments (Figures 3 and 4) indicate that streamflow is initially dominated by quickflow contributions. The groundwater contribution is low shortly after rainfall, but increases with time, eventually becoming the dominant fraction of the streamflow response. This is logical, since quickflow processes are much quicker than groundwater processes, but also decline more quickly (Reitz & Sanford, 2019).

The results show that in the Hupselse Beek, across the entire data series, 42% of precipitation is turned into streamflow. This is very similar to a previous study, which found a runoff coefficient of 39% for the Hupselse Beek (Brauer et al., 2023). For the Ramsbeek this is 22%. The results also show that the Hupselse Beek is more groundwater dominated than the Ramsbeek (71% groundwater contribution vs 40% respectively). One explanation for this could be that the Ramsbeek has a deeper average groundwater table compared to the Hupselse Beek, meaning that the groundwater processes are slower and less important in the Ramsbeek in comparison to the Hupselse Beek. This deeper groundwater table in the Ramsbeek is also a possible explanation for the lower runoff coefficient in the Ramsbeek. A deeper groundwater table leads to more storage space in the unsaturated zone, and therefore more time and space for evapotranspiration of water from the unsaturated zone, leading to a lower runoff coefficient (Eslami et al., 2025). Conversely, shallower groundwater tables provide less space for water retention, leading to rapid saturation of the unsaturated zone, and therefore promoting more runoff (Steenhuis et al., 2005).

4.2 Nonlinear and nonstationary responses of the catchments

This study demonstrates that both the Hupselse Beek and the Ramsbeek respond strongly nonlinear and nonstationary to both precipitation intensity and antecedent groundwater level. Specifically we found that runoff only occurs with higher precipitation intensities and higher antecedent groundwater levels. With high precipitation intensities, the precipitation intensity can exceed the infiltration capacity, leading to overland flow, which will increase the runoff coefficient (Machadeo et al., 2022). This is the reason why in figures 5 and 6, a streamflow response is only generated with high precipitation intensities. With low precipitation intensities, water can infiltrate in the unsaturated zone, where it can be taken up by plant roots, and might not be converted to streamflow. This also explains why the streamflow response to high intensity precipitation is more quickflow dominated. When the infiltration capacity is reached, a large portion of the precipitation will turn into overland flow, and be transported directly in to the stream as quickflow. This increases both the runoff coefficient and the quickflow contribution.

With low groundwater levels, the unsaturated zone is larger, leading to more space for the storage of rainfall, thereby decreasing the runoff coefficient. This is why we don’t see a streamflow response with low antecedent groundwater levels, except with extreme precipitation intensities. With high antecedent groundwater levels, the unsaturated zone is filled quickly, and therefore facilitating saturation excess runoff (Steenhuis et al., 2005).

When antecedent groundwater levels are high and precipitation intensity is low, a small streamflow response is still visible, consisting almost entirely of groundwater. This occurs because the precipitation intensity is too low to exceed the infiltration capacity, so overland

flow does not occur. Instead, the rainfall infiltrates and contributes to groundwater recharge, which slowly makes its way to the stream via subsurface flow.
Figure 7 showed this nonlinearity and nonstationarity in a more visual way, illustrating that for both the Hupselse Beek and the Ramsbeek barely any streamflow response is generated under low precipitation intensities and low groundwater levels.

4.3 WALRUS model performance
When running the WALRUS model with the obtained precipitation and potential evaporation data, a Nash-Sutcliffe efficiency of 0.65 is reached, showing that the WALRUS model can be used fairly well in the Hupselse Beek. However, when applying the modelled discharge timeseries from the WALRUS model in ERRA, our study shows that the WALRUS model does capture nonlinearity and nonstationarity, but it underestimates the magnitude of the streamflow response generated, especially for high precipitation intensities. When comparing the runoff coefficient of the average response to precipitation in the Hupselse Beek with the WALRUS model compared to observational data, our study shows that the WALRUS model underestimates the streamflow response by approximately 38% relative to observational data. When looking at a scenario with very high precipitation and antecedent groundwater levels, our results shows that this underestimation rises to approximately 62%.

These results signify that even though the Nash-Sutcliffe efficiency is relatively high, the WALRUS model still deviates from real world observations. It is debated weather goodness-of-fit tests are sufficient to test model performance (Ritter & Muñoz-Carpena, 2013) . Our study shows that the WALRUS model might not capture the necessary real world precipitation-streamflow relationship needed to accurately model the Hupselse Beek.

ERRA is a good evaluation tool for models, because ERRA can directly compare observational data with model output. Furthermore, ERRA can also quantify the difference between the modelled result and the observational data, and can therefore be very useful in model improvement (Gao et al., 2025).

4.4 Limitations
In this study a catchment average groundwater level was estimated from three groundwater wells per catchment. The results would be more accurate if the density of these groundwater wells was larger, so a more accurate catchment average groundwater level could be calculated. Furthermore, this study was unsuccessful in evaluating the streamflow response composition obtained by the WALRUS model. The ability of the model to reproduce hydrological processes could have been better explored if we could compare the different flow paths in the observational data with the modelled data.
Finally, for this study, the WALRUS model has only been applied to the Hupselse Beek. Our results would have been stronger if we also used the WALRUS model to obtain a modelled discharge time series of the Ramsbeek.

4.5 Future research
Building on this study, future work could expand this study to multiple catchments in different regions and under varying climatological conditions to assess whether similar nonlinear and nonstationary streamflow patterns emerge.
Another interesting direction could be to use ERRA as a model evaluation tool in order to check where models fail to simulate real world dynamics, and how these models can be improved. This could also be interesting with respect to climate change, with an increasing need for accurate flood and drought prediction systems.

5. Conclusion

In conclusion, the RRD's generated in this study suggest that both the Hupselse Beek and the Ramsbeek react fast to a precipitation event. The peak response of the Hupselse Beek occurs after approximately 8.5 hours, and the Ramsbeek reacts slightly slower, with a peak runoff response after 11 hours. The composition of these catchments is very different however. The Hupselse Beek is a groundwater dominated system, with 71% of the runoff response to a unit of precipitation consisting of groundwater and 29% consisting of quickflow. The Ramsbeek is not groundwater dominated. The runoff response to one unit of precipitation consists for approximately 40% of groundwater and 60% of quickflow.

When we look at the composition of the runoff response over time, we can see that the response generated in the first couple of hours is quickflow dominated, and that the fraction of groundwater input increases as time passes.

The results show that both catchments behave non-linearly and non-stationary, with almost no runoff response generated at low precipitation intensities (<0.9 mm/h) and low antecedent groundwater levels. Runoff response increases with increasing precipitation intensity and antecedent groundwater level.

In the Hupselse Beek, runoff response with low precipitation intensity and high antecedent groundwater is groundwater dominated, while runoff response with high precipitation intensity and antecedent groundwater level is mainly quickflow dominated. The Ramsbeek, which is less groundwater dominated on average, shows a more quickflow-heavy response even under wetter conditions.

The WALRUS model, applied to the Hupselse Beek, successfully captures the general nonlinear and nonstationary patterns in streamflow response. However, it underestimates the magnitude of runoff, particularly under high precipitation intensities. On average, WALRUS underestimates streamflow response by about 38%, and this underestimation increases to 62% under extreme conditions. These results indicate that, although the model performs reasonably well (NSE = 0.65), it does not fully replicate real-world runoff dynamics, particularly during high-flow events.

Acknowledgements

I would like to thank my supervisor Wouter Berghuijs for the feedback and the constant help for the duration of my thesis. Furthermore I would like to thank James Kirchner and Wouter Knoben for the feedback midway during my thesis, and for providing me with directions to take the thesis into. Lastly I want to thank Claudia Brauer for providing me with the potential evaporation data, and for developing the WALRUS model used in my study.

Sources

- Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, *34*(26), 5699–5703. https://doi.org/10.1002/hyp.13958

- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

- Beven, Keith J. *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2012.

- Beven, K. J. (2018). On hypothesis testing in hydrology: Why falsification of models is still a really good idea. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1278.

- Bormann H., & Elfert S. (2010). Application of WaSiM-ETH model to Northern German lowland catchments: model performance in relation to catchment characteristics and sensitivity to land use change. *Advances in Geosciences*, *27*, 1–10. https://doi.org/10.5194/adgeo-27-1-2010

- Botter, G., Basso, S., Rodriguez-Iturbe, I., & Rinaldo, A. (2013). Resilience of river flow regimes. *Proceedings of the National Academy of Sciences*, *110*(32), 12925–12930. https://doi.org/10.1073/pnas.1311920110

- Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., & 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. https://doi.org/10.1002/wrcr.20320

- Brauer C. C., Teuling A. J., Torfs P. J. J. F., & Uijlenhoet R. (2014). The Wageningen Lowland Runoff Simulator (WALRUS): a lumped rainfall-runoff model for catchments with shallow groundwater. *Geoscientific Model Development*, *7*(5), 2313–2332. https://doi.org/10.5194/gmd-7-2313-2014

- Brauer C. C., Van der Velde Y., Teuling A. J., & Uijlenhoet R. (2018). The Hupsel Brook Catchment: Insights from Five Decades of Lowland Observations. *Vadose Zone Journal*, *17*(1). https://doi.org/10.2136/vzj2018.03.0056

- Chang, D., Li, S., & Lai, Z. (2023). Effects of extreme precipitation intensity and duration on the runoff and nutrient yields. *Journal of Hydrology*, *626*. https://doi.org/10.1016/j.jhydrol.2023.130281

- Dion, P., Martel, J. L., & Arsenault, R. (2021). Hydrological ensemble forecasting using a multi-model framework. *Journal of Hydrology*, *600*, 126537.

- Dooge, J. C. I. (1959). A general theory of the unit hydrograph. *Journal of Geophysical Research*, *64*(2), 241–256. https://doi.org/10.1029/JZ064i002p00241

- Eslami, Z., Seybold, H., & Kirchner, J. W. (2025). Climatic, topographic, and groundwater controls on runoff response to precipitation: evidence from a large-sample data set. *EGUsphere*, *2025*, 1-15.

- Gao, H., Ju, Q., Zhang, D., Wang, Z., Hao, Z., & Kirchner, J. W. (2025). Quantifying Dynamic Linkages Between Precipitation, Groundwater Recharge, and Streamflow Using Ensemble Rainfall-Runoff Analysis. Water Resources Research, 61(1). https://doi.org/10.1029/2024WR037821

- Geertsema, T. J., Torfs, P. J. J. F., Eekhout, J. P. C., Teuling, A. J., & Hoitink, A. J. F. (2020). Wood-induced backwater effects in lowland streams. *River Research and Applications*, *36*(7), 1171–1182. https://doi.org/10.1002/rra.3611

- Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., & Wisser, D. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences*, *111*(9), 3251–3256. https://doi.org/10.1073/pnas.1222475110

- Healy, R. W., & Cook, P. G. (2002). Using groundwater levels to estimate recharge. *Hydrogeology Journal*, *10*(1), 91–109. https://doi.org/10.1007/s10040-001-0178-0

- Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, *29*(8), 2637–2649. https://doi.org/10.1029/93WR00877

- James, W., & Johanson, R. C. (1999). A note on an inherent difficulty with the unit hydrograph method. *Journal of Water Management Modeling*. https://doi.org/10.14796/JWMM.R204-01

- Khatami S., Peel M.C., Peterson T.J., & Western A.W. (2019). Equifinality and Flux Mapping: A New Approach to Model Evaluation and Process Representation Under Uncertainty. *Water Resources Research*, *55*(11), 8922–8941. https://doi.org/10.1029/2018WR023750

- Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, *45*(2). https://doi.org/10.1029/2008WR006912

- Kirchner, J. W. (2022). Impulse Response Functions for Nonlinear, Nonstationary, and Heterogeneous Systems, Estimated by Deconvolution and Demixing of Noisy Time Series. *Sensors (Basel, Switzerland)*, *22*(9). https://doi.org/10.3390/s22093291

- Kirchner, J. W., Benettin, P., & van Meerveld, I. (2023). Instructive surprises in the hydrological functioning of landscapes. *Annual Review of Earth and Planetary Sciences*, *51*(1), 277-299.

- Kirchner, J.W. (2024). Characterizing nonlinear, nonstationary, and heterogeneous hydrologic behavior using ensemble rainfall-runoff analysis (ERRA): proof of concept. *Hydrology and Earth System Sciences*, *28*, 4427–4454. https://doi.org/10.5194/hess-28-4427-2024

- Knapp, J. L., Berghuijs, W. R., Floriancic, M. G., & Kirchner, J. W. (2024). Catchment hydrological response and transport are affected differently by precipitation intensity and antecedent wetness. *Hydrology and Earth System Sciences Discussions*, *2024*, 1-19.

- Machado, R. E., Cardoso, T. O., & Mortene, M. H. (2022). Determination of runoff coefficient (C) in catchments based on analysis of precipitation and flow events. *International Soil and Water Conservation Research*, *10*(2), 208–216. https://doi.org/10.1016/j.iswcr.2021.09.001

- Mathias, S. A., McIntyre, N., & Oughton, R. H. (2016). A study of non-linearity in rainfall-runoff response using 120 UK catchments. *Journal of Hydrology*, *540*, 423-436.

- McDonnell, J. J., Spence, C., Karran, D. J., Van Meerveld, H. J., & Harman, C. J. (2021). Fill-and-spill: A process description of runoff generation at the scale of the beholder. *Water Resources Research*, *57*(5), e2020WR027514.

- Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., ... & Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, *2*(9), 592-609.

- Nippgen, F., McGlynn, B. L., Emanuel, R. E., & Vose, J. M. (2016). Watershed memory at the Coweeta Hydrologic Laboratory: The effect of past precipitation and storage on hydrologic response. *Water Resources Research*, *52*(3), 1673–1695. https://doi.org/10.1002/2015WR018196

- Perrini, P., Iacobellis, V., Gioia, A., Cea, L., Savenije, H. H. G., & Fenicia, F. (2025). Can Dominant Runoff Generation Mechanisms Be Disentangled Through Hypothesis Testing? Insights From Integrated Hydrological-Hydrodynamic Modeling. *Water Resources Research*, *61*(4). https://doi.org/10.1029/2024WR039394

- Reitz, M., & Sanford, W. E. (2019). Estimating quick-flow runoff at the monthly timescale for the conterminous United States. *Journal of Hydrology*, *573*, 841–854. https://doi.org/10.1016/j.jhydrol.2019.04.010

- Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, *480*, 33–45. https://doi.org/10.1016/j.jhydrol.2012.12.004

- Royal Netherlands Meteorological Institute (KNMI). (2025). Hourly precipitation data (2012–2024) [Data set]. KNMI. https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens

- Steenhuis, T.S., Agnew, L., G´erard-Marchant, P., Walter, M.T.: Overland flow. In: Hillel, D. (Ed.), Encyclopedia of Soils in the Environment, Elsevier., 130–133, https://doi.org/10.1016/B0-12-348530-4/00568-3, 2005.

- Tromp-van Meerveld, H. J., & McDonnell, J. J. (2006). Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water resources research*, *42*(2).

- Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews Water*, *2*(4), 359–392.

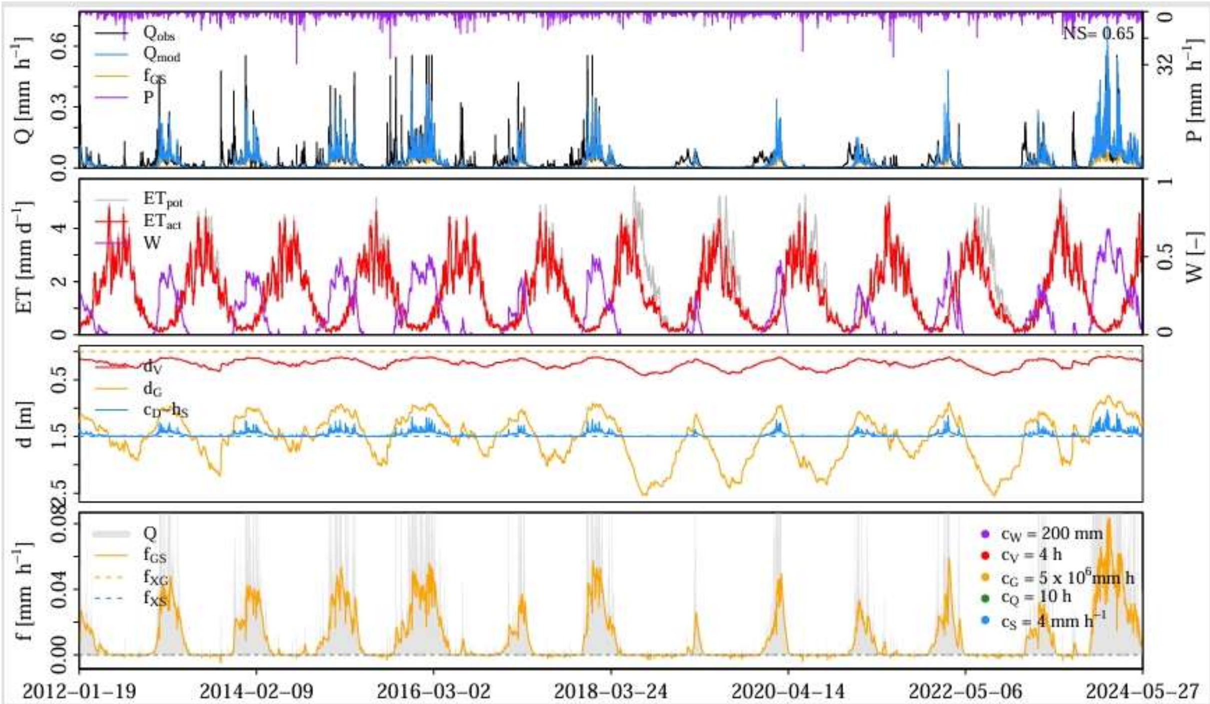
- Waterschap Rijn en IJssel. (2024). Hourly discharge data for the Hupselse Beek (2012–2024) [Data set]. https://waterdata.wrij.nl/index.php?wat=overzicht&type=alle

- Yan R., Gao J., & Li L. (2016). Modeling the hydrological effects of climate and land use/cover changes in Chinese lowland polder using an improved WALRUS model. *Hydrology Research*, *47*(S1), 84–101. https://doi.org/10.2166/nh.2016.204

Appendix

Model Parameters	
Wetness index parameter (cW)	100 (mm)
Vadose zone relaxation time (cV)	4 (h)
Groundwater reservoir constant (cG)	30e6 (mm h)
Quickflow reservoir constant (cQ)	80 (h)
Surface water parameter (cS)	10 (mm/h)
Catchment Characteristics	
fraction of the initial discharge originating from drainage (Gfrac)	0.4
Chanel depth (cD)	1500 (mm)
Surface water area fraction (aS)	0.01
Soil type (st)	Loamy sand

Appendix table 1: Table showing the parameters and catchment characteristics used in the WALRUS model



Appendix figure 2: Output of the WALRUS model of the Hupselse Beek. The figure shows a relatively good match between the observed discharge and the modelled discharge (NS = 0.65).