https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Simulating Subsurface Stormflow Across Catchments: Parameter Estimation and Uncertainties

Tamara Leins[1], Nikolai Späth[2], Jan Seibert[3], Francesca Pianosi[4], Christian Reinhardt-Imjela[2], Andreas Hartmann[1]

[1]Institute of Groundwater Management, TU Dresden, Dresden, Germany
[2]Institute of Geographical Sciences, FU Berlin, Berlin, Germany
[3]Department of Geography, University of Zurich, Zurich, Switzerland
[4]Department of Civil Engineering, University of Bristol, Bristol, United Kingdom

Correspondence to: Tamara Leins (tamara.leins@tu-dresden.de)

Abstract. Subsurface Stormflow (SSF) is an important runoff-generation process, especially in temperate and mountainous regions. It can play a substantial role in catchment-scale flood generation and nutrient or contaminant transport. Due to process heterogeneity and subsurface occurrence, it is difficult to be quantified in the field. Since hydrological models are often calibrated using only discharge data at the catchment outlet, SSF-relevant parameters are often affected by equifinality, leading to large uncertainties in simulated SSF. In this study, we used the widely used bucket-type HBV model to analyse simulated SSF uncertainty across four catchments. We performed Monte Carlo simulations with 1,000,000 parameter sets and selected the top 5% as behavioural. We subsequently analysed SSF parameter sensitivity in a regional sensitivity analysis and SSF simulation uncertainty in a GLUE-type approach. We examined whether more information on SSF is hidden in specific discharge conditions using a parameter-estimation approach based on percentiles of the flow-duration curve. In addition, we investigated simulated SSF uncertainty by disentangling simulated SSF volume and occurrence uncertainties. We found that there are high uncertainties in simulated SSF volume and occurrence across all catchments, with mostly low sensitivity of SSF-relevant parameters. We further found that SSF parameter identifiability could be improved by percentile parameter estimation, and more so in snow-insensitive catchments than in snow-dominated catchments. However, SSF simulation uncertainty still remained high. In the SSF uncertainty decomposition we could show that SSF occurrence uncertainty contributes a substantial part to total SSF simulation uncertainty. This study therefore emphasises the need to include SSF data into model calibration to ensure a realistic simulation of SSF and thereby achieve better process representation in lumped hydrological models and thus yield more robust prediction models. Our findings suggest that SSF proxy data indicating SSF occurrence in the catchment might already help to reduce SSF simulation uncertainty.

1 Introduction

Subsurface Stormflow (SSF) is an important runoff-generation process, especially in temperate and mountainous regions (Barthold and Woods, 2015; Weiler et al., 2006; McMillan et al., 2025). It occurs when there is a contrast in subsurface

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

hydraulic conductivities, with a higher-permeability layer overlaying a lower-permeability layer, e.g., the bedrock or a less permeable soil layer. Water is then (partly) deflected to flow laterally along the slope (Chifflard et al., 2019; Blume and van Meerveld, 2015). SSF can also occur when groundwater rises into more permeable layers and is laterally transported to the stream ("transmissivity feedback") (Bishop et al., 1990). SSF can play an important role in catchment-scale runoff generation, especially in flood generation (Zillgens et al., 2007; Markart et al., 2015; McGuire et al., 2024) as well as in nutrient and contaminant transport (Zhao et al., 2013; Wang et al., 2024).

As SSF is a heterogeneous and subsurface process, its observation is difficult and requires extensive instrumentation. Studies that measured SSF are often limited to only one or a few sites or hillslopes (e.g. Freer et al., 2002; Uchida et al., 2004; Tromp-van Meerveld and McDonnell, 2006; Du et al., 2016). Despite attempts to synthesise experimental findings across sites worldwide (e.g., McMillan et al., 2025), quantitative inter-site comparisons remain challenging because different experimental methods are hard to compare and the resulting experimental data are often not publicly available. This makes it difficult to establish general rules about SSF generation and its controlling factors, beyond the scale of a few studied hillslopes (Chifflard et al., 2019; Blume and van Meerveld, 2015). In addition, while hillslope-scale models have been shown to be useful for understanding SSF processes at the local scale, catchment-scale models are still needed to extrapolate process understanding to a larger scale (Chifflard et al., 2019; Blume and van Meerveld, 2015).

Hydrological models at the catchment scale can be classified into different groups according to their model-internal representation of hydrological processes: blackbox, greybox (conceptual/ bucket-type models) and whitebox (physically-based models). In addition, they can be classified according to their spatial discretisation into lumped and distributed models. Lumped models represent the catchment as a single zero-dimensional unit and thus cannot explicitly account for the spatial heterogeneity of hydrological properties across the catchment. However, the idea is that they can still represent the dominant hydrological processes in the catchment (Wagener and Gupta, 2005; Hartmann et al., 2023). SSF implementations in different hydrological models range from linear reservoirs with outflow triggered by a threshold (e.g. HBV, Lindström et al., 1997, Seibert and Vis, 2012) to complex physical approaches such as the Richards equation (e.g. WaSiM-ETH, Schulla and Jasper, 2007). In addition, especially in bucket-type models, linking model-internal fluxes to real-world hydrological processes might not be entirely obvious.

Parameters of hydrological catchment-scale models are usually calibrated, because the catchment-scale parameters have effective values, which summarise the subsurface heterogeneity of catchments or sub-catchments and as such cannot be measured in the field (Kirchner, 2006). This calibration process is usually based solely on discharge observations at the catchment outlet, given the limited availability of other flux and storage observations. However, the information content of these discharge time series is limited (Wheater et al., 1986; Ye et al., 1997), and can be further reduced when calibration is based on an overall performance measure, that aggregates model residuals over the entire calibration period (Wagener et al., 2003). This can lead to equifinality of calibrated models, where different parameterisations produce very similar discharge simulations while the model-internal fluxes vary considerably (Beven and Binley, 1992). In particular, the distinction among model-internal lateral streamflow components is typically affected by large uncertainties (Seibert and McDonnell, 2002).

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Models that can simulate discharge correctly, but cannot represent the dominant hydrological processes, might "give us the right results, but for the wrong reasons", making their predictions outside calibration conditions questionable (Kirchner, 2006). Therefore, a reliable process representation in model structures and parameterisations is highly important, for instance, for the simulation of future conditions or for flood modelling (Brunner et al., 2021).

All the above discussion also applies to SSF-relevant parameters. Therefore, it remains unclear how accurately and precisely existing catchment-scale hydrological models can simulate the volume and occurrence of SSF. Some studies have sought to improve SSF representation in hydrological models, though most rely on extensive field measurements at individual hillslopes or in small catchments. For instance, Hopp and McDonnell (2009) employed a three-dimensional, physically-based finite element model of a well-studied hillslope, calibrated against trench-flow data, to investigate dominant controls on SSF generation. Shen et al. (2025) refined SSF simulations in the Penn State Integrated Hydrological Model (PIHM) by improving the representation of preferential flow in macropores, calibrating the model using extensive observations, including trench-flow data from a zero-order catchment. Scanlon et al. (2000) modified TOPMODEL to more realistically simulate SSF generation from transient perched water tables, using piezometer data to detect the formation of saturated conditions above the groundwater table. Again relying on trench-flow measurements for calibration, Song et al. (2024) developed the Three-Stage Subsurface Stormflow-based Model (TSSM) to simulate SSF-related flash floods at the hillslope and micro-watershed scale. While these site-specific studies have advanced process understanding, they require observational infrastructure that is rarely available at broader scales.

In contrast, only a few studies have examined model-internal SSF uncertainties when calibration relies solely on discharge data. Hughes and Farinosi (2021) assessed uncertainties of hydrological processes in the semi-distributed Pitman model. They found that interflow contributions varied substantially across sub-basins, with uncertainty ranges spanning up to 50% in some cases. Hartmann et al. (2023) similarly reported large uncertainties in simulated streamflow contributions, including SSF, when using only discharge data for calibration. In their study, they found that incorporating hydrograph-separation-derived information improved the precision of simulated flow components. Despite these advances, studying and simulating SSF at the catchment scale beyond individual, data-intensive case studies remains challenging (McGuire et al., 2024; Chifflard et al., 2019).

In this study, we explored the limits of SSF parameterisation in lumped hydrological models across study sites when using discharge data alone for the parameter estimation. We investigated the following research questions: 1) Can parameter estimation of a lumped model be improved by using specific percentiles of the flow-duration curve instead of the entire discharge dataset? 2) How large is the model-internal SSF simulation uncertainty arising from parameter equifinality, and is this uncertainty dominated by occurrence or volume components? To do so we used the HBV model, a bucket-type model at the catchment scale, to simulate SSF and catchment discharge at four study sites in Central Europe. We divided discharge data according to the flow-duration curve (FDC) percentiles and performed the parameter estimation several times, each time using only discharge data within a single FDC percentile. Using a Monte Carlo based parameter estimation, we selected the same number of behavioural parameter sets for each FDC percentile based on the Kling-Gupta-Efficiency as an objective function.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Applying a regional sensitivity analysis combined with a GLUE-type uncertainty estimation, we analysed and compared the parameter sets and discharge and SSF simulations of the different percentile parameter estimations. According to the Law of Total Variance, we divided the total SSF simulation uncertainty into SSF volume uncertainty and SSF occurrence uncertainty. Using this framework, we were able to analyse the identifiability of SSF-relevant parameters across the study sites and FDC percentiles and to disentangle SSF simulation uncertainties.

2 Methods

In our approach, we used the widely used bucket-type HBV model to analyse simulated SSF uncertainty in a lumped hydrological model across four catchments. First, we examined whether more information on SSF is hidden in specific discharge conditions using a parameter-estimation approach based on percentiles of the flow-duration curve. Second, we further investigated simulated SSF uncertainty by disentangling simulated SSF volume and simulated SSF occurrence and their respective shares in total simulated SSF uncertainty (Figure 1). To do so, Monte Carlo simulations with 1,000,000 parameter sets were performed for each catchment. We estimated the behavioural model parameters by selecting the top 5% of parameter sets based on the fit of simulated to observed discharge, as measured by KGE (Gupta et al., 2009). This parameter estimation procedure was repeated for 1) all available discharge measurements ("global parameter estimation") and 2) the percentiles of the flow-duration curve in 10% increments ("percentile parameter estimation") by using only the fragments of the observed discharge timeseries that have an exceedance probability between, e.g., 10% and 20% in the 10-20% percentile parameter estimation. In this way, for each catchment 11 sets of top 5% behavioural parameter sets were produced. These were then used in a regional sensitivity analysis to determine SSF parameter sensitivities, globally and for each percentile. The percentile sensitivities were also attributed to the times of the respective percentile flows in the observed discharge timeseries to unveil temporal sensitivity patterns. For each catchment, the percentile parameter estimation which yielded the largest increase in SSF parameter sensitivity was selected for the following uncertainty analysis, along with the global parameter estimation. Simulated SSF uncertainty and simulated SSF occurrence probability were derived from a GLUE-type uncertainty estimation. In addition, the shares of SSF volume uncertainty and SSF occurrence uncertainty of total simulated SSF uncertainty were computed according to the Law of Total Variance.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

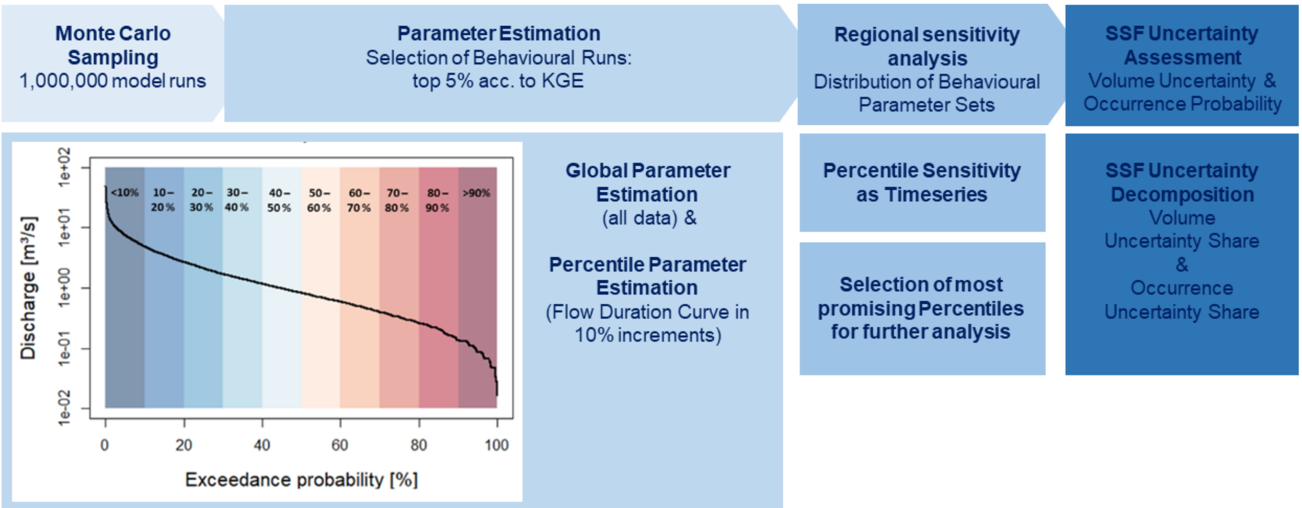


Figure 1 Methodology Flowchart

2.1 Hydrological simulation model, parameter and uncertainty estimation

For our analysis, we used the HBV model (Bergström, 1976, 1992; Lindström et al., 1997) in the version HBV-light (Seibert and Vis, 2012) as implemented in the SAFE toolbox (Pianosi et al., 2015), a widely used bucket-type hydrological model here run at hourly temporal resolution. In the lumped model structure (Figure 2), catchment discharge is simulated from input timeseries of precipitation, air temperature and potential evapotranspiration.

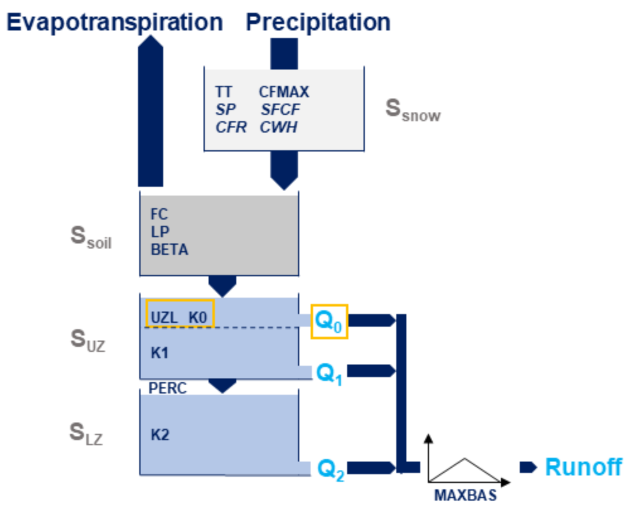


Figure 2 HBV structure with storage variables in grey, parameters in dark blue and fluxes in light blue. SSF (Q_0) and SSF-relevant parameters (UZL and K_0) are highlighted in yellow. Adapted from Seibert (2000)

In the HBV model, the snow routine is simulated using the degree-day method (Lindström et al., 1997). Snowmelt $M(t)$ [mm] is calculated based on the degree-day factor $CFMAX$ [mmd^{-1}°C^{-1}] and the threshold temperature TT [°C] according to equation

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

(1). TT separates precipitation into rain or snow depending on the current air temperature $T(t)$ [°C]. The snowpack retains rain and meltwater until the water holding capacity CWH [-] is exceeded. According to the refreezing coefficient CFR [-] the amount of refreezing liquid water within the snowpack R [mmd^{-1}] is computed by equation (2).

$$M(t) = P_{CFMAX} * (T(t) - P_{TT}) \tag{1}$$

$$R(t) = P_{CFR} * P_{CFMAX}(P_{TT} - T(t)) \tag{2}$$

In the soil moisture routine, a function of the ratio between soil water storage $S_{soil}(t)$ [mm] and the maximum soil moisture storage (field capacity, FC [mm]) determines the groundwater recharge $F(t)$ [mmd^{-1}] to the upper groundwater box $S_{UZ}(t)$ [mm] depending on the incoming water to the soil box $I(t)$ [mmd^{-1}], consisting of rain and snowmelt $M(t)$ (equation 3). Actual evapotranspiration $E_{act}(t)$ [mm] from the soil is computed from the potential evapotranspiration $E_{pot}(t)$ [mm] according to equation (4) (Seibert and Vis, 2012).

$$\frac{F(t)}{I(t)} = \left(\frac{S_{soil}(t)}{P_{FC}}\right)^{P_{BETA}} \tag{3}$$

$$E_{act}(t) = E_{pot}(t) * \min\left(\frac{S_{soil}(t)}{P_{FC} * P_{LP}}, 1\right) \tag{4}$$

With $BETA$ [-] Shape coefficient of the recharge function

LP [-] Soil moisture value above which $E_{act}(t)$ reaches $E_{pot}(t)$

The HBV model structure contains two groundwater boxes, the upper groundwater box with storage $S_{UZ}(t)$ [mm] and the lower groundwater box with storage $S_{LZ}(t)$ [mm], with percolation at the maximum rate $PERC$ [mmd^{-1}] connecting them. The groundwater boxes produce runoff $Q_{GW}(t)$ [mm] according to linear outflow equations with storage coefficients K_2 [d^{-1}] for the lower zone and K_1 for the upper zone. Additional outflow is produced from the upper groundwater box according to storage coefficient K_0 [d^{-1}], if $S_{UZ}(t)$ is above a threshold value UZL [mm] (equation 5). In the routing routine, the sum of all groundwater outflows is transformed using a triangular weighting function with routing parameter $MAXBAS$ [d] to simulate total catchment discharge $Q_{sim}(t)$ [mmd^{-1}] (equations 6 & 7) (Seibert and Vis, 2012).

$$Q_{GW}(t) = P_{K_2} * S_{LZ}(t) + P_{K_1} * S_{UZ}(t) + P_{K_0} * \max\left(S_{UZ}(t) - P_{UZL}, 0\right) \tag{5}$$

$$Q_{sim}(t) = \sum_{i=1}^{P_{MAXBAS}} c(i) * Q_{GW}(t - i + 1) \tag{6}$$

Where

$$c(i) = \int_{i-1}^{i} \frac{2}{P_{MAXBAS}} - \left|u - \frac{P_{MAXBAS}}{2}\right| * \frac{4}{P_{MAXBAS}^2} du \tag{7}$$

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

All HBV parameters along with their respective calibration ranges according to literature (Usman et al., 2024; Seibert, 1999; Abebe et al., 2010; Seibert and Vis, 2012; Bergström, 1992) are displayed in Table 1. For better comparability with literature values, the parameter values are displayed in units matching a daily temporal resolution. However, they were adjusted to an hourly temporal resolution within our model simulations.

In our analysis, we considered SSF to be conceptually represented through the overflow reservoir of the upper zone storage (Q_0) in the HBV model structure (Figure 2), making UZL and K_0 the SSF-relevant parameters of the model. However, as the HBV model structure is only a simplified representation of the real-world catchment, it is not clear which variable in the lumped model is the best proxy for the SSF process. Conceptually, SSF could also be represented as the combination of both fluxes from the upper zone storage ($Q_0 + Q_1$). To assess whether this ambiguity influences our main results, we performed a subsequent uncertainty analysis using both conceptual process allocations for the global parameter estimation.

Table 1 HBV parameter descriptions with calibration ranges. Parameter ranges are displayed in daily values, but were converted within the model simulations to fit an hourly timestep.

Routine	Parameter	Unit	Description	Calibration Range	
				Lower	Upper
Snow routine	TT	[°C]	Threshold temperature rain-snow	-2	2
	$CFMAX$	[$mmd^{-1}°C^{-1}$]	Degree-day factor	0.5	4
	$SFCF$	[-]	Snowfall correction factor	fixed (*)	
	CWH	[-]	Water holding capacity	fixed (*)	
	CFR	[-]	Refreezing coefficient	fixed (*)	
Soil moisture routine	FC	[mm]	Maximum soil moisture storage	50	550
	LP	[-]	Soil moisture value above which E_{act} reaches E_{pot}	0.3	1
	$BETA$	[-]	Shape coefficient of the recharge function	1	6
Response function	$PERC$	[mmd^{-1}]	Threshold parameter: Maximum percolation rate from the upper to the lower groundwater box	0	4
	UZL	[mm]	Threshold parameter: Activation of a second outflow from the upper groundwater box	0	70
	K_0	[d^{-1}]	Upper zone overflow storage coefficient	0.05	0.5
	K_1	[d^{-1}]	Upper zone storage coefficient	0.01	0.3
	K_2	[d^{-1}]	Lower zone storage coefficient	0.0001	0.1
Routing routine	$MAXBAS$	[d]	Length of triangular weighting function	1	2.5

* fixed according to literature values (as described below): Sauerland $\boldsymbol{SFCF}$ = 1, $\boldsymbol{CWH}$ = 0.0113, $\boldsymbol{CFR}$ = 0.0043; Ore Mountains $\boldsymbol{SFCF}$ = 1, $\boldsymbol{CWH}$ = 0.004, $\boldsymbol{CFR}$ = 0.0115; Black Forest $\boldsymbol{SFCF}$ = 1, $\boldsymbol{CWH}$ = 0.0044, $\boldsymbol{CFR}$ = 0.0113; Tyrolean Alps $\boldsymbol{SFCF}$ = 1, $\boldsymbol{CWH}$ = 0.0039, $\boldsymbol{CFR}$ = 0.0112 (Beck et al., 2020)

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

To reduce parameter interactions, the snow parameters CWH, CFR and $SFCF$ were fixed to the values derived in a global HBV parameter regionalisation by Beck et al. (2020). The snowfall correction factor $SFCF$, which is multiplied with all precipitation simulated as snow, was set to 1 across all catchments according to Beck et al. (2020). Although SFCF can be an influential parameter in alpine and forested catchments, a value of 1 implies no correction to the precipitation input and thus preserves the original water balance. We retained this fixed value across all catchments to avoid introducing water balance adjustments and to maintain consistency with the regionalised parameter set. All remaining HBV model parameters were calibrated via a Monte Carlo analysis. We used Latin-Hypercube sampling, as implemented in the SAFE toolbox (Pianosi et al., 2015), to draw 1,000,000 parameter sets from uniform distributions with ranges given in Table 1. The performance of each parameter set was evaluated based on the resulting model's ability to reproduce discharge measurements at the catchment outlets, as measured by the Kling-Gupta-Efficiency KGE (Gupta et al., 2009). We chose not to perform a split-sample test to maximise the information content of the available discharge timeseries, especially since the discharge timeseries needed to be split into 10 parts for the percentile parameter estimation. Studies have shown that this type of calibration, which skips model validation, can lead to more robust results (Shen et al., 2022). Prior to the calibration period, a warm-up period of at least 3 years was added, during which the state variables within the model can evolve, so initial conditions do not influence the model calibration.

The top 5% best performing parameter sets in each catchment were considered to be behavioural and used for simulation in further analysis in a generalised likelihood uncertainty estimation (GLUE) (Beven and Binley, 1992) -type approach. In this way, the equifinality of different parameter sets was recognised in the parameter estimation (Beven et al., 2008). All behavioural parameter sets were used to generate uncertainty bands for simulated model outputs. This uncertainty estimation was applied for simulated total discharge as well as for simulated model-internal flow components (e.g., Q_0, in Figure 2).

The behavioural threshold (top 5%) was set relatively low (still yielding 50,000 parameter sets per catchment), compared to other studies, e.g. Tibangayuka et al. (2022), where the threshold for behavioural runs was fixed at KGE>0.65. This threshold was chosen to leave some margins for further analysis and to account for minor data quality limitations (Beven, 2006) in some catchments, where, e.g., discharge observations were not quality-controlled over the entire time series or meteorological input data were derived from stations located outside the catchment boundaries. The threshold for behavioural runs was set as a percentage of all runs to get the same number of behavioural runs in every catchment. In addition, a fixed KGE threshold can work well for model calibration at one site but is dependent on flow variability and might be misleading when compared between sites (Williams, 2025). As a lower benchmark for the model performances in each catchment, we computed the efficiency of a random parameter set for each catchment, as proposed by Seibert et al. (2018).

2.2 Parameter identifiability across FDC percentiles

We assessed the results of the Monte Carlo parameter estimation using the regional sensitivity analyses (RSA) approach first proposed by Hornberger, Spear & Young (Hornberger and Spear, 1981; Spear and Hornberger, 1980) to examine the influence of the model parameters on model output and performance. In the RSA, the cumulative probability density function of the

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

behavioural parameter sets is plotted over the respective parameter calibration ranges. In these plots, parameter sensitivity is indicated by a strong deviation of a parameter's cumulative probability density function from the 1:1 line (Beven and Freer, 2001). This deviation can be quantified using the Kolmogorov-Smirnov test (KS-test), which determines the maximum distance (D) between the behavioural parameters' cumulative probability density function and the uniform distribution (1:1 line) according to equation 8 (Massey, 1951; Smirnov, 1939; Kolmogorov, 1933).

$$D = \max\left(F_{0\,(x)} - S_{N\,(x)}\right) \tag{8}$$

With $F_{0\,(x)}$ Prior cumulative distribution function of the parameters, i.e., uniform distribution

$S_{N\,(x)}$ Cumulative probability density function of the behavioural parameters

To assess whether we can gain more information on SSF from different parts of the hydrograph, we performed the parameter estimation using both all available discharge data ("global parameter estimation") and subsets corresponding to different percentiles of the flow-duration curve FDC ("percentile parameter estimations") (Figure 1). For percentile parameter estimations, discharge data were divided into 10 percentiles according to the FDC, e.g., those with exceedance probabilities <10%, 10-20%, 20-30%, etc. The parameter-estimation approach described in section 2.1 was then applied separately to each FDC percentile by calculating model performance (KGE) using only discharge observations within the respective exceedance probability range. For instance, in the 10-20% percentile parameter estimation, simulated discharge was evaluated exclusively against observed values with exceedance probabilities between 10% and 20%. This procedure yielded 11 sets of behavioural parameter sets per catchment: one from the global parameter estimation and one from each of the 10 percentile parameter estimations. RSA was then applied to each of these datasets. To gain more information on SSF, we selected the percentile parameter estimation that produced the largest increase in sensitivity indices (eq. 8) of SSF-relevant parameters (UZL and K_0) for further analysis in each catchment.

For a better visual representation of the changes in parameter sensitivity through the percentile parameter estimations, the sensitivity values were additionally mapped onto a continuous timeseries. More specifically, for every model parameter, the D-value from the KS-test (sensitivity measure) obtained in a given percentile parameter estimation was assigned to all time steps at which observed discharge lies within that percentile's exceedance probability range, revealing periods during which specific parameters exert greater influence on model performance.

2.3 Estimation and decomposition of simulated SSF uncertainty

The estimation of simulated SSF uncertainties was divided into two parts: SSF volume uncertainty and SSF occurrence probability. Both uncertainties were assessed using a GLUE-type approach (Beven and Binley, 1992), in which all behavioural parameter sets were used for simulation, leading to a range of simulated total discharge and a range of simulated model-internal flows for each time step. For better interpretation, the simulated SSF volume uncertainty is not expressed as a range of simulated total SSF volume, but as a range of simulated SSF contribution to total simulated discharge ($Q_0\,(t)/Q_{sim}(t)$), the SSF fraction at time t f_t [-]. To minimise the influence of outliers, SSF volume uncertainty was calculated based on the 95%

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

range of all simulated behavioural SSF contributions. The SSF occurrence probability p_t [-] was approximated by the frequency of SSF simulation, i.e., by the fraction of behavioural runs, that simulate SSF (i.e., $Q_0(t) > 0$) at any given timestep. Therefore, if all behavioural runs simulate SSF at a given point in time, the SSF occurrence probability is 100% ($p_t = 1$) and if all behavioural runs simulate that no SSF occurs, it is 0% ($p_t = 0$).

To disentangle the total simulated SSF uncertainty, it was divided into two parts: a volume uncertainty share ("How big is SSF, when SSF occurs?") and an occurrence uncertainty share ("Does SSF occur or not?"). These SSF uncertainty shares were calculated according to the Law of Total Variance (equation 9) (Casella and Berger, 2024; Champ and Sills, 2024). The equation was applied here using the SSF fraction at time t f_t [-] as output variable Y, and binary occurrence indicator I_t, which determines if SSF is occurring or not, as conditioning variable X (equation 10).

$$Var(Y) = E[Var(Y|X)] + Var(E[Y|X]) \tag{9}$$

$$Var(f_t) = E[Var(f_t|I_t)] + Var(E[f_t|I_t]) \tag{10}$$

The Law of Total Variance is composed of two terms: the Within-Occurrence term $E[Var(f_t|I_t)]$ and the Between-State term $Var(E[f_t|I_t])$, which we interpreted as SSF volume uncertainty share and SSF occurrence uncertainty share, respectively. They can be determined according to equations 11-12 and combined to equal the total unconditional variance across all runs (including zeros) $Var(f_t)$ (equation 13) (Casella and Berger, 2024).

$$E[Var(f_t|I_t)] = p_t\, \sigma_t^{2+} \tag{11}$$

$$Var(E[f_t|I_t]) = Var(\mu_t^+ I_t) = (\mu_t^+)^2 Var(I_t) = (\mu_t^+)^2 p_t(1 - p_t) \tag{12}$$

$$Var(f_t) = p_t\, \sigma_t^{2+} + p_t(1 - p_t)(\mu_t^+)^2 \tag{13}$$

When $I_t = 0$, SSF does not occur and therefore, $E[f_t|I_t] = 0$ and $\mathrm{Var}(f_t|I_t) = 0$. When $I_t = 1$, SSF occurs with $E[f_t|I_t] = \mu_t^+$ and $\mathrm{Var}(f_t|I_t) = \sigma_t^{2+}$, with μ_t^+ and σ_t^{2+} describing the mean and variance of the SSF fraction when SSF occurs.

For better comparability, the Within-Occurrence term and the Between-State term can be expressed as a share of total unconditional variance (equations 14-16).

$$S_{vol,t} = \frac{p_t \sigma_t^{2+}}{\mathrm{Var}(f_t)} \tag{14}$$

$$S_{\mathrm{occ},t} = \frac{p_t(1 - p_t)(\mu_t^+)^2}{\mathrm{Var}(f_t)} \tag{15}$$

$$S_{vol,t} + S_{\mathrm{occ,t}} = 1 \tag{16}$$

With	S_{vol}	Share of total variance due to volume uncertainty [-]
	S_{occ}	Share of total variance due to occurrence uncertainty [-]

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

The SSF uncertainty estimation and decomposition in every catchment were based on the global parameter estimation as well as the percentile parameter estimation, that yielded the largest increase in sensitivity of SSF-relevant parameters in each catchment. For our main analysis, we considered SSF to be conceptually represented through Q_0 in the HBV model structure (Figure 2). However, as process allocation is not completely clear in bucket-type models, we performed an additional analysis where we considered SSF to be represented through the sum of Q_0 and Q_1 and compared the results for the global parameter estimation.

2.4 Study sites and available data

We applied our approach in the four study sites of the RU FOR 5288: “Fast and Invisible: Conquering Subsurface Stormflow through an Interdisciplinary Multi-Site Approach” (Figure 3). The four study sites cover a range of topographic, climatic and land-use conditions (Table 2).

2.4.1 Study site descriptions

The Sauerland catchment is the largest catchment among the study sites (Table 2). The main land-use types are pasture, spruce forest and mixed forest with minor shares of arable land and settlement. The soil types in the catchment are dominantly loamy cambisols, leptosols and stagnasols and the geology is characterised by sandy-silty clay shale from the Lower and Middle Devonian on which periglacial cover beds have developed during the Late Quaternary (Chifflard et al., 2008). In the Ore Mountains catchment, the dominant land-use types are spruce forest, pasture and cropland, and the typical soil types are stagno-gleyic cambisols and dystric cambisols. The geology is characterised by gneiss and phyllite bedrock, overlain by periglacial cover beds (Heller and Kleber, 2016). The Black Forest catchment is the smallest catchment of all study sites (Table 2). The predominant land-use types are forest and grassland, and the soil types in the catchment are dominantly cambisols, with gleysols and colluvium along the stream. Geologically, the Black Forest catchment is characterised by crystalline bedrock overlain by periglacial drift covers (Bachmair and Weiler, 2012, 2014). The Tyrolean Alps catchment has the steepest slopes among the study sites (Table 2). The main land-use types are spruce forest, alpine pastures and alpine dwarf shrub. Calcareous schists form the bedrock, and the valley has been subject to glacial erosion. In the valley, there is a quaternary valley infill consisting of Holocene and Pleistocene sediments of different sediment types (Orsi et al., 2016). The climatic forcing data vary between catchments with the highest mean annual temperature and lowest mean annual precipitation in the Black Forest catchment and the lowest mean annual temperature and highest mean annual precipitation in the Tyrolean Alps catchment (Table 2).

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

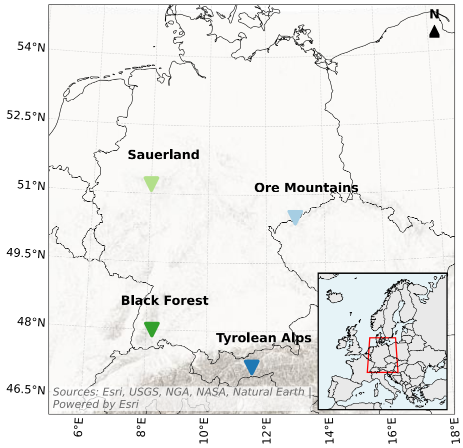


Figure 3 Study Site locations within Central Europe

Table 2 Study sites and their physio-geographical and hydrological properties

		Sauerland (SL)	**Ore Mountains (OM)**	**Black Forest (BF)**	**Tyrolean Alps (TA)**
Topography	Catchment area [km²]	47	31	0.25	8.84
	Mean elevation [m.a.s.l]	407.32	661.23	446.67	1975.39
	Elevation range [m.a.s.l]	312.95 - 517.37	498.13 - 805.6	329.56 - 577.76	1326.43 - 2595.21
	Slope [°]	0-83 (mean 9.24)	0-47 (mean 7.42)	0-56 (mean 18.67)	0-85 (mean 33.78)
	Sources	Geobasis NRW (2022)	GeoSN (2022)	LGL BaWü (2025)	Land Tirol (2022)
Stream	Stream name	Rose & Brachtpe	Gimmlitz	Rütlitobel	Padasterbach
	Sources	Ruhrverband (2025)	LTV (2022)	Universität Freiburg (2025)	BBT SE (2025)
Forcing	Climate *	Cfb	Cfb	Cfb	Dfc
	Mean annual temperature [°C] **	9.94	5.54	11.53	2.93
	Mean annual precipitation [mm] **	1115	1052	831	1293
	Sources	DWD (2025a, 2025b, 2025c, 2025d)	DWD (2025a, 2025b, 2025c, 2025d)	DWD (2025a, 2025b, 2025c, 2025d)	GeoSphere Austria (2021, 2020), Haiden et al. (2011), Haslinger and Bartsch (2016), Kann et al. (2011)

(*) Cfb: warm temperate, fully humid, warm summer; Dfc: snow, fully humid, cool summer; based on the Köppen-Geiger classification

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

(**) Calculated over the whole simulation timespan of the respective catchment

2.4.2 Datasets and pre-processing for simulation with the HBV model

The inputs to the HBV model are hourly timeseries of climatic forcing data including precipitation, temperature and potential evapotranspiration (sources are listed in Table 2). Hourly timeseries of potential evapotranspiration are derived from daily estimates of potential evapotranspiration E_{pot} provided by DWD (2025a) and GeoSphere Austria (2020) by disaggregation using hourly time series of solar data (DWD, 2025c; GeoSphere Austria, 2021). The daily sums of E_{pot} are distributed over 24 hours, creating a diurnal pattern according to the diurnal pattern of solar radiation, given that solar radiation is a major driver of E_{pot} and available in hourly resolution (Wang and Dickinson, 2012).

The HBV model was run for approximately 15 years in each catchment (Table 3). All available discharge observations were used for parameter estimation (see Table 2 for data sources). However, because data availability varied across catchments, calibration periods differed accordingly (Table 3).

Table 3 Simulated timespans and timespans of observed discharge data available for calibration

	Sauerland	Ore Mountains	Black Forest	Tyrolean Alps
Simulated timespan	2009-11-01 to 2025-06-29	2000-11-01 to 2015-11-01	2009-11-01 to 2025-06-10	2012-08-06 to 2025-06-29
Timespan of available discharge observations	2012-11-01 to 2025-04-01	2005-04-07 to 2011-08-04	2019-12-03 to 2021-03-04 & 2023-06-07 to 2025-06-10	2015-08-07 to 2020-01-08 & 2023-01-03 to 2024-11-10

3 Results

3.1 Hydrological simulation model, parameter and uncertainty estimation

For all catchments, the observed discharge timeseries mostly lies within the simulated uncertainty band (Figure 4) and the global parameter estimation yields sufficiently high KGE values (Table 4). The KGE of the model using a random parameter set, which we used as a lower benchmark for model performance, is below the 5% behavioural threshold for all catchments. For the Sauerland and Tyrolean Alps catchments, KGE values are generally higher compared to the Ore Mountains and Black Forest catchments. This is reflected in the higher KGE behavioural thresholds, higher mean behavioural KGEs, and higher KGE of the lower benchmark.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

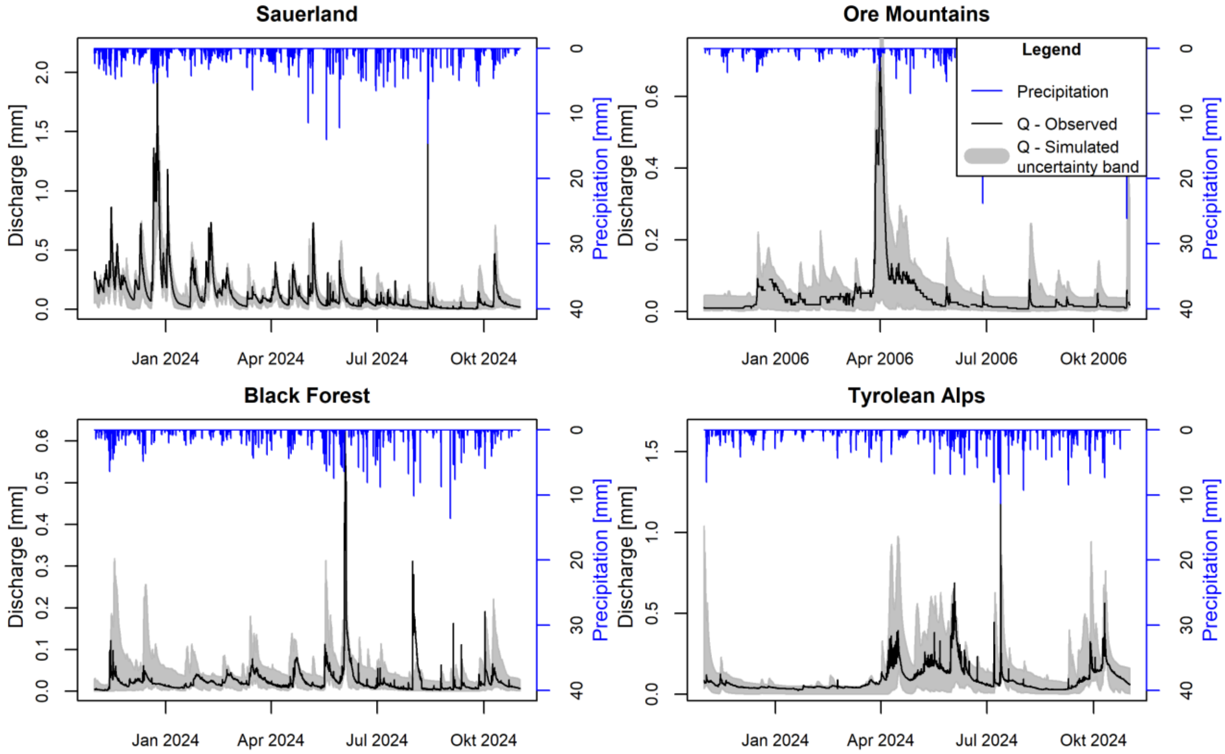


Figure 4 Discharge simulations for all catchments, exemplary for one hydrological year. The simulated discharge uncertainty bands are derived from the simulations of all behavioural parameter sets from the global parameter estimation.

310 **Table 4 KGE values [-] derived by global parameter estimation for all catchments**

	Sauerland	Ore Mountains	Black Forest	Tyrolean Alps
Behavioural Runs Threshold (top 5%)	0.61	0.56	0.55	0.70
Behavioural Runs Mean	0.65	0.60	0.60	0.73
Best Run	0.83	0.77	0.76	0.81
Lower Benchmark	0.51	-0.08	-0.43	0.57

3.2 Parameter identifiability across FDC percentiles

Figure 5 reports an analysis of behavioural parameter ranges through boxplots. Here, a distinct deviation of the median or a reduction in the interquartile range compared to the uniform distribution displayed in the background indicates that a parameter is better identifiable and shows the direction of the optimal parameter value. In turn, if a boxplot resembles the uniform
315 distribution, the parameter is difficult to identify. In Figure 5, the behavioural parameter ranges are compared across catchments, while also considering different parameter-estimation approaches: global parameter estimation (left), <10% high-flow parameter estimation (middle) and 10-20% high-flow parameter estimation (right). The snow parameters TT and $CFMAX$

are the most identifiable parameters in the Ore Mountains and Tyrolean Alps catchments (Figure 5), for global parameter estimation and especially for the <10% high-flow parameter estimation, indicating a very high relevance of these parameters for the model in these two catchments.

Among the four study sites, the SSF-relevant parameters K_0 and UZL are the most identifiable in the Sauerland catchment (Figure 5). The interquartile ranges are narrower and the boxplots are shifted towards lower values for UZL and higher values for K_0 compared to the uniform distribution. Within the HBV model structure, lower values of UZL conceptually represent a lower threshold for the activation of SSF, and higher values of K_0 show that the SSF storage empties faster. Therefore, with the observed parameter shifts, the model simulates more frequent and faster SSF. This shift is even more emphasised in the <10% high-flow parameter estimation, indicating that SSF becomes more important at peak flow and that low UZL values and high K_0 values are favoured, leading to a more frequent, fast SSF response. In the Black Forest catchment, behavioural ranges of the global parameter estimation are more uniformly distributed, while behavioural ranges of the <10% high-flow parameter estimation show a shift towards smaller values of UZL and larger values of K_0, indicating simulation of more frequent and faster SSF, like in the Sauerland catchment. In addition, there is a substantial reduction in the interquartile ranges, indicating improved parameter identifiability. In the Ore Mountains and Tyrolean Alps catchments, SSF-relevant parameters are less identifiable with wider interquartile ranges compared to the Sauerland and Black Forest catchments. For parameter UZL, interquartile ranges are slightly narrower and the boxplots are shifted towards higher values in both catchments. Unlike the Sauerland and Black Forest catchments, this indicates a higher activation threshold for SSF within the model structure resulting in less frequent simulation of SSF events.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

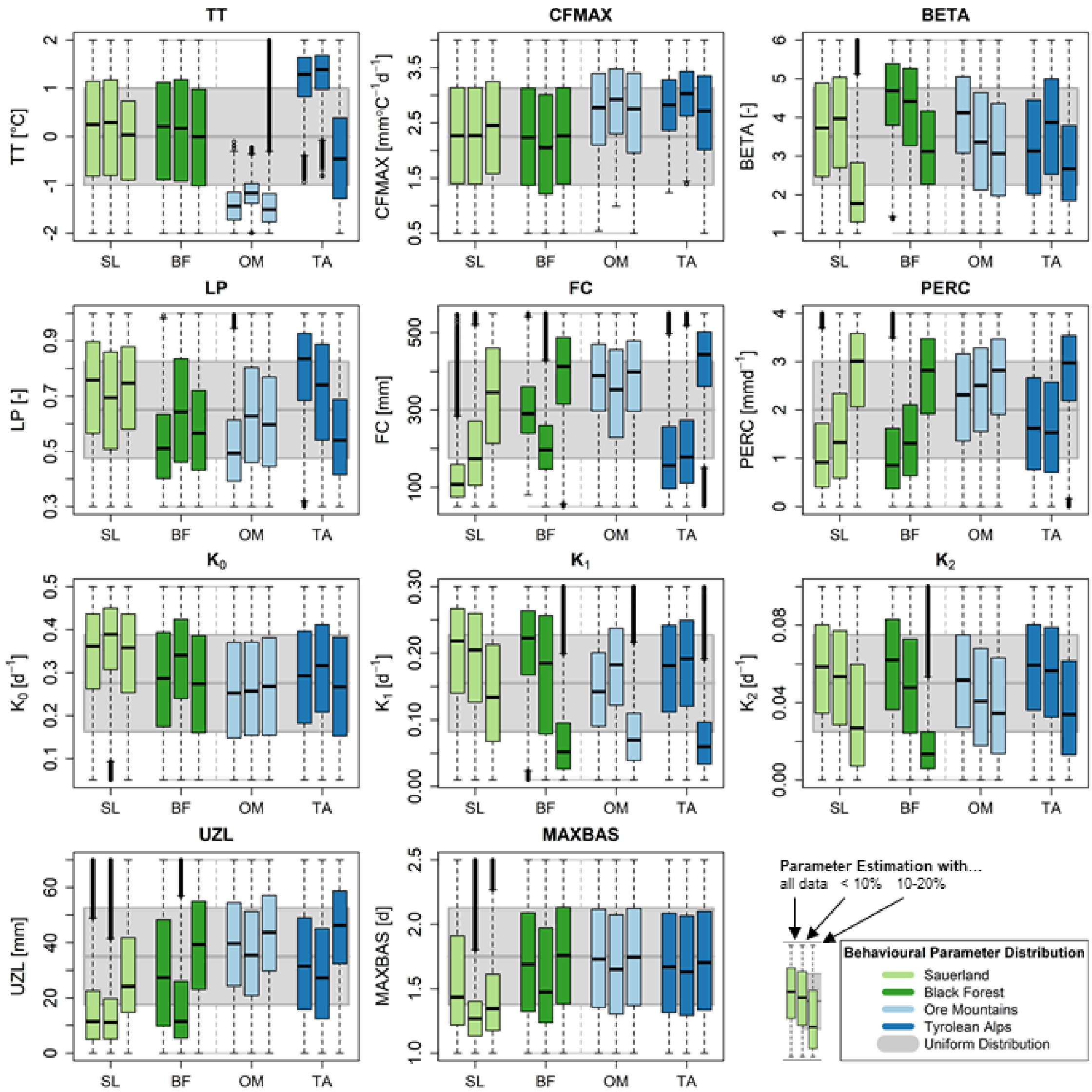


Figure 5 Behavioural parameter ranges for all catchments as boxplots (Sauerland – light green, Black Forest – dark green, Ore Mountains – light blue, Tyrolean Alps – dark blue) estimated based on all data (left), on the top 10% high-flow percentile (middle) and on the top 10-20% high-flow percentile (right). The grey background represents the uniform distribution from which the parameters were sampled.

In the RSA, cumulative distribution functions (CDFs) of behavioural parameter sets over the calibration range were compared to the 1:1 line, which represents the uniform distribution from which the parameters were sampled. Large deviations of the CDFs from the 1:1 line indicate stronger parameter influence. The maximum distance to the 1:1 line (i.e., the D-value in a Kolmogorov-Smirnoff test) can be taken as a quantitative measure of parameter influence. Larger D-values, therefore, indicate higher sensitivity of the output metric to that parameter. For the SSF-relevant parameters of the HBV model, K_0 and UZL,

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

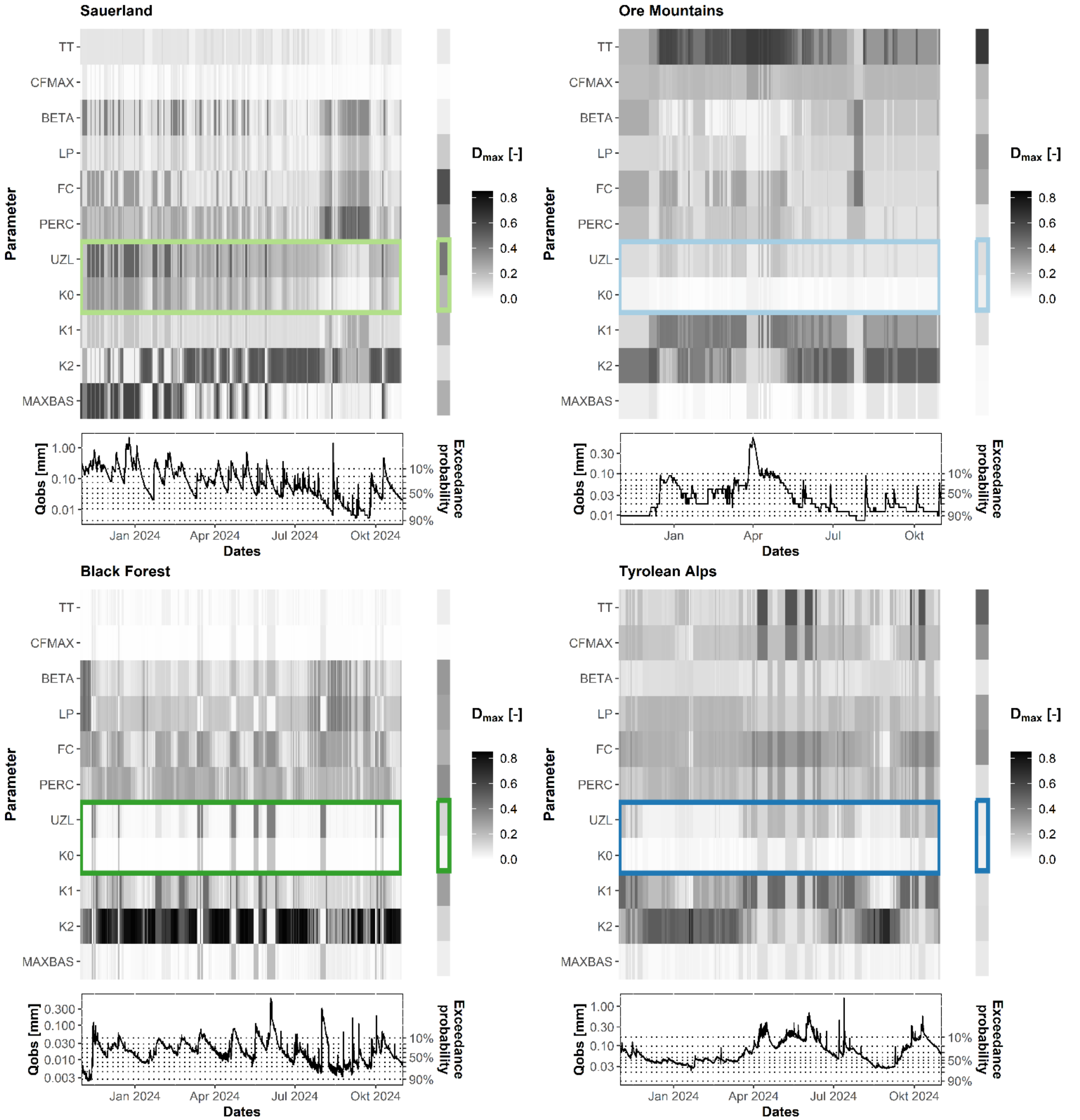
sensitivity in the global parameter estimation is generally low, with the exception of the Sauerland catchment. Specifically, D-values for K_0 are 0.22 for the Sauerland catchment, 0.05 for the Ore Mountains catchment, 0.03 for the Black Forest catchment, 0.05 for the Tyrolean Alps catchment, and for UZL are 0.42 for the Sauerland catchment, 0.10 for the Ore Mountains catchment, 0.13 for the Black Forest catchment, 0.06 for the Tyrolean Alps catchment. The RSA plots for the global and percentile parameter estimations for all model parameters can be found in the Appendix (Figures A.1 – A.4).

For some percentile parameter estimations, there are clear increases in the sensitivity of SSF-relevant parameters. For the Sauerland catchment, the largest increase is registered for the high-flow percentile with <10% exceedance probability (D-values: 0.32 for K_0 and 0.49 for UZL). The same applies to the Black Forest catchment (D-values: 0.17 for K_0 and 0.40 for UZL). For the Ore Mountains catchment, there is no increase in sensitivity through the percentile parameter estimations for the parameter K_0, but a slight increase for the parameter UZL with the 10-20% high-flow percentile parameter estimation (D-value 0.18). For the Tyrolean Alps catchment, there is a slight increase in sensitivity for the parameter K_0 for the high-flow percentile with <10% exceedance probability (D-value 0.10), but there is a larger increase in sensitivity for the parameter UZL for the high-flow percentile with 10-20% exceedance probability (D-value 0.22). Generally, the increase in sensitivity of SSF-relevant parameters is less pronounced in the Ore Mountains and Tyrolean Alps catchments than in the Sauerland and Black Forest catchments.

While there is an increase in SSF parameter sensitivity with the percentile parameter estimations, there is a simultaneous decrease in overall model performance (Table A.2). To maximise the information gained from the flow-duration curve approach, the percentile parameter estimation yielding the largest increase in SSF parameter sensitivity was selected for further analysis in each catchment. This was the <10% exceedance probability percentile for the Sauerland and Black Forest catchments, and the 10–20% exceedance probability percentile for the Ore Mountains and Tyrolean Alps catchments.

Figure 6 visualises the percentile parameter estimation in a timeseries and illustrates that across all catchments, the parameter K_2 tends to be more sensitive during median to low-flow periods while the parameters TT, $CFMAX$, K_0, UZL and $MAXBAS$ tend to be most sensitive during high-flow periods. For other parameters, the patterns are not as clear (Figure 6). In the Sauerland and Black Forest catchments, the display of sensitivity as a timeseries again accentuates the link between peak high-flow and the higher sensitivity of SSF-relevant parameters. In the Ore Mountains and Tyrolean Alps catchments, during peak flows (<10% percentile) the sensitivity of the snow parameters TT and $CFMAX$ becomes very large, while the sensitivities of all other parameters, except for $MAXBAS$, decrease during this time. The SSF-relevant parameters in both catchments, except K_0 in the Tyrolean Alps, reach peak sensitivity in the 10-20% high-flow percentile.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026



375 **Figure 6 Sensitivities (D-values) of model parameters from percentile parameter estimations illustrated as a timeseries for all catchments over 1 hydrological year, darker colour representing higher sensitivity. Vertical bars displayed to the right of the timeseries are the parameter sensitivities (D-values) from the global parameter estimation. The SSF-relevant parameters UZL and K_0 are highlighted in the respective catchment colour.**

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

3.3 Estimation and decomposition of simulated SSF uncertainty

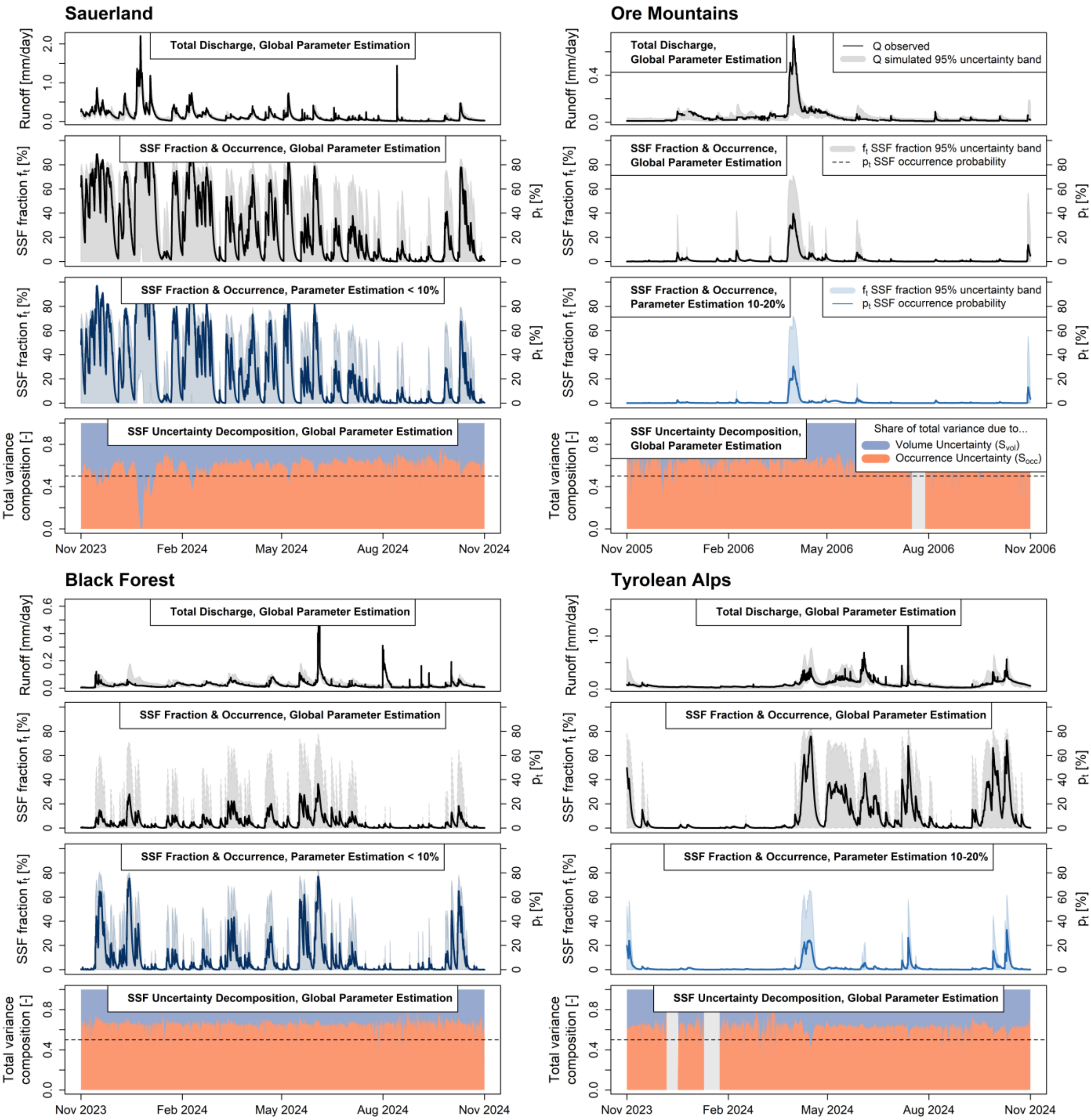
The simulated SSF occurrence probability is clearly linked to total observed discharge and increases with higher discharge values (Figure 7). The simulated SSF volume uncertainty, displayed here as the uncertainty in the SSF fraction of total discharge, is generally high, often spanning from 0% to over 60% SSF contribution to total simulated discharge. Like SSF occurrence probability, SSF volume uncertainty also increases with total observed discharge, with especially high volume uncertainties during peak flows. Within the HBV model structure, SSF appears to have the highest importance in the Sauerland catchment, where simulated SSF occurrence probability is the highest, with the Sauerland also being the only catchment where SSF volume uncertainty does not entail zero at all times (Figure 7). In the Black Forest, as indicated in previous results, SSF appears to be especially important for the simulation of high-flow peaks, with an increase in SSF occurrence probability for the <10% high-flow parameter estimation. The SSF occurrence probability is overall the lowest in the Ore Mountains catchment, where it is limited to discharge peaks.

Simulated SSF uncertainties can be slightly reduced, but remain high even with the percentile parameter estimations, where SSF-relevant parameters have shown to be more identifiable (for every catchment third plot from the top in Figure 7). In the Sauerland and Black Forest catchments, SSF occurrence probability increases with the <10% high-flow percentile parameter estimation. In the Ore Mountains and Tyrolean Alps catchments on the other hand, SSF occurrence probability decreases for the 10-20% high-flow parameter estimation.

The decomposition of total simulated SSF uncertainty shows that over almost the entire timeseries, simulated SSF occurrence uncertainty is larger than simulated SSF volume uncertainty across all catchments (Figure 7). The mean (± standard deviation) shares of SSF occurrence from total simulated SSF uncertainty across the whole timeseries are 0.63±0.09 for the Sauerland catchment, 0.64±0.08 for the Ore Mountains catchment, 0.68±0.08 for the Black Forest catchment and 0.65±0.11 for the Tyrolean Alps catchment. Only during very large peaks in discharge, SSF volume uncertainty outweighs SSF occurrence uncertainty, e.g., in the Sauerland catchment around January 2024.

When using the sum of Q_0 and Q_1 to conceptually represent SSF in a global parameter estimation, instead of Q_0, simulated SSF uncertainties become even larger (Figure A.5). SSF occurrence probability becomes more likely across catchments and the uncertainty decomposition shows that SSF volume uncertainty contributes a larger share of total simulated SSF uncertainty. However, still a considerable share of total simulated SSF uncertainty stems from SSF occurrence uncertainty (Table A.1).

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026



410 **Figure 7 Simulated SSF volume uncertainty, expressed as the uncertainty of SSF fraction of total discharge f_t, SSF occurrence probability p_t and SSF uncertainty decomposition for each catchment, over 1 hydrological year. The legends displayed in the plots for the Ore Mountains (top right) apply to all catchments.**

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

4 Discussion

4.1 Reliability of our uncertainty estimation and percentile parameter-estimation approach

In this study, we set up the HBV model for three Mid-Mountain catchments and for one Alpine catchment with varying climatic and topographic conditions. Similar to previous applications (e.g., Abebe et al., 2010; Zelelew and Alfredsen, 2013), the HBV model could capture the hydrological dynamics well across all catchments and the observed discharge timeseries mostly lie within the respective simulated uncertainty ranges (Figure 4). This is confirmed by the overall model performance with relatively high KGEs for the behavioural parameter sets across catchments (Table 4), and clearly higher than the lower benchmark (Seibert et al., 2018). Because we used all available discharge data for parameter estimation rather than splitting the data into separate calibration and validation periods, the reported performance values may overestimate the model`s predictive accuracy on independent datasets. However, since our objective was not to develop a predictive model but rather to analyse simulation uncertainty, we argue that the achieved model performance, though potentially optimistic, provides a sufficient basis for the subsequent uncertainty analysis. Average behavioural KGE values and thresholds vary between catchments, with overall higher model performances in the Sauerland and Tyrolean Alps and lower model performances in the Ore Mountains and Black Forest. However, this does not necessarily imply worse model fit in the Ore Mountains and Black Forest, as the same pattern is also reflected in the lower benchmarks, which are also higher for the Sauerland and Tyrolean Alps and lower for the Ore Mountains and Black Forest. Instead, the difference in KGE values between catchments might be linked to differences in data quality (especially of discharge data) or differences in discharge dynamics (Williams, 2025).

The ranges of the best 5% of model parameters were found within reasonable ranges compared to other studies in similar settings (Usman et al., 2024; Seibert, 1999; Abebe et al., 2010; Seibert and Vis, 2012). Most parameters, including the SSF-relevant ones, are insensitive and difficult to identify in the global parameter estimation. Such limitations in identifying all HBV model parameters through calibration based on discharge data alone have also been observed by other studies (Abebe et al., 2010; Tibangayuka et al., 2022; Zelelew and Alfredsen, 2013). The resulting parameter uncertainty is accounted for in our subsequent analysis.

The behavioural HBV parameters obtained in the global parameter estimation for each catchment can serve as effective representations of the respective catchment dynamics in a reasonable way. The estimated snow parameters TT and $CFMAX$ are insensitive in the Sauerland and Black Forest catchments, but highly sensitive in the Ore Mountains and Tyrolean Alps catchments. This is linked to the climatic forcing of catchments with lower elevations and higher average temperatures in the Sauerland and Black Forest and therefore a lower relevance of snow processes. The identifiable snow parameters in the Ore Mountains and Tyrolean Alps are within the same ranges as the parameters estimated in the HBV parameter regionalisation by Beck et al. (2020) for the respective area, except for the low value of TT in the Ore Mountains. One possible explanation is the location of the Ore Mountains climate station outside of the catchment area. The low value of TT in the Ore Mountains might compensate for a temperature difference between climate station and catchment location. The estimated soil parameters FC, LP and $BETA$ vary between catchments but all lie within the estimated ranges of Zelelew and Alfredsen (2013) in a similar setting. The estimated parameter for field capacity FC is considerably lower in the Sauerland and Tyrolean Alps catchments compared to the Black Forest and Ore Mountains. Both catchments also show a more dynamic discharge timeseries. As for the soil parameters, the response routine parameters K_0, K_1, K_2 and UZL also lie within the estimated parameter ranges of Zelelew and Alfredsen (2013), but vary across catchments. The routing parameter $MAXBAS$ only gains sensitivity in the Sauerland catchment, which is the largest catchment among the four study sites, and therefore links parameter sensitivity to the catchment area.

When interpreting differences in simulated SSF across the four study catchments, potential scaling effects were not explicitly accounted for. SSF contributions may vary with catchment size by either attenuating, e.g., through mixing or storage effects or amplifying, e.g., through the connection of contributing areas. Because the catchments examined in this study differ in size, some of the differences in SSF simulation uncertainty may partly reflect scaling effects rather than true differences in process representation. Future research could therefore extend this analysis to a larger number of catchments spanning a wider range of spatial scales, enabling a more systematic investigation of how simulated SSF contributions and their associated uncertainties vary with catchment size.

With our percentile parameter-estimation approach, we explored whether more information useful for the estimation of SSF-relevant parameters is hidden in specific discharge conditions that are lost when aggregating residuals over the entire calibration period (Wagener et al., 2003; Abebe et al., 2010). While the results show that especially for parameter estimation using high-flow conditions, SSF parameter identifiability could be improved to some degree, these percentile parameter-estimation approaches also led to a reduction of global model performances as indicated by lower KGE values (Table A.2). While the percentile parameter-estimation approach can be useful in our explorative analysis and might even be beneficial as an additional step in model calibration, it cannot be used as a sole metric for overall model parameter estimation.

In addition, the percentile parameter-estimation approach cannot fully capture the different catchment conditions and hydrological processes, e.g., high flows in spring, where snowmelt causes rather long events, cannot be distinguished from high flows in summer or autumn, where response on storm rainfall results in much shorter events. Still, the percentile parameter-estimation approach allows us to distinguish between the main discharge conditions across catchments and to derive specific behavioural parameter sets for the respective conditions to be used in further uncertainty analysis. This is reflected in the reasonable temporal distribution of parameter sensitivities (Figure 6), as sensitivity patterns align with the underlying hydrological processes. Parameters governing slow recession processes (e.g., K_2) exhibit higher sensitivity during low-flow periods, whereas parameters controlling faster processes (e.g., TT, $CFMAX$, K_0, UZL and $MAXBAS$) are more sensitive during high-flow conditions.

Our uncertainty estimation approach neglects other sources of uncertainty in hydrological modelling besides parameters, in particular model structure uncertainty and data uncertainty (Moges et al., 2021). The relatively low behavioural threshold for parameter estimation (yielding 50,000 behavioural runs) might partly compensate for potential data uncertainty arising from low-quality input or discharge data. The overall acceptable model performances across all study sites also supports this. Given

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

the large SSF simulation uncertainty, additional considerations of model structure and data uncertainty would only further amplify simulated uncertainties.

In addition, while we did not consider other model structures, we validated our uncertainty analysis results with two versions of conceptual process allocation. In the main analysis, we considered SSF to be conceptually represented through the model-internal flow Q_0. In this setting, SSF is simulated as the outflow from an overflow reservoir above a shallow groundwater box. But in an additional analysis, we considered SSF to be conceptually represented by the sum of model internal flows Q_0 and Q_1 for the global parameter estimation. In this setting, SSF is simulated with a slower outflow from a shallow reservoir and an additional faster outflow, that is activated when the shallow reservoir exceeds a threshold UZL (Figure 2). While there were some shifts in the resulting values, the main outcomes remained unchanged (Figure A.5, Table A.1): 1) Simulated SSF uncertainties are high in both cases, and 2) a considerable part of simulated SSF uncertainty originates from SSF occurrence uncertainty.

4.2 SSF identification across catchments and FDC percentiles

In the regional sensitivity analysis and SSF uncertainty estimation we found that SSF-relevant parameters are mostly insensitive in the global parameter estimation, with the exception of the Sauerland catchment (Figure 5 & 6). This low sensitivity of the SSF-relevant parameters within the HBV model in global parameter-estimation approaches was also found in previous studies (Tibangayuka et al., 2022; Zelelew and Alfredsen, 2013). Simulated SSF uncertainties are large across catchments, especially during high-flow periods, when the simulated uncertainty band of SSF contribution to total streamflow often ranges from 0% to >60% (Figure 7). Similarly high uncertainties of simulated SSF contributions to total streamflow were found by Hartmann et al. (2023) for a version of the HBV model and by Hughes and Farinosi (2021) for the Pitman model. This finding also aligns with previous research on equifinality, showing that different parameterisations can yield similar discharge simulations while model-internal processes can vary, making model-internal distinction between different lateral streamflow components highly uncertain (e.g., Beven and Binley, 1992; Seibert and McDonnell, 2002).

Therefore, in this study we investigated whether more information on SSF is hidden in parts of the flow-duration curve using our percentile parameter-estimation approach. We found that some additional information for parameter estimation can be extracted: the SSF-relevant parameters can become more sensitive through percentile parameter estimation (Figure 5) and simulated SSF uncertainty can be partially reduced. However, simulated SSF uncertainty still remains high (Figure 7).

The largest increase in SSF parameter sensitivity across all study sites could be achieved through parameter estimation using high-flow percentiles. This finding is consistent with previous research, also showing that SSF-relevant parameters in the HBV model are more sensitive when calibration metrics are applied that focus on high flow (Tibangayuka et al., 2022; Abebe et al., 2010). It also coincides with the conceptual understanding of SSF as an event process, which becomes especially important during high flows and even floods (e.g., Zillgens et al., 2007; Markart et al., 2015). The improved identifiability may be attributed to the gain in information from looking at individual discharge conditions and, therefore, at the different response modes of the catchment separately. In global parameter estimation with an overall measure of performance, the model residuals

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

are aggregated across multiple response modes, which can hide information about how different model components perform (Wagener et al., 2003; Pianosi and Wagener, 2016). For instance, low-flow periods might entail more information on slow recession processes, while high-flow periods might entail more information on faster processes like snow melt or SSF.

The comparison of parameter estimation results for the SSF-relevant parameters UZL and K_0 revealed two groups of catchments. The Sauerland and Black Forest catchments show a tendency towards higher values of UZL and lower values of K_0, whereas the Ore Mountains and Tyrolean Alps catchments show a tendency towards lower values of UZL and higher values of K_0 (Figure 5). Conceptually, lower values of UZL and higher values of K_0 indicate a more frequent occurrence of faster SSF. Therefore, the parameter values suggest faster, more frequent SSF in the Sauerland and Black Forest catchments than in the Ore Mountains and Tyrolean Alps catchments. In addition, SSF-relevant parameters in the Ore Mountains and Tyrolean Alps catchments exhibit lower sensitivities compared to the Sauerland and Black Forest catchments, particularly in the best percentile parameter estimation, indicating that SSF plays a less important role in these catchments. Instead, the Ore Mountains and Tyrolean Alps catchments showed the highest sensitivities for snow-related parameters, indicating that snow processes dominate model behaviour in these catchments.

We initially expected SSF to play an important role in all study sites, given their humid, mountainous characteristics and evidence from previous field studies. For instance, the studies of Chifflard et al. (2008) and Bachmair and Weiler (2014), using tracer and hydrometric data in the Sauerland and trench-flow data in the Black Forest respectively, showed that SSF plays an important role in these catchments, especially during periods with high soil moisture and during discharge peaks. Heller and Kleber (2016) analysed runoff generation processes using hydrometrical and soil temperature data in a similar catchment in the Ore Mountains and also found that especially at medium or high pre-moisture, SSF formation is frequent in the catchment. However, in our results, SSF processes appear less important in the snow-dominated Ore Mountains and Tyrolean Alps catchments. Such low SSF sensitivity within the model might reflect insufficient information in the available discharge data to identify SSF-relevant parameters and differentiate lateral flow contributions adequately. We argue that for the snow-dominated catchments (Ore Mountains and Tyrolean Alps) this is amplified by the strong snowmelt dynamics that dominate the discharge signal. Because both snowmelt and SSF can generate rapid runoff responses associated with discharge peaks, the stronger snow signal may superimpose any SSF-related information in the aggregated catchment discharge. This interpretation aligns with the understanding that hydrological models calibrated on discharge data can only identify a limited number of parameters (Wagener and Gupta, 2005; Zelelew and Alfredsen, 2013). This might also explain why the 10–20% high-flow percentile yielded the largest, but less pronounced improvement in SSF parameter identifiability in the snow-dominated catchments Ore Mountains and Tyrolean Alps, as opposed to the <10% high-flow percentile in the snow-insensitive catchments Sauerland and Black Forest. In the snow-dominated catchments, the highest flows are likely driven primarily by snowmelt, leaving the 10–20% percentile as a discharge condition where SSF signals are relatively more detectable. These results align with an earlier study by Zelelew and Alfredsen (2013), who investigated the contributions of HBV parameter sensitivity in 12 catchments in Southern Norway and also found that among the five most influential parameters three were snow parameters while the runoff response parameters were among the least sensitive parameters. The unidentifiability of

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

SSF-relevant parameters from discharge data might be further amplified in the Ore Mountains catchment, where the available discharge time series contains fewer pronounced peak events, which are conditions under which SSF would be expected to contribute substantially and thereby provide less information for parameter identification.

4.3 Transferability and impact of SSF simulation uncertainties

Across our four study sites, we found that SSF parameter identifiability can be increased to some extent by using our percentile parameter-estimation approach with high-flow percentiles and thus recovering some information on SSF lost through the aggregation of the entire residual timeseries into a single objective function (Wagener et al., 2005). Still, across all catchments and parameter-estimation approaches, we found high SSF simulation uncertainties. As our four study sites represent a range of topographical, climatic and land-use conditions, and these findings are consistent across all sites, we believe that our findings are representative for applications of HBV-like bucket-type models in other temperate, mountainous regions. In our SSF uncertainty decomposition, we found large contributions of SSF occurrence uncertainty to total simulated SSF uncertainty across catchments (Figure 7), percentile parameter estimations (Figure A.6) and conceptual representations of SSF within the HBV model structure (Figure A.5). Therefore, we assume that this finding might be similar for other bucket-type lumped model structures.

As SSF plays an important role in flood generation, and nutrient and contaminant transport (Zillgens et al., 2007; Zhao et al., 2013; Wang et al., 2024), these results have important implications for simulating these processes. When hydrological models are used to increase process understanding, test hypotheses and support decision-making (Moges et al., 2021), the large SSF simulation uncertainties must be considered. Not much information can be gained about process understanding at the catchment scale from models with large model-internal uncertainties as using such models for hypothesis testing might yield questionable results. Similarly, decisions based on such models may lack a robust foundation. When models fail to adequately represent dominant hydrological processes, as indicated by the high SSF simulation uncertainty, their predictions beyond calibration conditions become questionable (Kirchner, 2006; McGuire et al., 2024). Previous studies have therefore emphasised the critical importance of reliable process representation in hydrological models, particularly for simulating future conditions or flood events (Brunner et al., 2021).

One reaction to this problem of equifinality in hydrological modelling is the search for calibration methods that extract more information from available discharge time series (Wagener et al., 2003). Our percentile parameter-estimation approach represents such an attempt, targeting specific discharge conditions where SSF-related signals might be more pronounced. Our results indicate that this approach can improve SSF parameter identifiability, more so in snow-insensitive catchments than in snow-dominated catchments. Nevertheless, despite these improvements, substantial SSF simulation uncertainties persisted across all catchments. This finding suggests that even optimised use of discharge data reaches its limits in constraining internal model fluxes such as SSF. Another reaction to the equifinality problem would be to assess model performance with respect to additional measured variables beyond discharge (Wagener et al., 2003). The persistent uncertainties we found in simulated SSF therefore highlight the need to incorporate SSF-specific observations into parameter estimation to achieve a more realistic

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

process representation. This was already demonstrated successfully by Hartmann et al. (2023), who added additional SSF data derived from hydrograph separation in the calibration process. Hydrograph separation can be very dependent on assumed concentrations of the end-members. So ideally, even more data on SSF should be included in model calibration.

Unlike most previous research on uncertainty estimates with HBV-like models, we additionally disentangled simulated SSF uncertainty into SSF volume uncertainty and SSF occurrence uncertainty. While SSF parameter identifiability and SSF uncertainty simulations vary across catchments, the shares of SSF occurrence uncertainty and SSF volume uncertainty obtained by our decomposition led to consistent results across all catchments. For both conceptual representations (SSF as Q_0 and SSF as $Q_0 + Q_1$), SSF occurrence uncertainty contributes a considerable proportion to total simulated SSF uncertainty across catchments (Table A.1). In case of SSF representation through Q_0, SSF occurrence uncertainty even acts as the major source of total SSF uncertainty (Figure 7). We therefore conclude that for a substantial reduction in SSF simulation uncertainty, adding SSF proxy data into calibration, which indicate SSF occurrence in the catchment, might already be a substantial step forward at our study sites and in similar settings. This might be beneficial, as these proxies are much easier to observe in the field than volumetric measurements of SSF. Potential indicators include hillslope piezometer responses, which can reveal perched water table dynamics coinciding with SSF generation (Du et al., 2016), as well as soil moisture sensors and tensiometers, which have been used to detect SSF occurrence (Heller and Kleber, 2016). These observations could be incorporated into model calibration as additional constraints, for example by penalizing parameter sets that do not simulate SSF occurrence at times when field instruments indicate SSF. This finding also connects to McGuire et al. (2024), who point out the still-lacking ability to address questions about SSF using easily observable catchment characteristics rather than site-specific, data-intensive studies.

5 Conclusion

In this study, we found that across all catchments, SSF volume and occurrence are difficult to identify in the bucket-type HBV model using only catchment discharge for parameter estimation. The simulated uncertainty bands for SSF contribution to total discharge rarely did not include zero values and often reached >60%. With our percentile-based parameter-estimation approach, we found that additional information on SSF for parameter estimation can be extracted when high-flow percentiles are considered, but the information gain is limited. In two snow-insensitive catchments (Sauerland and Black Forest), SSF-relevant parameters became better identifiable by looking at the <10% high-flow percentile. In snow-dominated catchments (Ore Mountains and Tyrolean Alps) however, the improvement in SSF parameter identifiability was less pronounced. In addition, simulated SSF volume uncertainties remained high across all catchments, even after accounting for percentile parameter estimations.

We therefore conclude that, while specifically considering high-flow conditions in the parameter estimation process can yield improvements in SSF parameter identification, discharge data alone is not sufficient for robust parameterisation of SSF in hydrological models at the catchment scale. We therefore emphasise the need to include SSF data in model parameter

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

estimation to achieve more realistic SSF simulations, and thereby improve process representation in hydrological models, creating more robust prediction models.

We also found that across catchments, parameter-estimation approaches and conceptual representations of SSF, SSF occurrence uncertainty accounted for a large part of total simulated SSF uncertainty. This finding suggests that for a major reduction in SSF simulation uncertainties proxy data indicating the occurrence of SSF might already be sufficient.

Appendix

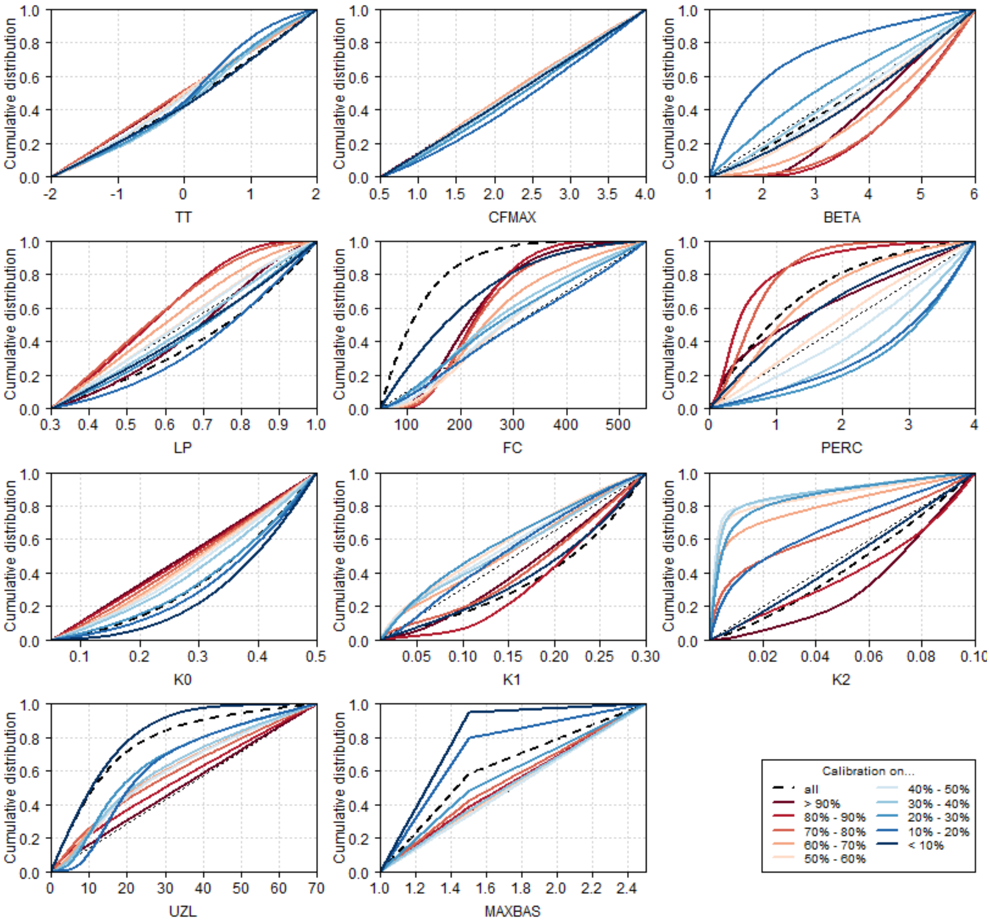


Figure A.1 Cumulative distribution functions of behavioural parameter sets (RSA) from global parameter estimation (black line, dotted) and percentile parameter estimations (colour coded– blue to red from high-flow parameter estimation to low-flow parameter estimation) for the Sauerland catchment

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

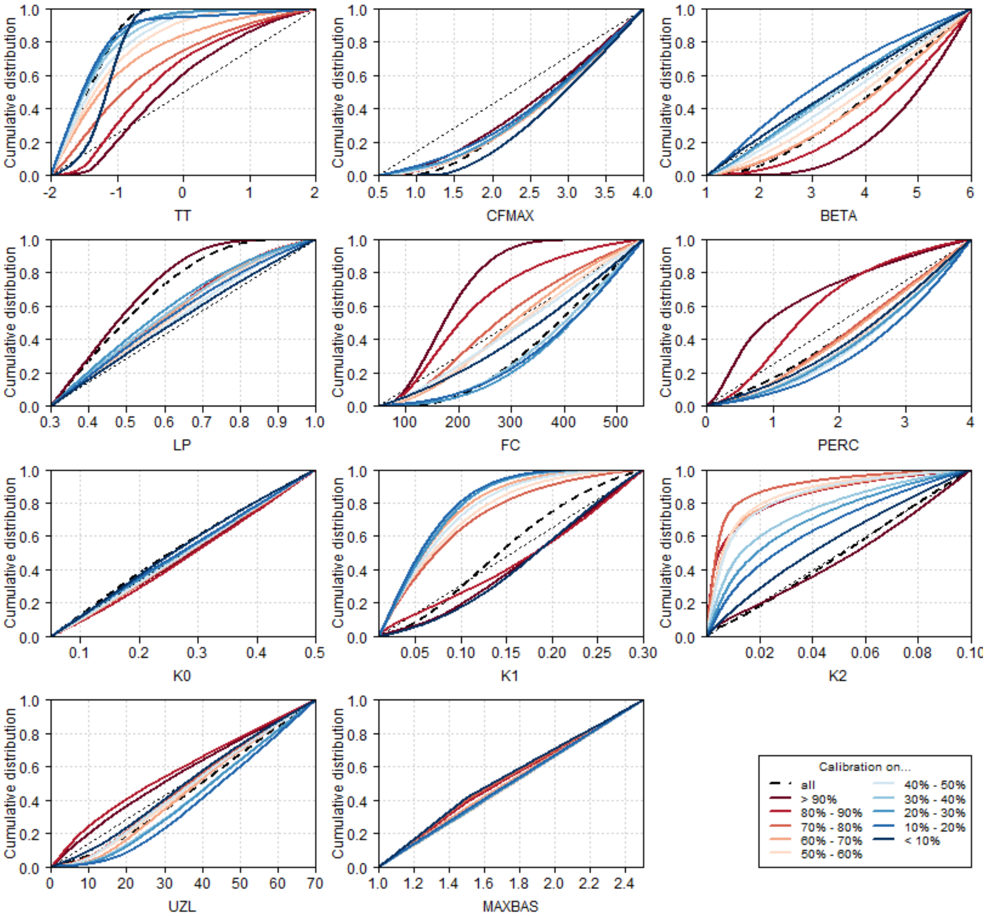


Figure A.2 Cumulative distribution functions of behavioural parameter sets (RSA) from global parameter estimation (black line, dotted) and percentile parameter estimations (colour coded– blue to red from high-flow parameter estimation to low-flow parameter estimation) for the Ore Mountains catchment

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

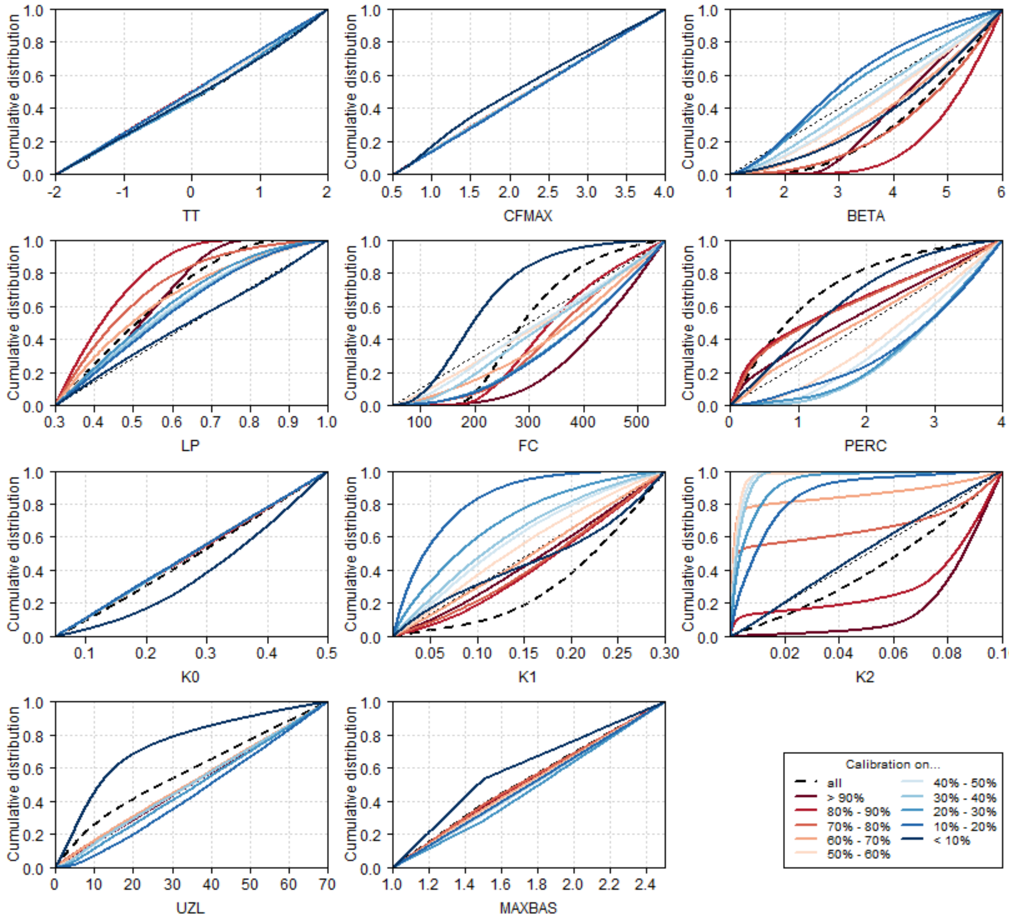
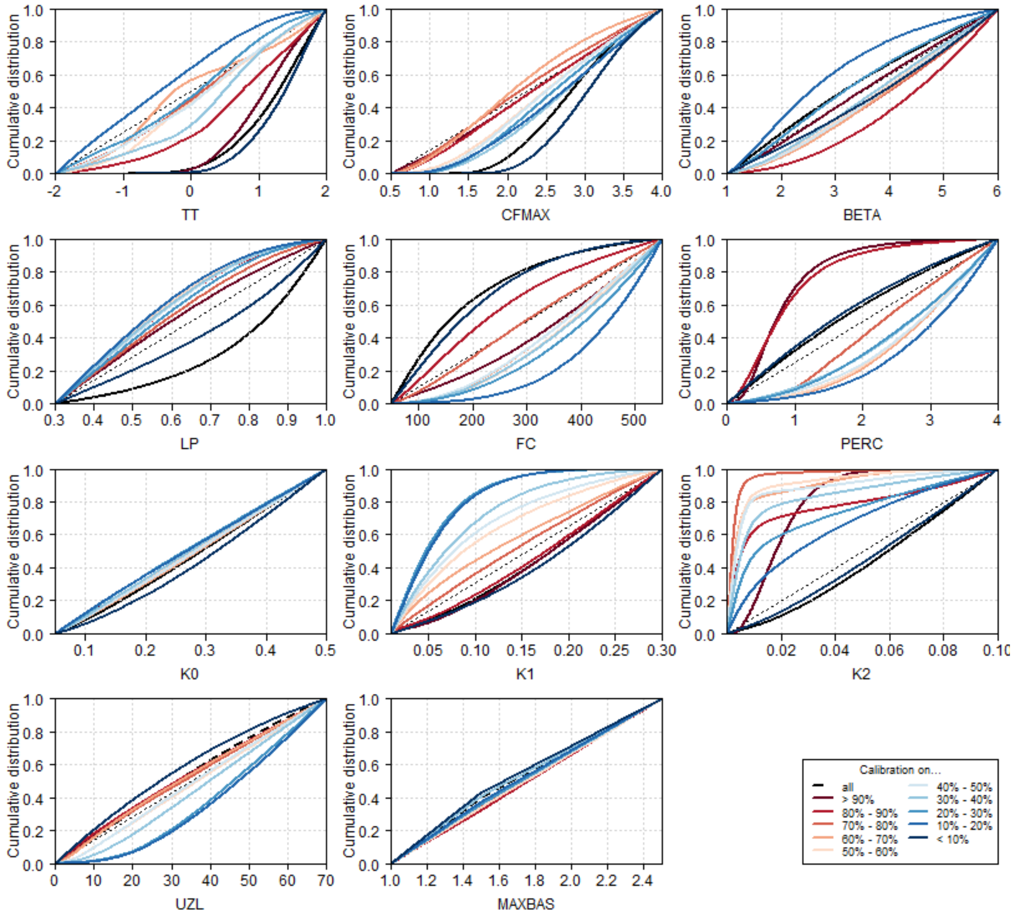


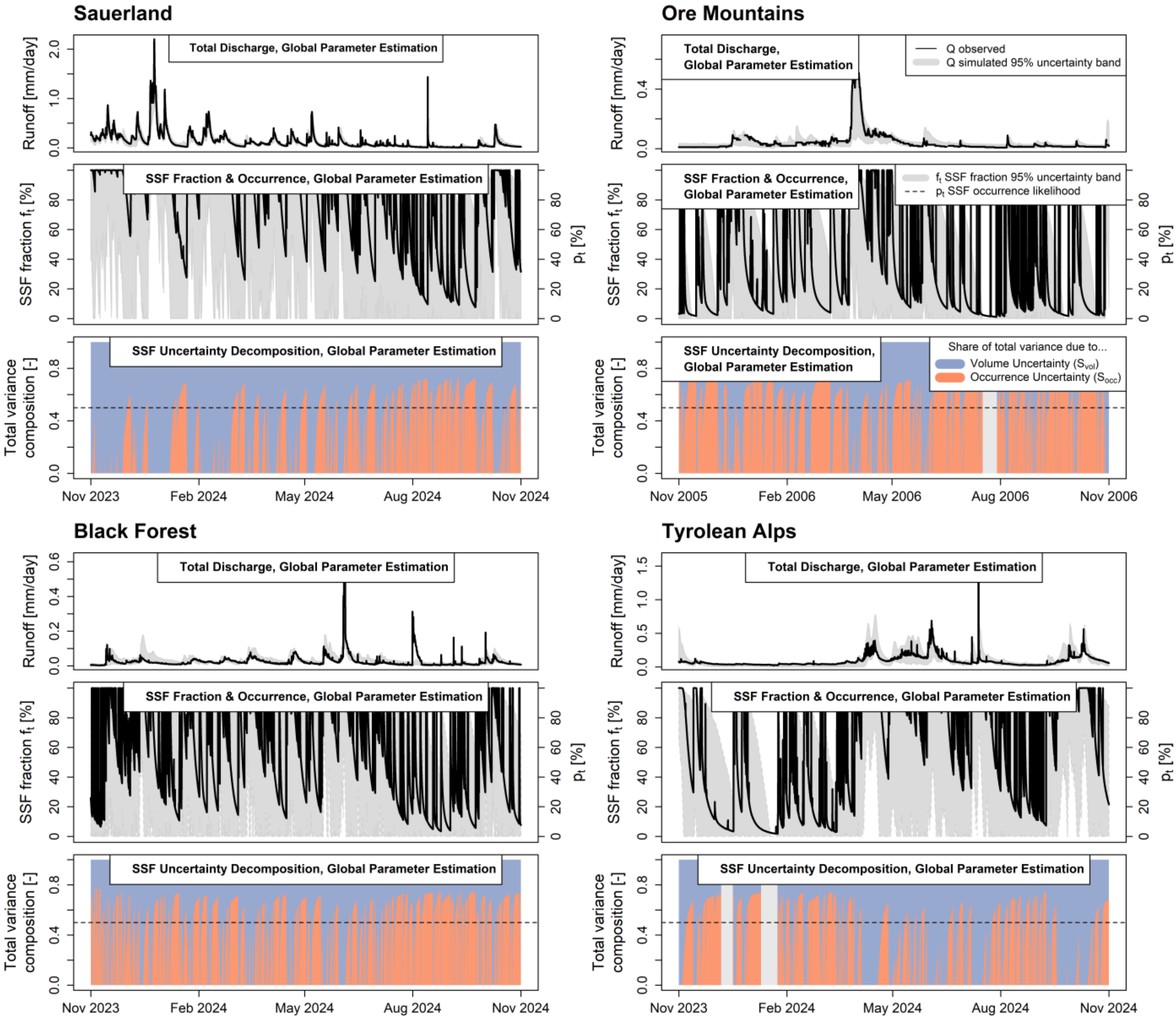
Figure A.3 Cumulative distribution functions of behavioural parameter sets (RSA) from global parameter estimation (black line, dotted) and percentile parameter estimations (colour coded– blue to red from high-flow parameter estimation to low-flow parameter estimation) for the Black Forest catchment

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026



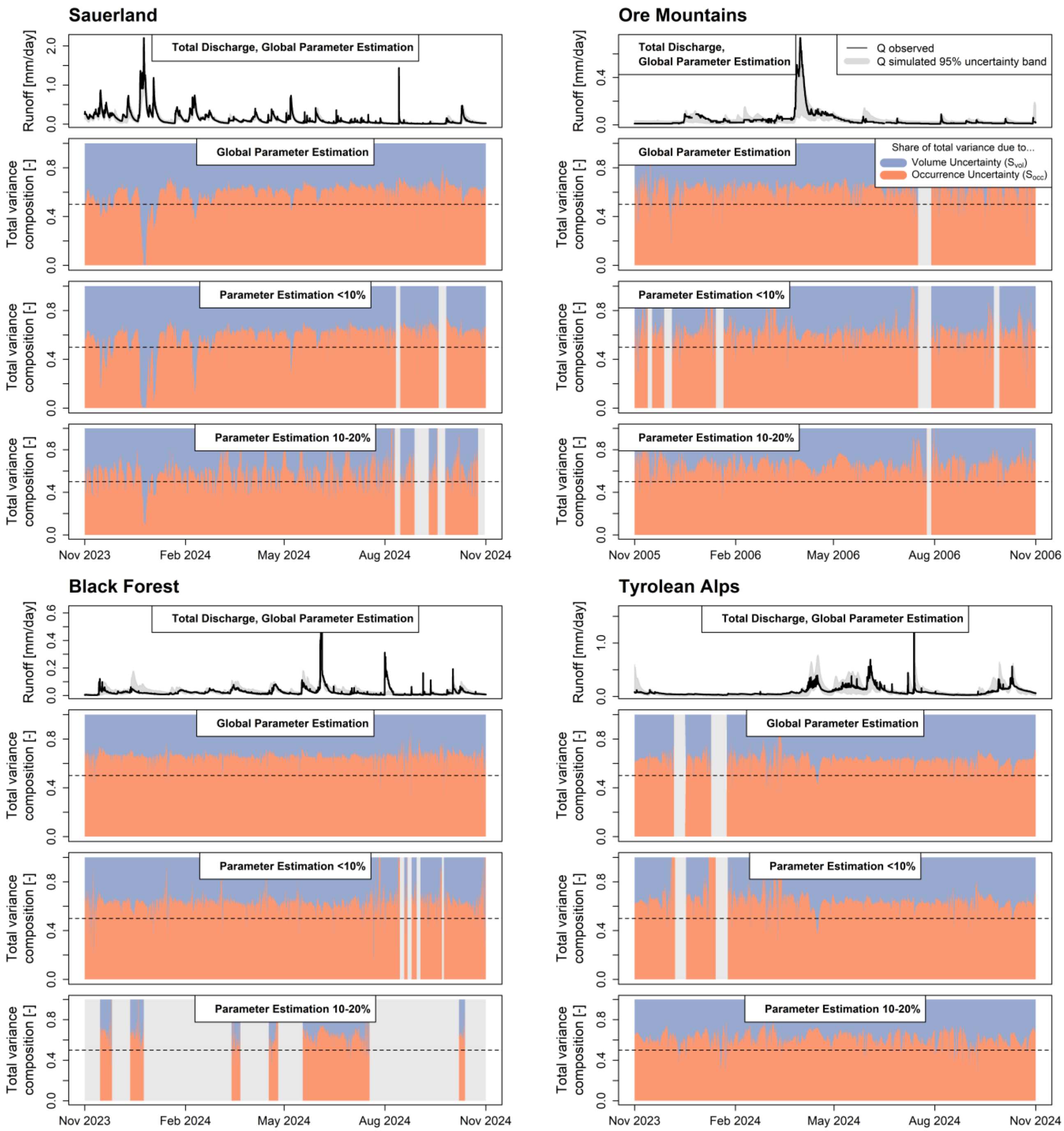
630 **Figure A.4 Cumulative distribution functions of behavioural parameter sets (RSA) from global parameter estimation (black line, dotted) and percentile parameter estimations (colour coded– blue to red from high-flow parameter estimation to low-flow parameter estimation) for the Tyrolean Alps catchment**

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026



635 **Figure A.5: Simulated SSF volume uncertainty, expressed as the uncertainty of SSF fraction of total discharge f_t, SSF occurrence probability p_t and SSF uncertainty decomposition for each catchment over one hydrological year when SSF is conceptually represented by the sum of Q0 and Q1, exemplary for the global parameter estimation.**

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026



640 **Figure A.6 SSF uncertainty decomposition for global parameter estimation, top 10% high-flow parameter estimation and top 10-20% high-flow parameter estimation over one hydrological year.**

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Table A.1 Occurrence uncertainty share (mean +- sd) over the entire calibrated timeseries

	Sauerland	Ore Mountains	Black Forest	Tyrolean Alps
Q_0 as SSF	0.63 +- 0.09	0.64 +- 0.08	0.68 +- 0.08	0.65 +- 0.11
Q_0 + Q_1 as SSF	0.39 +- 0.29	0.40 +- 0.29	0.50 +- 0.28	0.40 +- 0.31

Table A.2 Mean KGE of the behavioural runs from high-flow percentile parameter estimation

	Sauerland	Ore Mountains	Black Forest	Tyrolean Alps
Mean behavioural KGE <10% high-flow parameter estimation	0.31	0.33	0.28	0.54
Mean behavioural KGE 10-20% high-flow parameter estimation	-0.56	-1.88	-3.70	-2.83

Code availability

An exemplary code to reproduce the analyses presented in this study is available under https://github.com/iGW-TU-Dresden/SSF_uncertainty.

Author contributions

AH, TL, JS and FP contributed to the methodological development and conceptualization of the study. NS and TL pre-processed the data. TL performed the analyses and wrote the first draft of the manuscript. All authors revised the final manuscript.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

Acknowledgements

This research has been supported by the Deutsche Forschungsgemeinschaft as part of the Research Unit on Subsurface Stormflow (FOR 5288, HA 8113/10-1 & RE 4126/1-1). We thank the Ruhrverband, the Landestalsperrenverwaltung des Freistaates Sachsen, the Chair of Hydrology at the University of Freiburg and the Galleria di base del brennero - Brenner Basistunnel (BBT SE) for providing discharge data for our four catchments.

References

Abebe, N. A., Ogden, F. L., and Pradhan, N. R.: Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: Implications for parameter estimation, Journal of Hydrology, 389, 301–310, doi:10.1016/j.jhydrol.2010.06.007, 2010.

Abteilung Geoinformation (Land Tirol): Digitales Höhenmodell aus Laserscandaten, https://www.tirol.gv.at/sicherheit/geoinformation/geodaten-tiris/laserscandaten/, 2022.

Bachmair, S. and Weiler, M.: Hillslope characteristics as controls of subsurface flow variability, Hydrol. Earth Syst. Sci., 16, 3699–3715, doi:10.5194/hess-16-3699-2012, 2012.

Bachmair, S. and Weiler, M.: Interactions and connectivity between runoff generation processes of different spatial scales, Hydrological Processes, 28, 1916–1930, doi:10.1002/hyp.9705, 2014.

Barthold, F. K. and Woods, R. A.: Stormflow generation: A meta-analysis of field evidence from small, forested catchments, Water Resources Research, 51, 3730–3753, doi:10.1002/2014WR016221, 2015.

Beck, H. E., Pan, M., Lin, P., Seibert, J., van Dijk, A. I. J. M., and Wood, E. F.: Global Fully Distributed Parameter Regionalization Based on Observed Streamflow From 4,229 Headwater Catchments, JGR Atmospheres, 125, e2019JD031485, doi:10.1029/2019JD031485, 2020.

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

Bergström, S.: The HBV model—Its structure and applications, RH 4, SMHI Reports Hydrology, Norrköping, Sweden, 1992.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249, 11–29, doi:10.1016/S0022-1694(01)00421-8, 2001.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Beven, K. J., Smith, P. J., and Freer, J. E.: So just why would a modeller choose to be incoherent?, Journal of Hydrology, 354, 15–32, doi:10.1016/j.jhydrol.2008.02.007, 2008.

Bishop, K. H., Grip, H., and O'Neill, A.: The origins of acid runoff in a hillslope during storm events, Journal of Hydrology, 116, 35–61, doi:10.1016/0022-1694(90)90114-D, 1990.

Blume, T. and van Meerveld, H. J.: From hillslope to stream: methods to investigate subsurface connectivity, WIREs Water, 2, 177–198, doi:10.1002/wat2.1071, 2015.

Brunner, M. I., Slater, L., Tallaksen, L. M., and Clark, M.: Challenges in modeling and predicting floods and droughts: A review, WIREs Water, 8, e1520, doi:10.1002/wat2.1520, 2021.

Casella, G. and Berger, R.: Statistical Inference, Chapman and Hall/CRC, Boca Raton, 2024.

Champ, C. W. and Sills, A. V.: The Generalized Laws of Total Variance and Total Covariance, in: Applied Mathematical Analysis and Computations I, Wanduku, D., Zheng, S., Zhou, H., Chen, Z., Sills, A., Agyingi, E. (Eds.), Springer Proceedings in Mathematics & Statistics, Springer Nature Switzerland, Cham, 243–254, 2024.

Chifflard, P., Blume, T., Maerker, K., Hopp, L., van Meerveld, I., Graef, T., Gronz, O., Hartmann, A., Kohl, B., Martini, E., Reinhardt-Imjela, C., Reiss, M., Rinderer, M., and Achleitner, S.: How can we model subsurface stormflow at the catchment scale if we cannot measure it?, Hydrological Processes, 33, 1378–1385, doi:10.1002/hyp.13407, 2019.

Chifflard, P., Didszun, J., and Zepp, H.: Skalenübergreifende Prozess-Studien zur Abflussbildung in Gebieten mit periglazialen Deckschichten (Sauerland, Deutschland), Grundwasser, 13, 27–41, doi:10.1007/s00767-007-0058-1, 2008.

Climate Data Center CDC (DWD): Berechnete tägliche Werte von charakteristischen Elementen aus dem Boden und dem Pflanzenbestand., Version v19.3, 2019, https://opendata.dwd.de/climate_environment/CDC/derived_germany/soil/daily/historical/, 2025a.

Climate Data Center CDC (DWD): Stündliche Stationsmessungen der Lufttemperatur und Luftfeuchte für Deutschland, Version v24.03, https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/air_temperature, 2025b.

Climate Data Center CDC (DWD): Stündliche Stationsmessungen der Solarstrahlung (global/diffus) und der atmosphärischen Gegenstrahlung für Deutschland, Version v24.03, https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/solar/, 2025c.

Climate Data Center CDC (DWD): Stündliche Stationsmessungen des Niederschlags für Deutschland, Version v24.03, https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/precipitation, 2025d.

Du, E., Rhett Jackson, C., Klaus, J., McDonnell, J. J., Griffiths, N. A., Williamson, M. F., Greco, J. L., and Bitew, M.: Interflow dynamics on a low relief forested hillslope: Lots of fill, little spill, Journal of Hydrology, 534, 648–658, doi:10.1016/j.jhydrol.2016.01.039, 2016.

Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., and Kendall, C.: The role of bedrock topography on subsurface storm flow, Water Resources Research, 38, 5-1-5-16, doi:10.1029/2001WR000872, 2002.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Galleria di base del brennero - Brenner Basistunnel (BBT SE): Pegeldaten Padasterbach, 2025.

Geobasis NRW, B. K.: DGM1, https://www.opengeodata.nrw.de/produkte/geobasis/hm/dgm1_tiff/dgm1_tiff/, 2022.

GeoSphere Austria: WINFORE v2.1, https://data.hub.geosphere.at/dataset/winfore-v2-1d-1km, 2020.

GeoSphere Austria: INCA Stundendaten, https://data.hub.geosphere.at/dataset/inca-v1-1h-1km, 2021.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region, Weather and Forecasting, 26, 166–183, doi:10.1175/2010WAF2222451.1, 2011.

Hartmann, A., Payeur-Poirier, J.-L., and Hopp, L.: Incorporating experimentally derived streamflow contributions into model parameterization to improve discharge prediction, Hydrol. Earth Syst. Sci., 27, 1325–1341, doi:10.5194/hess-27-1325-2023, 2023.

Haslinger, K. and Bartsch, A.: Creating long-term gridded fields of reference evapotranspiration in Alpine terrain based on a recalibrated Hargreaves method, Hydrol. Earth Syst. Sci., 20, 1211–1223, doi:10.5194/hess-20-1211-2016, 2016.

Heller, K. and Kleber, A.: Hillslope runoff generation influenced by layered subsurface in a headwater catchment in Ore Mountains, Germany, Environ Earth Sci, 75, 943, doi:10.1007/s12665-016-5750-y, 2016.

Hopp, L. and McDonnell, J. J.: Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth, Journal of Hydrology, 376, 378–391, doi:10.1016/j.jhydrol.2009.07.047, 2009.

Hornberger, G. M. and Spear, R.: An Approach to the Preliminary Analysis of Environmental Systems, 12:1, 1981.

Hughes, D. A. and Farinosi, F.: Unpacking some of the linkages between uncertainties in observational data and the simulation of different hydrological processes using the Pitman model in the data scarce Zambezi River basin, Hydrological Processes, 35, doi:10.1002/hyp.14141, 2021.

Kann, A., Haiden, T., Emde, K. von der, Gruber, C., Kabas, T., Leuprecht, A., and Kirchengast, G.: Verification of Operational Analyses Using an Extremely High-Density Surface Station Network, Weather and Forecasting, 26, 572–578, doi:10.1175/WAF-D-11-00031.1, 2011.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, doi:10.1029/2005WR004362, 2006.

Kolmogorov, A.: Sulla determinazione empirica di una legge didistribuzione, Giornale Dell'instituto Italiano Degli Attuari, 4, 89–91, 1933.

Landesamt für Geobasisinformationen Sachsen (GeoSN): Digitale Höhenmodelle, https://www.geodaten.sachsen.de/downloadbereich-digitale-hoehenmodelle-4851.html, 2022.

Landesamt für Geoinformation und Landentwicklung Baden-Württemberg (LGL): Digitales Geländemodel (DGM1), https://www.lgl-bw.de, 2025.

Landestalsperrenverwaltung des Freistaates Sachsen (LTV): Pegledaten Burkersdorf, 2022.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 201, 272–288, doi:10.1016/s0022-1694(97)00041-3, 1997.

Markart, G., Römer, A., Bieber, G., Pirkl, H., Klebinder, K., Hörfarter, C., Ita, A., Jochum, B., Kohl, B., and Motschka, K.: Assessment of Shallow Interflow Velocities in Alpine Catchments for the Improvement of Hydrological Modelling, in: Engineering Geology for Society and Territory - Volume 3, Springer, Cham, 611–615, 2015.

Massey, F. J.: The Kolmogorov-Smirnov Test for Goodness of Fit, Journal of the American Statistical Association, 46, 68–78, doi:10.1080/01621459.1951.10500769, 1951.

McGuire, K. J., Klaus, J., and Jackson, C. R.: JAMES BUTTLE REVIEW: Interflow, subsurface stormflow and throughflow: A synthesis of field work and modelling, Hydrological Processes, 38, e15263, doi:10.1002/hyp.15263, 2024.

McMillan, H., Araki, R., Bolotin, L., Kim, D.-H., Coxon, G., Clark, M., and Seibert, J.: Global patterns in observed hydrologic processes, Nat Water, 3, 497–506, doi:10.1038/s44221-025-00407-w, 2025.

Moges, E., Demissie, Y., Larsen, L., and Yassin, F.: Review: Sources of Hydrological Model Uncertainties and Advances in Their Analysis, Water, 13, 28, doi:10.3390/w13010028, 2021.

Orsi, G., Burger, U., Marschallinger, R., and Nocker, C.: Geological model of an alpine lateral valley with implications for the design of a groundwater monitoring network – the example of the Padaster Valley (Eastern Alps, Austria), AJES, 109, doi:10.17738/ajes.2016.0003, 2016.

Pianosi, F., Sarrazin, F., and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environmental Modelling & Software, 70, 80–85, doi:10.1016/j.envsoft.2015.04.009, 2015.

Pianosi, F. and Wagener, T.: Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis, Hydrological Processes, 30, 3991–4003, doi:10.1002/hyp.10968, 2016.

Ruhrverband Essen - Abteilung WW: Wasserstands- und Durchflussganglinien, https://www.talsperrenleitzentrale-ruhr.de/online-daten/gewaesserpegel/, 2025.

Scanlon, T. M., Raffensperger, J. P., Hornberger, G. M., and Clapp, R. B.: Shallow subsurface storm flow in a forested headwater catchment: Observations and modeling using a modified TOPMODEL, Water Resources Research, 36, 2575–2586, doi:10.1029/2000WR900125, 2000.

Schulla, J. and Jasper, K.: Model description wasim-eth, 2007.

Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, Agricultural and Forest Meteorology, 98-99, 279–293, doi:10.1016/s0168-1923(99)00105-7, 1999.

Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrol. Earth Syst. Sci., 4, 215–224, doi:10.5194/hess-4-215-2000, 2000.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resources Research, 38, 23-1-23-14, doi:10.1029/2001WR000978, 2002.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol. Earth Syst. Sci., 16, 3315–3325, doi:10.5194/hess-16-3315-2012, 2012.

Seibert, J., Vis, M. J. P., Lewis, E., and van Meerveld, H. J.: Upper and lower benchmarks in hydrological modelling, Hydrological Processes, 32, 1120–1125, doi:10.1002/hyp.11476, 2018.

Shen, H., Tolson, B. A., and Mai, J.: Time to Update the Split-Sample Approach in Hydrological Model Calibration, Water Resources Research, 58, doi:10.1029/2021WR031523, 2022.

Shen, X., Liu, J., Han, X., Yang, H., Liu, H., and Ni, F.: Modelling Infiltration Based on Source-Responsive Method for Improving Simulation of Rapid Subsurface Stormflow, Water Resources Research, 61, e2024WR037487, doi:10.1029/2024WR037487, 2025.

Smirnov, H.: Sur les écarts de la courbe de distribution empirique, Recueil Mathématique (Matematiceskii Sbornik), 6, 3–26, 1939.

Song, Y., Zhang, Y., Chen, N., Chen, L., Zeng, X., and Qiu, A.: A field and modeling study of subsurface stormflow for Huanggou Hillslope, Journal of Hydrology: Regional Studies, 52, 101683, doi:10.1016/j.ejrh.2024.101683, 2024.

Spear, R. and Hornberger, G. M.: Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis, Water Research, 14, 43–49, doi:10.1016/0043-1354(80)90040-8, 1980.

Tibangayuka, N., Mulungu, D. M., and Izdori, F.: Performance evaluation, sensitivity, and uncertainty analysis of HBV model in Wami Ruvu basin, Tanzania, Journal of Hydrology: Regional Studies, 44, 101266, doi:10.1016/j.ejrh.2022.101266, 2022.

Tromp-van Meerveld, H. J. and McDonnell, J. J.: Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope, Water Resources Research, 42, doi:10.1029/2004WR003778, 2006.

Uchida, T., Asano, Y., Mizuyama, T., and McDonnell, J. J.: Role of upslope soil pore pressure on lateral subsurface storm flow dynamics, Water Resources Research, 40, doi:10.1029/2003WR002139, 2004.

Universität Freiburg, Professur für Hydrologie: Pegeldaten Rütlitobel, 2025.

Usman, M. N., Leandro, J., Broich, K., and Disse, M.: Multi-objective calibration and uncertainty analysis for the event-based modelling of flash floods, Hydrological Sciences Journal, 69, 456–473, doi:10.1080/02626667.2024.2322599, 2024.

Wagener, T. and Gupta, H. V.: Model identification for hydrological forecasting under uncertainty, Stoch Environ Res Ris Assess, 19, 378–387, doi:10.1007/s00477-005-0006-5, 2005.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V.: Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, Hydrological Processes, 17, 455–476, doi:10.1002/hyp.1135, 2003.

Wang, H., Bai, Y., and Huang, D.: Study of experimental and numerical simulation on the influence of gravel on the interflow of slope land, Environmental science and pollution research international, 31, 11716–11726, doi:10.1007/s11356-023-31808-7, 2024.

https://doi.org/10.5194/egusphere-2026-4470
Preprint. Discussion started: 25 August 2026

Wang, K. and Dickinson, R. E.: A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, Reviews of Geophysics, 50, doi:10.1029/2011RG000373, 2012.

Weiler, M., McDonnell, J. J., Tromp-van Meerveld, I., and Uchida, T.: Subsurface Stormflow, Encyclopedia of Hydrological Sciences, John Wiley & Sons, Ltd, 2006.

Wheater, H. S., Bishop, K. H., and Beck, M. B.: The identification of conceptual hydrological models for surface water acidification, Hydrological Processes, 1, 89–109, doi:10.1002/hyp.3360010109, 1986.

Williams, G. P.: Friends don't let friends use Nash-Sutcliffe Efficiency (NSE) or KGE for hydrologic model accuracy evaluation: A rant with data and suggestions for better practice, Environmental Modelling & Software, 194, 106665, doi:10.1016/j.envsoft.2025.106665, 2025.

Ye, W., Bates, B. C., Viney, N. R., Sivapalan, M., and Jakeman, A. J.: Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments, Water Resources Research, 33, 153–166, doi:10.1029/96WR02840, 1997.

Zelelew, M. B. and Alfredsen, K.: Sensitivity-guided evaluation of the HBV hydrological model parameterization, Journal of Hydroinformatics, 15, 967–990, doi:10.2166/hydro.2012.011, 2013.

Zhao, P., Tang, X., Zhao, P., Wang, C., and Tang, J.: Identifying the water source for subsurface flow with deuterium and oxygen-18 isotopes of soil water collected from tension lysimeters and cores, Journal of Hydrology, 503, 1–10, doi:10.1016/j.jhydrol.2013.08.033, 2013.

Zillgens, B., Merz, B., Kirnbauer, R., and Tilch, N.: Analysis of the runoff response of an alpine catchment at different scales, Hydrol. Earth Syst. Sci., 11, 1441–1454, doi:10.5194/hess-11-1441-2007, 2007.