AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR036420

Toward Flash Flood Modeling Using Gradient Resolving Representative Hillslopes

Ashish Manoj J[1], Ralf Loritz[1], Franziska Villinger[1], Mirko Mälicke[1], Mehdi Koopaeidar[1], Hans Göppert[2], and Erwin Zehe[1]

[1]Chair of Hydrology, Institute of Water and Environment (IWU), Karlsruhe Institute of Technology, Karlsruhe, Germany, [2]Wald + Corbe Consulting GmbH, Hügelsheim, Germany

Key Points:

- Physically based representative hillslope models can be used for flash flood predictions in small data-scarce and rural catchments
- Climate reanalysis data enable the initialization of a process-based model, helping to reduce the uncertainties in estimating antecedent soil conditions
- Transfer of model parameters within the same hydrological landscape is feasible

Correspondence to:
Ashish Manoj J,
ashish.jaseetha@kit.edu

Citation:
Manoj J, A., Loritz, R., Villinger, F., Mälicke, M., Koopaeidar, M., Göppert, H., & Zehe, E. (2024). Toward flash flood modeling using gradient resolving representative hillslopes. *Water Resources Research*, *60*, e2023WR036420. https://doi.org/10.1029/2023WR036420

Received 4 OCT 2023
Accepted 31 MAY 2024

Author Contributions:
Conceptualization: Ashish Manoj J, Franziska Villinger, Mehdi Koopaeidar, Hans Göppert, Erwin Zehe
Data curation: Ashish Manoj J
Formal analysis: Ashish Manoj J
Funding acquisition: Erwin Zehe
Investigation: Ashish Manoj J
Methodology: Ashish Manoj J, Ralf Loritz, Franziska Villinger, Hans Göppert, Erwin Zehe
Project administration: Erwin Zehe
Resources: Ashish Manoj J, Ralf Loritz, Franziska Villinger, Mirko Mälicke
Software: Ashish Manoj J, Mirko Mälicke
Supervision: Ralf Loritz, Erwin Zehe
Validation: Ashish Manoj J, Ralf Loritz
Visualization: Ashish Manoj J
Writing – original draft: Ashish Manoj J
Writing – review & editing: Ashish Manoj J, Ralf Loritz, Franziska Villinger, Mirko Mälicke, Mehdi Koopaeidar, Hans Göppert, Erwin Zehe

Abstract It is increasingly acknowledged that the acceleration of the global water cycle, largely driven by anthropogenic climate change, has a disproportionate impact on sub-daily and small-scale hydrological extreme events such as flash floods. These events occur thereby at local scales within minutes to hours, typically in response to high-intensity rainfall events associated with convective storms. In the present work, we show that by employing physically based representative hillslope models that resolve the main gradients controlling overland flow hydrology and hydraulics, we can get reliable simulations of flash flood response in small data-scarce catchments. To this end, we use climate reanalysis products and transfer soil parameters previously obtained for hydrological predictions in an experimental catchment in the same landscape. The inverted mass balance of flood reservoirs downstream is employed for model evaluation in these nearly ungauged basins. We show that our approach using representative hillslopes and climate data sets can provide reasonable uncalibrated estimates of the overland runoff response (flood magnitude, storm volume, and event runoff coefficients) in three of the four catchments considered. Given that flash floods typically occur at scales of a few km^2 and in ungauged places, our results have implications for operational flash flood forecasting and open new avenues for using gradient resolving physically based models for the design of small and medium flood retention basins around the world.

Plain Language Summary Flash floods have become increasingly common worldwide, with catastrophic damages to both human life and the economy. While the extent of global warming and climate change impacting these events is still under much debate, it is almost certain now that we need to be better equipped to understand and model these extremes to prevent and mitigate the possible risk to human life and infrastructure in a warming climate. To test, if we can use first principles derived from thermodynamic conservation laws and process based hydrological models for the same, we modeled flash flood response in four headwater catchments over Southern Germany using the concept of "representative hillslope." Since the regions considered in our work are poorly gauged, we made use of global climate reanalysis products and parameter transfer from past experiments. The encouraging results obtained in predicting the flood magnitude and volume speak to the overall applicability of our approach. We are able to get decent uncalibrated predictions in three out of the four catchments considered with minimum computational effort. Understanding and managing the adverse impacts of such extreme hydroclimatic events remains one of the crucial hurdles facing humanity toward the sustainable development goals (SDG17) in this decade.

1. Introduction

As early as 2008, the Organisation for Economic Co-operation and Development (OECD) highlighted climate change and hydro-meteorological extremes as among the most pressing challenges facing humanity. Flood events, a key component of these extremes, manifest at varying spatial and temporal scales, each driven by distinct meteorological conditions. Flash floods, for instance, occur on local scales within a span of minutes to hours. These events are triggered by high-intensity rainfall from convective storms, often resulting in strong infiltration excess and significant Hortonian overland flow (Bronstert et al., 2018; Marchi et al., 2010, 2016; Meyer et al., 2022; Ruiz-Villanueva et al., 2012). While these floods pose immediate risks, such as loss of human life, their consequences extend to long-term impacts like soil erosion, sediment transport, and subsequent deterioration of water quality and soil fertility, particularly in agricultural settings. On the other end of the spectrum are large-scale riverine floods, which occur due to synoptic scale low-pressure systems characterized by widespread and sustained precipitation. Unlike flash floods, these events are governed by capacity-controlled

runoff formation processes like saturation excess, known as Dunne overland flow (Dunne & Kirkby, 1978), and subsurface stormflow. Additionally, flood routing in channel networks and snowmelt contributions, play crucial roles (Blöschl et al., 2007). This stands in contrast to the Hortonian overland flow (Horton, 1933) typically observed in flash floods driven by convective storms and affected by preferential flow processes (Bronstert et al., 2023).

Flood forecasting and risk management have to cope with both types of flood events, and both are naturally highly sensitive to climate change (IPCC, 2021). The largest observed floods in many European rivers have occurred in the last three decades, which count among the most flood-rich periods in the past 500 years (Blöschl et al., 2020). The same holds true for local flash floods. For instance, 22 flash floods in southwest Germany occurring in the past 20 years, had estimated design return periods exceeding 500 years (Goppert, 2018). While return periods of that magnitude are uncertain, they indicate nevertheless that these are exceptional events. In a broader context, the finding is in line with the recent accumulation of flash floods in Europe (Meyer et al., 2022). This likely reflects the already ongoing acceleration of the hydrological cycle, with expected increasing frequencies of more intense convective rainstorms (Bürger et al., 2019; Mueller & Pfister, 2011) and flash floods due to Clausius-Clapeyron scaling (Pall et al., 2007). This is alarming, as the flash flood series in the summer of 2016 alone caused about €2.5 bn of damage in Germany (Munich Re, 2016). All this recent evidence calls for improving the current standards in (a) flood predictions and (b) methods for deriving hydrological extreme values for design. Flash floods are, unfortunately, rarely observed with conventional rain and discharge measurement networks (Borga et al., 2008), which implies that the sample for model learning and testing is small. While being rare, flash floods typically impact small catchments or even specific hillslopes, which are largely unobserved and have never experienced such events in the near observational past. We thus argue that the problem of flash flood predictions strongly relates to the classical "predictions in ungauged basins - PUB problem" (Hrachowitz et al., 2013; Sivapalan et al., 2003).

Despite their success, conceptual hydrological models have clear limitations in the context of local flash floods and Hortonian overland flow formation (Fatichi et al., 2016; Hrachowitz & Clark, 2017). Infiltration excess or Hortonian overland flow occurs when rainfall intensity exceeds the soil infiltration rate and is quite difficult to capture and characterize. This is because infiltration is a highly discontinuous and spatially variable process. Previous studies have reported a possible elated infiltration capacity by the emergence of macropore flow (Bronstert et al., 2023), which allows for rapid, purely gravity driven infiltration, bypassing the slow soil matrix. An increasing macropore density can thus drastically reduce Hortonian overland flow generation (Zehe & Blöschl, 2004). Rain splash may, the other way around, lead to aggregate breakdown and siltation of the surface and significantly reduce infiltration capacity during high intense rainfall. This may drastically enhance Hortonian overland flow production, especially on agricultural fields during stages of low canopy coverage (Niehoff et al., 2002; Villinger et al., 2022).

Here, we hypothesize that gradient resolving, physically based hydrological models (Fatichi et al., 2016; Paniconi & Putti, 2015) are better suited to address the challenges of flash flood predictions at small scales. By solving coupled partial differential equations that represent infiltration, soil moisture dynamics, overland flow and streamflow hydraulics and evaporation, such models offer better opportunities for capturing spatially distributed production and concentration of overland flow (Pérez et al., 2011; Steinbrich et al., 2016; Zehe et al., 2001).

The primary challenges in employing physically based models for flash flood predictions in largely unobserved areas lie in their demand for extensive input data and their computational expense. To address the latter, we favor the use of the recently proposed representative hillslope concept by Loritz et al. (2017), which has not yet been tested for uncalibrated predictions and extreme flood events. To overcome the data requirement challenge, we propose to leverage existing information from well-studied experimental catchments within the same hydrological landscape. Specifically, we suggest utilizing these well-instrumented catchments as 'donors' to transfer data on soil hydraulic properties and surface roughness to the poorly instrumented target catchments (donor models) for making a priori (uncalibrated) predictions. This could then give way for a posteriori (manual calibration) exercise once more relevant event information becomes available. It's also important to understand the significance of such a parameter transfer in comparison to estimating soil hydraulic functions using pedotransfer functions (Szabó et al., 2021; Zhang & Schaap, 2017) and readily available soil texture data.

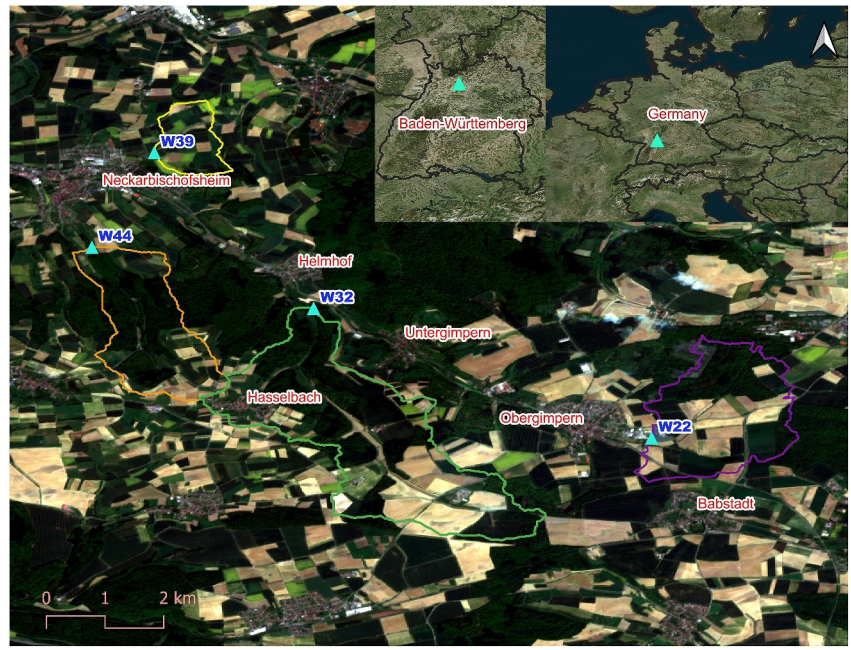

Figure 1. Overview of the location of the four headwater catchments considered in the study. Also, shown in the figure are the downstream flood control reservoirs, which afford protection to the towns in the region. The overlay layer depicts a Sentinel-2 (Drusch et al., 2012) multispectral true color composite image for the region during May–June 2016. Figure B1 in Appendix B shows the stream network of Krebsbach and Schwarzbach and also the total accumulated precipitation during the event.

Our main aim is thus to evaluate the potential of physically based hydrological models for predicting flash floods triggered by convective rainstorms in data-scarce regions. In this context, we put special emphasis on interactions between overland flow velocities, controlled by surface roughness, and Hortonian overland flow generation, controlled by the initial soil water content and infiltration characteristics, including macropores and surface sealing. To address the related challenges, we thus explore whether:

1. It is feasible to transfer model parameters from the well monitored experimental catchment to the data-scarce catchments for a priori uncalibrated flash flood predictions in response to convective storms, and how does this compare to simulations with parameters derived from soil pedotransfer functions?
2. Climate reanalysis data can be used to initialize process-based hydrological models in these data scarce regions and how much uncertainty is reduced using such a dynamic spin up compared to a random estimate of the antecedent soil conditions?
3. Reservoirs in headwater catchments (<5 km^2) and their inverted mass balance offer largely untapped potential to study flash floods, complementary to a rather low number of operational stream gauges that are available at this scale worldwide?

2. Venue and Model

2.1. Study Area

The four target headwater catchments ”W22,” “W32,” “W39” belong to the Elsenz-Schwarzbach catchment in the State of Baden Württemberg, Southern Germany (Figure 1 & Figure B1 in Appendix B). The catchments are located within the eastern “Kraichgau,” west of Bad Rappenau and around 50 km from the nearest cities - Heidelberg and Karlsruhe. Due to a series of catastrophic flooding episodes in 1993–1994, a comprehensive flood protection concept for the entire region was envisaged, which led to the development of local flood retention basins throughout the catchment area. The size of the catchments varies from 1 to 6 km^2; they all drain into the Krebsbach, which joins into the Schwarzbach near Waibstadt (the nearest gauging station—Eschelbronn Schwarzbach, being more than 12 km from our study area). The Elsenz-Schwarzbach finally merges into the Neckar, one of the Rhine's largest tributaries. From Figure 1, it is clear that even though the catchments are

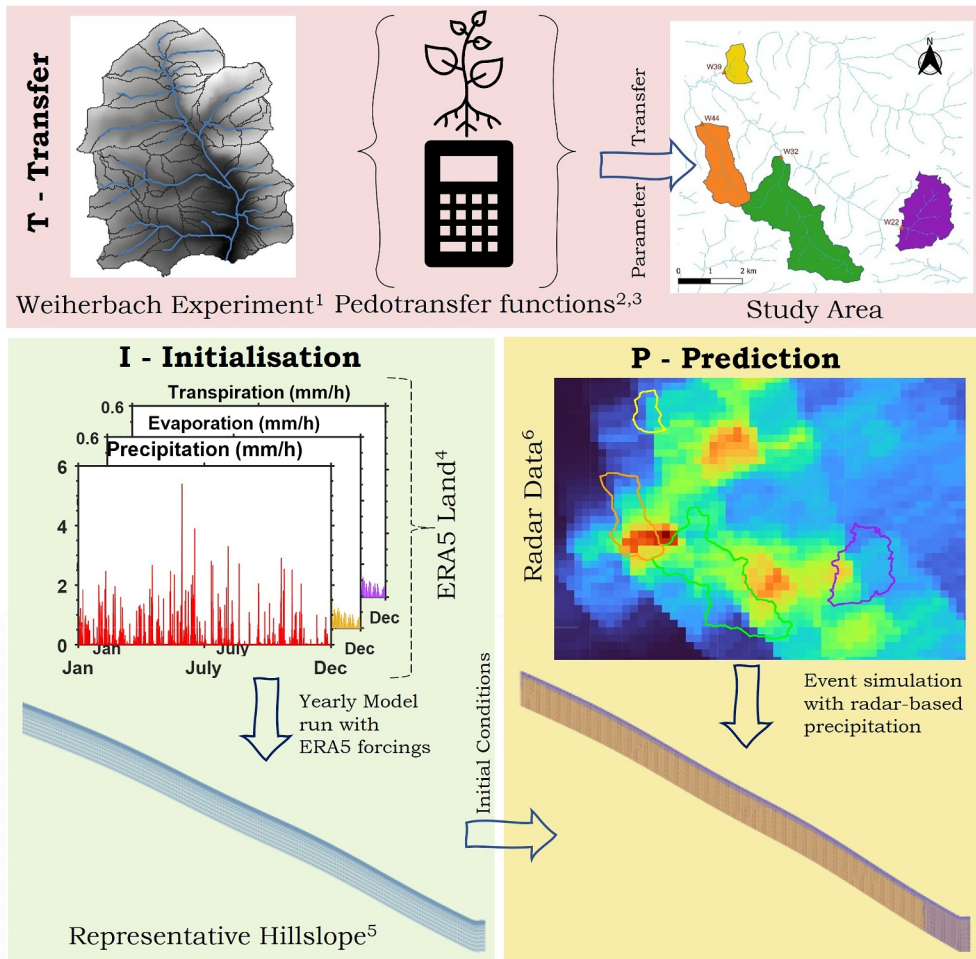

Figure 2. Illustration of our methodological approach– Transfer (T), Initialization (I), and Prediction (P). In Transfer (T), we transfer our knowledge of hillslope properties and soil parameters from the Weiherbach to our study area in Krebsbach. The Initialization (I) phase involves deriving the representative hillslope (detailed in Figure 3) for the catchments and using the ERA5 Land forcings (hourly and 0.1° × 0.1°) to run the hillslope model for an entire year. In the prediction phase, (P), the same model is run with the fine-resolution radar forcing and initial conditions from Initialization (I) for predicting the flash flood discharge (1—Zehe et al. (2001), 2—Zhang and Schaap (2017), 3—Szabó et al. (2021), 4—Muñoz-Sabater et al. (2021), 5—Figure 3 & Loritz et al. (2017), 6—Kachelmannwetter, n. d (Radar Data.).

primarily agricultural in nature (major crops being—cereals, maize, sugar beets and potatoes), they are situated upstream of the population centers of the region. While these catchments are in the same hydrological landscape as the previously monitored Weiherbach experimental headwater shed (Schierholz et al., 2000), they are, despite the available water level gauges in the reservoirs, completely unmonitored with respect to rainfall, streamflow, soil moisture and soil hydraulic properties.

During the end of May to early June 2016, several strong convective rainfall events clustered in Germany because of persistent atmospheric conditions (Bronstert et al., 2018; Meyer et al., 2022; Piper et al., 2016). Rain totals exceeding 100 mm were reported in a day (in some cases even within 2 hours), triggering flash floods in many small catchments over Southern Germany. The impacts in the Elsenz-Schwarzbach were also severe, with several of the flood control reservoirs being overtopped. To investigate the feasibility of our approach (Figure 2) and of the CAFLOW model to simulate such events, we focus our attention on the severe event of 08 June 2016 in the region (Appendix B). Since no streamflow gauges are available for the four headwater catchments, we use the water level measurements in the flood control reservoirs to estimate the runoff response based on the reconstructed reservoir inflow (W22, W32, W39, and W44). The storm runoff response is calculated based on inverting the reservoir mass balance with the knowledge of the reservoir geometry and stage-outflow relationship (Appendix C). Related uncertainties are accounted for by using a relative percentage error value (5%) in the stage level measurements.

2.2. CATFLOW in a Nutshell

The quest for accurately identifying and modeling the governing processes of the water balance and reactive pesticide transport in rural catchments motivated both the setup of the Weiherbach experimental catchment in the 90s (Plate & Zehe, 2008; Zehe et al., 2001) and the development of the process-based model CATFLOW (Zehe et al., 2001). The model relies on the subdivision of a catchment into several 2D hillslopes and an interconnected drainage (optional) and river network. However, each hillslope is modeled separately and hence this provides the opportunity to run each hillslope individually without the associated stream network. Hillslopes are discretized along a two-dimensional cross section using terrain following curvilinear orthogonal coordinates. Soil water dynamics within the hillslopes are characterized using the potential based form of the 2D Darcy–Richards equation, solved by a mass conservative Picard solver using adaptive time stepping (Celia & Bouloutas, 1990). Soil hydraulic properties can be parameterized according to van Genuchten (1980) and Mualem (1976), Tang and Skaggs (1977) or the recently proposed PDI model (Peters et al., 2021).

Overland flow is simulated using the diffusion wave approximation of the Saint-Venant equation and explicit upstreaming, in combination with the Gauckler-Manning-Strickler formula. CATFLOW can account for partially saturated and fully saturated subsurface flow as it solves the Darcy-Richards equation in potential-based form. The study of Loritz et al. (2017) explains how this, together with no-flow boundary conditions in the lower parts of the hillslope, can be used to create a riparian zone with a local active groundwater body to simulate baseflow. Evaporation and transpiration are usually simulated using a SVAT (Surface Vegetation—Atmosphere Transfer) module based on the Penman–Monteith equation, accounting for annual cycles of plant phenology, albedo, and roughness using tabulated data. Stomatal conductance is characterized after Jarvis (1976), or via the inversion of sap flow data (Loritz et al., 2022).

The model can represent macropores either as stochastically generated spatially connected pathways of very low flow resistance (similar to fractures) or via an effective macroporosity factor, which scales the ratio between infiltration into the macropore and the matrix domain (Loritz et al., 2017; Wienhofer & Zehe, 2014). For the simplified macroporosity factor approach, as detailed in Zehe et al. (2001), the bulk hydraulic conductivity k_s^B, is linearly increased by a relative scaling factor, f_m (if relative saturation (S) at a grid point exceeds a certain threshold (S_o)) as follows:

$$\begin{aligned} k_s^B &= k_s + k_s f_m \frac{S - S_o}{1 - S_o} \quad if \quad S_o \le S < 1 \\ k_s^B &= k_s \qquad\qquad \text{otherwise} \end{aligned} \tag{1}$$

3. Methodology

Setting up a process based model of any hydrological system mainly requires two types of information (Remson et al., 1971). The first one concerns the fundamental laws governing the dynamics of system state variables and fluxes and related process parametrizations (e.g., preferential macropore flow). Our selection of the appropriate model implicitly selects these conditions for the given problem.

The second involves the data representing the "landscape" in the equation set. A proper identification of these properties is crucial for reliable model performance, and they can be divided into (a) system geometry (topographic and river network properties), (b) system parameters (soil and vegetation characteristics) and (c) initial and boundary conditions (antecedent soil moisture, rainfall and meteorological forcings). The current section details the steps required for setting up the model in this respect. We first explain the concept of the representative hillslope and its derivation from digital topographical data, then elaborate on the transfer of soil and land use parameters from the Weiherbach, along with the derivation of parameters via pedotransfer functions. Finally, we explain the spin up of the model using ERA5 Land and the radar-based precipitation product used during the event simulation.

3.1. The Representative Hillslope Concept

Physically based hydrological models are renowned for their substantial computational demands, often impeding their broader application (Paniconi & Putti, 2015). As a result, catchment hydrology research has pivoted toward

simplifying these models, ensuring they retain their physical underpinnings (Güntner & Bronstert, 2004). Notable models that exemplify this approach include the hillslope storage Boussinesq model by Troch et al. (2003) and the representative elementary watershed model proposed by Reggiani et al. (1998). In this study, we adopt a gradient-based simplification termed "representative hillslopes," as introduced by Loritz et al. (2017). Their work demonstrated that the water balance and streamflow generation in the Colpach catchment (19 km^2) could be accurately simulated using a single representative hillslope, negating the need for an associated river network.

Here we provide a concise explanation on why this approach works. The concept behind a *representative hillslope* is that both surface and subsurface water fluxes are propelled by differences in potential energy (Loritz et al., 2017; Zehe et al., 2013). These differences emerge from rainfall distribution over varied topography. In the context of intense convective rainstorms, our focus narrows to the energy balance of overland flow. Here, the driving potential energy difference hinges on the relative elevation between a location and its corresponding flow outlet. It's crucial to recognize that only a minute portion of this potential energy is converted into overland flow kinetic energy, with the majority being dissipated, primarily influenced by factors such as Manning's roughness (Schroers et al., 2022).

Preserving this energy dynamic on average implies that the topography of the representative hillslope should be structured to maintain average topographic gradients along the flow path to the nearest drainage point. A viable method involves segmenting geo-potential energy by proximity to the river and averaging within each segment. Specifically, we consider the distribution of flow profile lines shown in Figure 3b for catchment W22. For any distance class (also shown in Figure A1: Appendix A), the total flow potential is the sum of the potentials of all the cells within the class, which is proportional to the relative elevation difference of the cells. For the catena profile, we require a representative value for this class so that the total energy remains on average conserved. We use the Linear Average Representative Slope Profile concept from Francke et al. (2008) for the same. The method involves a weighting factor based on the relative occurrence of each cell in a flow path (characterized by the flow accumulation values). Therefore, the value of the mean elevation (h_i) for a class at distance i:

$$h_i = \frac{\sum_{j=1}^{n} h_j \sqrt{f_j}}{\sum_{j=1}^{n} \sqrt{f_j}} \tag{2}$$

where h_j & f_j are the relative elevation and flow accumulation values for each cell in the class at a distance i and j: 1 *to* n be the total number of cells in the class. The representative value for any other attribute (say width) can also be calculated similarly.

3.2. System Geometry (Deriving the Hillslope Profile)

The representative hillslope topography is derived for each of the four catchments (Figure 1) as illustrated for the catchment W22 in Figure 3. First, the digital elevation model (DEM) is pre-processed to fill all depressions and sinks. We then derive the flow accumulation, aspect, and stream rasters from this filled DEM. The distance to the river and elevation to river rasters (which indicates the relative horizontal and vertical distance from a cell to the nearest river segment, respectively) are then extracted.

The distance from the nearest river segment and the corresponding relative elevation difference is plotted for all the cells within the catchment of interest (Figure 3b). In Figure 3b, each green dot denotes a 10×10 m^2 cell in the catchment. The representative hillslope catena is then derived based on the methodology explained in Section 3.1. The potential energy conservation along the direction of the flow profile by means of the weighted mean elevation values is validated using four different distance classes (100, 200, 300 & 400 m: Figure A1 in Appendix A). The representative hillslope profile obtained for W22 is shown as a pink overlay in Figure 3b. The catena length is chosen intuitively based on the relative elevation and distance from stream distribution plots (Figure 3b). In essence, the catena length depends on the distributions of distances and elevations to the river raster, as shown in Figure 3. We choose a suitable value for the catena length such that all major classes of elevation and distance to river cells are included (e.g., 530 for catchment W22 as shown in Figure 3). Based on the elevations in each distance class and the chosen catena length, the representative hillslope (pink overlay in 3b) geometry for the target catchment is calculated.

10.1029/2023WR036420

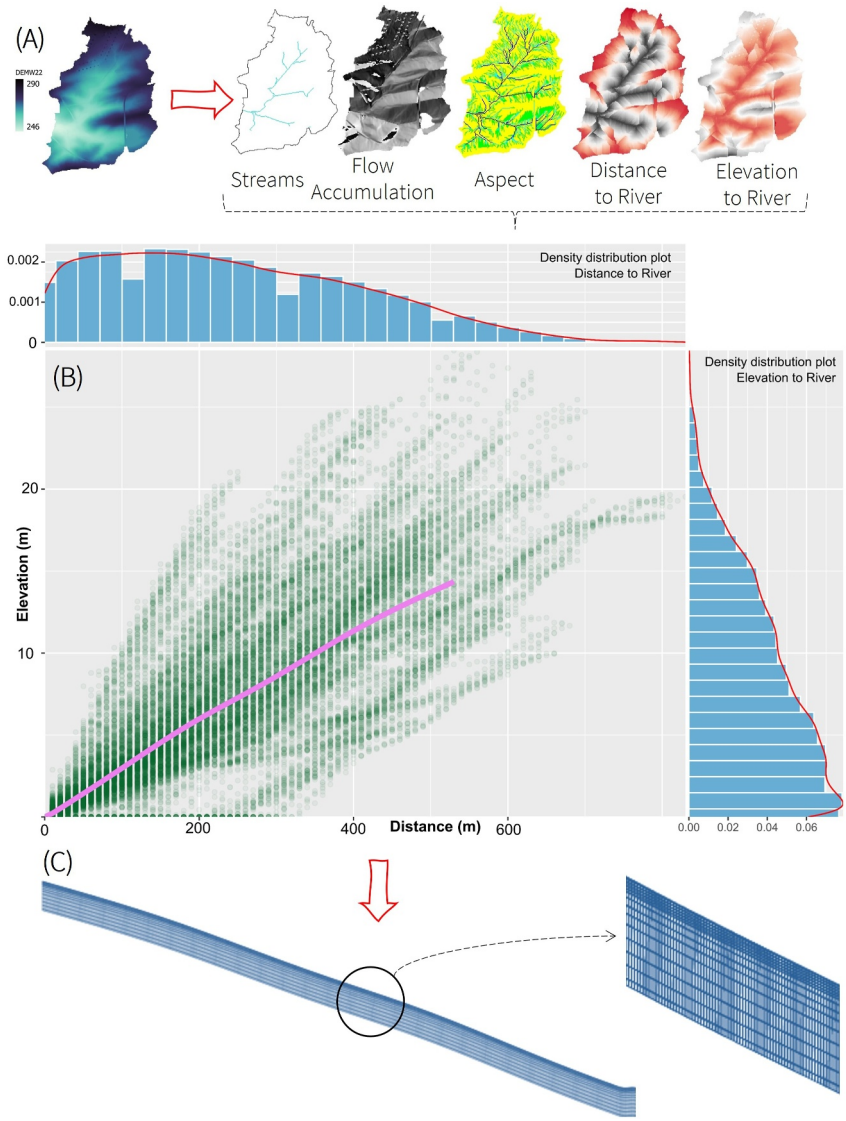

Figure 3. Workflow diagram illustrating the major steps involved in deriving the representative hillslope catena for the catchment W22. Derivation of raster maps (streams, flow accumulation, aspect, distance, elevation to river from the filled digital elevation model (a). Selection and binning of every distance and corresponding elevation to the nearest river segment (b). Calculation of mean distance using flow accumulation weights (see also Appendix A). Final derived representative hillslope (pink overlay line in panel (b)) in panel (c). The panel shows the grid mesh used to model the single representative hillslope. As detailed in Section 3.2, we opt for more finer grid spacing toward the surface as shown in the inlay figure.

The entire catchment is then represented by this single hillslope, the corresponding patterns of soils (Table 1) and landuse in CATFLOW (Figure 3c), when simulating the catchment water balance and flash flood runoff. To this end, the hillslope W22 was discretized into 531, 1 m wide horizontal segments and 15 vertical layers. The total hillslope depth was set to 2 m, based on the transfer of knowledge from Weiherbach (see Section 3.3). The vertical grid resolution varied from 0.05 m near the surface to 0.25 m toward the bottom node (Figure 3c). For ease of numerical simulation, we choose a uniform width (area of catchment/representative hillslope length) for all hillslope elements. Boundary conditions were set to the atmospheric boundary at the top, no flow boundary at the left margin. At the lower boundary a gravitational flow condition was established.

Table 1
Soil Hydraulic Parameters After van Genuchten (1980) and Mualem (1976) Used to Set Up the Soil Properties in the CATFLOW Model

Soil	$k_s\left(\frac{m}{s}\right)$	$\theta_s(\%)$	$\theta_r(\%)$	$\alpha(m^{-1})$	n
Calcaric Regosol on Loess (Zehe et al., 2001)	3.1e−7	0.44	0.06	0.40	2.06
Coluvisol on Loess (Zehe et al., 2001)	1.7e−6	0.40	0.04	1.90	1.25
Rosetta3_LGRB	1.66e−6	0.45	0.08	0.43	1.54
EUPTFv2_LGRB	7.12e−8	0.44	0.006	0.35	1.21
Rosetta3_LUCAS	2.44e−6	0.44	0.08	0.38	1.56
EUPTFv2_LUCAS	1e−7	0.44	0.004	0.43	1.21

Note. The first two parameter sets were obtained from undisturbed soil cores in the Weiherbach/donor catchment. The remaining parameters were derived from two soil pedotransfer functions, Rosetta3 and EUPTFv2, using soil textural data (shown in Table D1—Appendix D).

3.3. Transfer of System Parameters From Weiherbach

Since our present study area in the Elsenz-Schwarzbach consists mainly of agricultural loess catchments, which share similar geological and pedological characteristics and the same major crops with the Weiherbach/donor catchment (Zehe et al., 2001), we attempted to transfer the soil and land use parameters (particularly surface roughness) from the previous field experiments (Figure 2-T). We transfer, furthermore, the finding that a successful simulation of the water balance and stream flow generation could be achieved, assuming that all hillslopes in the area share the same relative soil catena (Zehe et al., 2001), consisting of Calcaric Regosol on Loess (FAO/UNESCO, 1988; *Pararendzina*) in the upper 90% and Coluvisol on Loess (FAO/UNESCO, 1988; *Kolluvium*) in the lower 10% of the hillslope.

Hence, we assigned the same relative catena to the representative hillslopes of the target catchments. In the donor model version, we also transferred the corresponding soil hydraulic functions for these soils after Van-Genuchten (1980) and Mualem (1976), which were determined by Schäfer (1999) and Delbrück (1997) for these soils in the Weiherbach using both field and laboratory experiments (Table 1 and also Table 3 in Zehe et al., 2001).

To investigate the value of this a priori information for predictions in our poorly gauged regions, we alternatively used pedotransfer functions to estimate the hydraulic parameters required for CATFLOW model from soil textural data. We employed two well known pedotransfer function –Rosetta3 (based on the UNSODA database: Zhang & Schaap, 2017) and EUPTFv2 (based primarily on European samples: Szabó et al., 2021). These functions require the relative proportions of sand, silt, and clay in the soil over our study region (Table D1 in Appendix D) as inputs. The corresponding van Genuchten- Mualem parameters are also shown in Table 1 (For e.g. Rosetta3_LGRB refers to the parameter set obtained using Rosetta3 model and LGRB Textural data). Please note that the corresponding model versions used a single soil for the entire representative hillslope.

Estimates of surface roughness after the Manning-Strickler coefficient, K_{st}, for different crop types and maturity stages were obtained from more than 60 rainfall simulation experiments conducted in the Weiherbach to characterize frictional losses for shallow overland flow (Scherer et al., 2012) (Throughout the remainder of this work, we use both Manning's roughness and Strickler values interchangeably to refer to the roughness coefficient (K_{st}) in the Gauckler-Manning-Strickler formula. Interested readers are referred to Hager (2015) for a historical anecdote). Previous studies (Lumbroso & Gaume, 2012) have shown that the traditional estimates of the Manning's coefficient obtained for riverine floods do not adequately represent flash flood conditions. Specifically, due to overbank flow during such extreme events, changes in the associated roughness properties are invariable. Hence, we use an ensemble approach for the surface roughness. Instead of running the model for a single value of Manning's roughness, we run the simulations for the range of roughness values within the informed experiments in the Weiherbach (6–12 $m^{1/3}/s$) and report the mean and spread of the ensemble predictions. This is of particular interest as it sheds light on the sensitivity of flash flood timing, peak and volume to changes in overland flow velocities.

As stated, CATFLOW also includes an advanced evapotranspiration subroutine, which enables time continuous simulations for a model spin up. However, the use of this module requires detailed information about the relative fraction of each crop, which is not available for the summer of 2016, as well as ground based data on radiation,

wind speed, air humidity and temperature, which are neither at hand for our study area nor for most regions in the world. Hence, we decided to circumvent this challenge because the model offers the option to use evaporation data as input as well. We thus ran the model using globally available climate data sets for the model spin up. Specifically, we drove the hillslope models with the climate reanalysis product ERA5 Land (Muñoz-Sabater et al., 2021), using precipitation and evapotranspiration during the event simulation and for model spin up as detailed in the next section.

3.4. Initial and Boundary Conditions

The problem of inferring the initial conditions is a key challenge in event-based modeling (Beauchamp et al., 2013; Zeimetz et al., 2018), because simulation results are highly sensitive to changes in antecedent wetness (Zehe & Blöschl, 2004; Zehe et al., 2005). The challenge is usually not estimating the "actual" soil moisture state but establishing an initial state that is coherent with the land atmosphere interactions and the model physics (Koster et al., 2009). In essence, we seek an initialization identical to the dynamics being captured by our model.

In the present work, we use the ERA5 Land hourly precipitation and evapotranspiration reanalysis data for such a dynamic initialization of our representative hillslope model (Figure 2-I) using a spin up period of 1 year. The model was run using the mean catchment values of forcing data from ERA5 Land until the event of interest (8.06.2016 00:00 UTC); the corresponding soil moisture patterns were saved and then used as initial conditions for a-priori event simulations with a radar based precipitation estimate (temporal resolution of 5 min).

During the event of 08 June 2016 (Appendix B), there were no operational rainfall gauges in the target catchments. The nearest gauge operated by the Baden-Württemberg State Institute for the Environment, Survey and Nature Conservation (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg—LUBW) lay toward the southeast of catchment W22 in Bad Rappenau—Bonfeld (LUBW Station ID—76730: Figure B1 in Appendix B). The gauge recorded a total precipitation sum of around 28 mm on 08 June. The German Weather Services (Deutscher Wetterdienst -DWD) operates a nearby gauge in Waibstadt (DWD Station ID—13674), west of catchment W44. The DWD gauge reported a total daily precipitation of 11 mm. Considering the mismatch between the two gauges (Figure B3 in Appendix B), and the need for a finer spatiotemporal estimate of the convective storm activity, we opt for a radar product (temporal resolution—5 min, spatial resolution −221 m × 221 m) provided by Kachelmannwetter (Kachelmannwetter, 2023) as the forcing boundary condition for the model. The Kachelmann product merges DWD radar data with official and private rain gauges, including clutter correction, and so on. This product is also used by the federal flood forecasting center in the state of Baden-Württemberg. Appendix B depicts radar images of the storm on 08.06.2016 over our study region. Overall, it can be seen that the storm activity is captured quite well by the fine resolution radar product. The direction of the storm explains the smaller magnitude of total precipitation reported by the DWD gauge compared to the LUBW gauge (which seems to be nearer to the storm center: Figures B1 and B3 in Appendix B). Note that the total sum registered by the DWD gauge corresponds well to the accumulated precipitation from ERA5 Land, while the radar product indicates almost 50 mm of rainfall for catchment W22.

4. Results

In the following section, we first show the initialization using ERA5 Land and evaluate the simulated hillslope scale soil moisture compared to globally available soil moisture products (Section 4.1). We then detail the a priori event based flash flood simulations with the donor model versions using radar based precipitation forcing in Section 4.2. In Section 4.3, we present an a posteriori effort to improve the model in two out of the four catchments by including spatially variable precipitation forcings. Lastly (Section 4.4), we explore the sensitivity of the simulated flash floods to varying infiltration capacities due to macropore/surface sealing, soil hydraulic properties inferred from pedo-transfer functions and variations in antecedent soil water content compared to the dynamic initialization.

4.1. Model Initialization With ERA5 Land

Figure 4 shows the top soil (0–5 cm) water content simulated with the representative hillslope that was forced by ERA5 Land precipitation and evapotranspiration for catchment W22. Note that the variations of Manning Strickler, K_{st} ($m^{1/3}/s$) lead to a variation in soil water content during the summer period. To characterize the coherence of these soil moisture simulations with the gridded ERA5 reanalysis product, we calculated the Kling-

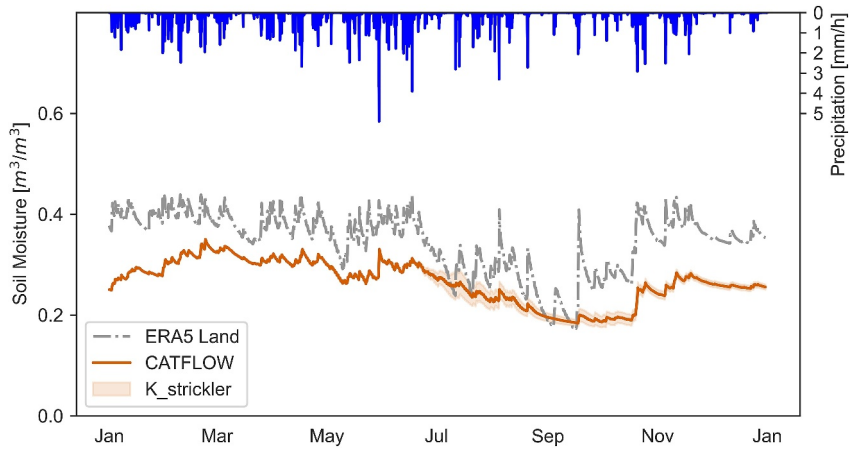

Figure 4. Time series of ERA5 Land surface soil moisture (0–7 cm) averaged over the entire catchment (gray). The precipitation data from ERA5 Land is shown in blue, while the soil moisture simulations with CATFLOW (0–5 cm) are depicted in red, representing the ensemble mean. The shaded regions illustrate the uncertainties (± the standard deviation) corresponding to different values of the Strickler coefficient (Kst = 6–12).

Gupta Efficiency (KGE) (Gupta et al., 2009) between the averaged CATFLOW top layer soil moisture ensemble with the spatially averaged ERA5 Land surface soil moisture (0–7 cm) (Figure 4 and Table 2).

While this revealed high KGE values, CATFLOW simulations were consistently drier than the ERA5 Land reanalysis product and the yearly CATFLOW runs (Figure 4). This mismatch likely reflects the different soil parameterizations, vertical grid depths and scale disparities in the two models. It is important to note that we do not expect a perfect fit between the two modeled soil moisture products, our interest is in capturing the overall local dynamics in soil moisture changes, as explained below.

In order to gain a more comprehensive understanding of the impact of bias on the overall KGE calculations, we calculated the three main components of the modified KGE (Kling et al., 2012). These included the Pearson correlation (r), the bias ratio (beta), and the variability ratio (gamma), as outlined in Table 2. As expected, we obtained high pearson correlation values (around 0.80) for all the different runs (varying Kst values). The high correlation shows that our approach reproduces the yearly dynamics of soil moisture changes in the region well (using the coarse resolution globally available ERA5 Land data as a benchmark). The values of beta and gamma indicate the overall bias and variability of the modeled values compared to the ERA5 Land data.

As we used ERA5 Land forcing variables (precipitation and evapotranspiration) to run the CATFLOW model, and then again ERA5 Land soil moisture states to evaluate the model performance, it remains to be seen whether the correlation is not only inherited from the forcing product. While a comparison with direct soil moisture observations would be preferable, unfortunately, these are not available for the catchments considered in this work, which are essentially ungauged except for the water levels in the reservoir. Due to the absence of in-situ soil observations, we use the following reanalysis and remote sensing soil moisture products (Figure D1 in Appendix D) for comparison: (a) GLEAM (Global Land Evaporation Amsterdam Model: Miralles et al., 2011) (b) ERA5 Land: Muñoz-Sabater et al., 2021) (c) GLDAS (NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019) (d) MERRA (Modern-Era Retrospective analysis for Research and Applications version 2—tavg1_2d_lnd_Nx: Global Modeling And Assimilation Office, 2015), and (e) SMAP (Soil Moisture Active Passive L-Band radiometer Level-3 Product: Chan et al., 2018). The pairwise correlation for the different products for catchment W22 is shown in Figure 5.

We stress here that we don't expect a direct match of the simulated time series in CATFLOW and other products due to varying spatial resolution, vertical grid sizes and hydraulic parameters. However, this is not necessary for initialization purposes. What we are aiming for is a strong temporal rank

Table 2
Goodness of Fit Measures Between the Modeled Soil Moisture Values of Different CATFLOW Runs (Varying Manning Strickler Coefficient K_{st}) With ERA5 Land Surface Soil Moisture for Catchment W22

K_{st} ($m^{1/3}/s$)	KGE	r	Gamma	Beta
6	0.651	0.798	0.812	0.786
7	0.654	0.797	0.821	0.784
8	0.662	0.800	0.836	0.782
9	0.691	0.822	1.050	0.752
10	0.690	0.822	1.051	0.751
11	0.671	0.800	0.861	0.779
12	0.699	0.824	1.019	0.756

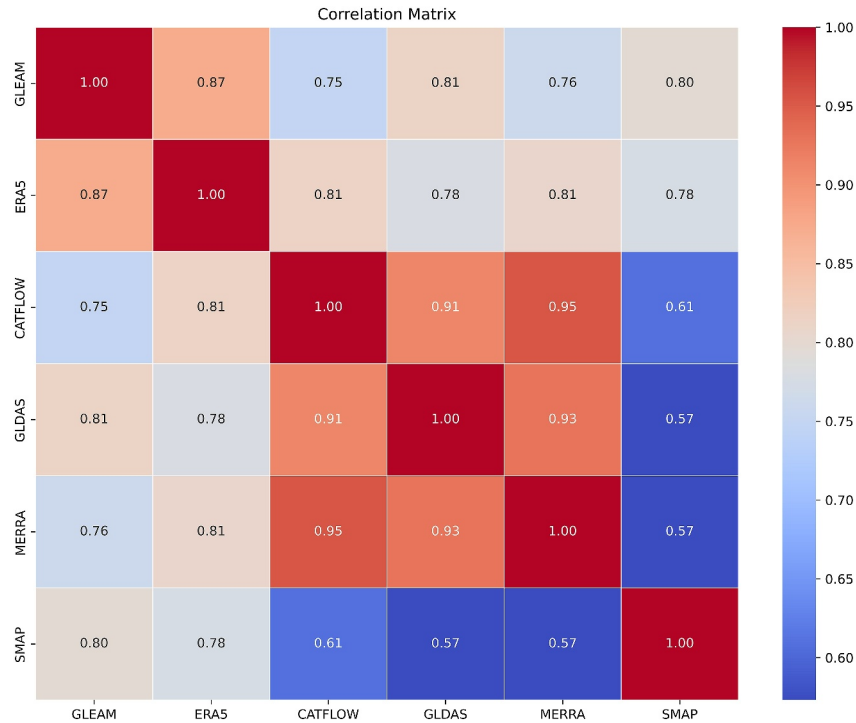

Figure 5. Correlation Matrix Plot illustrating the pairwise correlations between the different soil moisture products.

correlation, which means that the non-exceedance probabilities match well. The high correlation values obtained between CATFLOW and the different products (0.75—GLEAM, 0.81—ERA5, 0.91—GLDAS, 0.95—MERRA, & 0.61—SMAP) point to the feasibility of our approach to robustly estimate initial conditions using such reanalysis products in the absence of in situ observations. The multimodal comparison exercise also indicates that the choice of the reanalysis product (ERA5 Land) is not important. This is a valuable takeaway for setting up the model in a forecast mode. The initialization could be done with a forecast product (compared to reanalysis products, these are available at shorter lead times), and then radar forcing forecasts could be employed to provide event-based predictions.

4.2. A Priori Flash Flood Modeling Using Representative Hillslopes

The "donor" representative hillslope models were then used to simulate the runoff response for the convective storm event on 08.06.2016 in the four catchments in the study area using the dynamical initial conditions obtained from the yearly scale runs using ERA5 Land. This approach of initializing the models using the reanalysis data sets helps in avoiding a random guess of the initial states, which leads to considerable uncertainty, as detailed in Section 4.4. Figure 6 displays the simulated catchment response modeled using a spatially uniform precipitation input—the spatially averaged radar precipitation over each catchment. The simulated runoff in the four catchments is evaluated against the reconstructed inflow hydrograph obtained from the reservoir mass balance (Appendix C), assuming relative measurement error measures for peak flow, volume, and time to peak, as given in Table 3.

The donor model captures the steep ascent of the rising limb of the flood hydrograph, albeit with a time lag, and matches the magnitude of peak discharge values in at least two out of the four catchments (W22 and W39). Note, that the uncertainties in the simulated response due to variable surface roughness are almost identical to the possible observational errors in the gauge level measurements (5%) that propagated into the estimated storm hydrograph.

The peak flow errors (Table 3) in W22 (6%) and W39 (11%) are remarkably small, given the fact that this are uncalibrated predictions and the high uncertainties involved in local flash flood predictions (Bronstert et al., 2018; Marchi et al., 2010). It is also interesting to note that the simulations slightly underestimate the peak flow but overestimate the flow volume for both the catchments. Note that the simulated peak is delayed in W39, while it occurs earlier in W22.

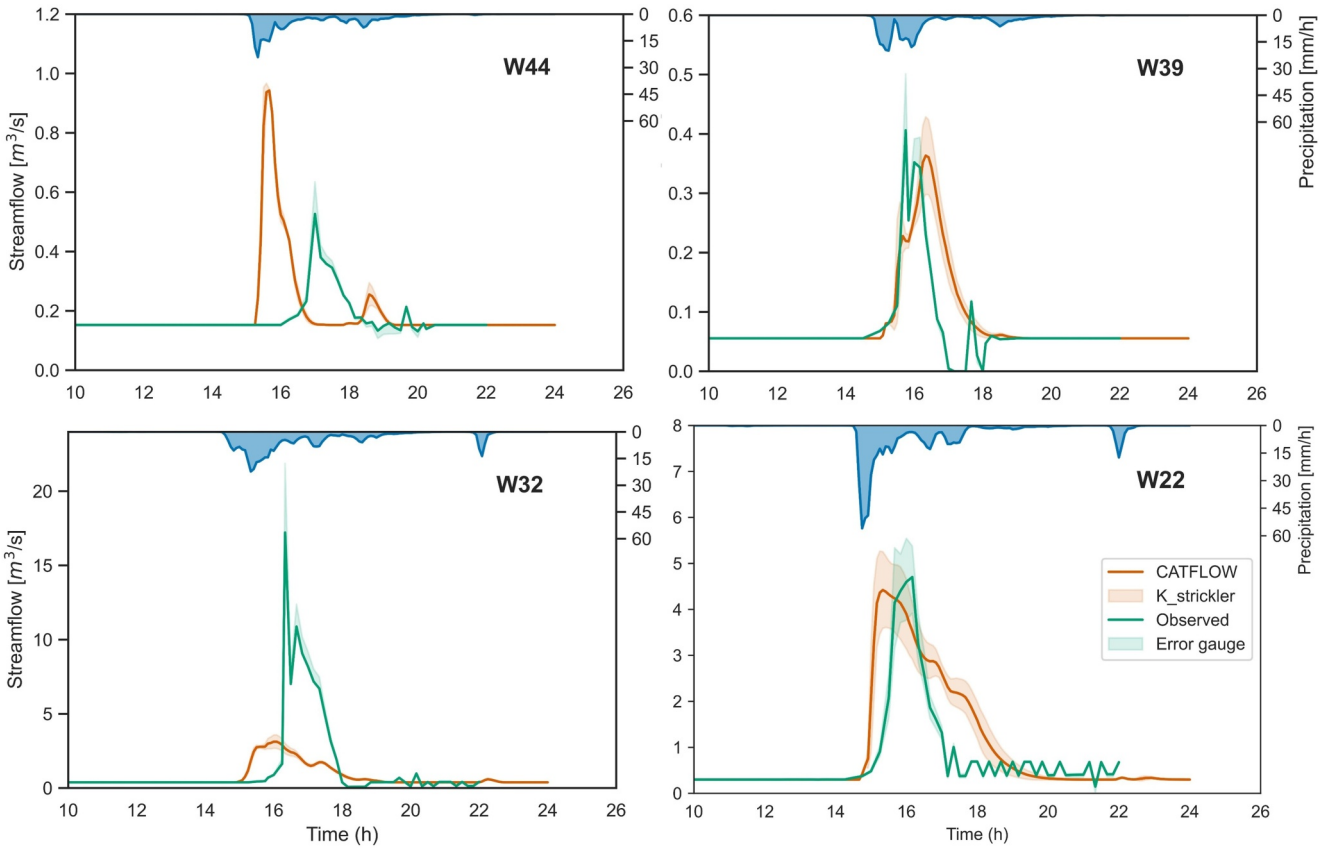

Figure 6. Rainfall—Runoff hydrographs for the flash flood event on 08.06.2016 at the four headwater catchments (W22, W32, W39, and W44). Green curve indicates the reconstructed inflow to the flood defense reservoir (Appendix C), assuming measurement uncertainties of 5%. Red curve indicates the mean values (±SD) of the predicted flood discharge by the CATFLOW model ensemble (varying Strickler coefficient values). All simulation times are in UTC time zone.

In the two remaining catchments, the performance is clearly worse. In catchment W32, the model severely underestimates the storm response, while in W44, it slightly overpredicts the discharge values. To better understand the underlying reasons for this mismatch, we closely examined the storm pattern compared to the relative shape, LULC and orientation of the catchments.

Table 3
Characteristics of Simulated and Reconstructed Storm Hydrographs

	W22 (2.91 km^2)		W32 (5.6 km^2)		W39 (0.73 km^2)		W44 (2.44 km^2)	
Flood characteristics	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
Storm Precipitation (mm)	49	–	35	–	26	–	24	–
Peak Discharge, Q (m^3/s)	4.70	4.42	17.21	3.12	0.41	0.36	0.53	0.94
Time of Peak, t (s)	58,200	55,200	58,800	57,900	56,700	58,800	61,200	56,400
Flood Volume, V (m^3)	45,637	57,978	72,209	49,868	5,189	5,959	13,466	14,537
Flood Volume, V (mm)	15.7	19.9	12.9	8.9	7.1	8.2	5.5	5.9
Runoff Coefficient, R	0.32	0.41	0.37	0.25	0.27	0.31	0.23	0.25
Percentage Error in Peak Discharge, P_Q (%)	–	6	–	82	–	11	–	−79
Error in time to Peak, P_t (s)	–	3,000	–	900	–	−2,100	–	4,800
Percentage Error in Flood Volume, P_V (%)	–	−27	–	31	–	−15	–	−8

Note. The error values are calculated between the mean values of the ensemble CATFLOW predictions and the inverted flood hydrograph for each catchment. Area of each catchment is indicated in brackets.

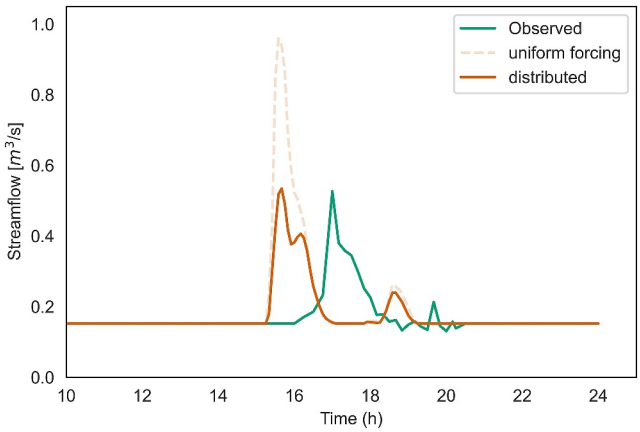

Figure 7. Flood hydrographs for catchment W44. Green curve indicates the reconstructed reservoir inflow, dotted red curve stands for model run using uniform precipitation forcing for the entire representative hillslope, solid red line denotes the model run with distributed forcing. All simulation times are in UTC time zone.

4.3. Role of LULC, Distributed Rainfall Forcing and a Posteriori Model Updating

From Figure 1, we can observe that the catchments W32 and W44 appear to be more elongated and fan-shaped in contrast to the broader shaped catchments W22 and W39. A look at the storm's direction (Figure B2 in Appendix B) suggests that our initial assumption of uniform precipitation across the representative hillslopes might be too simple for these elongated catchments (W32 & W44).

On the other hand, the sharp discharge increase in W32 of around 15 m^3/s within 15 min seems astonishingly high, given the overall precipitation input and the responses in other catchments. One possible explanation could be an obstruction in the flow path perhaps due to debris like wood or sediment from the agricultural upstream areas of W32, which, as indicated by Figure B1 in Appendix B, was closer to the storm center. This might have inadvertently created a temporary retention area, which then burst after a certain point, mimicking the effects of a dam break and resulting in a sudden inflow to the reservoir. Magnitude amplification due to such debris flows and driftwood blockages during flash floods have been reported in regions around the world (Chen et al., 2021; Schalko et al., 2018; Spreitzer et al., 2019).

The total event runoff coefficients calculated for each catchment (Table 3) also shows that while the approach slightly overestimates the response in all the other three catchments (W22, W39, and W44), it clearly underrepresents the runoff response by around 12% in Catchment W32. This apparently stronger runoff production could be explained by the presence of a larger fraction of impervious sealed built-up surface in W32. From Figure 1, it is seen that the small town of Haselbach lies within the catchment area, this contrasts with the other three catchments which are mostly only of agricultural or forest type. Note that this imposes limitations on the parameter transfer from the agricultural rural Weiherbach catchment. Another interesting point is that there is a well defined distribution of agricultural and forested areas along the stream profile in W32: crops at the upstream plateaus and forest along the tributaries or near the outlet. These regions could hence behave like sub catchments having distinct concentration times.

However, it is worthwhile to note that out of all the four catchments the timing of the peak is most accurately captured in W32, which also has relevant implications for flood warning systems.

To investigate whether a distributed forcing input could help in better characterization of response in these elongated catchments, we re-ran the simulations for catchments W32 and W44 using a variable precipitation along the representative hillslopes. Intuitively, we divided the catchment as having two different precipitation forcings over the upslope and downslope regions, to better reflect the storm pattern over the region (Appendix B). While this didn't lead to major changes for catchment W32 (apart from a minor increase in the peak flood), this lead to a clear improvement for W44 (Figure 7). The a posteriori predicted discharge values for catchment W44 match the observation remarkably better, as relative peak errors reduce from around 80% to just 2% (Table 4). The relative volume error decreases from 8% to 2% while the time to peak error remains nearly constant.

4.4. Sensitivity to Preferential Infiltration, Estimated Soil Hydraulic Functions and Initial Conditions

In order to test the sensitivity of the simulated flood discharge to preferential flow processes in the catchment, we employed the simplified macroporosity approach described in Section 2.2 (Equation 1) for Catchment W22 (having a significant amount of agricultural coverage). Essentially, we simulated the increased infiltration due to macropore flow and reduced infiltration due to surface sealing by relative scaling of the bulk hydraulic conductivity along the representative hillslope element. We applied a scaling factor of 10, which is consistent with the results of previous studies (Niehoff et al., 2002). We then ran simulations changing the longitudinal fraction (10%–30%) of the hillslope where surface sealing and a reduction in infiltration occurred. The multiplicative scaling factor for saturated hydraulic conductivity was then applied to the remaining region (90%–70%). As a result, the top parts of the hillslope experience higher macropore flow, while the foot of the hillslope experiences stronger ponding and less infiltration.

Table 4
Goodness of Fit Measures for the Representative Hillslope Modeled Discharges ($K_{st} = 9$) With the Reservoir Stage Inverted Streamflow Measures for Distributed Forcings Over Catchment W44

	W44 (2.44 km^2)		
Flood characteristics	Observed	Uniform forcing	Distributed forcing
Peak Discharge, Q (m^3/s)	0.53	0.96	0.53
Time of Peak, t (s)	64,800	56,100	56,400
Flood Volume, V (m^3)	13,465.9	14,567.4	13,714.54
Runoff Coefficient, R	0.23	0.25	0.25
Percentage Error in Peak Discharge, P_Q (%)	–	−82.2	−1.5
Error in time to Peak, P_t (s)	–	53.1	50
Percentage Error in Flood Volume, P_V (%)	–	−8.2	−1.8

The corresponding sensitivity is shown in Figure 8a. By increasing the extent of sealed surface from 10% to 30% we get a clearly better matching to the observed hydrograph for catchment W22. The storm volume errors are considerably reduced (around 10%) compared to the uncalibrated simulation, which did not consider preferential flow. This is in line with the observation by Niehoff et al. (2002) for the Lein Catchment (also part of the Kraichgau region in South West Germany) of an overestimation of direct surface runoff when preferential flow processes are neglected during convective events. Villinger et al. (2022) tested the same idea when revisiting the flash flood of 1996 in the Weiherbach. Although the significance of such preferential flow processes is widely acknowledged in literature (Bronstert et al., 2023; Zehe et al., 2001), they are not yet incorporated in conventional flood forecasting methods.

The flood simulations for catchment W22 have been illustrated in Figure 8b using the derived parameters from pedotransfer functions (Table 1 & Table D1). The results show that while the model can indeed be set up using such derived parameters, the relative error values for a priori uncalibrated predictions are significantly large. The simulations derived from Rosetta3 drastically underestimate the runoff generation, while those from EUPTFv2 strongly overestimate the storm response. This deviation is in line with the difference in the saturated hydraulic conductivity estimated by the two pedotransfer models (Table 1) and speaks clearly for Hortonian overland flow.

To understand the sensitivity of the flood simulations to variations in antecedent soil water content, we designed a virtual experiment in which the initial soil conditions were taken from a random perturbed sample (30%–90% relative saturation). These experimental runs show a considerable spread both in flood and volume values compared to the ERA5 Land Initialization (Figure 8c). The scatterplot presented in Figure D2 also shows a wide range of errors, highlighting the benefits of utilizing ERA5 Reanalysis for making a priori model predictions in regions with limited data availability.

5. Discussion

In line with our main hypothesis, we demonstrated the feasibility of the physically based CATFLOW model for a priori flash flood predictions in data-scarce, small (<6 km^2) headwater catchments in response to convective storm events. We used the representative hillslope concept and transferred soil hydraulic and roughness parameters using the well observed Weiherbach as donor catchment. Due to the absence of a coherent observational network in the area, we employed initial conditions from a spin up driven by climate reanalysis data. The key to evaluating the event-based simulations was to inversely estimate the storm inflow using water level data and reservoir geometries. This comparison allowed us to quantify the relative simulation error values.

5.1. Toward Short Term Predictability in Ungauged Basins

The "Predictions in Ungauged Basin" problem (IAHS PUB Initiative: Hrachowitz et al., 2013) and the prediction of local flash floods has much in common, as the latter, unfortunately, usually occur within poorly instrumented areas. The PUB initiative attempted to solve the problem of hydrologic predictions in ungauged basins using the concept

10.1029/2023WR036420

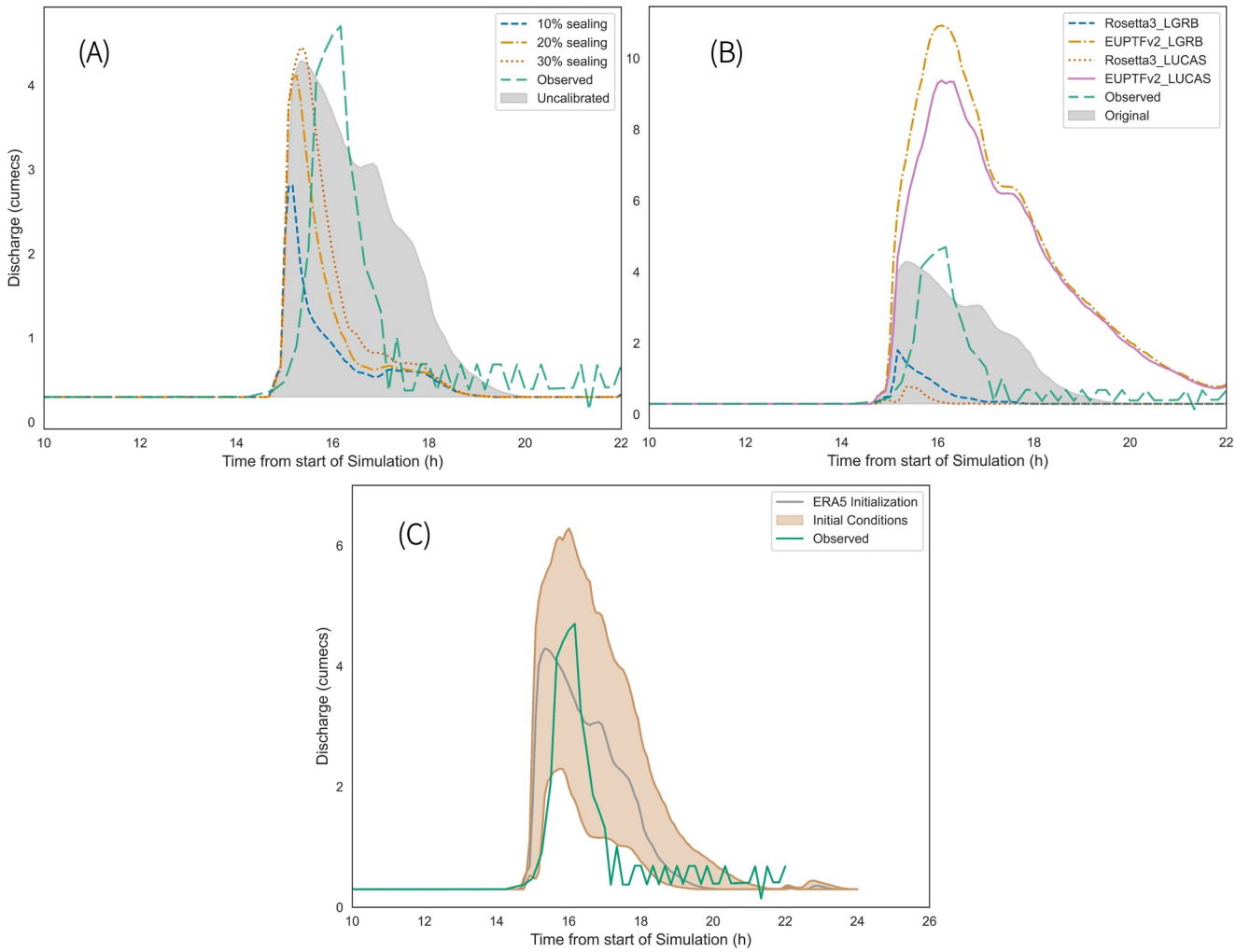

Figure 8. Sensitivity of the simulated flash floods (Catchment W22) to varying infiltration capacities due to macropore/surface sealing (a), soil hydraulic properties inferred from pedo-transfer functions (b) and variations in antecedent soil water content compared to the dynamic initialization (c).

of regionalization, that is, to undertake a transfer of hydrological understanding from gauged to ungauged environments. Spatial proximity is, in this context, one of the most widely used and simple regionalization techniques.

While parameter transfer attempts are not uncommon, using specially trained regionalization functions for conceptual models (Hundecha & Bardossy, 2004), we have shown that this works decently well with parameters that are directly measured without the need for posterior parameter tuning. We were able to demonstrate that parameters and entire model structures of a physically based model can be transferred (and in a sense generalized) from well observed "donor" catchments to largely unobserved regions. This is an important finding, particularly for setting up such process-based models in an uncalibrated setting. And it suggests that experimental catchments provide information that is useful beyond the single catchment case. In consequence, the relevance of hydrological observatories like the Weiherbach (Zehe et al., 2001), the HOAL (Blöschl et al., 2016) or the Attert experimental basin (Pfister et al., 2017) continues undebated.

It is important to have information about behavioral hillslope structures and soil hydraulic parameters before transferring them. This information can be obtained from experimental watersheds within the same hydrological landscape. However, such data may not be available in all areas, so we explored simulations (Figure 8b) using soil pedotransfer functions based on standard soil maps. We compared these simulations to the original donor simulations and found significant errors in the pedotransfer simulations. This is not surprising as pedotransfer functions are based on a wide range of soil samples and may not be suitable for localized predictions without

posterior parameter calibration. However, the fact that the model can indeed be set up using globally available soil texture data along with open source pedotransfer functions, implies that our approach is generally applicable globally. However, a posteriori calibration exercises may likely be needed to ensure that the model predicts Hortonian overland flow response in such new regions. However, this might not necessarily require a large set of events, as previous studies have shown that physically based models can be trained on a few, informative events (Villinger et al., 2022; Zehe et al., 2005).

5.2. Tackling Data Scarcity for Initialization

One way to overcome the challenge of estimating initial conditions for event simulations is by using climate reanalysis products such as ERA5 Land. The importance of such antecedent soil moisture conditions in constraining the flood response cannot be overemphasized (Manoj J et al., 2022, 2023), as has been shown for many catchments across Europe (Berghuijs et al., 2019; Blöschl et al., 2019; Marchi et al., 2010) and also here in Section 4.4. Global climate models have delivered commendable outcomes when it comes to capturing climate and weather extremes on regional scales. However, they remain largely unusable for estimating the impacts of smaller scale hydrological events (IPCC, 2021; Poschlod, 2022).

Our method of utilizing the ERA5 Land as a spin up for the CATFLOW model is beneficial in initialization over data-sparse, small headwater catchments. The strong correlation to the other soil moisture products indicate, that the choice of the product is not necessarily crucial. It is crucial to perform a dynamic spin up because a random variation in initial soil saturation caused a large uncertainty in the a-priori runoff simulations. Using the reanalysis product for deriving the event initial conditions significantly reduces the large uncertainty in the spread of peak flood and volume, as depicted in Figure 8c.

5.3. Managing Observational Uncertainties

Flash floods usually come as (bad) surprises, often impacting regions when and where we least expect them (Borga et al., 2008). Hence, strategies that provide robust warnings are essential. However, since they are also quite rare in nature, there lacks a coherent motivational starting point to invest time and resources into them (Montz & Gruntfest, 2002).

Marchi et al. (2010) analyzed around 25 major flash floods over Europe and showed that proper observational records didn't exist for more than half of the investigated events. During such intense flash floods, direct current meter measurements are often not possible due to safety and technical considerations. Furthermore, these events usually occur in remote ungauged regions with limited accessibility (Borga et al., 2008). It is important to stress here that even in cases with flow measurement gauges, prediction of discharge values during such convective events usually involves lot of uncertainties due to faulty devices, dynamical riverbed changes and floating debris in the stream (Lumbroso & Gaume, 2012).

As is common in such poorly gauged catchments (Bronstert et al., 2018), we didn't have a streamflow gauge to compare our model performance. Hence, we made use of the reservoir geometry and downstream flood retention reservoirs to obtain a crude estimate of the storm characteristics. This strategy should create a win-win situation because local water resource managers are natural end users of such a warning system, and we tremendously increase the sample of historical test cases and complement the small sample that is available from the few gauging stations that observe catchments <10 km^2.

To investigate the plausibility of the reservoir inverted estimates for the four headwater catchments of the Krebsbach, we compare those to streamflow record from the nearest downstream gauging station on the Schwarzbach River (Eschelbronn Schwarzbach, Figure B1 in—Appendix B—LUBW Gauge ID 461). Figure 9 indicates that the streamflow increased quite abruptly from around 2 m^3/s in the evening of 8th June to 27 m^3/s (increase of around 1,250%) at midnight. The abrupt increase occurred within 6 hr, followed by the recession by the end of the next day, pointing out the flashy nature of the flood. It is also interesting to note that the peak flow of 27 m^3/s (Figure 9) matches quite well with the sums of individual flood values (W22–4.7 m^3/s, W32—17.2 m^3/s, W39—0.4 m^3/s, and W44—0.5 m^3/s) obtained for the four headwater catchments by inverting the flood reservoir mass balance. This indicates that the individual flood values obtained are well plausible.

While it is true that the LUBW stream gauge in Eschelbronn drains a much larger area (around 200 km^2), the precipitation values recorded by the DWD precipitation gauge (11 mm: Figure B3) underpin, in line with the

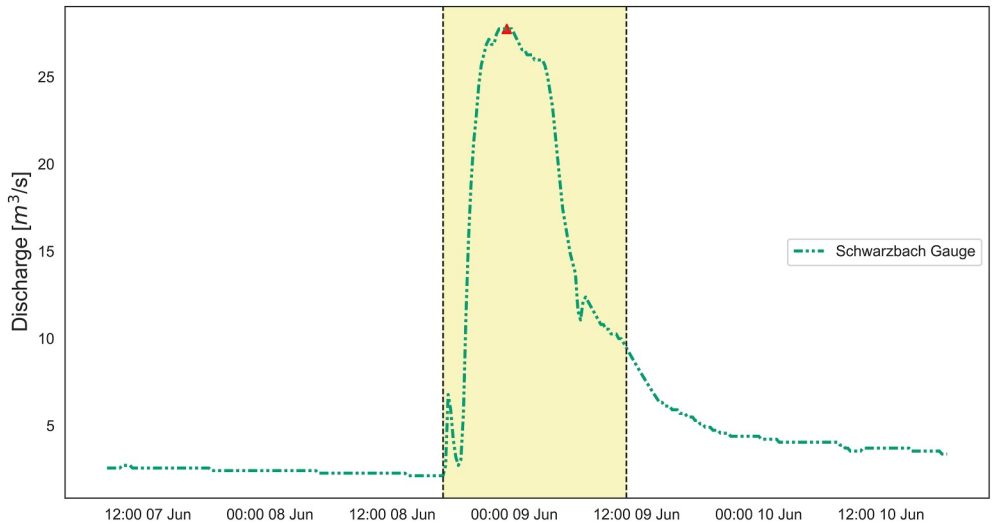

Figure 9. Discharge record at the nearest LUBW gauging station to the Krebsbach (Eschelbronn Schwarzbach –Gauge ID 461) denoting the instantaneous streamflow (temporal resolution—15 min) values during the event of 08–09 June 2016 (All times in MEZ).

reports on the flash flood summer in 2016 (Goeppert, 2018), the that these floods were caused by localized convective events. Moreover, the timing of the peak at the gauging station is also generally in line with the observed timings in the four catchments (after accounting for the time required for the flood wave to reach the gauge along the river network from Krebsbach to Schwarzbach (Figure B1) and possible time conversions between UTC and MEZ).

6. Limitations and Outlook

The perils of applying physically based models at large scales, while the governing equations were developed for much smaller spatial scales, are reviewed in literature (Hrachowitz & Clark, 2017). Such distributed, process-based approaches are also criticized for their complexity and larger data requirement compared to mathematically simpler conceptual models. While conceptual models are without doubt helpful, they are not so intuitive to be combined with a perceptional model of a catchment, as argued by Loritz et al. (2017). Top down (conceptual) models do not perform well in regions and scenarios which deviate from their well-calibrated range of conditions (Fatichi et al., 2016; Hrachowitz & Clark, 2017), and they have challenges in representing Hortonian overland flow.

We believe that, under the threat of a non-stationary climate (Milly et al., 2008) and unprecedented flow regime changes (Pérez Ciria et al., 2019), strategies which involve a convergence of both modeling philosophies are called for. The representative hillslope approach for flash flood modeling is a venture in this regard. And the related use of a gradient resolving, physically based model corroborated well known and provided surprising sensitivities to overland flow. The former is the importance of macropores and infiltration characteristics; the latter is the sensitivity of surface roughness to flood volumes. These findings suggest that the mass and momentum balance of overland flow during flash floods are strongly coupled and thus need coupled modeling, which, however, is not the case in standard flood forecast models.

However, limitations remain that need to be properly understood and accounted for. The 2D effective representative hillslope used to represent the catchments implies the assumption of symmetry where the runoff production is controlled by hillslope parallel and vertical fluxes and their driving gradients (Loritz et al., 2017). The derived effective catena profile depicts our best guess based on the available topographical data (DEM). Furthermore, we used only one transferred (typical) soil catena for modeling. Any uncertainties and errors in the terrain representation will invariably propagate to our model geometry. Another point is the sensitivity to different DEM resolutions, raster filling and flow direction algorithms (Loritz et al., 2019). The hillslope scale again emerges as an interesting sub-unit within a catchment (*virtual laboratories*: Fatichi et al., 2016) for testing the impact of such uncertainties on runoff response.

So far, the direct flood simulations were essentially event based with no separate baseflow component (a constant baseflow was considered from start till end). Moreover, in our case, we do not attempt to fit the model response to water level curve obtained from the reservoir level measurements. Our main aim is to a priori simulate the catchment response during high intensity events in a simple, parsimonious manner. Since the overarching focus of this study was to test if the representative hillslope models and the parameter transfer could work in getting predictions right in rather poorly gauged catchments, we opted for an experimental design in which the approach was tested by spatial sub-sampling rather than temporal. The approach should indeed be extended to investigate other storm events in individual catchments. This provides for an interesting research avenue worth exploring in future works, particularly on the value of dynamic initialization in different years.

We also endeavored to consider the uncertainties in our modeled response (by varying the surface roughness) and the observational benchmark (relative error in gauge measurements). It is indeed true that the choice of process based model implies that we deal with a much larger number of system parameters and boundary conditions, compared to conceptual models. The strength is that these parameters are interpretable and, as shown in this study, transferable.

The forcings and soil moisture simulated by any land surface model (ERA5 Land, in our case) is model-dependent and direct transfer of one model product to another can lead to inconsistencies due to deviations in formulations (Koster et al., 2009). Attempting such a switch of forcing from a coarse gridded reanalysis product (ERA5 Land) to a fine resolution radar precipitation product would usually entail a re-engineering of the model and associated variables. However, we show here that a process based, spatially distributed model can capture the dynamics due to their mechanistic description of the flow system (conserving both the energy and mass balances). Moreover, as we show in Section 4.1, we are more interested in the temporal soil moisture variability rather than the absolute values predicted by the models. Hence, we expect very less model bias due to the choice of the reanalysis product.

One argument frequently put against the use of process-based models in flash flood modeling and forecast strategies is their higher computational times. In the current attempt, we reiterate that by employing a representative approach which spatially averages along the main driving gradient of flow, we can preserve the total flow potential of the catchment without significant computational effort For reference, the spin up phase for the entire year had run time of less than 10 min while the event simulation for each catchment took around 2 min, in a normal Windows PC with 32 GB RAM only.

7. Conclusions

Based on the provided evidence, we conclude that physically based modeling in data-scarce, small headwaters using representative hillslopes, supplemented by climate reanalysis products, appears to be a viable pathway for achieving dependable rainfall-runoff simulations during high intensity storm events. By ensuring that these representative hillslopes align with the principles of energy conservation, we strike a balance between the intricacies required by physically based models and the desired simplicity rooted in parsimony considerations. Integrating with global climate reanalysis products effectively addresses the persistent challenges of data availability, a crucial aspect when modeling extreme events in data-limited regions globally. Our endeavor to model and understand flood dynamics in these specific regions, despite the data limitations, gave crucial insights which presents a step forward in mitigation and preparation for such extreme events. The findings indicate that the modeled hydrograph aligns well with the observed flood curve, derived from reservoir gauge level measurements, in three of the four studied catchments. We also looked at the role played by localized soil surface conditions and preferential flow during convective storms which are quite important for improving our understanding of small scale but high impact flash floods. While the approach demonstrated limitations in one of the region's larger catchments, further exploration and research, as outlined in the subsequent text, could provide more insights into modeling elongated catchments, especially those with urban developments.

Appendix A: Energy Considerations in the Derivation of the Representative Hillslope Catena

From Newtonian mechanics, flow potential at a relative elevation (h) is defined as

$$E = m \times g \times h \tag{A1}$$

where E is the potential energy of the water on the hillslope (J), m is its mass (kg), g represents the gravitational acceleration (m s^{-2}), and h is the relative height of the water above a reference (m).

For each class (say $x = l$ m), the average flow potential due to elevation values is related to the sum of the individual flow potential of all the cells $(j : 1\ to\ n)$ within the class

$$E_{avg}^{x=l} = \frac{E_{total}^{x=l}}{n} = \frac{\sum_{j=1}^{n} E_j^{x=l}}{n} \tag{A2}$$

The flow potential in the representative hillslope element at $x = l$ m is given by:

$$\overline{E}^{x=l} = \overline{m} \times g \times \overline{h} \tag{A3}$$

Where $\overline{h}$ is the estimate of weighted mean elevation for a class at distance l, calculated using Equation 2 in Section 3.1. Table A1 shows these different energies for all the four classes illustrated in Figure A1. On average, the relative errors between flow potential in the classes and in the derived representative catena are seen to decrease as the distance from the stream increases.

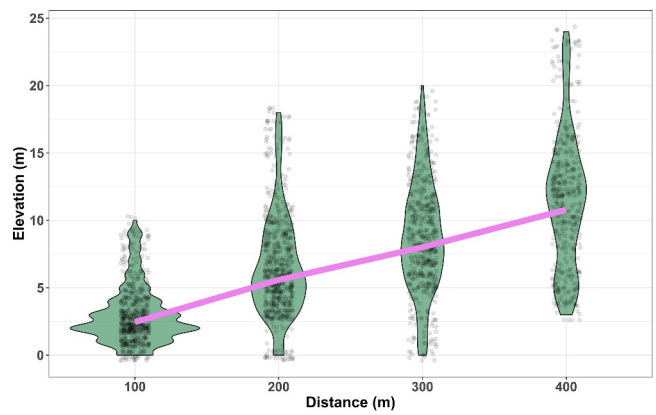

Figure A1. Plots showing the distribution of elevation values of each cell within four distinct distance classes from Figure 3b. The pink line denotes the representative hillslope profile derived from the mean elevation values using the approach detailed in Section 3.1.

Table A1
The Difference Between the Total Flow Potential in Each Class (Figure A1) and in the Derived Representative Hillslope in Terms of Density, ρ and Gravitational Acceleration, g

Classes	l = 100 m	l = 200 m	l = 300 m	l = 400 m
Energy in class, $E_{avg}^{x=l}$	301ρg	660ρg	877ρg	1152ρg
Energy in profile, $\overline{E}^{x=l}$	245ρg	558ρg	801ρg	1079ρg
Relative Error (%)	−22.8	−18.28	−9.4	−6.7

Appendix B: Storm Event on 08.06.2016

See Figure B4.

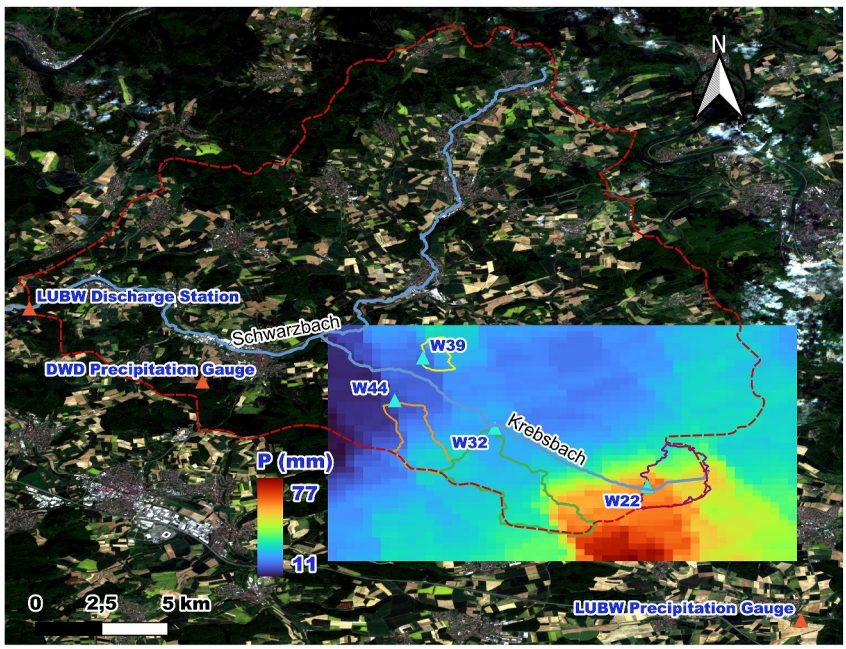

Figure B1. Overview of the Schwarzbach catchment till the downstream streamflow station at Eschelbronn. An overlay layer displays the total accumulated precipitation (in mm) during the event across the four catchments. Additionally, the DWD and LUBW precipitation gauges are shown.

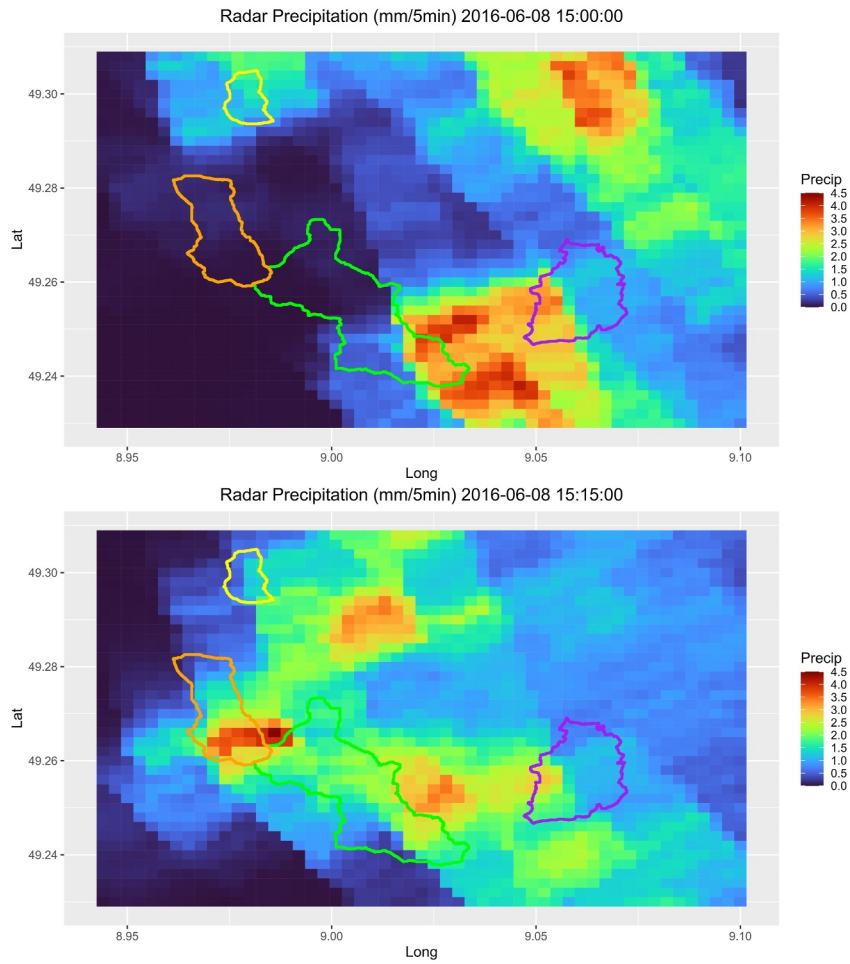

Figure B2. Evolution of the convective storm event on 08.06.2016 over the Krebsbach as captured in the chosen radar based precipitation product (Kachelmannwetter, 2023).

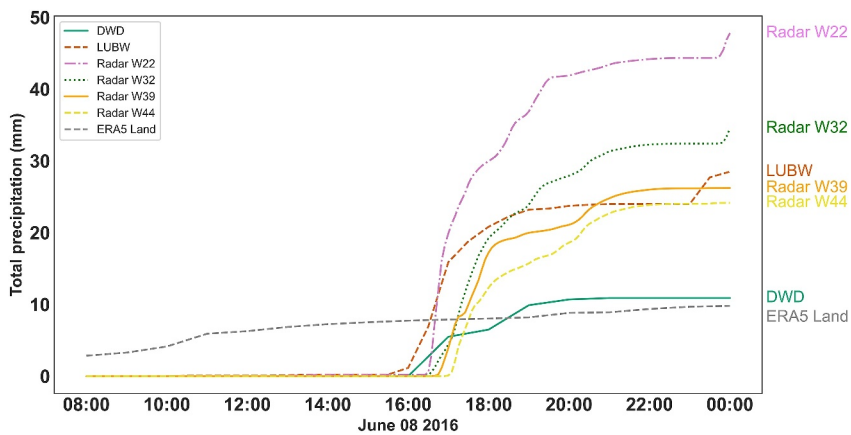

Figure B3. Cumulative timeseries plots showing the hourly dynamics captured by the different precipitation estimates for the 08 June 2016 event.

Figure B4. Impact of flash floods on 08.06.2016 over Catchment W22 (Zweckverband Hochwasserschutz Elsenz-Schwarzbach).

Appendix C: Flood Estimation Using Reservoir Mass Balance

Mass conservation has long been the foundation of hydrological modeling. This basic physical law is usually expressed (for hydrological systems) in the form of:

$$\frac{dS}{dt} = I(t) - O(t) \tag{C1}$$

where the change of a system's mass storage (S) with respect to time (t) is equal to total mass input, $I(t)$ minus total mass output, $O(t)$. This represents one of the most basic physical constraints placed on the functioning of any hydrological system.

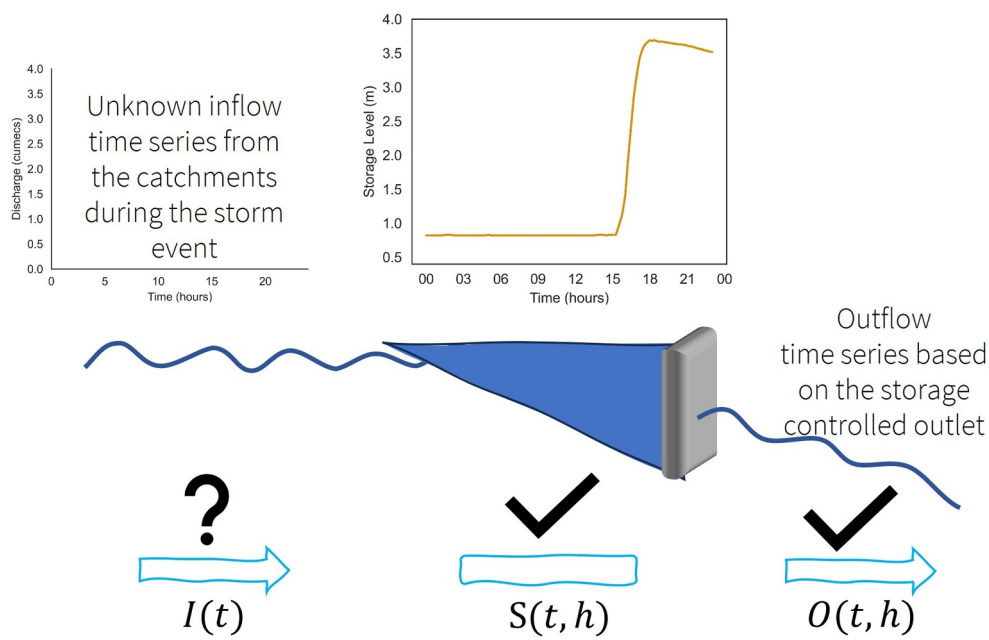

Figure C1. Schematic representation of the reservoir mass balance inversion.

Considering the mass balance of the downstream flood reservoirs in the four catchments (Figures 1 and B1) as shown in Figure C1, the storage in the reservoir at any time t being a function of the level (h). An automatic recorder measures the water level in the reservoir as shown in Figure C1. The outflow being again a function of the water level in the reservoir. Having knowledge of the reservoir geometry relations ($S = f(h)$) and the stage-discharge relationship of the outlet ($O = g(h)$), we now need to estimate the inflow to the reservoir from the catchment due to the convective storm activity. Again, from Equation C1:

$$\frac{S(t+\Delta t, h+\Delta h) - S(t,h)}{\Delta t} = I(t) - O(t) \tag{C2}$$

Hence, the inflow is given by,

$$I(t) = \frac{S(t+\Delta t, h+\Delta h) - S(t,h)}{\Delta t} + O(t) \tag{C3}$$

Now for the uncertainty analysis, we consider a relative error of 5% in the reservoir level measurements and again calculate the inflows using Equation C3. The inflow hydrograph obtained, and calculations are further shown for catchment W22 in Figure C2 and Table C1 respectively.

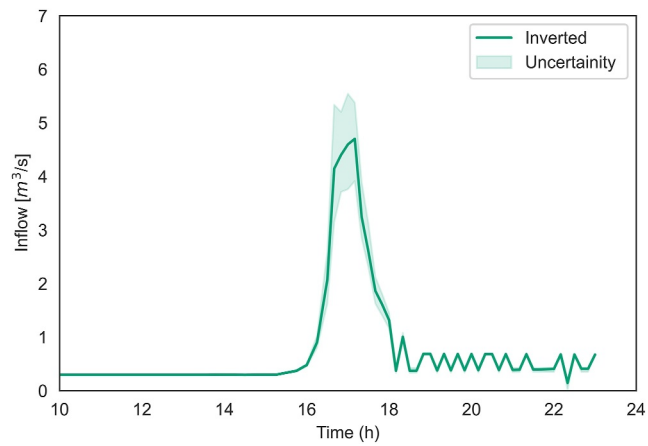

Figure C2. Reconstructed inflow time series for catchment W22. All times in CET.

Table C1
Reservoir Mass Balance Calculations for Catchment W22

S No	Time	Level (m)	Storage (m^3)	Change in storage (m^3)	Outflow (m^3/s)	Dt	Inflow (m^3/s)
63	08-06-2016 15:45	1.12	41.94883	16.42903	0.354373	900	0.372627
64	08-06-2016 16:00	1.4	107.4919	65.54307	0.407547	900	0.480373
65	08-06-2016 16:15	1.98	471.3575	363.8656	0.500051	900	0.904346
66	08-06-2016 16:30	2.48	1,828.668	1,357.311	0.565968	900	2.074091
67	08-06-2016 16:40	2.82	3,952.833	2,124.165	0.605154	600	4.145429
68	08-06-2016 16:50	3.06	6,215.17	2,262.337	0.631057	600	4.401619
69	08-06-2016 17:00	3.26	8,583.66	2,368.49	0.651033	600	4.598516
70	08-06-2016 17:10	3.43	11,005.15	2,421.491	0.666923	600	4.702742
71	08-06-2016 17:20	3.53	12,547.01	1,541.86	0.676144	600	3.245911
72	08-06-2016 17:30	3.6	13,690.94	1,143.932	0.68265	600	2.589203

10.1029/2023WR036420

Appendix D: CATFLOW Soil Moisture Simulations

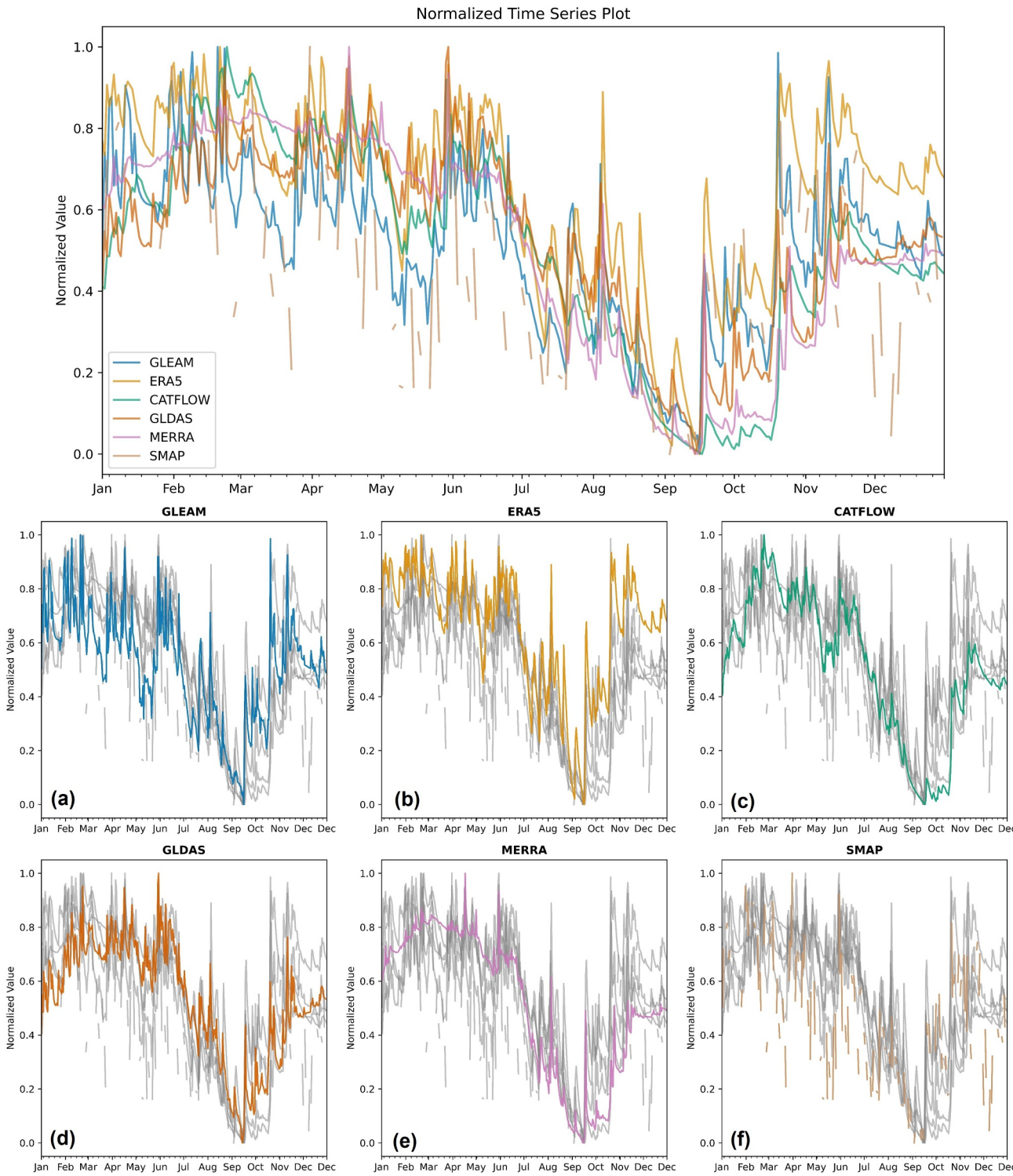

Figure D1. Normalized time series plot for catchment W22 comparing the yearly dynamics of CATFLOW simulated soil moisture with the various soil moisture products. (a) GLEAM (Global Land Evaporation Amsterdam Model: Miralles et al., 2011) (b) ERA5 Land: Muñoz-Sabater et al., 2021) (c) CATFLOW, (d) GLDAS (NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019) (e) MERRA (Modern-Era Retrospective analysis for Research and Applications version 2—tavg1_2d_lnd_Nx: Global Modeling And Assimilation Office, 2015), and (f) SMAP (Soil Moisture Active Passive L-Band radiometer Level-3 Product: Chan et al., 2018).

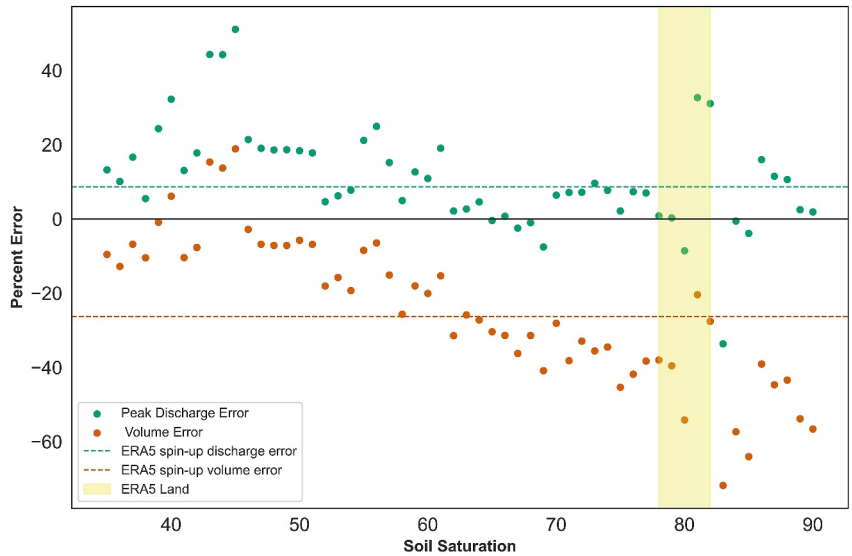

Figure D2. Scatterplot demonstrating how changes in the initial soil saturation values affect the variation in peak discharge and volume errors for catchment W22. The dotted lines represent the errors in the original set-up, where soil moisture conditions were inferred from the ERA5 Land Initialization. The yellow-shaded region indicates the possible absolute value of soil saturation from the ERA5 Land directly.

Table D1
Textural Data Required for Estimating the Hydraulic Functions Using the Soil Pedotransfer Functions

Source	Sand (%)	Silt (%)	Clay (%)
LGRB (State Ministry for Geology, Natural Resources and Mining, Baden Württemberg (Landesamt für Geologie Rohstuff und Bergbau (LGRB), 2024)—Sheet No 6820, Schwaigern	1.7	83.4	14.9
EU LUCAS (European Union Land Use/Cover Area frame Survey (Ballabio et al., 2016), POINT_ID—42462906)	5	82	13

Acknowledgments
This research contributes to the Implementation of an InfraStructure for dAta-BasEd Learning in environmental sciences funded by the German Research Foundation (DFG). We would like to thank the Elsenz-Schwarzbach Water Board (Zweckverband Hochwasserschutz Elsenz-Schwarzbach) for providing the flood reservoir information and other relevant event data. We would like to acknowledge all the stakeholders and original members of the Weiherbach project. The authors would also like to thank the Editor, Associate Editor, and three reviewers for their critical but very insightful comments, which have significantly improved the overall structure and clarity of the manuscript. We gratefully acknowledge the Open Access Publishing Fund of the Karlsruhe Institute of Technology (KIT). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The soil moisture data (SMAP—Chan et al., 2018; MERRA2—Global Modeling And Assimilation Office, 2015; GLDAS—Li et al., 2019; ERA5 Land—Muñoz-Sabater et al., 2021) was accessed and pre-processed using Google Earth Engine (Gorelick et al., 2017). The GLEAM product (Martens et al., 2017) was accessed via the GLEAM homepage (Global Land Evaporation Amsterdam Model, 2024). We acknowledge all the original data providers for making their data sets freely and publicly available. The simulation results are archived in Zenodo (Manoj J, 2024). All other data and codes used in this study are available on request from the corresponding author, Ashish Manoj J (ashish.jaseetha@kit.edu).

References

Ballabio, C., Panagos, P., & Monatanarella, L. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, *261*, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006

Beauchamp, J., Leconte, R., Trudel, M., & Brissette, F. (2013). Estimation of the summer-fall PMP and PMF of a northern watershed under a changed climate. *Water Resources Research*, *49*(6), 3852–3862. https://doi.org/10.1002/wrcr.20336

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The relative importance of different flood-generating mechanisms across Europe. *Water Resources Research*, *55*(6), 4582–4593. https://doi.org/10.1029/2019WR024841

Blöschl, G., Blaschke, A. P., Broer, M., Bucher, C., Carr, G., Chen, X., et al. (2016). The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: A hypothesis-driven observatory. *Hydrology and Earth System Sciences*, *20*(1), 227–255. https://doi.org/10.5194/hess-20-227-2016

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., et al. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111. https://doi.org/10.1038/s41586-019-1495-6

Blöschl, G., Kiss, A., Viglione, A., Barriendos, M., Böhm, O., Brázdil, R., et al. (2020). Current European flood-rich period exceptional compared with past 500 years. *Nature*, *583*(7817), 560–566. https://doi.org/10.1038/s41586-020-2478-3

Blöschl, G., Merz, R., & Reszler, C. (2007). Floods in Austria. In *Extreme hydrological events: New concepts for security* (pp. 81–90). Springer. https://doi.org/10.1007/978-1-4020-5741-0_6

Borga, M., Gaume, E., Creutin, J. D., & Marchi, L. (2008). Surveying flash floods: Gauging the ungauged extremes. *Hydrological Processes*, *22*(18), 3883–3885. https://doi.org/10.1002/hyp.7111

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., et al. (2018). Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany. *Science of the Total Environment*, *630*, 977–991. https://doi.org/10.1016/j.scitotenv.2018.02.241

Bronstert, A., Niehoff, D., & Schiffler, G. R. (2023). Modelling infiltration and infiltration excess: The importance of fast and local processes. *Hydrological Processes*, *37*(4), 1–20. https://doi.org/10.1002/hyp.14875

Bürger, G., Pfister, A., & Bronstert, A. (2019). Temperature-driven rise in extreme sub-hourly rainfall. *Journal of Climate*, *32*(22), 7597–7609. https://doi.org/10.1175/JCLI-D-19-0136.1

Celia, M. A., Bouloutas, E. T., & Zarba, R. L. (1990). A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resources Research*, *26*(7), 483–1496. https://doi.org/10.1029/wr026i007p01483

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., et al. (2018). Development and assessment of the SMAP enhanced passive soil moisture product [Dataset]. *Remote Sensing of Environment*, *204*, 931–941. https://doi.org/10.1016/j.rse.2017.08.025

Chen, J., Liu, W., Zhao, W., Jiang, T., Zhu, Z., & Chen, X. (2021). Magnitude amplification of flash floods caused by large woody in Keze gully in Jiuzhaigou National Park, China. *Geomatics, Natural Hazards and Risk*, *12*(1), 2277–2299. https://doi.org/10.1080/19475705.2021.1961882

Delbrück, M. (1997). *Großflächiges Bromidtracerexperiment zur räumlichen und zeitlichen Variabilität des Wassertransports an einem Lößhang* (Dissertation). Universität Heidelberg.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Dunne, T., & Kirkby, M. J. (1978). Field studies of hillslope flow processes (pp. 227–293).

FAO/UNESCO. (1988). Soil map of the world.

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., et al. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, *537*, 45–60. https://doi.org/10.1016/j.jhydrol.2016.03.026

Francke, T., Güntner, A., Mamede, G., Müller, E. N., & Bronstert, A. (2008). Automated catena-based discretization of landscapes for the derivation of hydrological modelling units. *International Journal of Geographical Information Science*, *22*(2), 111–132. https://doi.org/10.1080/13658810701300873

Global Land Evaporation Amsterdam Model. (2024). GLEAM [WWW document]. Retrieved from https://www.gleam.eu

Global Modeling And Assimilation Office. (2015). MERRA-2 tavg1_2d_lnd_Nx: 2d, 1-hourly, time-averaged, single-level, assimilation, land surface diagnostics V5.12.4. [Dataset]. https://doi.org/10.5067/RKPHT8KC1Y1T

Goppert, H. (2018). Evaluation of expired heavy rain events via radar measurements. *Wasserwirtschaft*, *108*(11), 44–50. https://doi.org/10.1007/s35147-018-0223-8

Göppert, H. (2018). Auswertung von abgelaufenen Starkregenereignissen über Radarmessungen. *Wasserwirtschaft*, *108*(11), 44–50. https://doi.org/10.1007/s35147-018-0223-8

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone [Software]. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Güntner, A., & Bronstert, A. (2004). Representation of landscape variability and lateral redistribution processes for large-scale hydrological modelling in semi-arid areas. *Journal of Hydrology*, *297*(1–4), 136–161. https://doi.org/10.1016/j.jhydrol.2004.04.008

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hager, W. H. (2015). Albert Strickler: His life and work. *Journal of Hydraulic Engineering*, *141*(7), 1–5. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001000

Horton, R. E. (1933). The role of infiltration in the hydrological cycle. *Transactions - American Geophysical Union*, *14*, 446–460.

Hrachowitz, M., & Clark, M. P. (2017). HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences*, *21*(8), 3953–3973. https://doi.org/10.5194/hess-21-3953-2017

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., et al. (2013). A decade of Predictions in Ungauged Basins (PUB)-A review. *Hydrological Sciences Journal*, *58*(6), 1198–1255. https://doi.org/10.1080/02626667.2013.803183

Hundecha, Y., & Bardossy, A. (2004). Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of Hydrology*, *292*(1–4), 281–295. https://doi.org/10.1016/j.jhydrol.2004.01.002

IPCC. (2021). *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen] (Vol. 3949). Cambridge University Press.

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *273*, 593–610.

Kachelmannwetter. (2023). Kachelmannwetter [WWW document]. Retrieved from https://kachelmannwetter.com/de

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424–425*, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. *Journal of Climate*, *22*(16), 4322–4335. https://doi.org/10.1175/2009JCLI2832.1

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., et al. (2019). Global GRACE data assimilation for groundwater and drought monitoring: Advances and challenges [Dataset]. *Water Resources Research*, *55*(9), 7564–7586. https://doi.org/10.1029/2018WR024618

Loritz, R., Bassiouni, M., Hildebrandt, A., Hassler, S. K., & Zehe, E. (2022). Leveraging sap flow data in a catchment-scale hybrid model to improve soil moisture and transpiration estimates. *Hydrology and Earth System Sciences*, *26*(18), 4757–4771. https://doi.org/10.5194/hess-26-4757-2022

Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., & Zehe, E. (2017). Picturing and modeling catchments by representative hillslopes. *Hydrology and Earth System Sciences*, *21*(2), 1225–1249. https://doi.org/10.5194/hess-21-1225-2017

Loritz, R., Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H., & Zehe, E. (2019). A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation. *Hydrology and Earth System Sciences*, *23*(9), 3807–3821. https://doi.org/10.5194/hess-23-3807-2019

Lumbroso, D., & Gaume, E. (2012). Reducing the uncertainty in indirect estimates of extreme flash flood discharges. *Journal of Hydrology*, *414–415*, 16–30. https://doi.org/10.1016/j.jhydrol.2011.08.048

Manoj J, A. (2024). Simulation results of Manoj J et al (2023, in review) [Dataset]. https://doi.org/10.5281/zenodo.10958813

Manoj J, A., Guntu, R. K., & Agarwal, A. (2022). Spatiotemporal dependence of soil moisture and precipitation over India. *Journal of Hydrology*, *610*, 127898. https://doi.org/10.1016/j.jhydrol.2022.127898

Manoj J, A., Pérez Ciria, T., Chiogna, G., Salzmann, N., & Agarwal, A. (2023). Characterising the coincidence of soil moisture – Precipitation extremes as a possible precursor to European floods. *Journal of Hydrology*, *620*, 129445. https://doi.org/10.1016/j.jhydrol.2023.129445

Marchi, L., Borga, M., Preciso, E., & Gaume, E. (2010). Characterisation of selected extreme flash floods in Europe and implications for flood risk management. *Journal of Hydrology*, *394*(1–2), 118–133. https://doi.org/10.1016/j.jhydrol.2010.07.017

Marchi, L., Cavalli, M., Amponsah, W., Borga, M., & Crema, S. (2016). Upper limits of flash flood stream power in Europe. *Geomorphology*, *272*, 68–77. https://doi.org/10.1016/j.geomorph.2015.11.005

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Meyer, J., Neuper, M., Mathias, L., Zehe, E., & Pfister, L. (2022). Atmospheric conditions favouring extreme precipitation and flash floods in temperate regions of Europe. *Hydrology and Earth System Sciences*, *26*(23), 6163–6183. https://doi.org/10.5194/hess-26-6163-2022

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Climate change: Stationarity is dead: Whither water management? *Science*, *319*(5863), 573–574. https://doi.org/10.1126/science.1151915

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Montz, B. E., & Gruntfest, E. (2002). Flash flood mitigation: Recommendations for research and applications. *Environmental Hazards*, *4*(1), 15–22. https://doi.org/10.3763/ehaz.2002.0402

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Mueller, E. N., & Pfister, A. (2011). Increasing occurrence of high-intensity rainstorm events relevant for the generation of soil erosion in a temperate lowland region in Central Europe. *Journal of Hydrology*, *411*(3–4), 266–278. https://doi.org/10.1016/j.jhydrol.2011.10.005

Munich Re. (2016). Munich Re [WWW document]. Retrieved from https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2016/2016-07-12-media-information.html

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications [Dataset]. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Niehoff, D., Fritsch, U., & Bronstert, A. (2002). Land-use impacts on storm-runoff generation: Scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany. *Journal of Hydrology*, *267*(1–2), 80–93. https://doi.org/10.1016/s0022-1694(02)00142-7

Pall, P., Allen, M. R., & Stone, D. A. (2007). Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming. *Climate Dynamics*, *28*(4), 351–363. https://doi.org/10.1007/s00382-006-0180-2

Paniconi, C., & Putti, M. (2015). Physically based modeling in catchment hydrology at 50: Survey and outlook. *Water Resources Research*, *51*(9), 7090–7129. https://doi.org/10.1002/2015WR017780

Pérez, A. J., Abrahão, R., Causapé, J., Cirpka, O. A., & Bürger, C. M. (2011). Simulating the transition of a semi-arid rainfed catchment towards irrigation agriculture. *Journal of Hydrology*, *409*(3–4), 663–681. https://doi.org/10.1016/j.jhydrol.2011.08.061

Pérez Ciria, T., Labat, D., & Chiogna, G. (2019). Detection and interpretation of recent and historical streamflow alterations caused by river damming and hydropower production in the Adige and Inn river basins using continuous, discrete and multiresolution wavelet analysis. *Journal of Hydrology*, *578*, 124021. https://doi.org/10.1016/j.jhydrol.2019.124021

Peters, A., Hohenbrink, T. L., Iden, S. C., & Durner, W. (2021). A simple model to predict hydraulic conductivity in medium to dry soil from the water retention curve. *Water Resources Research*, *57*(5). https://doi.org/10.1029/2020WR029211

Pfister, L., Martinez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. *Hydrological Processes*, *31*(10), 1828–1845. https://doi.org/10.1002/hyp.11134

Piper, D., Kunz, M., Ehmele, F., Mohr, S., Mühr, B., Kron, A., & Daniell, J. (2016). Exceptional sequence of severe thunderstorms and related flash floods in May and June 2016 in Germany - Part 1: Meteorological background. *Natural Hazards and Earth System Sciences*, *16*(12), 2835–2850. https://doi.org/10.5194/nhess-16-2835-2016

Plate, E., & Zehe, E. (2008). Hydrologie und Stoffdynamik kleiner Einzugsgebiete. In *Prozesse und Modelle*. Schweizerbart'sche Verlagsbuchhandlung.

Poschlod, B. (2022). Attributing heavy rainfall event in Berchtesgadener Land to recent climate change – Further rainfall intensification projected for the future. *Weather and Climate Extremes*, *38*, 100492. https://doi.org/10.1016/j.wace.2022.100492

Reggiani, P., Hassanizadeh, S. M., & Sivapalan, M. (1998). A unifying framework for watershed thermodynamics: Balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Advances in Water Resources*, *22*(4), 367–398. https://doi.org/10.1016/s0309-1708(98)00012-8

Remson, I., Hornberger, G. M., & Molz, F. J. (1971). *Numerical methods in subsurface hydrology, with an introduction to the finite element method*. Wiley-Interscience.

Ruiz-Villanueva, V., Borga, M., Zoccatelli, D., Marchi, L., Gaume, E., & Ehret, U. (2012). Extreme flood response to short-duration convective rainfall in South-West Germany. *Hydrology and Earth System Sciences*, *16*(5), 1543–1559. https://doi.org/10.5194/hess-16-1543-2012

Schäfer, D. (1999). *Bodenhydraulischen Funktionen eines Kleineinzugsgebiets –Vergleich und Bewertung unterschiedlicher Verfahren* (Dissertation). Institute of Hydromechanics, University of Karlsruhe.

Schalko, I., Schmocker, L., Weitbrecht, V., & Boes, R. M. (2018). Backwater rise due to large wood accumulations. *Journal of Hydraulic Engineering*, *144*(9), 1–13. https://doi.org/10.1061/(asce)hy.1943-7900.0001501

Scherer, U., Zehe, E., Träbing, K., & Gerlinger, K. (2012). Prediction of soil detachment in agricultural loess catchments: Model development and parameterisation. *Catena*, *90*, 63–75. https://doi.org/10.1016/j.catena.2011.11.003

Schierholz, I., Schäfer, D., & Kolle, O. (2000). The Weiherbach data set: An experimental data set for pesticide model testing on the field scale. *Agricultural Water Management*, *44*(1–3), 43–61. https://doi.org/10.1016/S0378-3774(99)00083-9

Schroers, S., Eiff, O., Kleidon, A., Scherer, U., Wienhöfer, J., & Zehe, E. (2022). Morphological controls on surface runoff: An interpretation of steady-state energy patterns, maximum power states and dissipation regimes within a thermodynamic framework. *Hydrology and Earth System Sciences*, *26*(12), 3125–3150. https://doi.org/10.5194/hess-26-3125-2022

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., et al. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880. https://doi.org/10.1623/hysj.48.6.857.51421

Spreitzer, G., Tunnicliffe, J., & Friedrich, H. (2019). Using Structure from Motion photogrammetry to assess large wood (LW) accumulations in the field. *Geomorphology*, *346*, 106851. https://doi.org/10.1016/j.geomorph.2019.106851

Steinbrich, A., Leistert, H., & Weiler, M. (2016). Model-based quantification of runoff generation processes at high spatial and temporal resolution. *Environmental Earth Sciences*, *75*(21), 1–16. https://doi.org/10.1007/s12665-016-6234-9

Szabó, B., Weynants, M., & Weber, T. K. D. (2021). Updated European hydraulic pedotransfer functions with communicated uncertainties in the predicted variables (euptfv2). *Geoscientific Model Development*, *14*(1), 151–175. https://doi.org/10.5194/gmd-14-151-2021

Tang, Y. K., & Skaggs, R. W. (1977). Experimental evaluation of theoretical solutions for subsurface drainage and irrigation. *Water Resources Research*, *13*(6), 957–965. https://doi.org/10.1029/WR013i006p00957

Troch, P. A., Paniconi, C., & Emiel van Loon, E. (2003). Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response. *Water Resources Research*, *39*(11). https://doi.org/10.1029/2002WR001728

van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Villinger, F., Loritz, R., & Zehe, E. (2022). Physically based simulation of an occurred flash flood using "representative hillslopes" in a rural catchment. *Hydrologie und Wasserbewirtschaftung*, *66*, 286–297. https://doi.org/10.5675/HyWa_2022.6_1

Wienhofer, J., & Zehe, E. (2014). Predicting subsurface stormflow response of a forested hillslope - The role of connected flow paths. *Hydrology and Earth System Sciences*, *18*(1), 121–138. https://doi.org/10.5194/hess-18-121-2014

Zehe, E., Becker, R., Bárdossy, A., & Plate, E. (2005). Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. *Journal of Hydrology*, *315*(1–4), 183–202. https://doi.org/10.1016/j.jhydrol.2005.03.038

Zehe, E., & Blöschl, G. (2004). Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions. *Water Resources Research*, *40*(10). https://doi.org/10.1029/2003WR002869

Zehe, E., Ehret, U., Blume, T., Kleidon, A., Scherer, U., & Westhoff, M. (2013). A thermodynamic approach to link self-organization, preferential flow and rainfall-runoff behaviour. *Hydrology and Earth System Sciences*, *17*(11), 4297–4322. https://doi.org/10.5194/hess-17-4297-2013

Zehe, E., Maurer, T., Ihringer, J., & Plate, E. (2001). Modelling water flow and mass transport in a Loess catchment. *Physics and Chemistry of the Earth, Part B*, *26*(7–8), 487–507. https://doi.org/10.1016/S1464-1909(01)00041-7

Zeimetz, F., Schaefli, B., Artigue, G., García Hernández, J., & Schleiss, A. J. (2018). New approach to identifying critical initial conditions for extreme flood simulations in a semicontinuous simulation framework. *Journal of Hydrologic Engineering*, *23*(8), 1–9. https://doi.org/10.1061/(asce)he.1943-5584.0001652

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, *547*, 39–53. https://doi.org/10.1016/j.jhydrol.2017.01.004