The role of antecedent conditions in translating precipitation events into extreme floods at catchment scale and in a large basin context

Maria Staudinger[1], Martina Kauzlaric[2], Alexandre Mas[3], Guillaume Evin[3], Benoit Hingray[3], and Daniel Viviroli[1]

[1]Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[2]Mobiliar Lab for Natural Risks, University of Bern, Bern, Switzerland
[3]Univ. Grenoble Alpes, INRAE, CNRS, IRD, Grenoble INP, IGE, Grenoble, 38000, France

Correspondence: Maria Staudinger (maria.staudinger@geo.uzh.ch)

Abstract.

In this study, we analyze how precipitation, antecedent conditions, and their spatial patterns and interactions lead to extreme floods in a large catchment. The analysis is based on 10,000 years of continuous simulations from a hydro-meteorological model chain for a large catchment, the Aare river basin, Switzerland. To account for different flood-generating processes, we

5   based our work on simulations with hourly time resolution. The hydro-meteorological model chain consisted of a stochastic weather generator (GWEX), a bucket-type hydrological model (HBV), and a routing system (RS Minerve), providing the hydrological basis for flood protection management in the Aare river basin.

From the long continuous simulations of runoff, snow, soil moisture and dynamic storage, we were able to assess which combinations of antecedent conditions and triggering precipitation lead to extreme floods in the sub-basins of the Aare catchment.

10   We found that only about 18 to 44% (depending on the sub-catchment) of annual maximum precipitation (AMP) and simulated annual maximum flood (AMF) events occurred simultaneously, highlighting the importance of antecedent conditions for the generation of large floods. For most sub-catchments in the 200-500 km^2 range, after return periods greater than 500 years we found only AMF caused by a triggering AMP, which is notably higher than the return periods typically used in design.

Spatial organization within a larger area is complicated. After routing the simulated runoff, we analyzed the important

15   patterns and drivers of extreme flooding at the outlet of the Aare river basin using a random forest. The different return period classes had distinct key predictors and showed specific spatial patterns of antecedent conditions in the sub-catchments leading to different degrees of extreme flooding. While precipitation and soil moisture conditions from almost all sub-catchments were important for more frequent floods, for rarer events only the conditions in specific sub-catchments were important. Snow conditions were important only from specific sub-catchments and for more frequent events.

20 ## 1 Introduction

Floods in general and extreme floods in particular are a threat to infrastructure and human life. With the awareness that it is not feasible to protect everything and everybody from damage caused by floods, the goal today is to minimize damage. This is achieved by a combination of structural and technical management as well as regional development planning (e.g., hazard

maps). In this context, the return period of floods, i.e., the expected time interval between flood events that exceed a specific

25 magnitude (Makkonen, 2006), is a crucial metric. The standard approach for flood frequency analysis is to use observations of floods and fit a theoretical extreme value distribution to estimate return periods for specific flood peaks and ultimately derive design floods for safety assessments. However, there is generally a lack of observations for extreme events such as floods, which leads to uncertain estimations of the associated return periods. The uncertainty inherent in the return periods is directly linked to the measurement accuracy and quantity of observed extreme events. For instance, established rating curves might not

30 be appropriate anymore for very large floods (Westerberg et al., 2011, 2020), which leads to particularly large uncertainties in the estimation of return periods for larger floods, i.e., rarer events. Recently, there has been an emerging body of literature on so-called heavy tails describing the effects of the lack of observations (Merz et al., 2022; Klemeš, 2000; Klemeš, 2000, and references therein). The issue becomes even more exacerbated because of the increasing non-stationarity of observations, which are caused by climate change but also by the impacts of human infrastructure such as water reservoirs on the streamflow

35 and the extremes (e.g. Hingray et al., 2010).

There are some alternative approaches to overcome the lack of flood observations using historical data from archives or paleo floods (Schulte et al., 2019; Castellarin et al., 2012; Merz and Blöschl, 2008). For very rare flood events, approaches that link extreme precipitation to floods are often used (e.g, Naghettini et al., 1996) rather than performing flood frequency analysis based on discharge observations. Not only limited data contribute to the uncertainty of the estimated return periods, but also

40 the choice of a suitable theoretical extreme value distribution as well as the optimization method to find appropriate parameters describing the theoretical extreme value distribution (?Klemeš, 2000).

To overcome the lack of data in flood frequency analysis, hydro-meteorological modelling chain approaches have been proposed and applied in a scientific context as an alternative to design flood approaches (Lamb et al., 2016; Falter et al., 2015; Hundecha and Merz, 2012; Viviroli et al., 2022). In these hydro-meteorological modelling chain approaches, meteorological

45 simulations from a weather generator force a hydrological model which performs continuous simulation (CS) (Beven, 1987) of discharge including floods. These approaches provide a considerably larger pool of realistic weather combinations potentially leading to floods compared to the pool of observed events. This approach thus (1) allows estimating rarer flood events and (2) provides a more robust base for floods of medium to high return periods than possible with observations only. There are several additional advantages when combining CS with weather generators: It is applicable in ungauged catchments (using parame-

50 ters derived from regionalization) and allows to better link flood estimation with physical processes. Also, flood antecedent conditions are provided without the need for explicit assumptions about them (Calver and Lamb, 1995; Pathiraja et al., 2012; Viviroli et al., 2022). Further, this approach treats the processes in a spatially consistent manner and captures the space–time interactions of the relevant processes (Falter et al., 2015). It provides the possibility to extract not only flood peaks but any other feature of the hydrograph, such as for instance flood volume, and allows thus for a bi- or multivariate flood frequency

55 analysis for safety assessment (e.g., Blazkova and Beven, 2004; Brunner et al., 2016). Moreover, this type of modelling chain can be easily extended using discharge simulations as input to flood plains (see e.g. Lamb et al., 2016) or in a further step to a damage model (see e.g., Falter et al., 2015). Finally, this approach also allows testing future scenarios under changed conditions

for instance regarding climate or land use change (see e.g., Köplin et al., 2013, 2014) or regarding different regulation in the catchment.

60   The sources of uncertainty shift when using such a hydro-meteorological modelling chain for flood frequency analysis. While this approach reduces the uncertainties inherent in conventional flood frequency analysis, it adds various model uncertainties. The weather generator is a statistical model that includes numerous parameters estimated with uncertainty, and which provides a simplified representation of weather dynamics both temporally and spatially. The weather generator is used in a specific temporal resolution and parameterization and makes use of a given underlying meteorological station density for a specific

65   catchment. These components introduce uncertainty in the representativity of the rainfall distribution types generated that can lead to floods in a specific region. Also the hydrological model and the routing system are subject to uncertainty. Here, the main sources of uncertainty stem from model structure and parameter uncertainty. These uncertainties could be estimated by using an ensemble of simulations using different parameter sets and this way helping in the decision-making process for flood safety management (Todini, 2004; Blazkova and Beven, 2004; Wood and Lettenmaier, 2008).

70   Floods can be generated by different processes, and a specific amount of rainfall might trigger a flood in some cases. In other cases, the same amount might barely increase the discharge. The outcome depends on the intensity, duration, and spatial distribution and the localization of the precipitation event as well as on the antecedent conditions within the catchment.

The antecedent conditions of the catchment are shaped by its history of drying and wetting over time. These dynamics depend on catchment properties that allow for a large or only a small storage capacity, and on spatio-temporal interactions in

75   the catchment (activated or not, see Tarasova et al. (2019) and references therein) such as how the stream network is connected during a precipitation event and other aspects of functional connectivity (Blume and van Meerveld, 2015) within a catchment. Knowledge about the relationship between antecedent hydrological catchment conditions and meteorological conditions during the event can help better estimate and manage floods (Nied et al., 2017; Brunner et al., 2021). Many studies showed the importance of the antecedent moisture conditions for flood generation in catchments of various scales, revealing a notable

80   influence for the streamflow response of a catchment to a preceding extreme rainfall event (e.g. Michele and Salvadori, 2002; Berthet et al., 2009; Brocca et al., 2008; Bennett et al., 2018). For the large catchment scale, Nied et al. (2013) studied the role of soil moisture conditions prior to flood events and found it to be a major driver for flood generation on the large catchment scale (Elbe catchment, Germany). When linking these relationships between antecedent conditions, triggering precipitation and catchment response to flood frequency, more accurate estimations can be made, such as those for inundated areas (Sikorska

85   et al., 2015; Brunner et al., 2017).

Many flood generation processes can be captured by observations of daily streamflow, and the link between antecedent conditions and generated flood was studied previously at large catchment scale (Nied et al., 2013; Falter et al., 2015). However, some flood-generating processes, including those leading to the most devastating floods, occur on a very short temporal scale and require hourly or even finer data resolution to capture the potentially critical space-time dynamics within a catchment. The

90   final flood at a specific site depends on the antecedent conditions, the triggering precipitation but also on the spatial interplay of processes occurring at different scales. Lakes and flood plains might buffer the flood peaks, while floods coinciding from tributaries of different sub-catchments could increase the overall flood peak due to superposition.

Objectives

In this study we assess the role of antecedent conditions for floods of different return periods including extreme floods. The
95 floods and associated antecedent conditions are simulations from a hydro-meteorological modelling chain consisting of a
stochastic weather generator optimized for the large catchment scale, a hydrological model and a routing system. We specif-
ically assess 1) the link between precipitation, antecedent conditions and return periods for the sub-catchments of the Aare
river catchment and 2) their temporal and spatial interaction, accounting also for retention and confluences leading to extreme
floods at the large catchment outlet.

100 **2 Methods and data**

2.1 Catchments

We studied the large-scale Aare river basin in Switzerland. It was divided into 127 sub-catchments, of which we selected 20
(Table 1) for a more detailed analysis of antecedent conditions, triggering precipitation and flood return periods. These selected
sub-catchments are larger than 200 km^2 and vary in elevation range, slope, aspect, and hence also in hydrological regime. For
105 their geographic setting in Switzerland see the map of the Aare river basin and its sub-catchments (Figure 1).
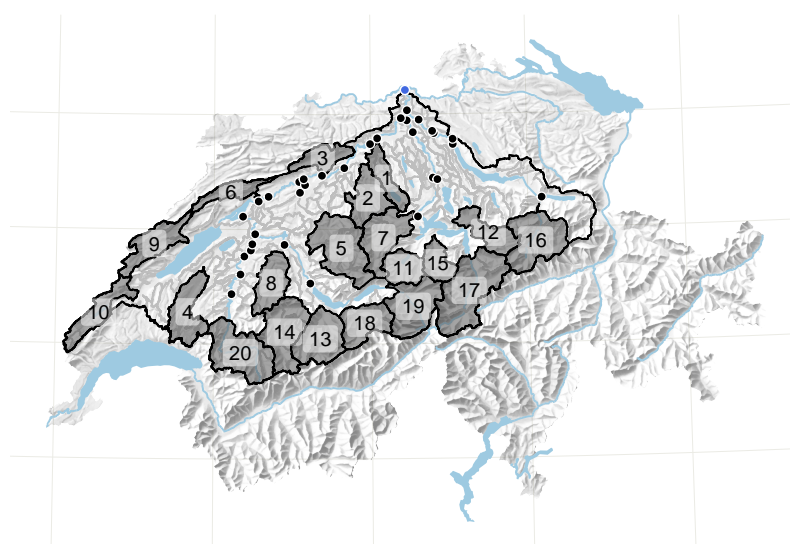

Figure 1. Location of the selected sub-catchments in the Aare river basin, Switzerland. The number labels are indices that can be found in
Table 1. The black points are the routing system nodes that are considered in this study, and the blue point is the outlet of the Aare river
basin. The source of the underlying relief map is the Swiss Federal Office of Topography.

Table 1. Selection of sub-catchments of the Aare river basin with a size larger than 200 km^2. Average catchment response time (ACRT) in hours was calculated from the maximum cross-correlation between precipitation and discharge.

River	Site	No. in Fig. 1	Area	Glacier [%]	Karst [%]	Lakes [%]	Regime	ACRT [h]
Suhre	Suhr	1	243	0.00	0.02	5.91	pluvial inférieur	8
Wigger	Zofingen	2	366	0.00	0.00	0.23	pluvial inférieur	10
Dünnern	Olten, Hammermühle	3	234	0.00	30.02	0.00	pluvial jurassien	9
Broye	Payerne, Caserne d'aviation	4	416	0.00	0.49	0.00	pluvial inférieur	12
Emme	Burgdorf, Lochbach	5	661	0.00	2.14	0.00	pluvial supérieur	9
Suze	Biel-Bienne, Hauserwehr	6	209	0.00	69.00	0.00	nivo-pluvial jurassien	9
Kleine Emme	Emmen	7	478	0.00	7.04	0.00	nivo-pluvial préalpin	13
Sense	Thörishaus, Sensematt	8	351	0.00	8.07	0.13	nivo-pluvial préalpin	9
Areuse	Boudry	9	378	0.00	77.60	0.14	nivo-pluvial jurassien	17
L'Orbe	Orbe, Le Chalet	10	343	0.00	72.03	3.05	nivo-pluvial jurassien	14
Sarner Aa	Sarnen	11	269	0.00	38.72	3.79	nival de transition	49
Muota	Ingenbohl	12	317	0.00	35.63	0.95	nival de transition	10
Kander	Hondrich	13	491	6.69	27.74	0.26	nivo-glaciaire	8
Simme	Latterbach	14	563	1.75	22.94	0.00	nival alpin	14
Engelberger Aa	Buochs, Flugplatz	15	228	3.45	39.17	0.13	b-glacio-nival	10
Linth	Mollis, Linthbrücke	16	600	3.61	28.09	0.81	nivo-glaciaire	11
Reuss	Seedorf	17	833	8.54	8.93	0.23	b-glacio-nival	11
Lütschine	Gsteig	18	381	18.43	26.32	0.00	a-glacio-nival	6
Aare	Brienzwiler	19	555	20.54	12.50	1.08	a-glacio-nival	7
Sarine	Broc, Château d'en bas	20	636	0.49	21.24	0.40	nival de transition	12

2.2 Hydro-meteorological modelling chain

We used the CS approach paired with a stochastic weather generator producing very long time series (here 10,000 years), which expands the pool of possible flood events and encompasses more extreme events than observations alone. This way, we could analyze many possible but unobserved meteorological conditions causing a wide range of antecedent conditions. This enabled us to study the effect of antecedent conditions on the generation of extreme floods. To account for many flood-generating configurations relevant in our study catchments, we based our work on simulations at hourly resolution. This also allows for a more comprehensive pool of flood events than using daily data. Moreover, the hourly time step enables a realistic simulation and examination of the interplay of flood peaks coming from different parts of the large river catchment. Nevertheless, some runoff generation processes leading to floods that happen on a finer temporal resolution are not included and some smaller scale flood processes are not covered with the structure of the hydrological model.

Each sub-catchment may then be treated as spatial sub-unit of the whole system. The sub-catchments were selected using an appropriate discretisation level regarding the study goals, i.e., to identify large floods for the entire Aare river basin. An even finer discretisation might not lead to more insights regarding this goal, because regionally confined floods were hardly ever observed to contribute to notable floods in the Aare river basin.

2.2.1 Weather generator

This study exploits 10,000 year simulations of mean areal precipitation (MAP) and temperature (MAT) for each sub-catchment of the the Aare river basin described in Viviroli et al. (2022). These long synthetic time series were generated by the stochastic weather generator GWEX (Evin et al., 2018, 2019), which reproduces the statistical behaviour of weather events at different temporal and spatial resolutions, focusing on extremes. GWEX is a multi-site, two-part stochastic weather generator, relying on the structure proposed by Wilks and Wilby (1999) for precipitation. Observations of precipitation from 300 stations and observations of temperature from 77 stations were used to fit the weather generator. First a data base of hourly "pseudo-observed" precipitation and temperature was built. For the stations and time period which only had data recorded on a daily basis, the hourly values for precipitation and temperature were obtained by disaggregation. Then GWEX was fitted and simulated at the daily time step. Finally, the daily simulations were disaggregated using the temporal structure of the observed precipitation, either from the same day if data were available from a nearby station, or from a similar day within the period for which continuous hourly observations were available (see details in Viviroli et al., 2022).

The weather generator was optimized for the entire Rhine river basin, rather than specifically for the Aare river basin or its individual sub-catchments. It very well reproduced the cumulative distribution functions of at-site precipitation for different temporal aggregation levels, and of all persistence properties of precipitation including the cumulative distribution functions of wet and dry spell lengths (not shown). Due to the hourly basis of the disaggregated precipitation and temperature observations from meteorological stations, very high-intensity but strongly localized events that occur at sub-hourly resolution are not reliably simulated.

However, the advantage of a regionally applied weather generator is the possibility to analyze the hydrological effects emerging at the regional scale, for instance, for a large hydrological catchment such as the Aare river basin.

2.2.2 Hydrological model

We modelled all sub-catchments within the Aare river basin using the bucket-type hydrological model HBV (Bergström et al., 1995; Seibert and Vis, 2012) and simulated discharge, soil moisture storage, upper and lower groundwater storage as well as snow accumulation and melt. The model was calibrated to observed discharge using a performance metric based on the Kling-Gupta efficiency (Gupta et al., 2009), with more weight on the bias in the upper quantiles (50-80%) than in the classic metric, i.e. 0.25 weight to correlation between simulated and observed discharge, 0.25 weight to variability, 0.25 to bias of the full discharge range and 0.25 weight to the upper quantiles of discharge. This was done to focus more on flood events and at the same time not to give too much weight to the uncertain peaks.

For the simulations, we forced the hydrological model with MAT and MAP from the weather generator. The generated daily temperature lapse rate was used to allocate the temperature conditions to the elevations of each catchment. To consider the precipitation lapse rate, a constant adjustment factor of 5%/100m was applied.

The HBV model comprises different model routines to simulate snow accumulation and melt with a degree-day approach, evapotranspiration and soil moisture storage, the drainage from the groundwater storage (response routine) and routing to finally simulate streamflow at the outlet of the catchment. In the snow routine, the forcing precipitation is adjusted for gauge under-catch and other errors in the gauging if classified as snowfall. The extent of this correction can vary considerably depending on the hypsography of a catchment. In this study the response routine was chosen in the configuration of a non-linear drainage equation (Lindström et al., 1997), assuming an exponential increase in groundwater response with increasing water stored in the groundwater bucket. The model had ultimately 16 parameters, of which 13 were used in the calibration and 3 were fixed to values that were used in previous studies (Viviroli et al., 2022). For glaciated catchments, the model has five additional parameters of which three are set to default values. An overview of the model parameters and limits for calibration can be found in Table 2.

For each catchment we derived 100 plausible parameter sets with a genetic algorithm calibration procedure (Seibert, 2001). For this study, only one representative parameter set was selected. It is representative in the sense that it represents the median floods from the ensemble (100 parameter sets) of exceedance curves (relationships between annual maximum flood and return periods) using a percentile approach and choosing the median as proposed by Sikorska-Senoner et al. (2020). For considering the parameter uncertainty we could use all the ensemble or some members representing the range of possible parameter sets, but in this study we focused on the median representative parameter set and the antecedent conditions created using this.

2.2.3 Routing system

The simulated discharge from the HBV model was then combined and routed using the RS Minerve hydrological routing system (García Hernández et al., 2020). This system is fast to run and is well suited for application in topographically and hydraulically complex regions (regulated lakes, hydropower) such as Switzerland (Horton et al., 2022). The main impacts of

Table 2. Parameters of the hydrological bucket type model HBV.

Routine	Parameter	Lower limit	Upper limit	Fixed	Description
Glacier	KGmin	0.0001	0.2		minimum outflow coefficient
Glacier	CFGlacier	1	2		correction factor glacier
Glacier	KSI			5E-05	snow to ice conversion factor [1/h]
Glacier	RangeKG			0	max. minus min. outflow coefficient [1/h]
Glacier	CFSlope			1	correction factor slope [-]
Snow	TT	-2.5	2.5		threshold temperature [° C]
Snow	CFMAX	0.001	5		degree day factor [mm/h ° C]
Snow	SFCF	0.4	1.6		snow correction factor[-]
Snow	CFR			0.05	refreezing coefficient
Snow	CWH			0.1	snow water holding capacity
Soil	FC	50	100		maximum storage in soil box [mm]
Soil	LP	0.3	1		threshold reduction ETP [-]
Soil	BETA	1	5		shape coefficient [-]
Soil	PERC	0	1		max. flow from upper to lower gw box [-]
Response	Alpha	0	1		shape coefficient [-]
Response	K1	0.0001	0.1		recession coefficient (upper gw bucket) [1/h]
Response	K2	0.00001	0.05		recession coefficient (lower gw bucket) [1/h]
Routing	MAXBAS	1	100		factor triangular weighting [h]
Precipitation redistribution	PCALT			5	lapse rate precipitation [%/100m]

bank overflow and floodplain retention were considered for a wide range of peak flows by adding parallel channels at relevant sites, both in series and in parallel. These channels account for estimated channel flow capacity and inundated areas. Levee breaks were not considered. Stage-area-volume relationships were extracted from digital terrain information (Swisstopo, 2005) for the nine larger lakes in the Aare river basin. Six of these lakes are regulated, and the regulation rules are usually expressed as

175   stage-discharge relationships with seasonal, monthly or even daily variations. These rules have been digitized and implemented in RS Minerve, with simplifications made where necessary. Where available and feasible, the rules were adapted for flood events (i.e., deviating from normal operation). The output nodes have been placed at sites that correspond to selected gauging sites of the Swiss Federal Office for the Environment. For the 10,000 years simulation we assumed no changes in the current regulation and general stationarity of the system.

180   **2.3   Event selection**

For each of the selected sub-catchments the annual maximum flood (maximum hourly discharge), AMF, and the annual maximum precipitation sum over a fixed time window, AMP, were extracted from the full simulation period of 10,000 years. A fixed
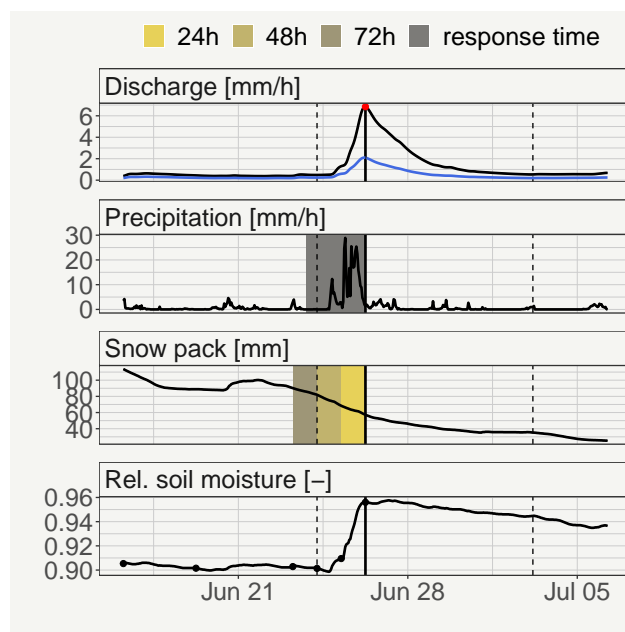
Figure 2. Sketch of the extracted antecedent conditions and the triggering precipitation for one exemplary flood event. The reference is the peak of the annual mean flood (AMF), the dashed lines indicate the flood event as found by the recursive filter (Eckhardt, 2005). The grey area in the precipitation panel shows the average catchment response time (ACRT) for this example catchment. The triggering precipitation is the sum over a catchment-specific time window considering the ACRT before the beginning of the flood event. The blue line in the hydrograph (upper panel) indicates the contribution of snowmelt to the discharge. Relative soil moisture (lower panel) is extracted at different points in time (24, 48, and 72 hours before the AMF) as well as at the beginning of the flood event.

window for the precipitation sums is common in meteorological studies for extreme value statistics. However, we adjusted the fixed window size from the commonly used meteorological windows to a catchment-specific hydrologically more meaningful

185 window, namely the average response time of the catchment. The average catchment response time (ACRT) was estimated by calculating the maximum cross-correlation between precipitation and discharge making a seasonal distinction, because of possible delay due to snow accumulation and melt in winter and spring (see Keller et al., 2018; Tarasova et al., 2019). That means that some catchments may have a fixed window of 12 hours to find the AMP, other 24 hours etc. The flood event belonging to the AMF peak is estimated using the recursive Eckhardt filter (Eckhardt, 2005) defining the start and end of the flood event

190 where baseflow and discharge converge. Therefore, beside the flood peak, also flood volume was estimated. The filter works well for most flood events in our catchment selection, but it tends to underestimate the flood volume during double-peak events.

Flood-triggering precipitation was defined as the precipitation that fell between ACRT before the start of the flood event and the flood peak. Hence, all precipitation that fell before that start point in time (ACRT before the start of the AMF event) was assumed to alter the antecedent conditions but to not directly trigger the flood event. The ACRTs for each catchment

195 are listed in Table 3. Snow has a twofold functioning as both a temporary storage of precipitation and delayed precipitation

9

input in the form of snowmelt. Like rain, snow influences the antecedent conditions and the melt water might contribute to triggering a flood event. We excluded annual maximum precipitation events where precipitation was likely snow and not rain, considering the substantial time gap between snow accumulation and snow melt. Rain-on-snow events are thus not covered with this approach and would also not be adequately simulated by the hydrological model used.

200   Starting from the AMF peak, we extracted a number of characteristics including simulated soil moisture, dynamic catchment storage (comprising soil water and groundwater storage (see Staudinger et al., 2019)), snow pack at time points before the AMF and snow melt contributing to the flood, as well as the associated triggering precipitation. These characteristics collectively describe the conditions during and preceding the flood. The same characteristics were also extracted before the AMP but here instead of the preceding precipitation, the discharge response to each AMP was extracted. All considered characteristics
205   regarding flood events, antecedent conditions and triggering precipitation as well as antecedent conditions and reaction to the AMPs are listed in Table 3.

The soil moisture conditions of the catchment were computed as relative soil moisture filling. For this the simulated soil moisture [mm] at any time was compared to the absolute maximum soil moisture that was simulated during the full simulation period [mm] in each sub-catchment. Snow conditions in the catchment were included in terms of snow pack before a flood
210   event and in terms of relative snow melt water contribution to discharge during the flood event. The latter was computed as fraction of simulated snow melt in simulated discharge during the flood event.

Table 3. Characteristics extracted for the annual maximum flood (AMF) and the annual maximum precipitation (AMP) event to describe the antecedent conditions, triggering precipitation event and streamflow response. ACRT = average catchment response time.

Type	AMF	AMP
Reference	flood peak	max. sum of P over ACRT window
Precipitation (P)	sum of P over ACRT plus flood start to peak	
Discharge (Q)	flood peak	max. Q within ACRT after AMP
	flood volume	
	flood duration	
	fast component volume	
	snow melt water volume	
Snow conditions	snow melt in Q	snow melt in Q
	snow pack ACRT before	snow pack ACRT before
Storage state	soil moisture at ACRT before	soil moisture (ACRT)
Seasonality	month of occurrence	month of occurrence

2.4 Return period estimation

Return periods were calculated for all annual maximum flood events and all annual maximum precipitation events. We calculated the empirical return periods of each AMF and AMP based on Weibull plotting positions. The return periods were categorized to return period classes "10 years" (between 0 and 10 years), "100 years" (between 10 and 100 years), "300 years" (between 100 and 300 years), "500 years" (between 300 and 500 years), "1000 years" (between 500 and 1000 years), and "1000+ years" (more than 1000 years). These classes are based on the different stages of flood safety assessment and form by definition an unbalanced stratification of the full sample of annual events.

2.5 Occurrence of annual maxima of precipitation and flood

When the AMF was not caused by the AMP, we expect the greatest influence of wet catchment antecedent conditions. In these "non-matching" cases, annual maximum precipitation does not trigger the annual maximum flood and hence during and before the AMP there might be antecedent conditions that allow for the rain to be stored in the catchments and do not lead to an immediate large streamflow response. For the AMFs that were not triggered by the largest precipitation events, there are two possible cases: 1) the catchment was considerably wetter compared to the conditions during the AMP – pointing at the decisive role of antecedent catchment conditions or 2) the precipitation amount triggering the AMF was very similar to the AMP, but did not quite reach the AMP amount. In other words, in these non-matching cases a rainfall event that is slightly or markedly lower than the AMP event leads to much higher runoff production efficiency and thus ultimately to the annual maximum flood.

We considered cases as "non-matching" when the AMP did not overlap with the window of the flood event plus the preceding average response time before the flood event. Since we want to focus on hydrologically effective precipitation we excluded AMPs where precipitation presumably was snow and accumulated.

2.6 Critical flood conditions for the large scale catchment

From a regional management point of view, what matters are floods occurring at the outlet of the river basin or at a point of interest within it. Critical conditions at these points are formed by antecedent conditions in specific regions (spatial patterns of wetness for contributing sub-catchments), by the phase, amount, and location of precipitation, and by the combined effect of individual space-time dynamics.

By only examining many catchments individually and how precipitation and antecedent conditions shape the streamflow response at their outlets, the link to the regional importance of floods at the sub-catchment scale would be missing. For instance, if the extreme flood of one catchment is always occurring out of sync with the other sub-catchments, then there might be no great impact expected in the large catchment context. However, if two or more catchments usually exhibit strong response to precipitation inputs and their flood peaks combine and reinforce one another, then these cases become crucial in the regional flood risk context.

In this study, critical floods at the large catchment scale were defined by the return periods of the floods at the outlet of the Aare river basin. To model the return period classes of these critical floods, we set up a classification type random forest. The

spatial pattern of antecedent conditions of the sub-catchments, the triggering precipitation within these sub-catchments, and

245 the conditions at the critical points of the routing were used as features for this random forest.

The conditions of precipitation and antecedent conditions for each sub-catchment were extracted for the individual sub-catchments trying to capture the seasonality of streamflow to travel from the outlet of each sub-catchment to the basin outlet. The conditions of the routing system that were considered as features in the random forest were the discharge values at critical locations preceding the floods at the outlet of the Aare river basin. Also here we accounted for travel times from the outlet to

250 each potentially relevant routing system location. The distributions of the features for the precipitation, antecedent soil moisture conditions and conditions at potentially relevant routing system locations can be found in the supplementary material.

The random forest was grown using a stratified sampling to improve the detection of the rarer return period classes given the very biased distribution of the number of flood events per return period class. The stratified sampling was set to 26, the size of flood events of the rarest class (500+). For the random forest, 5000 trees were grown and we applied 26 variables at each split,

255 which is more than the default square root of the number of features, but as recommended in Genuer et al. (2008) better for high dimensional classification type data sets. The optimal number of trees was determined by incrementally testing different numbers of trees and evaluating the overall out-of-bag estimate of the error rate for misclassifying the return period class of flood events. We examined the variable importance of the different predictors for each flood return period class using the mean decrease in the Gini index (MDI), which measures node impurity, i.e. how well the random forest trees split the data.

260 3 Results

3.1 Matching and non-matching AMP and AMF - sub-catchment scale

We found that only about 18 to 44% (depending on the sub-catchment) annual maximum precipitation (AMP) events and annual maximum flood (AMF) events occurred simultaneously in the simulations, highlighting the importance of antecedent conditions for the generation of large floods. When looking more closely into the non-matching events, we found that numer-

265 ous AMPs occurred after the AMF of that year. That means that these AMPs neither directly triggered the flood event nor contributed to wetting up the catchment before the flood event (Figure 3). For the rain and snow dominated catchments, 60% and more of the events of the AMP occurred after the AMF, while for the glacier influenced catchments about 40% of the AMPs occurred after the AMF.

In addition to the general decrease of the number of events for increasing return periods, the matching and non-matching

270 annual maxima of precipitation and discharge are not evenly distributed per return period class (Figure 4). With higher return periods there were more matching events. This indicates that increasingly extreme flood events are primarily explained by large precipitation amounts and less by existing antecedent conditions. Nevertheless, even for very large return periods, the antecedent conditions still seemed non-negligible and for single sub-catchments even large fractions of AMF were not explained by AMP (>25% up to 75% class 1000+). The points indicating high percentage of non-matching events in the large return

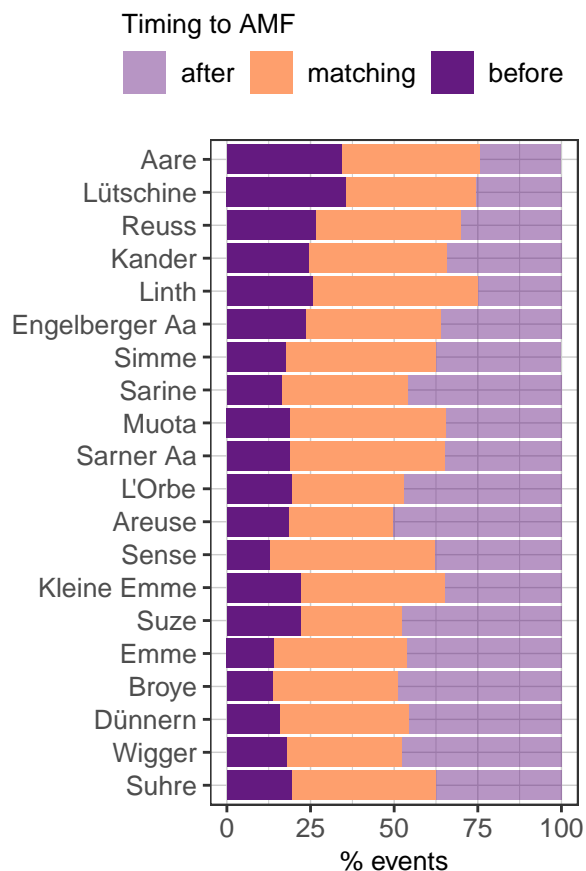275 periods come from the Wigger river catchment (Figure 4).

Figure 3. Timing of AMP to AMF events, i.e. occurrence of annual maximum precipitation (AMP) before, with or after the annual maximum flood (AMF).

Figure 5 shows the distribution of soil moisture conditions for matching and non-matching events separately and grouped for the different return period classes. Applying this separation reveals the soils are wetter during the non-matching events than in the matching events (Figure 5). This implies that even smaller precipitation events can lead to large floods if the antecedent conditions are wetter. When comparing different return period classes, it seems that soil moisture filling increases

280 for the more extreme events (higher return periods) in the non-matching event years. In some catchments, only matching events were found in the higher return period classes, meaning that the highest precipitation triggered the highest flood. In most studied catchments, for a return period class of 500 years or higher, only matching events were noted. This indicates the diminishing influence of soil moisture antecedent conditions on the occurrence of rarer flood events. The Wigger river catchment is distinctive among the selected catchments due to the presence of non-matching events across all return period

285 classes. Furthermore, only one matching event occurred for each return period class of 500 years or higher. This catchment exhibits a specific seasonality with numerous floods driven by snow melt that result in AMF, rather than summer rainfall events.
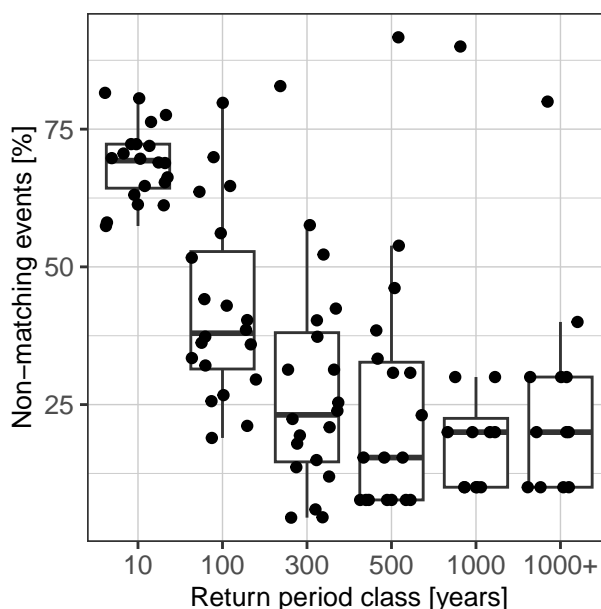
Figure 4. Percentage of non-matching events per return period class, excluding years with presumed snow AMPs. The higher the return period class is, the lower is the percentage of non-matching events. The dots show the percentage of non-matching events for the individual selected sub-catchments.

For the glaciated catchments, the ranges of antecedent soil moisture conditions were very similar between matching and non-matching events, indicating generally more persisting wet soil moisture conditions in these catchments. This can be also seen in the reference daily soil moisture distribution from all years. The difference in soil moisture conditions between matching and non-matching events is becoming smaller as we move from rain-dominated to snow-influenced and glaciated catchments. In rain-dominated catchments, antecedent soil moisture conditions vary widely for both matching and non-matching events. However, there is a tendency towards wetter soil moisture conditions for the non-matching events. The ranges of soil moisture conditions tend to be narrower in the snow influenced catchments compared to the rain dominated catchments. Here, the difference between matching and non-matching events in terms of soil moisture conditions diminishes and all events occur under wetter conditions compared to the rain dominated regime type catchments. The glaciated catchments show the narrowest range and are characterized by consistently wetter soil moisture antecedent conditions for both matching and non-matching events as they do also for the daily reference throughout the year. Generally and for all regime types, we see a more pronounced difference between the matching and non-matching moving towards higher return periods. It is important to keep in mind that as the return period increases, the depicted density functions are constructed from a decreasing number of events. For the more frequent events, a larger pool of events is available, with a wider spread for both matching and non-matching events. Conversely, for the rarer flood events we can compare only few to very few events.

Table 4. Confusion matrix of the classification type random forest with stratified sampling. The elements of the matrix indicate how many events of one return period class where also modelled into this return period class by the random forest.

	10	100	300	500+	class error
10	6278	1325	50	7	0.18
100	108	511	114	17	0.32
300	0	18	24	9	0.53
500+	0	4	6	16	0.38

We analyzed the contribution of snow melt to both matching and non-matching annual maxima of precipitation and discharge by calculating the volume of simulated snow melt in relation to the flood volume (Figure 6). As for the soil moisture antecedent conditions, there was more snow melt contribution in non-matching events than in matching events. In the rain dominated catchments we found a rather large difference for the snow melt contribution to the streamflow when comparing matching and non-matching events. For the snow melt dominated catchments this difference is smaller, reflecting that these catchments frequently experience snow melt influenced floods. Also the Wigger catchment, which has a rain dominated regime, aligns with the description provided in the soil moisture antecedent conditions part above. It shows important snow melt contributions in the AMF up to the highest return period class. The catchments with glacier melt influenced regimes display similar distributions of snow melt contributions to floods for both matching and non-matching events.

3.2 Drivers and spatial patterns leading to floods at the large basin outlet

The confusion matrix of the random forest (Table 4) displays how accurately the random forest attributed the floods of a specific return period class to that same based on the provided features. For instance, in the confusion matrix (Table 4), the return period class of ten years was matched 6278 times. However, it was erroneously put in the return period class of 100 years 1325 times, and in the return period class of 300 years 50 times. The random forest model appeared to have difficulty classifying the 300-year return period class, with a classification error of 52% (see Table 4), compared to the other return period classes.

The maps in Figure 7 illustrate the variable importance of the antecedent conditions regarding snow and soil moisture, triggering precipitation and routing node conditions split up for the different return period classes. When the mean decrease in the Gini impurity (MDI) is higher, also the variable importance is higher. From these maps, it appears that for lower return periods (100 years), soil moisture is the most important feature and for all analyzed sub-catchments. Snow pack conditions are assigned small importance in general but slightly higher for the glacier and snow influenced alpine catchments. However, triggering precipitation is only important for some sub-catchments and not important for all the glacier influenced catchments from the sub-catchment selection. The triggering precipitation of the sub-catchments Simme, Emme, and Suhre gets a slightly higher importance than the other considered sub-catchments. For the discharge conditions at all the routing system locations some but small importance is indicated by the MDI.

Moving to higher return periods the pattern of the attribution of the individual sub-catchments is changing and precipitation becomes generally more important in explaining the attribution to the very high return period class and her also for the catchments with a glacier influenced regime. The attribution level assigned by the MDI to the triggering precipitation for the 300-year return period class is very homogeneous. While the pattern of soil moisture antecedent conditions variable importance remains about the same as in the lower return period class, triggering precipitation gains importance compared to the soil moisture conditions when looking at the 300-year return period. Snow pack instead loses importance in classifying the return period class. For the 300-year return period class the locations in the routing system cascading downstream of the Sarine sub-catchment gain variable importance, while there was barely any assigned to these routing system locations for the return period class 100 years.

In the return period class 500+ years, the variable importance is mainly attributed to the triggering precipitation of all sub-catchments and particularly to Reuss and Muota. Also in the Broye sub-catchment, triggering precipitation becomes more prominent in assigned variable importance for these events, while it was negligible for the lower return period classes. Soil moisture does not help much in determining this return period class, with some importance being assigned to the Kander sub-catchment and some but even less to the Sense, Wigger, Suhre and Duennern sub-catchments. The Muota sub-catchment did have very little variable importance in the lower return period classes, but becomes more important in the classification of events for the 1000+ return period class.

While snow melt contribution to the flood at the small sub-catchment level showed a distinct difference between the matching and the non-matching events, underscoring their importance for some events of this return period class, it did barely contribute to the flood return period classification at the river basin outlet, particularly for the rarer flood events. This can be explained by the different seasonality of floods occurring at the sub-catchment scale compared to the floods at the outlet, which are mainly summer floods. For both snow and the flow conditions at locations in the routing system there were no noteworthy variable importance assigned. Interestingly, when looking only at the change of variable importance for the conditions at the routing system locations, some attribution changes for the different return period classes can be noticed. The attribution of variable importance is for the lower return period class (100 years) at locations closest to the Aare river basin outlet, for the higher return period class (300 years) it moves to the locations in the cascade downstream of the Sarine catchment, and for the highest return period class to the outlet of Lake Thun.

4 Discussion

The result that the AMP and AMF are to a large degree not occurring together are also found in other climatic zones for instance in the contiguous United States Do et al. (2020) found a low correlation between changes in precipitation extremes and floods and attributed that to the small fraction of co-occurrence of these events.

As found by Nied et al. (2017) for the Elbe catchment at a large catchment scale, antecedent soil moisture conditions are important for both frequent and rare flood events observed in our study on the sub-catchments. In contrast to the finding of Nied et al. (2017), we discovered that these conditions are important up to surprisingly high return periods. The soil moisture

360 conditions of the sub-catchments were also identified as important variables to describe the return period classes at the river basin outlet, particularly for the medium range frequent floods (100 and 300 years return period class). Only for the rarest flood events the role of the soil moisture condition could not be distinguished anymore when comparing the matching and non-matching events. The decreasing effect of antecedent conditions was also found in comparisons of trends in floods and extreme precipitation events (Woldemeskel and Sharma, 2016; Tramblay et al., 2019; Bennett et al., 2018). Wasko and Nathan

365 (2019) found the threshold of when the importance of antecedent soil moisture conditions was negligible compared to the triggering precipitation for Australian catchments using a flood streamflow elasticity approach, already at events of around a 10 year return period.

There was a difference for the different regime types of the sub-catchments regarding the influence of soil moisture antecedent conditions. Notably, the rain dominated and slightly snow influenced catchments exhibited the most important dif-

370 ference in antecedent conditions between matching and non-matching events. This points at the importance of soil moisture antecedent conditions, particularly for catchments with these specific regimes. For the snow antecedent condition and its influence on the flood events, although they were to a certain degree important for the floods at the outlets of the sub-catchments, they did not emerge as important triggering factors at the large catchment scale. Also, the routing was not found to be important for modelling the return period classes of the floods at the outlet. This does not mean that they are unimportant for single

375 events, but rather suggests that they are not as critical to classify the events into return period classes.

Having very long CS available allowed looking at many more flood events than it would be possible with observations alone. For instance, it would not have been possible to analyze the space-time patterns that are important to describe floods on a large catchment scale with observations alone. Nied et al. (2017) used reshuffling meteorological and soil moisture conditions to get more insights into the importance of hydro-meteorological processes on floods. In our approach, using the stochastic weather

380 generator at the beginning of the hydro-meteorological modelling chain, the CS was even more stretched, allowing to analyze the space-time patterns at the large catchment scale for the rare flood events as well and providing a more robust basis for the more frequent floods. Based on the CS approach, which includes a stochastic weather generator, extreme floods with return periods of more than 300 years had still relevant variable importance assigned to soil moisture conditions for almost all sub-catchments. Nevertheless, when trying to model the return period classes of the 500-year return period, the data set may have

385 been still too small, and the floods included in the 500-year return period might have been too diverse to be properly classified with the features provided to the random forest.

4.1 Limitations of the study design

With our study set up we could not make an analysis if the precipitation event was patchy or not (spatial analysis of the rain event) as for instance Tarasova et al. (2020, 2019) did within the hydrological catchments. However, we could analyze the

390 spatial interplay of the sub-catchments of the Aare river basin with regard to large floods at its outlet. We used a lumped hydrological model for each sub-catchment with mean areal input. That means that we could not find patterns within the sub-catchments that are particularly critical for the large system. These patterns might have been informative for some flood events in specific sub-catchments, since the relationship between performance of the streamflow simulation and spatial resolution of

precipitation is both scale-dependent and catchment-dependent as shown for instance by Lobligeois et al. (2014) for France.
However, from the regional perspective, we could analyze how the interplay of sub-catchment antecedent conditions and
precipitation input as well as buffering and timing upstream the outlet of the Aare river basin influenced the floods at the outlet.

The robustness of the results depends, in part, on the type of precipitation events that is used to force the hydrological model.
This assigns a crucial role to the weather generator at the beginning of the modelling chain. However, for this study the main
goal was to find conditions that lead to extreme flood events and the role of antecedent conditions therein. The way in which
the weather generator was set up and optimized might not represent the full range of possible events at each sub-catchment in
the region. Nevertheless, we do not expect major changes in the antecedent conditions prior to a large precipitation event since
these are generally building up over a longer time period.

4.1.1 Event definition

The flood event was defined based on Q_{peak} and the precipitation event based on the sum of precipitation over the catchment-
specific window (average response time), which does not include any information about other precipitation event properties
that are potentially of interest, such as precipitation intensity or storm advancement. These catchment specific windows for the
definition of the precipitation events, however, is important when considering varying time scales of flood generating processes
in different catchments. For Switzerland, for instance, Froidevaux et al. (2015) found a rather short discharge memory for
catchments in pre-Alpine, Alpine and South Alpine regions, and that considering more than three to four days of antecedent
precipitation was not relevant for flood generation. However, antecedent conditions of four or more days before the flood were
found to be relevant in the Jura Mountains, in the western and eastern Swiss Plateau, and at the outlet of large lakes.

When considering return periods, a more process based event definition is not possible because the events must originate from
the same population. It is not reasonable to assume that longer and shorter precipitation sums belong to the same population.
This leaves us with the approach of choosing fixed time windows as basis. While this is not problematic for flood peaks, which
represent a single discharge value per year, it becomes an issue for precipitation events. For instance, when precipitation events
are defined over a fixed window, which is the standard approach in meteorology, we might not capture the entire precipitation
event including its start and end. This can result in the loss of information about storm intensity, storm advancement, and
other important factors. Moreover, we might miss the full information on effective precipitation if we look at the precipitation
amounts only.

The event definition in this study relies solely on information about precipitation and discharge, and this can be done both
with simulated time series (as in our hydro-meteorological modelling chain approach) and observed time series. Even if we
have additional simulated variables in our approach, we could pretend to not have them and see how far we get in predicting
floods using only precipitation. Precipitation events and their return periods are often used to estimate the flood return periods
(Naghettini et al., 1996). Having the additional simulated variables to analyze the antecedent conditions to the flood events
reveals cases where this approach is not sufficient, i.e. the cases were the annual largest precipitation event did not lead to the
largest annual flood event. We found that such cases are rather common within our catchments. This challenges the assumptions

made in design approaches transferring AMP to AMF on a statistical rather than at an event based basis. Hence, these results point at the important role of antecedent conditions even for relatively large return periods.

The definition of when a flood and a precipitation event match could influence the role assigned to antecedent conditions.

430   Many AMPs occurred after the AMF and had no effect on triggering or preparing antecedent conditions. However, for AMPs before the AMFs, the definition could play a role. After conducting a sensitivity analysis by systematically altering the definition of ACRT before the flood from 0.5 to 1.5 to 2 times, we found that our findings remained unchanged (supplementary material).

4.1.2   Characterization of the antecedent conditions

435   At times, precipitation accumulates as snow, and the subsequent snow melt later in the season contributes to the liquid water input into a catchment. To accurately select the largest water input event and compare it to the largest flood response event, it becomes necessary to consider snow accumulation and melt processes. However, the buildup of the snow pack varies spatially within each sub-catchment, influenced by factors such as elevation, aspect, vegetation distribution, and wind re-distribution. This is only very roughly covered in our hydrological model by distributing precipitation into elevation zones. This spatial het-

440   erogeneity, for instance, results in varying routing times for snow melt and will probably vary strongly between sub-catchments. Moreover, the snow melting process in the hydrological model is based on a degree day approach, neglecting for instance advective rain-on-snow events, which can be substantially contributing to the generation of floods.

A future approach to studying the antecedent conditions leading to floods could be to look directly at snow melt and rain rather than precipitation to understand the processes and antecedent conditions that lead to floods, seasonal differences in soil

445   moisture etc. By additionally looking at the annual maximum total water input to the system, i.e., the sum of liquid precipitation and snow melt, we might be able examine the soil moisture and other catchment storage antecedent conditions more closely. However, snow melt can be different depending on the processes involved (Sikorska-Senoner and Seibert, 2020), and in a simple snow routine, advective rain events that fall on a snow pack and bring energy to melt it faster would not have been included (rain-on-snow). In this study we focused on the return periods from an almost classic AMP approach versus the return

450   periods of AMFs and found that not all AMPs necessarily lead to AMFs. One of the motivations for using this almost classic approach of a fixed window for the AMP extraction was that the AMPs usually come without further information about other possible inputs to the catchment but are derived directly from the meteorological station data.

4.1.3   Flood frequency analysis and flood generating processes

Hydrologically, one can ask the question whether all these maximum annual flood events can be treated as if they would origi-

455   nate from the same population since often they are created by different flood generating processes (see e.g. Merz and Blöschl, 2008). Also from a management point of view, floods originating from different generating processes might be expected more in one season than in another. In addition, they might and behave differently on flood inundation areas (Sikorska et al., 2015; Brunner et al., 2017). In the statistical analysis, floods types that dominate upper tails of the distribution may not be adequately represented, often treated as a single sample along with more frequent floods Tarasova et al. (2020). A flood type wise model

460 was proposed by Fischer (2018); Fischer et al. (2019); Fischer and Schumann (2021), where floods from a peak-over-threshold approach were first separated into flood types and then combined into a mixture model to calculate the return period from the joint function.

With the annual maximum flood approach, only one flood per year is analyzed regarding the antecedent conditions. If we would have chosen a peak over threshold (POT) approach instead, we could have sampled more relevant events per year.

465 However, this approach has the downside that events could be dependent. In statistical flood frequency analysis, the assumption of event independence is however crucial, allowing these events to be treated as random variables. Moreover, employing a POT approach often involves subjective choices, such as determining the appropriate threshold (Fischer, 2018; Fischer and Schumann, 2021) and selecting which events to pool.

When we compare AMF and AMP and assess their linkage, we might arrive at conclusions that stem from a lack of clear

470 differentiation regarding the size of the event. For instance, the precipitation sum preceding the AMF can be nearly as large as the AMP, or it can be much smaller. The conclusion regarding antecedent conditions might be different depending on which time window we used. Similarly, the streamflow response to an AMP might be large but not quite as large as the streamflow response that contributed to the actual AMF. Since we selected the AMP using a fixed window and adjusted the precipitation triggering the AMF based on the onset of the flood, peak and the fixed window preceding the flood onset, this comparison

475 could not be done in a straightforward manner. The relative difference of the precipitation preceding the AMF with regard to AMP ranged, on average, from 22% to 85% across all catchments.

4.2 Broader impact

As outlined in the introduction, the approach of using a weather generator within a regional (large catchment) context, in combination with a hydrological model and routing, has several advantages. It implicitly "reshuffles" the initial conditions

480 and combines them with plausible weather events for this region. This approach results in a larger and more diverse pool of flood events compared to only using observations or making assumptions about the antecedent conditions. The approach could also be applied to different large catchments within a comparable climate, dominated by similar regime types. One important prerequisite for generating long weather time series using a weather generator is, however, the availability of an adequate number of weather stations with sufficiently long records to robustly estimate plausible weather.

485 Like in other studies that compare flood and precipitation events such as GRADEX (e.g. Guillot and Duband, 1969; Naghettini et al., 1996), or future (Brunner et al., 2021) and past (Wilhelm et al., 2022) frequency distributions of the two variables, there are some thresholds or tipping points that emerge. These thresholds are associated with the influence of the antecedent conditions and appear to remain important even for remarkably high flood return periods in our study. This underlines their importance, emphasizing that they should not be neglected. Consequently, the transfer from precipitation frequency distributions

490 to flood frequency distributions should be checked for appropriateness in each specific case.

5 Conclusions

In this study, we assessed the role of antecedent conditions for floods of different return periods using simulations from a hydro-meteorological modelling chain, which includes a stochastic weather generator, a hydrological model, and a routing system as basis. We focused on the link between precipitation, antecedent conditions and return periods for the sub-catchments of the Aare river basin. The availability of very long CS enabled the analysis of a greater number of flood events than would have been possible with observations alone. For example, it would have been impossible to analyze the space-time patterns that are crucial for describing floods on a large catchment scale using only observations. his way, we could investigate the temporal and spatial interactions among conditions in these sub-catchments that lead to floods at the outlet of the Aare river basin.

In the case of sub-catchments, antecedent conditions play an important role for floods with large return periods up to 500 years. This role diminishes and becomes negligible for very high return periods only of more than 500 years. The regime type of the sub-catchments played a crucial role: In rain-dominated catchments, the soil moisture antecedent conditions led to the most substantial difference between matching and non-matching events of AMP and AMF. For the snow-influenced and the glacier-influenced catchments, this difference diminished.

At the large catchment scale, soil moisture antecedent conditions are crucial for correctly classifying the lower return periods but diminish in importance as we consider higher return periods of 500 and more years. Neither snow antecedent conditions nor confluence and flow time were found to be important for classification at the outlet of the river basin when using a random forest classification type model.

Hence, it is important to check the appropriateness of transferring from precipitation frequency distributions to flood frequency distributions, as the antecedent catchment conditions are usually not negligible.

References

Bennett, B., Leonard, M., Deng, Y., and Westra, S.: An empirical investigation into the effect of antecedent precipitation on flood volume, Journal of Hydrology, 567, 435–445, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.10.025, 2018.

Bergström, S., Singh, V. P., and others: The HBV model., Computer models of watershed hydrology., pp. 443–476, 1995.

520 Berthet, L., Andréassian, V., Perrin, C., and Javelle, P.: How crucial is it to account for the antecedent moisture conditions in flood fore-casting? Comparison of event-based and continuous approaches on 178 catchments, Hydrology and Earth System Sciences, 13, 819–831, https://doi.org/10.5194/hess-13-819-2009, 2009.

Beven, K.: Towards the use of catchment geomorphology in flood frequency predictions, Earth Surface Processes and Landforms, 12, 69–82, https://doi.org/10.1002/esp.3290120109, 1987.

525 Blazkova, S. and Beven, K.: Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic, Journal of Hydrology, 292, 153–172, https://doi.org/10.1016/j.jhydrol.2003.12.025, 2004.

Blume, T. and van Meerveld, H. I.: From hillslope to stream: methods to investigate subsurface connectivity, Wiley Interdisciplinary Reviews: Water, 2, 177–198, https://doi.org/10.1002/wat2.1071, 2015.

530 Brocca, L., Melone, F., and Moramarco, T.: On the estimation of antecedent wetness conditions in rainfall–runoff modelling, Hydrological Processes, 22, 629–642, https://doi.org/https://doi.org/10.1002/hyp.6629, 2008.

Brunner, M. I., Seibert, J., and Favre, A.: Bivariate return periods and their importance for flood peak and volume estimation, WIREs Water, 3, 819–833, https://doi.org/10.1002/wat2.1173, 2016.

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A., and Seibert, J.: Flood type specific construction of synthetic design 535 hydrographs, Water Resources Research, 53, 1390–1406, https://doi.org/10.1002/2016WR019535, 2017.

Brunner, M. I., Swain, D. L., Wood, R. R., Willkofer, F., Done, J. M., Gilleland, E., and Ludwig, R.: An extremeness threshold determines the regional response of floods to changes in rainfall extremes, Communications Earth & Environment, 2, 173, https://doi.org/10.1038/s43247-021-00248-x, 2021.

Calver, A. and Lamb, R.: Flood frequency estimation using continuous rainfall-runoff modelling, Physics and Chemistry of the Earth, 20, 540 479–483, https://doi.org/10.1016/S0079-1946(96)00010-9, 1995.

Castellarin, A., Kohnová, S., Gaál, L., Fleig, A., Salinas, J. L., Toumazis, A., Kjeldsen, T., and Macdonald, N.: A review of applied-statistical methods for flood-frequency analysis in Europe., Tech. rep., The Centre for Ecology & Hydrology (FloodFreq COST Action ES0901)., http://nora.nerc.ac.uk/id/eprint/19286/, 2012.

Do, H. X., Mei, Y., and Gronewold, A. D.: To What Extent Are Changes in Flood Magnitude Related to Changes in Precipitation Ex-545 tremes?, Geophysical Research Letters, 47, e2020GL088 684, https://doi.org/https://doi.org/10.1029/2020GL088684, e2020GL088684 2020GL088684, 2020.

Eckhardt, K.: How to construct recursive digital filters for baseflow separation, Hydrological Processes, 19, 507–515, https://doi.org/10.1002/hyp.5675, 2005.

Evin, G., Favre, A.-C., and Hingray, B.: Stochastic generation of multi-site daily precipitation focusing on extreme events, Hydrology and 550 Earth System Sciences, 22, 655–672, https://doi.org/10.5194/hess-22-655-2018, 2018.

Evin, G., Favre, A.-C., and Hingray, B.: Stochastic generators of multi-site daily temperature: comparison of performances in various applications, Theoretical and Applied Climatology, 135, 811–824, https://doi.org/10.1007/s00704-018-2404-x, 2019.

Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hundecha, Y., Apel, H., and Merz, B.: Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain, Journal of Hydrology, 524, 182–193, https://doi.org/10.1016/j.jhydrol.2015.02.021, 2015.

Fischer, S.: A seasonal mixed-POT model to estimate high flood quantiles from different event types and seasons, Journal of Applied Statistics, 45, 2831–2847, https://doi.org/10.1080/02664763.2018.1441385, 2018.

Fischer, S. and Schumann, A. H.: Multivariate Flood Frequency Analysis in Large River Basins Considering Tributary Impacts and Flood Types, Water Resources Research, 57, https://doi.org/10.1029/2020WR029029, 2021.

Fischer, S., Schumann, A., and Bühler, P.: Timescale-based flood typing to estimate temporal changes in flood frequencies, Hydrological Sciences Journal, 64, 1867–1892, https://doi.org/10.1080/02626667.2019.1679376, 2019.

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., and Martius, O.: Flood triggering in Switzerland: the role of daily to monthly preceding precipitation, Hydrology and Earth System Sciences, 19, 3903–3924, https://doi.org/10.5194/hess-19-3903-2015, 2015.

García Hernández, J., Foehn, A., Fluixá-Sanmartín, J., Roquier, B., Brauchli, T., and Paredes Arquiola, J.: RS MINERVE – Technical manual, v2.25, Tech. rep., CREALP, Sion, https://crealp.ch/wp-content/uploads/2021/09/rsminerve_technical_manual_v2.25.pdf, 2020.

Genuer, R., Poggi, J.-M., and Tuleau, C.: Random Forests: some methodological insights, arXiv preprint arXiv:0811.3619, 2008.

Guillot, P. and Duband, D.: La méthode du gradex pour le calcul de la probabilité des crues à partir des pluies, in: Floods and their Computation, in: Proceedings of the Leningrad Symposium,, edited by International Association of Hydrological Sciences and World Meteorological Organization, pp. 560–569, Gentbrugge, Paris, 1969.

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

Hingray, B., Schaefli, B., Mezghani, A., and Hamdi, Y.: Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments, Hydrological Sciences Journal, 55, 1002–1016, https://doi.org/10.1080/02626667.2010.505572, 2010.

Horton, P., Schaefli, B., and Kauzlaric, M.: Why do we have so many different hydrological models? A review based on the case of Switzerland, WIREs Water, 9, https://doi.org/10.1002/wat2.1574, 2022.

Hundecha, Y. and Merz, B.: Exploring the relationship between changes in climate and floods using a model-based analysis, Water Resources Research, 48, https://doi.org/10.1029/2011WR010527, 2012.

Keller, L., Rössler, O., Martius, O., and Weingartner, R.: Delineation of flood generating processes and their hydrological response, Hydrological Processes, 32, 228–240, https://doi.org/10.1002/hyp.11407, 2018.

Klemeš, V.: Tall Tales about Tails of Hydrological Distributions. I, Journal of Hydrologic Engineering, 5, 227–231, https://doi.org/10.1061/(ASCE)1084-0699(2000)5:3(227), 2000.

Klemeš, V.: Tall Tales about Tails of Hydrological Distributions. II., Journal of Hydrologic Engineering, 5, 232–239, https://doi.org/10.1061/(ASCE)1084-0699(2000)5:3(232), 2000.

Köplin, N., Schädler, B., Viviroli, D., and Weingartner, R.: The importance of glacier and forest change in hydrological climate-impact studies, Hydrology and Earth System Sciences, 17, 619–635, https://doi.org/10.5194/hess-17-619-2013, 2013.

Köplin, N., Schädler, B., Viviroli, D., and Weingartner, R.: Seasonality and magnitude of floods in Switzerland under future climate change, Hydrological Processes, 28, 2567–2578, https://doi.org/10.1002/hyp.9757, 2014.

Lamb, R., Faulkner, D., Wass, P., and Cameron, D.: Have applications of continuous rainfall–runoff simulation realized the vision for process-based flood frequency analysis?, Hydrological Processes, 30, 2463–2481, https://doi.org/10.1002/hyp.10882, 2016.

23

590 Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 201, 272–288, https://doi.org/10.1016/S0022-1694(97)00041-3, 1997.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C.: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, Hydrology and Earth System Sciences, 18, 575–594, https://doi.org/10.5194/hess-18-575-2014, 2014.

595 Makkonen, L.: Plotting Positions in Extreme Value Analysis, Journal of Applied Meteorology and Climatology, 45, 334–340, https://doi.org/10.1175/JAM2349.1, 2006.

Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., Guse, B., Viglione, A., Vorogushyn, S., Macdonald, E., Wietzke, L., and Schumann, A.: Understanding Heavy Tails of Flood Peak Distributions, Water Resources Research, 58, https://doi.org/10.1029/2021WR030506, 2022.

600 Merz, R. and Blöschl, G.: Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information, Water Resources Research, 44, https://doi.org/10.1029/2007WR006744, 2008.

Michele, C. and Salvadori, G.: On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition, Journal of Hydrology, 262, 245–258, https://doi.org/https://doi.org/10.1016/S0022-1694(02)00025-2, 2002.

Naghettini, M., Potter, K. W., and Illangasekare, T.: Estimating the upper tail of flood-peak frequency distributions using hydrometeorological
605 information, Water Resources Research, 32, 1729–1740, https://doi.org/10.1029/96WR00200, 1996.

Nied, M., Hundecha, Y., and Merz, B.: Flood-initiating catchment conditions: a spatio-temporal analysis of large-scale soil moisture patterns in the Elbe River basin, Hydrology and Earth System Sciences, 17, 1401–1414, https://doi.org/10.5194/hess-17-1401-2013, 2013.

Nied, M., Schröter, K., Lüdtke, S., Nguyen, V. D., and Merz, B.: What are the hydro-meteorological controls on flood characteristics?, Journal of Hydrology, 545, 310–326, https://doi.org/10.1016/j.jhydrol.2016.12.003, 2017.

610 Pathiraja, S., Westra, S., and Sharma, A.: Why continuous simulation? the role of antecedent moisture in design flood estimation, Water Resources Research, 48, https://doi.org/10.1029/2011WR010997, 2012.

Schulte, L., Wetter, O., Wilhelm, B., Peña, J. C., Amann, B., Wirth, S. B., Carvalho, F., and Gómez-Bolea, A.: Integration of multi-archive datasets for the development of a four-dimensional paleoflood model of alpine catchments, Global and Planetary Change, 180, 66–88, https://doi.org/10.1016/j.gloplacha.2019.05.011, 2019.

615 Seibert, J.: On the need for benchmarks in hydrological modelling, Hydrological Processes, 15, 1063–1064, https://doi.org/10.1002/hyp.446, 2001.

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

Sikorska, A. E., Viviroli, D., and Seibert, J.: Flood-type classification in mountainous catchments using crisp and fuzzy decision trees, Water
620 Resources Research, 51, 7959–7976, https://doi.org/10.1002/2015WR017326, 2015.

Sikorska-Senoner, A. E. and Seibert, J.: Flood-type trend analysis for alpine catchments, Hydrological Sciences Journal, 65, 1281–1299, https://doi.org/10.1080/02626667.2020.1749761, 2020.

Sikorska-Senoner, A. E., Schaefli, B., and Seibert, J.: Downsizing parameter ensembles for simulations of rare floods, Natural Hazards and Earth System Sciences, 20, 3521–3549, https://doi.org/10.5194/nhess-20-3521-2020, 2020.

625 Staudinger, M., Stoelzle, M., Cochand, F., Seibert, J., Weiler, M., and Hunkeler, D.: Your work is my boundary condition ! Challenges and approaches for a closer collaboration between hydrologists and hydrogeologists, Journal of Hydrology, 571, 235–243, https://doi.org/10.1016/j.jhydrol.2019.01.058, 2019.

Swisstopo: DHM25: Das digitale Höhenmodell der Schweiz, Tech. rep., Wabern, https://www.swisstopo.admin.ch/content/
swisstopo-internet/de/geodata/height/dhm25/_jcr_content/contentPar/tabs_copy/items/dokumente/tabPar/downloadlist/downloadItems/
630   868_1464696772548.download/dhm25infode.pdf, 2005.

Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., Viglione, A., Plötner, S., Guse, B., Schumann, A., Fischer, S., Ahrens, B.,
Anwar, F., Bárdossy, A., Bühler, P., Haberlandt, U., Kreibich, H., Krug, A., Lun, D., Müller-Thomy, H., Pidoto, R., Primo, C., Seidel, J.,
Vorogushyn, S., and Wietzke, L.: Causative classification of river flood events, WIREs Water, 6, https://doi.org/10.1002/wat2.1353, 2019.

Tarasova, L., Basso, S., and Merz, R.: Transformation of Generation Processes From Small Runoff Events to Large Floods, Geophysical
635   Research Letters, 47, https://doi.org/10.1029/2020GL090547, 2020.

Todini, E.: Role and Treatment of Uncertainty in Real-Time Flood Forecasting, Hydrological Processes, 18, 2743–2746,
https://doi.org/10.1002/hyp.5687, 2004.

Tramblay, Y., Mimeau, L., Neppel, L., Vinet, F., and Sauquet, E.: Detection and attribution of flood trends in Mediterranean basins, Hydrology
and Earth System Sciences, 23, 4419–4431, https://doi.org/10.5194/hess-23-4419-2019, 2019.

640   Viviroli, D., Sikorska-Senoner, A. E., Evin, G., Staudinger, M., Kauzlaric, M., Chardon, J., Favre, A.-C., Hingray, B., Nicolet, G., Raynaud,
D., Seibert, J., Weingartner, R., and Whealton, C.: Comprehensive space–time hydrometeorological simulations for estimating very rare
floods at multiple sites in a large river basin, Natural Hazards and Earth System Sciences, 22, 2891–2920, https://doi.org/10.5194/nhess-
22-2891-2022, 2022.

Wasko, C. and Nathan, R.: Influence of changes in rainfall and soil moisture on trends in flooding, Journal of Hydrology, 575, 432–441,
645   https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.05.054, 2019.

Westerberg, I., Guerrero, J.-L., Seibert, J., Beven, K. J., and Halldin, S.: Stage-discharge uncertainty derived with a non-stationary rating
curve in the Choluteca River, Honduras, Hydrological Processes, 25, 603–613, https://doi.org/10.1002/hyp.7848, 2011.

Westerberg, I. K., Sikorska-Senoner, A. E., Viviroli, D., Vis, M., and Seibert, J.: Hydrological model calibration with uncertain discharge
data, Hydrological Sciences Journal, pp. 1–16, https://doi.org/10.1080/02626667.2020.1735638, 2020.

650   Wilhelm, B., Rapuc, W., Amann, B., Anselmetti, F. S., Arnaud, F., Blanchet, J., Brauer, A., Czymzik, M., Giguet-Covex, C., Gilli, A., Glur,
L., Grosjean, M., Irmler, R., Nicolle, M., Sabatier, P., Swierczynski, T., and Wirth, S. B.: Impact of warmer climate periods on flood hazard
in the European Alps, Nature Geoscience, 15, 118–123, https://doi.org/10.1038/s41561-021-00878-y, 2022.

Wilks, D. S. and Wilby, R. L.: The weather generation game: a review of stochastic weather models, Progress in Physical Geography: Earth
and Environment, 23, 329–357, https://doi.org/10.1177/030913339902300302, 1999.

655   Woldemeskel, F. and Sharma, A.: Should flood regimes change in a warming climate? The role of antecedent moisture conditions, Geophys-
ical Research Letters, 43, 7556–7563, https://doi.org/https://doi.org/10.1002/2016GL069448, 2016.

Wood, A. W. and Lettenmaier, D. P.: An ensemble approach for attribution of hydrologic prediction uncertainty, Geophysical Research
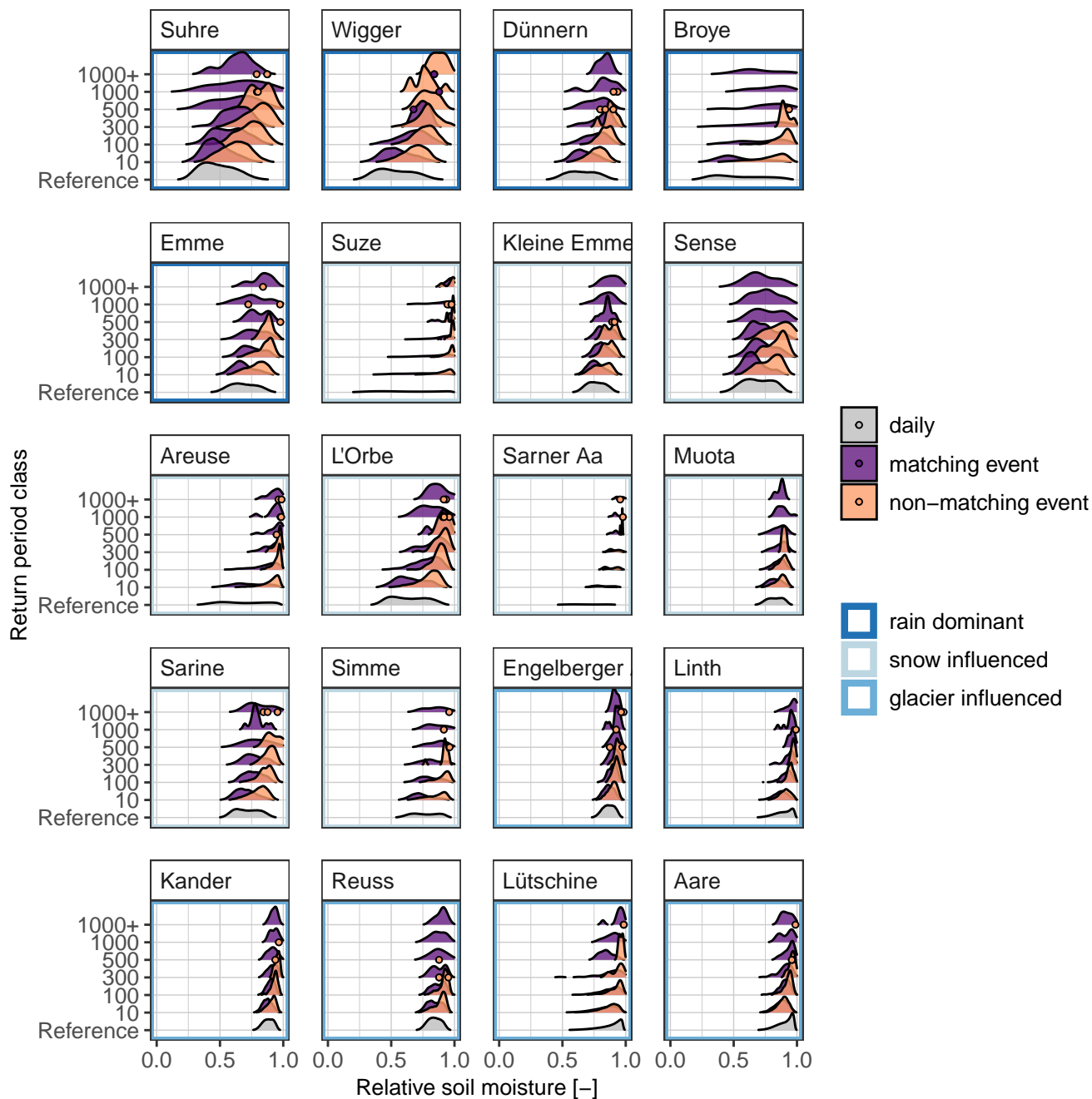Letters, 35, 2008.

Figure 5. Comparison between antecedent soil moisture conditions for matching and non-matching annual maxima of precipitation and discharge and the reference of daily soil moisture conditions throughout the year (grey). In cases with two or fewer events per class, no density distribution is plotted. Instead, the soil moisture values are shown as single points. The colour frame around the panel indicates the discharge regime of the catchment.
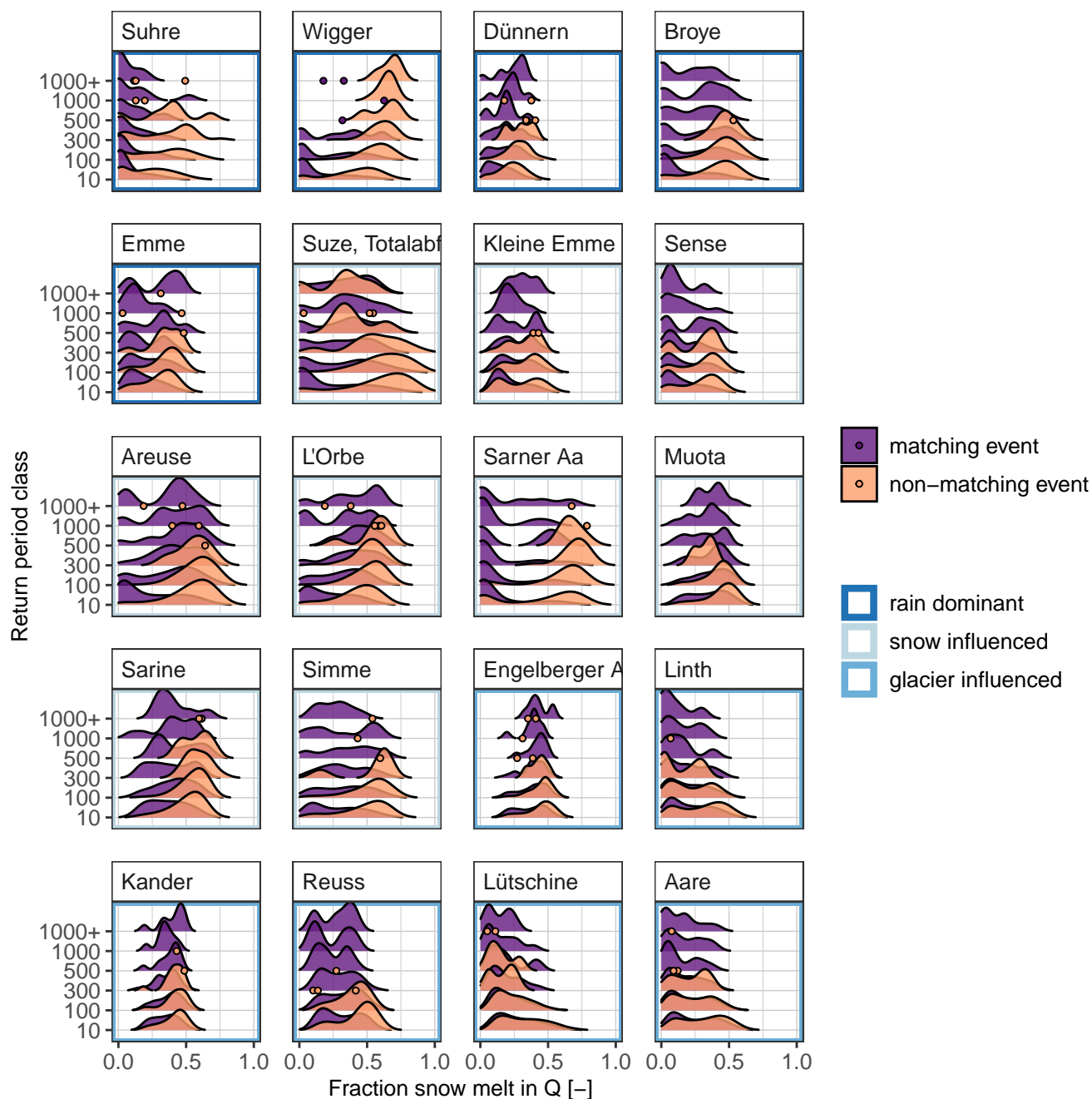
Figure 6. Comparison between contributing snow melt at matching and non-matching annual maxima of precipitation and discharge. In cases with two or fewer events per class, no density distribution is plotted. Instead, the contributing snow melt values are shown as single points. The colour frame around the panel indicates the discharge regime of the catchment.
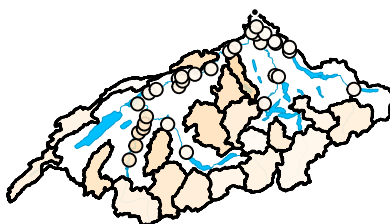
MDI [–]

0   148
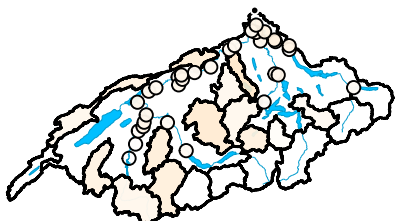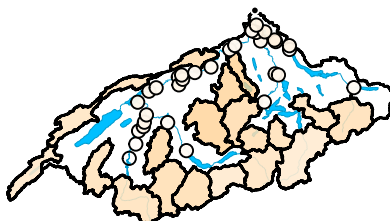
100 years

PREC   SM   SNOW
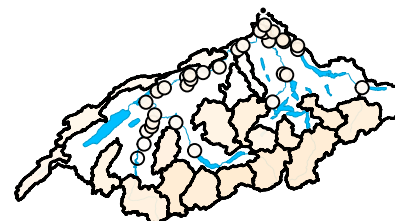
300 years
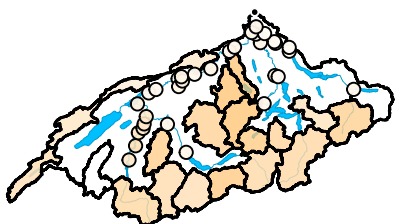
PREC   SM   SNOW
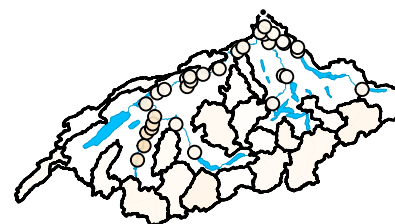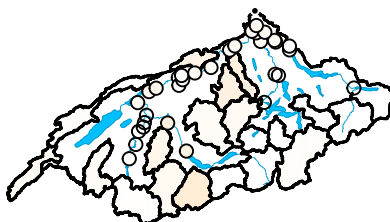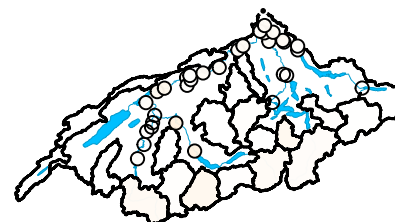
500+ years
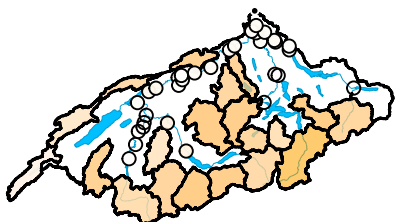
PREC   SM   SNOW

Figure 7. Maps of the Aare river basin and its sub-catchments showing the variable importance of triggering precipitation, soil moisture and snow conditions in the sub-catchments as well as the flow conditions at important nodes in the routing system for the return period classes 100, 300 and 500+ years. The mean decrease in the Gini impurity (MDI) is used to measure the variable importance, the darker the colour the more importance was assigned to the respective variable.