Journal of Hydrology 456–457 (2012) 30–43

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Runoff models and flood frequency statistics for design flood estimation in Austria – Do they tell a consistent story?

M. Rogger [a,*], B. Kohl [b], H. Pirkl [c], A. Viglione [a], J. Komma [a], R. Kirnbauer [a], R. Merz [d], G. Blöschl [a]

[a] Institute for Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Karlsplatz 13, 1040 Vienna, Austria
[b] Federal Research and Training Centre for Forests, Natural Hazards and Landscape, Rennweg 1, 6020 Innsbruck, Austria
[c] Technical Office for Geology Dr. Herbert Pirkl, Plenergasse 5/27, 1180 Vienna, Austria
[d] Department for Catchment Hydrology, Helmhotz Centre for Environmental Research UFZ, Theodor-Lieser-Straße 4, 06120 Halle (Saale), Germany

ARTICLE INFO

Article history:
Received 2 February 2012
Received in revised form 18 May 2012
Accepted 30 May 2012
Available online 7 June 2012
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Attilio Castellarin, Associate Editor

Keywords:
Design flood
Design storm method
Flood frequency statistics
Rainfall runoff models

SUMMARY

Design floods for a given location at a stream can be estimated by a number of approaches including flood frequency statistics and the design storm method. If applied to the same catchment the two methods often yield quite different results. The aim of this paper is to contribute to understanding the reasons for these differences. A case study is performed for 10 alpine catchments in Tyrol, Austria, where the 100-year floods are estimated by (a) flood frequency statistics and (b) an event based runoff model. To identify the sources of the differences of the two methods, the 100-year floods are also estimated by (c) Monte Carlo simulations using a continuous runoff model. The results show that, in most catchments, the event based model gives larger flood estimates than flood frequency statistics. The reasons for the differences depend on the catchment characteristics and different rainfall inputs that were applied. For catchments with a high storage capacity the Monte Carlo simulations indicate a step change in the flood frequency curve when a storage threshold is exceeded which is not captured by flood frequency statistics. Flood frequency statistics therefore tends to underestimate the floods in these catchments. For catchments with a low storage capacity or significant surface runoff, no step change occurs, but in three catchments the design storms used were larger than those read from the IDF (intensity duration frequency) curve leading to an overestimation of the design floods. Finally, also the correct representation of flood dominating runoff components was shown to influence design flood results. Geologic information on the catchments was essential for identifying the reasons for the mismatch of the flood estimates.

1. Introduction

The estimation of design floods has been at the heart of hydrological research since its beginnings. Design floods are the basis for building flood protection measures and performing integrated flood management to protect people's lives and property. In the last two decades major flood events have further raised the awareness of national and international authorities to the importance of reducing flood risks. Estimating design floods accurately is crucial for all these tasks. To this purpose the hydrological literature proposes a number of different methods including: (a) statistical methods, (b) deterministic methods and (c) methods that are a combination of the two.

(a) Statistical methods estimate the design flood by fitting a flood frequency distribution to observed flood peaks. If a long flood record is available this approach is commonly applied and has been established as the "standard approach" to flood frequency analysis (Klemeš, 1993). In case no flood records are available the flood frequency distribution can be obtained by using different regionalization methods (see an overview in Blöschl et al. (in press)). The greatest limitation of the statistical approach is the need for representative flood records. Large errors can occur when short records are extrapolated to estimate low probability floods (Katz et al., 2002; Klemeš, 1993).
(b) Alternatively, deterministic methods can be applied. The classical deterministic approach is the design storm method (ARR, 1987; ASCE, 1996; DVWK, 1999; FEH, 1999) that consists of selecting a design storm from the IDF (intensity duration frequency) curve of rainfall with a given duration and using it as an input to an event based runoff model to estimate the design flood hydrograph. The duration of the storm is varied and the storm that gives the largest flood peak is considered to be the representative storm. The main advantage of the approach is that it is reasonably easy to implement and catchment processes can be taken into account. Yet, the method simplifies the physical process of rainfall

* Corresponding author. Address: Vienna University of Technology, Karlsplatz 13/222-2, A-1040 Vienna, Austria. Tel.: +43 1 58801 22327; fax: +43 1 58801 22399.
E-mail address: rogger@hydro.tuwien.ac.at (M. Rogger).

0022-1694/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2012.05.068

runoff transformation and is based on three critical assumptions: the choice of the design rainfall hyetograph (shape and duration); the equality between the rainfall and discharge return periods; and the selection of soil moisture conditions before the storm event (Camici et al., 2011). It has been argued that the assumption of equality of return period is generally not valid (Pilgrim and Cordery, 1975; Viglione et al., 2009) and that the proper selection of antecedent conditions and the design storm is of paramount importance if the probabilities are considered to be equal (Packman and Kidd, 1980; Viglione et al., 2009).

(c) Other approaches are the ones that combine statistical and deterministic methods which are generally referred to as derived flood frequency approaches. First proposed by Eagleson (1972) it combines the probability density function of rainfall with a basin response function to obtain the flood frequency distribution. Due to increasing computational power the approach was further developed and is now often applied by coupling long stochastically generated rainfall series with continuous rainfall runoff models to generate a long discharge series that can be used for flood frequency estimation (Blazkova and Beven, 2002, 2009; Brath et al., 2002; Sivapalan et al., 2005; Faulkner and Wass, 2005). In contrast to the event based design storm method these continuous simulations have the advantage that no assumption on the return period of the design rainfall, its duration and intensity and the antecedent soil moisture have to be made (Boughton and Droop, 2003; Koutsoyiannis, 1994). A disadvantage for practical applications may be long computational times and the complexity of the stochastic precipitation model.

Which of the above listed methods is applied to a specific design problem most often depends on the data availability and whether catchments have been modified (FEH, 1999). For the choice of an appropriate method the modeller has to keep the strength and weaknesses of the different methods in mind. It is though unclear, how and to what extent the assumptions in the approaches translate into differences in the flood estimates. So far there are only very few studies that have tried to compare the different methods with each other. In an Australian case study Boughton and Hill (1997) compared estimates from a continuous approach and flood frequency statistics for a 108 km^2 catchment in Victoria. Their results emphasised the importance of using long flood records for flood frequency statistics since they showed that statistical estimates might yield low values if only a part of the observed record was used. This result was also confirmed in a second case study where Boughton et al. (2002) compared design flood estimates by a continuous simulation system with estimates from flood frequency analysis and a design storm approach in three small catchments (62 km^2, 108 km^2, 259 km^2). The use of very short flood records led to flood estimates that were up to 50% smaller than the results from the other two methods. For large return periods the results from the rainfall runoff methods that were based on long rainfall observations were deemed to be more reliable. In another Australian case study McKerchar and Macky (2001) compared design flood estimates from flood frequency analysis of six catchments to estimates from regional flood frequency analysis and a design storm approach. They concluded that design flood estimates generated by design storm methods often tend to be too large with differences of more than 100% compared to other estimates. Similarly Gutknecht et al. (2006) concluded in an Austrian case study that design floods from the design storm approach yield larger results than estimates from flood frequency statistics and regional methods for very low probability floods (return period of 5000-years). They proposed a multi-pillar approach to design flood estimation suggesting that the comparison of different methods can help reduce uncertainties in the flood estimates. In a British case study Calver et al. (2009) compared the statistical and design storm approaches, of the Flood Estimation Handbook (FEH, 1999), to a continuous estimation approach. They performed the study on 107 gauged British catchments treating them as ungauged in order to be able to evaluate the performance. All methods had errors of up to ∓35% for flood estimates with return periods between 2 and 50 years. For the FEH approaches underestimation of peaks was slightly more common, while for the continuous approach no obvious bias towards over- or underestimation could be identified.

Although these studies have compared estimates from different methods they have focused on quantifying the magnitude of difference in the estimates rather than trying to identify the *reasons* for the differences. The aim of this paper therefore is to contribute to a better understanding of the reasons for the differences in the estimates. A case study is performed for 10 alpine catchments in Tyrol, Austria, where the 100-year design floods are estimated by (a) flood frequency statistics and (b) a design storm approach (event based model). To identify the sources of the differences of the two methods, the 100-year floods are also estimated by (c) a continuous approach (Monte Carlo simulations). Field surveys were performed in all catchments to facilitate the parameter choice of the runoff models.

Sections 2 and 3 give information on the characteristics of the case study catchments as well as on the field surveys. Section 4 presents the results of the event model used in the design storm method. Section 5 outlines the method of the Monte Carlo simulations using the continuous runoff model and Section 6 gives the results of the flood frequency statistics. Section 7 compares the methods and Section 8 presents the conclusions.

2. Study area

The 10 catchments of this study are located in the Alps of Tyrol in Western Austria and range in size from 4 to 98 km^2 (Fig. 1 and Table 1). All catchments are gauged with a runoff record of at least 20 years. For all catchments rainfall and temperature records from stations within or in the vicinity of the catchments were available. The mean precipitation ranges from 870 to 1850 mm per year and the mean annual discharge from 0.2 to 3.9 m^3/s. The catchments are sparsely populated with large areas of rock and talus material in the upper altitudes followed by alpine vegetation influenced by pasturing in the middle altitudes and large forest and meadow areas in the lower altitudes (see Fig. 2 for examples). Extensive field surveys were performed to understand the hydrogeologic and hydrologic characteristics of the catchments.

3. Field surveys

3.1. Hydrogeologic processes

The hydrogeologic assessment in this study was based on a classification method that combines hydrogeologic and geomorphologic information (Pirkl et al., 2000; Kohl et al., 2008). The method distinguishes five hydrogeologic runoff process classes: deep groundwater flow, shallow groundwater flow, interflow, overland flow on rocks, glaciers or saturated areas and karstic areas (Fig. 3).

In a first step, geologic maps, hydrogeologic maps, maps on unconsolidated sediments, orthophotos and a digital elevation model were examined. The orthophotos were used to identify the river network including non-permanent micro channels and saturated areas. A dense river network or net of micro channels

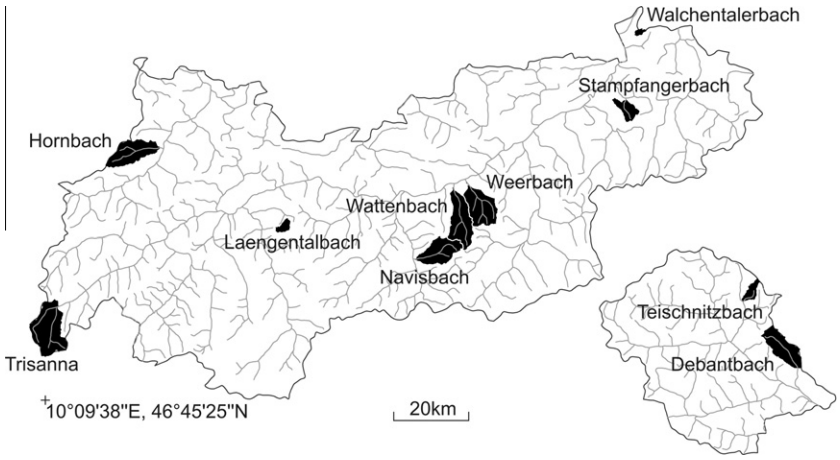


Fig. 1. Catchments in Tyrol used in this study.

Table 1
Characteristics of the catchments used in this study (mean annual precipitation and discharge estimated from at least 20 years of data, mean runoff coefficient c_r from rule-based method of Markart et al. (2004) see Section 3.2).

Stream	Catch. size (km^2)	Max. altitude (m a.s.l.)	Min. altitude (m a.s.l.)	Mean annual prec. (mm/y)	Mean annual discharge (m^3/s)	Mean slopes (°)	Mean length of hillslopes (m)	Mean runoff coefficient c_r
Trisanna	98	3295	1522	1280	1.45	27	723	0.49
Wattenbach	73	2722	537	1150	2.05	27	701	0.26
Weerbach	73	2565	534	1240	2.25	26	558	0.25
Hornbach	64	2480	971	1845	3.91	38	588	0.41
Navisbach	62	2745	1101	1082	1.63	28	472	0.27
Debantbach	57	3065	1037	869	1.94	32	654	0.29
Stampfangerbach	21	1679	633	1230	0.56	22	466	0.36
Teischnitzbach	14	3551	1650	1090	0.51	33	675	0.47
Längentalbach	9	1905	2954	1100	0.36	35	437	0.42
Walchentaler Bach	4	692	1238	1550	0.18	17	475	0.21

and saturated areas point towards areas where mainly surface runoff takes place. Furthermore the orthophotos and digital terrain model were used to delineate areas with deep creeping. These areas can be identified by the form of hillslopes and micromorphologic details such as trenches parallel to the slope and small terraces in the slope. They are a result of the pressure relief on hillslopes after the ice of the last ice ages melted off. Some of them are still in movement (in a depth from 10 to more than 150 m) with very slow velocities of a few cm/year, while others are not moving anymore. Areas with deep creeping are mainly characterised by deep groundwater flow.

The geologic and hydrogeologic maps were used to support the analysis since they provide information on the distribution and composition of bedrock, which gives an idea about the rocks tendency to weather and form fissures. This can for instance be an important source for identifying karstic areas. The maps on unconsolidated sediments provide information on the lithology which hints at the hydrogeologic conditions such as the grain size distribution and conductivity of the sediments. Based on this first evaluation, preliminary runoff process maps were compiled.

The preliminary maps were then refined and corrected in extensive field trips to each catchment. In the field trips the river network was checked and geomorphologic structures such as moraines were verified. Additionally, runoff measurements at selected points in the river network were performed in order to roughly estimate the specific discharge of the area to see whether it coincides with what one would expect given the chosen hydrogeologic runoff process class.

Fig. 4a shows a section of the orthophoto of the Wattenbach catchment with the identified river network and surface edges of deep creeping areas and the corresponding hydrogeologic runoff process map (Fig. 4b). The hydrogeologic classification provides qualitative information for the rainfall runoff models, in particular about the storage capacities, the dominant processes and the depth at which runoff processes mainly take place.

3.2. Surface runoff processes

The surface runoff processes were estimated in the field by the rule-based method of Markart et al. (2004) which is based on more than 700 artificial rainfall experiments in the Austrian, German and Italian Alps. In this approach vegetation, soil characteristics and land use information are used as indicators for event runoff coefficients and retardation coefficients for surface runoff or more specifically overland flow. The event runoff coefficient (r_c) is defined here as the ratio of event rainfall and overland flow at a rainfall intensity of 100 mm/h. The retardation coefficient (c_r) is the one proposed by Zeller (1981) used for estimating the time of concentration on the hillslope (see Appendix Eq. (A.1)).

For the identification of the event runoff coefficient the results of the artificial rainfall experiments were used to distinguish seven main hydrologic vegetation/land cover classes such as bare soil/soil with pioneer species, grassland/hay meadow and forests. These classes are divided into subclasses which differentiate different types of forests or grassland species. The subclasses generally occur under certain typical conditions which are defined by providing complementary information for each subclass such as soil texture (six classes from fine to coarse soil), typical land use forms (e.g. pasturing, skiing), distinct features (e.g. erosion properties) and the predominant soil moisture content. In order to easily identify

Fig. 2. (a) Hornbach catchment and (b) Wattenbach catchment.

the predominant soil moisture content in the field the rule-based approach includes a number of indicator plants (Fig. 5) that hint at different soil moisture conditions from dry to wet. The blueberry is for instance a typical indicator of dry to slightly moist soil conditions, while the king cup only occurs on wet soils.

Table 2 gives an example of how the rule based approach can be used to identify the runoff coefficients in the field. First a hydrological vegetation class is identified for the area of interest. Assume we deal with a dwarf shrub vegetation area that is mainly covered with alpine rose. Table 2 shows the rule-based information for the subclass alpine rose with the complementary information that describes the typical situations of occurrence. Assume the soil on the area of interest is coarse. Coarse soil is always an indicator for low runoff coefficients and drier areas. This can be confirmed by indicator plants such as the blueberry. Coarse soils with low moisture content do not provide sufficient grass cover to be preferred pastures. Therefore usually none or only a low pasturing activity occurs. Assume no pasturing activity takes place on the area of interest. This then results in a runoff coefficient between 0% and 10% (Table 2 bolded cells).

The retardation coefficient is identified similarly to the runoff coefficient but by a more simplified scheme that only takes vegetation/land cover classes and land use properties into account. The scheme is shown in Table 3. According to this scheme the example of the alpine rose has a retardation coefficient between 0.10 and 0.12.

For all 10 pilot catchments the rule-based approach was used to estimate the runoff and retardation coefficients. To simplify the field work, preliminary maps were prepared where areas belonging to similar vegetation/land cover classes were delineated using orthophotos and land use information. The maps were then refined in field trips to each pilot catchment. Fig. 6 shows the maps of the runoff coefficients and retardation coefficients for the Hornbach catchment.

4. Event based approach (design storm approach)

4.1. The Zemokost model

The Zemokost model (Kohl, 2011) is an event based rainfall runoff model commonly used for design flood estimation in small alpine catchments in Tyrol. The design flood is determined by applying the model in a design storm procedure. In this approach a design storm with a certain return period is chosen from the IDF curve of rainfall (from the Austrian standard design tables: Weilguni (2009)) and used as an input to the event model. The duration of the storm is varied and the storm that gives the largest flood peak is considered to be the representative storm. The design flood for the entire catchment is obtained by routing the discharge contributions from the subcatchments to the catchment outlet along a simplified river network. The model accounts for overland flow and interflow.

In each subcatchment surface runoff generation and concentration is described by a combination of the rational method (Kuichling, 1889) and a linear storage model (Nash, 1958). The estimated hydrograph is characterised by three values: the peak flow, the time of initial delay and the time of concentration. The peak flow is calculated with the rational method using the average mapped runoff coefficients from field surveys (Fig. 6). The mapped runoff

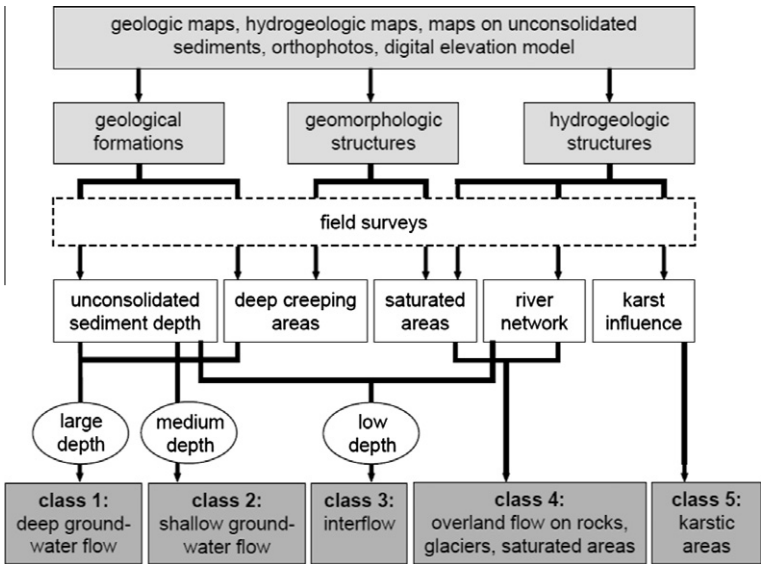


Fig. 3. Schematic of hydrogeologic classification of runoff processes on the basis of geologic, geomorphologic and hydrogeologic information.

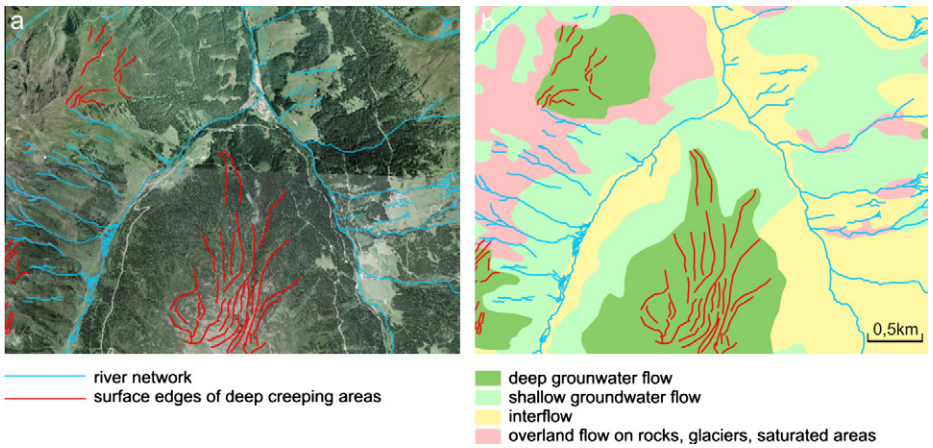
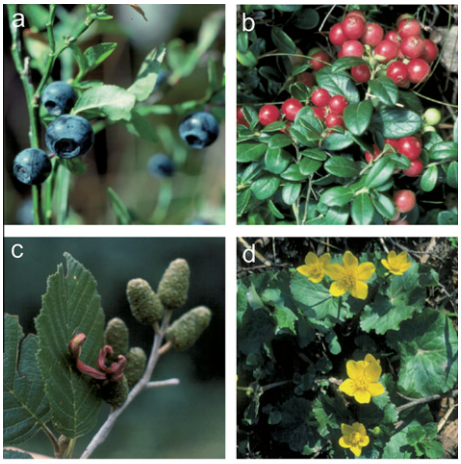


Fig. 4. Detail of Wattenbach catchment: (a) orthophoto with hydrogeologic and geomorphologic analysis and (b) hydrogeologic runoff process map (there are no karst areas in this catchment).



indicator plants	soil moisture
a) blueberry *(Vaccinium myrtillus)*	dry - slightly moist
b) lingonberry *(Vaccinium vitis-idaea)*	dry - slightly moist
c) gray alder *(Alnus incana)*	moist-wet
d) kingcup *(Calta palustris)*	wet

Fig. 5. Examples for indicator plants showing soil moisture listed in Markart et al. (2004).

coefficients refer to at a rainfall intensity of 100 mm/h, but are reduced for smaller rainfall intensities according to an empirical relationship between rainfall intensity and runoff coefficient. This relationship was estimated in field experiments were different sites were irrigated with rainfall intensities between 10 and 120 mm/h and runoff generation was observed (Kohl and Markart, 2002). The runoff coefficient partitions the design storm input into overland flow and subsurface runoff. Surface runoff generation is assumed to be delayed during the time of initial delay since the soil has to get saturated first before overland flow evolves. The design storm falling on the catchment during the time of initial delay represents the amount of water needed to reach soil saturation. The initial delay is defined as average delay observed during a number of artificial rainfall experiments (see Section 4.2). The time of concentration of each subcatchment is estimated by an empirical relationship proposed by Zeller (1981; see Appendix Eq. (A.1)) that accounts for hillslope length, gradient and flow retardation. As a retardation coefficient, the average of the mapped retardation coefficients of the subcatchments is used (Fig. 6). The estimated peak flow, time of initial delay and time of concentration describe a very simplified trapezoidal hydrograph. This hydrograph is transformed into a more realistic surface runoff response by passing it through a linear storage with the time of concentration as storage coefficient. This gives the surface flow hydrograph for each subcatchment.

The part of the design storm that does not contribute to overland flow can become interflow in the model. The information from the hydrogeologic assessment (see Section 4.2) is used to determine the fraction of areas on which interflow occurs in each subcatchment. By multiplying the remaining design storm depth with these areal fractions, the interflow contribution of the subcatchment is obtained. The rest of the design storm that does not contribute to overland flow or interflow is considered deep groundwater flow that does not contribute to the event hydrograph. The hydrograph of the interflow is estimated by the same

Table 2
Rule-based approach of Markart et al. (2004) to obtain event runoff coefficients r_c for the dwarf shrub vegetation subclass alpine rose. Bolded cells relate to example in text.

Vegetation/ land cover class	Vegetation/ land cover subclass	Soil texture	Soil moisture content	Land use properties/distinct features	r_c (%)
Dwarf shrub vegetation	**Alpine rose**	**Coarse soil, loose**	**Dry-slightly moist**	**None**	**0–10**
				Low pasturing activity, small areas with erosion	11–30
		Coarse soil with fines	Moist	Alpine rose with sphagnum moss in areas with snow accumulation	31–50
		Fine soil	Dry-slightly moist	Mat grass, medium pasturing influence and erosion on areas up to 25%	31–75
		Fine soil, cohesive	Slightly moist-moist	Saturated areas with rill network	51–75
			Wet	Water logging with dense rill network	>75

Table 3
Estimation of the retardation coefficient (c_r) (Markart et al. (2004). Bolded cells relate to example in text.

Vegetation/land cover class	Vegetation/land cover subclass	c_r
Areas without vegetation	Asphalt, concrete, rock, ice	0.00–0.02
Bare soil	Low occurrence of initial vegetation	0.00–0.02
	High occurrence of initial vegetation	0.02–0.04
Grassland	Mat grass	0.00–0.02
	Skiing	0.02–0.04
	Hay meadow	0.04–0.06
Moist areas	Low occurrence of moss	0.04–0.06
	High occurrence of moss	0.06–0.08
Bush vegetation	Common heather	0.06–0.08
	Alder, birch, blueberry	0.08–0.10
	Alpine rose	**0.10–0.12**
Forest	Bare soil as forest floor	0.00–0.02
	Plant litter on forest floor	0.02–0.04
	Grass on forest floor	0.04–0.06
	Moss on forest floor	0.06–0.08
	Low occurrence of dwarf shrubs on forest floor	0.08–0.10
	High occurrence of dwarf shrubs on forest floor	0.10–0.12

Table 4
Retardation coefficient c_{rint} in dependence of unconsolidated sediment conductivity (Kohl (2011), conductivities adapted from DIN 18130-1 (1998)).

Unconsolidated sediment texture classes estimated in field	(conductivity [m/s])	c_{rint}
Very strong permeability	($> 10^{-2}$)	0.1
Very strong to strong permeability	(5.5×10^{-3})	0.6
Strong permeability	(10^{-3})	1
Strong to medium permeability	(5.5×10^{-4})	6
Medium permeability	(10^{-5})	100
Low permeability	(10^{-7})	10,000
Very low permeability	($>10^{-8}$)	100,000

concept as surface flow. It is assumed that near surface runoff behaves similarly to surface runoff as it mainly takes place in macropores and pipes from plant roots and is hence different from the subsurface matrixflow in lower layers. In this case no initial delay is accounted for as initial delay is already considered in surface runoff generation and interflow is assumed only to take place in the uppermost part of the soil. The time of concentration of the interflow is determined in a similar way as the surface flow velocity, taking hillslope length, gradient and interflow retardation into account (Appendix Eq. (A.2)). The interflow retardation coefficient can only be estimated very roughly during field visits based on unconsolidated sediment texture. The sediments encountered in the field are assigned to one of the unconsolidated sediment texture classes listed in Table 4. For these classes mean hydraulic conductivities (adapted from DIN 18130-1 (1998) from laboratory experiments) are available that are used as indicative values for the interflow retardation coefficient. The retardation coefficients c_{rint} listed in Table 4 were estimated by trial and error simulating subsurface hydrographs and analysing whether the delay of the hydrograph peak was consistent with what you would expect by the given conductivity. This way, based on the sediment classification in the field, a first relative value for the retardation coefficient c_{rint} can be obtained. Usually this value is then adapted during the plausibility check (Section 4.3). The trapezoidal hydrograph obtained for the interflow is also passed through a linear storage with the time of concentration of the interflow as storage coefficient in order to determine the final interflow hydrograph of the subcatchment.

For each subcatchment surface flow and interflow are summed up and form the event hydrograph of the subcatchment. The subcatchment discharges are routed to the catchment outlet using the empirical flood routing formula proposed by Rickenmann (1996), Appendix Eq. (A.3) which is particularly suited for pool-riffle channels and was estimated based on 400 runoff measurements in the Swiss Alps. The formula estimates flow celerity as a function of channel slope, characteristic $d90$ retardation of the channel bed (90% quantile of particle size distribution) and mean of discharge (mean of discharge at subcatchment outlet and discharge contributing from upper subcatchments at subcatchment inlet). The $d90$ retardation of the channel bed was roughly estimated during the field trips.

4.2. Parameter estimation and input data

The main advantage of the model is that the model parameters such as the runoff coefficient and retardation coefficient can be determined in field surveys by the method of Markart et al. (2004). Most of the remaining parameters are geometric values (e.g. slope and length of hillslope) that can be derived from a digital elevation model.

An important parameter in the design storm procedure is the choice of the initial soil moisture conditions. In the Zemokost model the initial soil moisture conditions are represented indirectly by the time of initial delay of surface runoff generation. The time of delay is estimated depending on the runoff coefficient of the subcatchment area. A relationship between the runoff coefficient and mean initial time of delay was determined by Kohl (2011) based on the outcome of more than 250 artificial rainfall experiments were different alpine vegetation plots were irrigated with an intensity of 100 mm/h and the initial time of delay was measured. The experiments were carried out during the spring and summer season where the largest floods occur. These measurements also include the initial soil moisture conditions since they represent the soil moisture conditions encountered in the

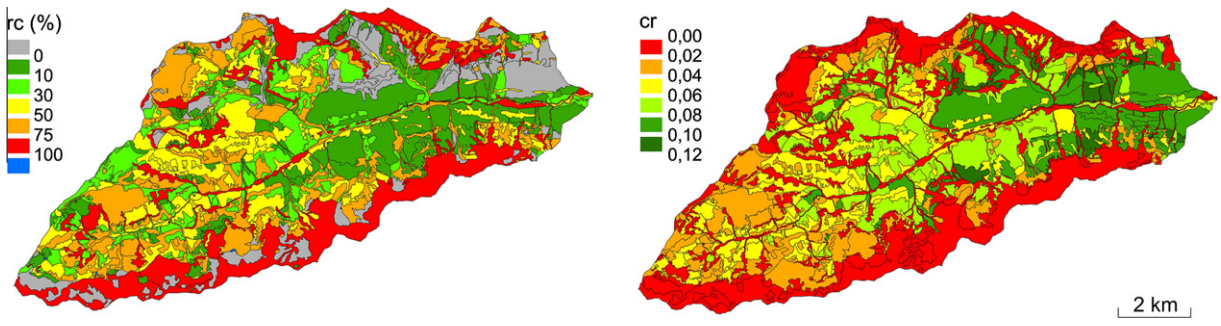
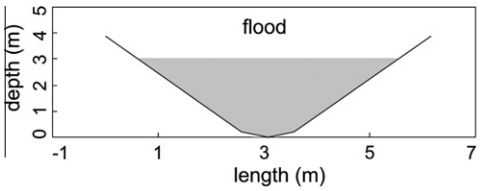


Fig. 6. Maps of runoff coefficients (r_c) and retardation coefficients (c_r) of the Hornbach catchment.



	Q (m^3/s)	v (m/s)
field estimation	4.0	1.0
model	4.9	1.1

Fig. 7. Hornbach catchment: examples of model plausibility check by comparing modelled 100-year flood discharge to largest occurred flood identified by flood marks in the field.

field which varied from very dry to very wet during the experiments. By choosing the mean initial delay for each runoff coefficient class as input to the model also the mean soil moisture conditions are accounted for.

The fraction of the subcatchment area where interflow occurs in the model was estimated using the hydrogeologic runoff process maps (Fig. 4). This fraction includes all areas that belong to the hydrogeologic runoff process class interflow or overland flow on rocks where runoff coefficients are larger than 10% and the distance from the river network is less than 200 m.

As a rainfall input, block rainfall with constant intensity with the return period of interest was used. While time patterns are often used in the design storm method to account for temporal rainfall variability within an event, the Tyrolean procedure uses block rainfall. 100-year rainfall was taken from the Austrian standard design tables (Weilguni, 2009) which are based on a combination of statistical analyses of raingauge data (Ökostra, 1992) with results from an atmospheric model (Lorenz and Skoda, 2000). The latter, typically, give larger values than the statistics. Weilguni (2009) argues that the larger values are more realistic as the raingauges may under-represent convective storms and also involve some gauge catch deficit. While the exact magnitude of the T-year storms may be disputed, they are the values that are used in Austria by consensus. The spatial distribution of rainfall was accounted for by assuming that the rainfall centre occurs in the upper part of the catchment and drapers off according to an empirical areal reduction factor (Blöschl, 2009). The duration of the design storm was varied from 5 min to 6 days and the largest flood peak calculated was considered to be the 100-year flood on the basis of the design storm method.

4.3. Plausibility check of design flood estimates

The Zemokost model was developed for the use in ungauged catchments. Extensive plausibility checks were therefore performed. In case simulation results were not plausible the model parameters were adapted to obtain a better representation of catchment processes.

A first hydrologic assessment was carried out during the field surveys. Specifically, the characteristics of the stream channels and of the catchment were compared with other channels and catchments that had been modelled in previous studies on the basis of comparative hydrology (Falkenmark and Chapman, 1989; Gaál et al., in press). For example, mossy stream channels without traces of floods suggested little hydrologic activity, so flow velocities and sediment activity were likely less than in incised channels where scour was visible. Similarly, erosion rills on the hillslopes suggested overland flow and hence larger specific floods than on hillslopes where this was not the case (Merz and Blöschl, 2008a). Additionally, interviews with local residents were carried out during the field visits to get a qualitative understanding of the magnitude and the characteristics of past floods. Also, a description of the major past (historic) flood events of each catchment was provided by the Torrent Control services for eight of the 10 pilot catchments (Trisanna, Wattenbach, Weerbach, Hornbach, Navisbach, Debantbach, Stampfangerbach, Walchentalerbach). Discharge estimates of the historic floods were often not available but the description of the events and related damages gave an impression of the possible magnitude of flood events in the catchments. This was used to assess the order of magnitude of the modelled 100-year flood. As another plausibility check the modelled floods were compared with the highest flood events in neighbouring catchments and with regionalisation results.

Fig. 7 illustrates an assessment of the model results conducted for several cross sections in each catchment. During field trips the largest occurred flood that was identifiable by flood marks was determined at chosen cross sections. Also the velocity of that flood event was roughly estimated. These values were then compared to the modelled 100-year flood event. Obviously the values can be quite different, but the idea of the assessment was to understand whether the magnitude of estimated design flood was similar to large occurred flood events.

While this plausibility check was considered important, in this study discharge measurements were also available. The model was therefore also compared to observed runoff for the two biggest events in each catchment. Rainfall was approximated by a block rainfall since block rainfall is part of the design procedure. Fig. 8 shows the 2005 flood event in the Hornbach catchment as an example. Given that the model was not calibrated, the simulation matches the observed hydrograph quite well. For most of the events of the 10 pilot catchments Nash–Sutcliffe coefficients of 0.90 were obtained. Only in a few cases the simulated events had lower Nash–Sutcliffe coefficients because the recession or base

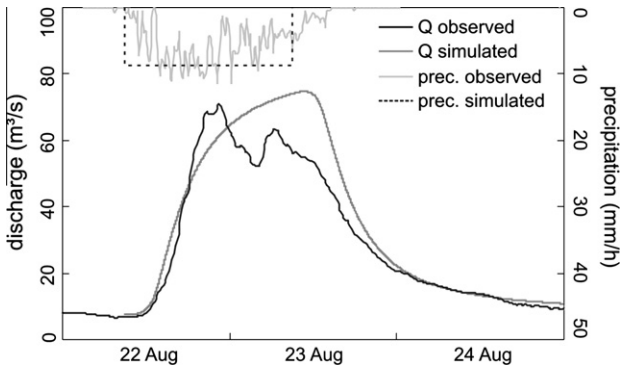


Fig. 8. Hornbach catchment: simulated and observed flood event in August 2005.

Table 5
100-year design floods estimated by the event based approach, Monte Carlo approach and flood frequency statistics (Gumbel distribution incl. confidence bounds).

Catchment	100-year flood (m^3/s)			Flood frequency statistics		
	Event based approach	Monte Carlo approach	Flood frequency statistics	5% conf. bound (m^3/s)	95% conf. bound (m^3/s)	Years of data
Trisanna	101	78	72	60	94	25
Wattenbach	54	43	28	24	36	22
Weerbach	61	50	32	27	40	22
Hornbach	153	109	85	75	104	26
Navisbach	75	51	28	24	35	25
Debantbach	70	63	47	39	62	21
Stampfangerbach	82	32	21	17	28	19
Teischnitzbach	26	17	10	9	13	25
Längentalbach	17	8	6	5	7	24
Walchentaler Bach	13	16	12	10	16	22

flow after the events was overestimated, but in all cases the flood peak was simulated well.

During the plausibility check, the interflow retardation coefficients c_{rint} and the $d90$ retardation of the channel bed were changed in some catchments. The final 100-year floods estimated with the event model after performing the plausibility check are listed in Table 5.

5. Continuous Monte Carlo approach

5.1. The soil moisture accounting model

In the second approach the design floods were estimated by a derived flood frequency approach that combines a stochastic rainfall model with a continuous rainfall runoff model. The main advantages over the event based approach are that antecedent soil moisture can be simulated and the return period is clearly defined. The runoff model is a continuous spatially distributed water balance model (Blöschl et al., 2008a). It consists of a snow routine, a soil moisture routine and a flow routing routine. The snow routine represents snow accumulation and snow melt by a simple degree-day concept that divides precipitation into snow and rainfall and accounts for snowmelt. Rainfall and snowmelt are partitioned into a component that increases soil moisture and a component that contributes to runoff by a nonlinear function depending on the maximum soil moisture storage. Soil moisture can only decrease by evapotranspiration which is estimated from potential evapotranspiration and air temperature. Runoff routing on the hillslope is represented by an upper zone and two lower reservoirs. Rainfall and snowmelt that contribute to runoff enter the upper zone reservoir and leave this reservoir through three paths: percolation to the lower reservoirs defined by a percolation rate, outflow from the reservoir with a fast storage coefficient that represents interflow and, additionally, when a defined threshold is exceeded, outflow through a further outlet with a very fast storage coefficient that represents surface or near surface runoff. Percolation into the two lower reservoirs is split into two components by a defined percentage. The two lower reservoirs represent groundwater and deep groundwater flow. Runoff is calculated on a pixel as the sum of the outflows from all reservoirs and aggregated to subcatchments. Subcatchment runoff is routed through the stream network by a cascade of linear reservoirs. The model has been tested in several Austrian catchments (Blöschl et al., 2008b; Komma et al., 2008). A detailed description of the model is given in Blöschl et al. (2008a).

5.2. Input data, parameter estimation and model validation

The model requires rainfall, air temperature and potential evapotranspiration as inputs. For all catchments at least 20 years of rainfall and temperature data were available at a resolution of 15 min. Catchment rainfall was assumed to be equal to the rainfall measured at the raingauges in order to avoid any smoothing effects of interpolation since the main interest was in simulating extremes. Air temperatures were interpolated from 20 surrounding stations accounting for elevation by a regional regression. Potential evapotranspiration (also called reference crop evapotranspiration) was estimated by the modified Blaney Criddle method (DVWK, 1996) as a function of air temperature. In order to represent the fast runoff processes in the catchments a temporal resolution of 15 min was chosen, and the spatial resolution was set to 200 m $\times$ 200 m.

The model parameters were first set on the basis of the information obtained from the field surveys. For the a priori choice of the of the upper zone reservoir and soil storage parameters the catchments were subdivided into six classes of surface runoff response based on the runoff coefficients from the field surveys (Fig. 6), orthophotos and land use information. An example is shown in Fig. 9. Areas that belong to the low runoff – forest response unit, for instance, represent low runoff production, so high storage coefficients were chosen. Very fast response units such as very fast runoff – rock, on the other hand, have a low storage capacity and fast runoff response, so small storage coefficients were chosen. The a priori parameters for the lower zone and deep groundwater storage were estimated using the hydrogeologic runoff process maps (Fig. 4). Areas with deep groundwater flow were assigned high percolation rates into the groundwater reservoirs and high contributions to the deep groundwater reservoir while areas with predominantly surface runoff were set to have very low percolation to the groundwater reservoirs and no contribution to the deep groundwater reservoir so that mainly overland flow can occur on those pixels.

The parameters were fine-tuned by comparing simulated and observed runoff on a seasonal and an event scale. For most catchments the runoff simulations had Nash–Sutcliffe coefficients around 0.80 (calibration and validation period) with a good representation of the seasonal runoff dynamics and water balance. Only for the Stampfangerbach, Debantbach and Trisanna catchment lower coefficients were obtained. In the case of the Stampfangerbach catchment the measured temperatures in the winter and spring were often too low so that base flow during winter and snow melt were not simulated well, but the model performed very well during summer time where the largest floods occur. For the Debantbach catchment on the other hand the measured rainfall input and measured discharge did not match for some big events since the rainfall station was located outside the catchment. For the Trisanna catchment base flow conditions were hard to simulate since water is diverted from the river for power generation. A detailed analysis of the strategy of parameter selection and model performance for two of the catchments, the Weerbach and Wattenbach, can be found in Rogger et al. (in press).

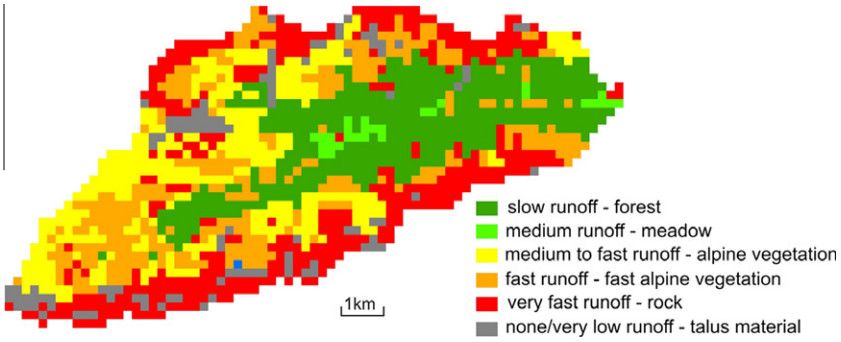


Fig. 9. Hornbach catchment with subdivision of the catchment into surface runoff response units.

5.3. Monte Carlo simulations

The calibrated continuous rainfall runoff model was used for each catchment for running 10,000 years of Monte Carlo runoff simulations at a resolution of 15 min. As input for these simulations precipitation was generated by the stochastic point rainfall model of Sivapalan et al. (2005) calibrated to the rainfall data from stations in or in the vicinity of the catchments. Details on calibration and validation of the precipitation model can be found in Viglione et al. (2012). For the Monte Carlo simulations the observed air temperatures (20 years or more) were used repeatedly to cover the entire simulation period. In general a joint generation of rainfall and air temperature is useful, provided the model closely resembles the joint correlation structure of the site in the context of a given weather pattern. In the present context, no joint generation was needed as the major storms (and hence the major floods) are driven by rainfall only, so snow melt does not play an important role, and the occurrence of particular air temperatures during large storms is not very relevant. From the stochastic runoff simulations the maximum annual floods were isolated and subjected to frequency analysis. For all catchments, the simulated flood statistics were consistent with the observed flood statistics. However, the Monte Carlo simulations provide additional insight into the shape of the flood frequency curves and the simulated rainfall characteristics. The Monte Carlo estimates of the 100-year floods for the 10 catchments are listed in Table 5.

6. Flood frequency statistics

For all 10 catchments, the 100-year floods were finally estimated from the flood peak data by flood frequency statistics. A Gumbel distribution, which cannot be rejected by the Anderson–Darling goodness-of-fit test at 5% significance in 9 out of 10 cases (Laio et al., 2009), was used to fit the maximum annual peak data and parameters and uncertainty bounds were estimated by maximum likelihood and through a Bayesian MCMC (Markov Chain Monte Carlo) methods respectively (Viglione et al., in preparation). Since the flood records were not very long (between 20 and 30 years), the 5% and 95% confidence bounds show considerable scatter (Table 5).

7. Comparison of methods and discussion

The estimates from the three methods were now compared to shed light on the differences between the event based results and the flood frequency statistics (Fig. 10). In some catchments, as for the Trisanna and the Debantbach catchment, the flood estimates are very similar, while in other catchments, such as the Stampfangerbach, the differences are larger than 100%. The value of the Monte Carlo approach now is to assist in understanding (i) whether the shape of the flood frequency curve assumed in the flood frequency statistics is appropriate, (ii) whether the magnitude of the design storms of the event model are appropriate and (iii) whether there are differences in the assumed runoff components. For the assessment the results have been split into four groups with group (a) being the group with similar estimation results while groups (b–d) relate to the different reasons for the discrepancies between the event model and the flood frequency statistics (Fig. 10).

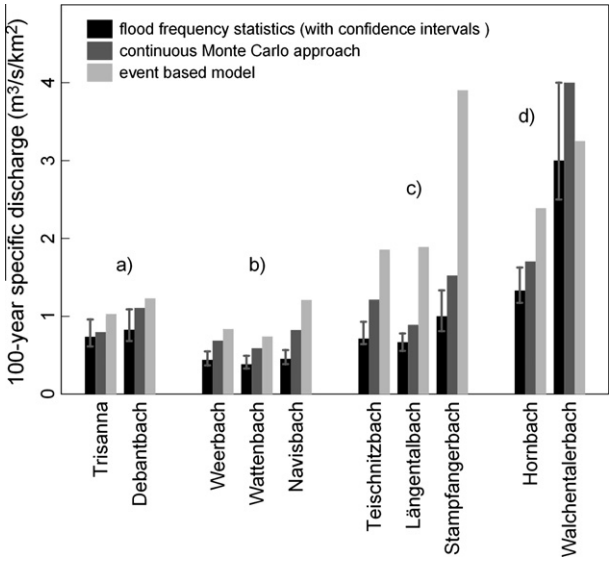


Fig. 10. Comparison of 100-year specific discharge estimated by the three methods. The catchments have been grouped into (a–d) according to their similarities or main reasons for discrepancies (see main text).

(a) *Trisanna and Debantbach – similar flood estimates:* For the Trisanna and Debantbach catchments the design flood estimates from the three methods (Table 5) are very similar with relative differences smaller than 50%. The estimates from the event based approach are larger than the flood frequency estimates, but, in contrast to the other catchments, lie very close to the confidence intervals of the statistical estimates. The input data were consistent (in contrast to group c)) and no large differences in the model assumptions occurred (in contrast to group (b) and (d)). Which means that hydrological processes were represented in a similar way in both rainfall runoff models. For these two catchments, the three methods lend credence to each other which greatly facilitates the choice of a design flood value.

(b) *Weerbach, Wattenbach and Navisbach – shape of the flood frequency curve:* For the Weerbach, Wattenbach and Navisbach catchments (Fig. 10b) the event model estimates are much larger than

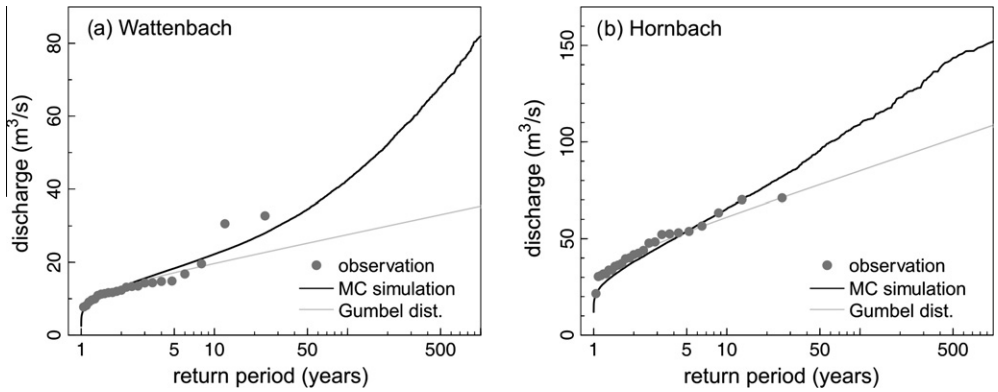


Fig. 11. Results of the Monte-Carlo (MC) simulations compared to maximum annual observed floods and a fitted Gumbel distribution for (a) the Wattenbach and (b) the Hornbach catchments.

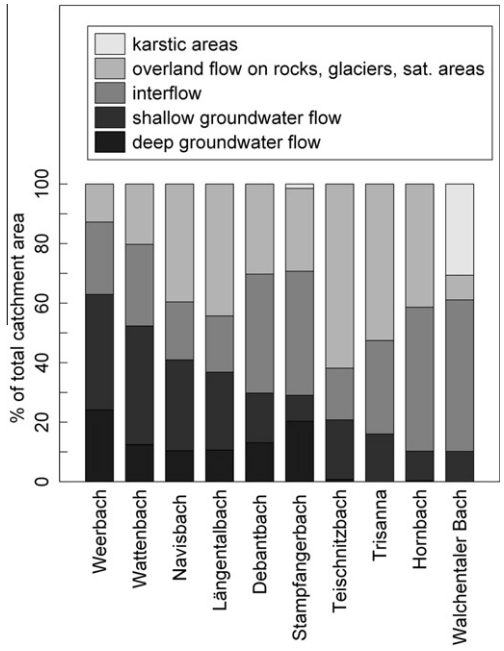


Fig. 12. Areal contributions of hydrogeologic runoff process classes (as in Fig. 4b) in the catchments. Catchments are ordered by decreasing groundwater contribution.

Table 6
Ratio of the 100-year flood peak of the Monte Carlo (MC) approach and the 100-year flood peak of from flood frequency statistics (Gumbel distribution).

Catchment	100-year flood peak ratio
Navisbach (step change)	1.82
Weerbach (step change)	1.56
Wattenbach (step change)	1.54
Teischnitzbach	1.70
Stampfangerbach	1.52
Debantbach	1.34
Längentalbach	1.33
Walchentalerbach	1.33
Hornbach	1.28
Trisanna	1.08

the flood frequency estimates with relative differences larger than 90% (Table 5). Also, the estimates of the Monte Carlo approach are similarly larger and are closer to the event model estimates. To understand these differences, the shapes of the simulated flood frequency curves were examined.

Fig. 11 shows the flood frequency curves obtained from the continuous Monte Carlo simulations for the Wattenbach and the Hornbach catchment. For comparison the plotting positions (Weibull) of the observed annual maxima and the fitted Gumbel distribution is also shown. For the Wattenbach catchment a sudden change in the slope of the flood frequency curve takes place at a return period of around 30 years. Rogger et al. (in press) suggested that such a step change occurs when a threshold of storage capacity is exceeded. For events associated with smaller return periods, much of the rainfall infiltrates and does not contribute to event runoff. However, once the threshold is exceeded fast surface runoff is produced in large parts of the catchment. This changeover occurs for return periods of around 30 years. At much larger return periods the flood frequency curves flatten out. The flood frequency curves of the Weerbach and Navisbach show similar shapes. Since detailed hydrogeologic information was available in all catchments, the interpretation for the Wattenbach can be extended to all catchments. Fig. 12 shows the areal contribution of the hydrogeologic runoff process classes for all catchments. Areas where mainly groundwater flow takes place are also those with high storage capacities. The figure shows that the Weerbach, Wattenbach and Navisbach catchments are those with the highest storage capacities. These are also those where a change in slope of the flood frequency curve was identified. In contrast, the flood frequency curve of the Hornbach catchment, for example, (Fig. 11b) does not show a change in slope. This is consistent with the hydrogeologic process classes in the catchments. While, in the Wattenbach, the two groundwater flow classes cover 52% of the area, these are only 10% in the Hornbach catchment.

The analyses suggest that in the Weerbach, Wattenbach and Navisbach catchments the main reason for the differences between the event model and the flood frequency statistics are the large storage capacities of the catchments along with the inability of the flood frequency approach to account for their effects. If the step change occurs at a return period of around 30 years and around 30 years of flood data are available, as is the case in the three catchments examined here, it will not be apparent in the data. Flood frequency statistics will therefore not account for it and underestimate the 100-year flood. The underestimation can be very significant as illustrated here, more than 82% for the Navisbach, and 56% and 54% for the Weerbach and Wattenbach, with the estimated design floods being by far larger than the confidence intervals of the statistical estimates. The findings confirm the criticism of many authors that the use of too short data records can lead to larger errors (Katz et al., 2002; Klemeš, 1993) as shown also in the case study by Boughton and Hill (1997) and Boughton et al. (2002). This is particularly also the case when, as in the case of the

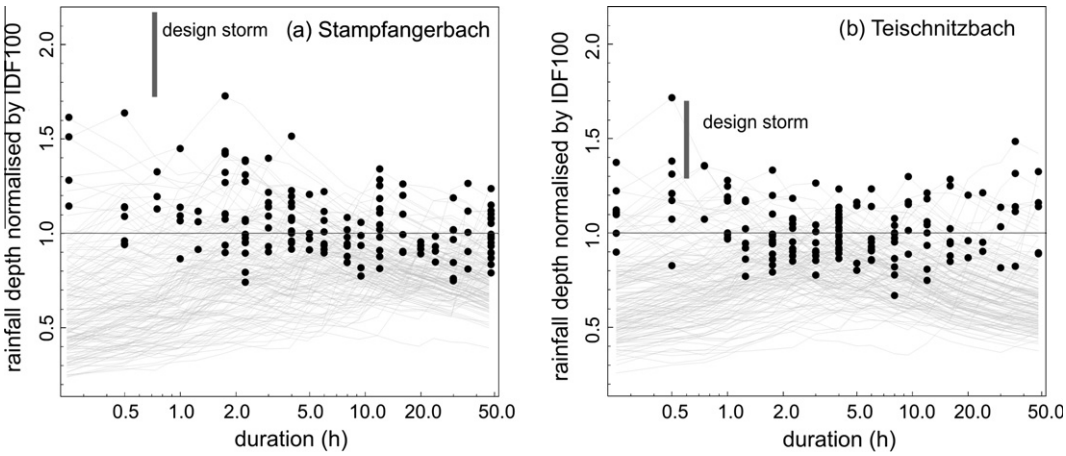


Fig. 13. Maximum rainfall intensities of storms that produce 100-year floods in the continuous Monte Carlo simulations (grey lines show events intensities and black dots mark the maximum intensity during the event) and rainfall depths used in the design storm approach (bars indicate the range between mean and maximum rainfall intensity of the catchment). All values are normalised by the 100-year rainfall depths read off the IDF curves from local rainfall data. Duration refers to the aggregation time of rainfall. (a) Stampfangerbach and (b) Teischnitzbach catchment.

presented catchments, the flood records contain none or only very few large events.

Besides the effect of the step change process, in general the design flood estimates from the Monte Carlo approach are slightly larger than the statistical estimates. Table 6 shows the ratio in the 100-year flood estimates from the Monte Carlo approach and the 100-year floods from flood frequency statistics. The largest differences occur for catchments where a step change in the flood frequency curve could be identified. But large differences also occur for the Teischnitzbach and Stampfangerbach catchment. This difference can as well be attributed to a better estimation of the storage capacity of the catchments which causes a straighter extrapolation of the observed data in the Monte Carlo approach compared to the fit with a Gumbel distribution in flood frequency statistics. An example for this effect is shown in Fig. 11b for the Hornbach catchment.

(c) *Teischnitzbach, Längentalbach und Stampfangerbach – choice of design storm:* The largest differences between the event model and flood frequency estimates occur in the Teischnitzbach, Längentalbach and Stampfangerbach catchments (Fig. 10c) with relative differences between 160% and 290% (Table 5). However, in contrast to the results of group (b), the event model estimates are significantly larger than the Monte Carlo estimates while the Monte Carlo estimates are similar to the estimates of flood frequency statistics. To understand these differences and to see whether differences in the model rainfall inputs could be identified, the design storms used in the event model were compared to the rainfall input of the Monte Carlo approach and statistical analyses of measured rainfall data.

First, the rainfall input of the continuous Monte Carlo approach was analysed. To this end the maximum annual flood events that produced flood peaks of about 100 years (50–200 years) were selected from the results of the 10,000 years of runoff simulations. These are 150 flood events. For each of these events, simulated rainfall intensities inside a 48 h window before the flood peak were analysed by extracting the maximum intensities corresponding to different aggregation levels (15 min to 24 h). In Fig. 13 the maximum intensities for different aggregation levels of the 150 events are shown as grey lines and the overall maximum intensity of each event is marked by a black dot. All intensities have been normalised by the 100-year rainfall intensities read from the IDF curves for the same duration. The IDF curves were obtained by statistical analyses of measured rainfall data. The scatter of the intensities in Fig. 13 is due to the randomness of rainfall time-patterns and the antecedent soil moisture of the runoff model. An interesting point to identify is the duration where most floods occur. In the Teischnitzbach catchment, a very small catchment with low storage capacity that is partially glaciated, the flood events are accumulated around a duration of 3–4 h which corresponds with the time of concentration of the catchment. For the Stampfangerbach durations around 4 h are important, but also longer rainfall events can trigger large floods since some areas in the catchment have a large storage capacity (see Fig. 12) and only start contributing to the floods events when storages are full.

For comparison the design storm intensities used in the event model are also shown in Fig. 13. The bars represent the range of mean to maximum rainfall intensities in the catchment. Especially important are the maximum rainfall intensities as the design storm contribution to the design flood events is not linear with maximum rainfall intensities having a stronger influence on the calculated flood hydrograph. The duration of the design storm shown in Fig. 13 is equivalent to the critical design storm duration that produces the largest flood. In both catchments the design storm intensities are much higher than the 100-year IDF values. This is because the design storms in Austria are based on a combination of statistical raingauge data analyses and results from an atmospheric model. The design storms therefore often tend to be larger than the statistics themselves. In the case of the Stampfangerbach, Teischnitzbach and Längentalbach catchments the design storms are particularly large compared to the other catchments where the mean and maximum design storms intensities are closer to the 100-year intensities from the IDF curves (Table 7). For the Stampfangerbach catchment this is most probably caused by the fact that the catchment is located at the northern rim of the Alps where the influence of orographic lifts can cause large rainfall intensities in the atmospheric model (Lorenz and Skoda, 2000). In case of the Teischnitzbach catchment, a very small catchment, the values are higher since the critical design storm has a very short duration (only a little larger than half an hour) and the increase in intensity with shorter duration from the atmospheric model is much larger than that of the statistical raingauge data analyses causing the design storm values to be significantly larger than the 100-year IDF values. The same applies for the Längentalbach catchment.

Also, in the design storm method, the storm durations that produce the largest flood peaks are between half an hour and 1 h for

Table 7
First column: ratio of mean catchment 100-year design storm depth and 100-year IDF rainfall depth. In brackets: ratio of maximum 100-year design storm depth in the catchment and 100-year IDF rainfall depth). Second column: ratio of the 100-year flood peak of the event model and the 100-year flood peak of Monte Carlo approach.

Catchment	Design storm ratio: catchment average (local maximum)	100-year flood peak ratio
Stampfangerbach	1.72 (2.17)	2.56
Längentalbach	1.51 (1.61)	2.13
Teischnitzbach	1.28 (1.70)	1.53
Walchentalerbach	1.43 (1.56)	0.81
Trisanna	1.10 (1.69)	1.29
Weerbach	0.95 (1.45)	1.13
Wattenbach	0.78 (1.33)	1.20
Navisbach	0.78 (1.32)	1.47
Hornbach	0.67 (1.34)	1.40
Debantbach	0.64 (0.97)	1.11

the two catchments in Fig. 13, while for the Monte-Carlo simulations the largest peaks (on average) are produced by rain bursts with durations around 3 h for the Teischnitzbach catchment and 5 h for the Stampfangerbach catchment (where even larger durations can also be significant). This difference in the critical rainfall duration can again be explained by the fact that Austrian design storm values have large intensities for short durations compared to the IDF curves. While for the Monte Carlo approach the rainfall input corresponds with the IDF curves of the measured data, the design storm inputs for the Zemokost model have significantly larger intensities than the IDF curves for short durations causing shorter rainfall events to produce the largest floods in the event model.

Table 7 shows the double relation of the mean design storm intensity of the Zemokost model in relation to the 100-year IDF values with the 100-year design flood estimates of the Zemokost model in relation to the design flood estimates of the Monte Carlo approach. In the case of the Stampfangerbach, Längentalbach and Teischnitzbach catchments the mean (and maximum) design storm depths are significantly larger than the 100-year IDF depths. The resulting floods are also significantly larger than the 100-year flood estimates of the Monte Carlo approach (factors between 1.5 and 2.6). In the case of the Walchentalerbach catchment the large difference in the rainfall inputs is not reflected in the design flood estimates. This can be explained by the fact that this is a karstic catchment and flood runoff in the event model is dominated by slower interflow causing a strong attenuation of the design storm in the subsurface (see below point d).

Table 7 also shows that for about half of the catchments the mean design storm intensities (and maximum intensities) are closer to the 100-year IDF values and more similar to the mean rainfall intensity that produces the 100-year design flood in the Monte Carlo approach. In these cases the differences in the design flood estimates cannot mainly be attributed to differences in the rainfall input.

These analyses suggest that in the Teischnitzbach, Längentalbach and Stampfangerbach catchments the main reason for the differences between the event model and the flood frequency statistics are the rainfall inputs. More importantly, the design storms used in Austria are significantly larger than the IDF estimates from raingauge statistics. While the IDF curves may have some biases, e.g. due to raingauge catch deficit, the design storm intensities do seem to be large in the light of the flood estimate comparisons. It appears that the design flood method overestimates the 100-year floods substantially in these catchments and this is mainly due to the choice of the design storm depths. More work is needed to ascertain the error bounds of the rainfall estimates.

The findings in this set of catchments suggest an additional reason why estimates from the design storm approach can be larger than flood frequency estimates. In the three presented catchments with a rather low storage capacity and large areas that contribute to overland flow (Fig. 12) no step change could be identified. In these cases the differences in the estimates were mainly caused by large design storm values which shows that the right choice of design storm in the event model has a crucial impact on the design flood estimates as suggested by a number of authors (Viglione et al., 2009; Packman and Kidd, 1980). A similar behaviour can probably also be observed in catchments with a larger storage capacity.

(d) *Hornbach and Walchentalerbach – runoff components:* The remaining catchments are the Hornbach and the Walchentalerbach catchments (Fig. 10d). The estimates for the Walchentalerbach catchment are similar (Table 5) which is surprising since the analysis of the rainfall input showed that the design storm of the event model is significantly larger than the input of the Monte Carlo approach (Table 7). The results of the Hornbach catchment show a large difference between the event model and flood frequency statistics (about 80%) although no large differences in the rainfall input could be identified (Table 7). It is worth mentioning that the two catchments differ from the other catchments in the study as they are set in a limestone geology that can be influenced by karstic effects. Under these conditions subsurface behaviour is harder to characterise. To understand the differences or unexpected similarities in the results, the runoff components that dominate the design flood event were examined for both catchments.

For the Walchentalerbach catchment the hydrogeologic assessment suggests large areas strongly influenced by karst (~30%, see Fig. 12) as well as interflow and shallow groundwater flow (~60%, see Fig. 12), while only 8% of the catchment area is characterised by fast runoff on rocks. In this case a good characterisation of subsurface processes and runoff components is crucial for estimating design floods well. A comparison of the flood events simulated with the event model and the continuous model showed that runoff components were represented differently in the two approaches. The event model is primarily based on field data so that design floods are mainly composed of interflow as suggested by the hydrogeologic assessment (Fig. 12). In the continuous model, on the contrary, parameter choice is based on field data but also more on runoff observations compared to the event model. Discharge data of the Walchentalerbach catchment suggest that a fast runoff component is present (maybe due to karstic effects). The design flood of the continuous model is hence dominated by faster overland flow. As a consequence the difference in the rainfall inputs is not visible in the flood estimates as the large rainfall input in the event model transforms into strongly attenuated interflow contributing slowly to the design event. In the continuous model on the contrary rainfall contributes much faster to the design event so that flood estimates from both methods are similar despite the differences in rainfall input.

The Hornbach catchment differs in its hydrogeology from the Walchentalerbach catchment as no strongly karstic areas were explicitly identified, but a large amount of areas with overland flow (~40%) are present that can be influenced by karstic effects. In this case the design flood estimates of the event and the continuous model both consist of overland flow and interflow. In hindsight after comparing the results from the different methods and taking the geologic catchment characteristics into account it appears that the overland flow contribution in the event model has been overestimated. This is not a problem when simulating events with long rainfall durations that are dominated by interflow (event 2005, Fig. 8). For the estimation of the design flood though, a short rainfall event (<1 h) with a high intensity was identified as critical

rainfall. In this case, a high contribution of fast overland flow occurs in the event model, although the continuous model suggests a larger interflow contribution and hence slower runoff.

These analyses suggest that appropriate representation of dominating runoff components in the applied event based and continuous rainfall runoff model is important in design flood estimation. In the catchments analysed this is difficult as they are set in a limestone geology where subsurface processes are difficult to identify.

8. Conclusions

The aim of this paper was to contribute to better understanding the differences in the estimates of flood frequency statistics and the design storm method. A case study was performed for 10 alpine catchments in Tyrol, Austria, where the two methods were applied and continuous Monte Carlo simulations were used to identify the sources of the differences of the two methods. Detailed maps on surface and subsurface properties were available from field surveys to facilitate the model parameterization.

The findings of this study show that the reasons for the differences can be related to both the rainfall and catchment characteristics. In this context the information from the hydrogeologic assessment proved to be extremely valuable in assessing the catchments overall runoff behaviour and storage capacities.

The results for the study can be summarised as follows:

- In some of the catchments the flood estimates of the two methods are consistent. In these cases the different methods lend credence to each other which greatly facilitates the choice of a design flood value.
- In study catchments with a high storage capacity, a step change in the slope of the flood frequency curve can occur which is not accounted for in flood frequency statistics if flood records are too short and a smooth distribution function is used. In these cases flood frequency statistics underestimates the design floods with differences up to 80%.
- In study catchments with a low storage capacity or a significant amount of areas that contribute to surface runoff no step change in the flood frequency curve occurs. In these catchments large differences in the flood estimates can be caused by extremely large design storm values. It was shown that in some cases 100-year Austrian design storms (according to the Austrian design storm method) are considerably larger than 100-year IDF values. Due to this difference design floods are overestimated by up to 290% compared to flood frequency estimates.
- Furthermore also a good representation of dominating runoff components in the event model influences design flood estimation. Two cases were presented that are set in a karstic environment, where in one case subsurface flow attenuation is overestimated and in the other case the overland flow component is overestimated causing differences of up to 80% in the flood estimates from the design storm approach compared to flood frequency statistics.

The continuous Monte Carlo simulations proofed to be a very useful tool to understand the reasons for the differences in the two methods since it allows for a detailed description of catchment processes, but also a clear determination of the return period of the flood event of interest.

In general the study suggests that the comparison of different methods for design flood estimation as proposed by Gutknecht et al. (2006) can help to identify uncertainties in the inputs and model assumptions. In case design flood estimates from different methods disagree, the comparison of the results may enhance discussions among modellers and experts which can help clarify ideas about the flood generating processes in the catchment. The study also shows that process based approaches that take catchment information into account, are very valuable as we can infer reasons for mismatches of flood estimates. It is therefore very useful to include as much catchment information as possible. Methods for doing this have been proposed by Merz and Blöschl in their flood frequency hydrology framework (2008a,b).

Acknowledgments

We would like to acknowledge financial support from the Hydrographic Service Tyrol, the Austrian Forest Engineering Service in Torrent and Avalanche Control, Section Tyrol and the AdaptAlp project (Flash Floods Tyrol project, HOWATI) as well as the Austrian Science Funds FWF (Doctoral Programme on Water Resource Systems, DK-plus W1219-N22). We would also like to thank Marco Borga for his useful comments on the paper.

Appendix A. Main flow equations used in the Zemokost model

Time of concentration of overland flow on the hillslope Zeller (1981):

$$t_c = c \cdot c_r \cdot L_h^{\frac{1}{3}} \cdot S_h^{-\frac{1}{3}} (r_c \cdot i)^{-\frac{2}{3}} \quad (A.1)$$

where t_c is the time of concentration of surface flow (s); c is the unit conversion factor (527); c_r is the retardance coefficient of surface flow; L_h is the hillslope length (m); S_h is the hillslope slope (–); r_c is the runoff coefficient (–); i is the mean rainfall intensity of design storm (mm/h).

Time of concentration of interflow:

$$t_{\text{int}} = c \cdot c_{\text{rint}} \cdot L_h^{\frac{1}{3}} \cdot S_h^{-\frac{1}{3}} ((1 - r_c) \cdot i \cdot f_{\text{int}})^{-\frac{2}{3}} \quad (A.2)$$

where t_{int} is the time of concentration of interflow (s); c is the unit conversion factor (527); c_{rint} is the retardance coefficient of interflow; L_h is the hillslope length (m); S_h is the hillslope slope (–); r_c is the runoff coefficient (–); i is the mean rainfall intensity of design storm (mm/h); f_{int} is the fraction of subcatchment area contributing to interflow.

Channel velocity by Rickenmann (1996):

$$c_e = a \cdot g^{0.33} \cdot Q_{mean}^{0.34} \cdot S_c^{0.20} \cdot d_{90}^{-0.35} \quad (A.3)$$

where c_e is the celerity (m^3/s); a is the empirical constant (0,37); g is the gravity (m/s^2); Q_{mean} is the mean discharge from discharge at subcatchment outlet Q_{out} and discharge from upper subcatchments Q_{in} (m^3/s); S_c is the channel slope; $d90$, 90% quantile of grain size distribution of stream bed.

References

ARR, Australian Rainfall and Runoff, 1987. A Guide to Flood Estimation. The Institution of Engineers, Australia.

ASCE, 1996. American Society of Civil Engineers Task Committee on Hydrology Handbook. Hydrology Handbook. ASCE Publications.

Blazkova, S., Beven, K., 2002. Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). Water Resour. Res. 38 (8), 1–14. http://dx.doi.org/10.1029/2001WR000500, 1139.

Blazkova, S., Beven, K., 2009. A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. Water Resour. Res. 45, W00B16. http://dx.doi.org/10.1029/2007WR006726.

Blöschl, G., 2009. Skriptum Ingenieurhydrologie (lecture notes hydrology, lecture nr. 223.027). Institute for Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Austria. <www.hydro.tuwien.ac.at/lehre/lva/downloads.html>.

Blöschl, G., Reszler, C., Komma, J., 2008a. A spatially distributed flash flood forecasting model. Environ. Modell. Softw. 23 (4), 464–478.

Blöschl, G., Reszler, C., Komma, und, J., 2008b. Hydrologische Hochwasservorhersage für den Kamp – Erfahrungen mit den Ereignissen 2006 und 2007 (Hydrological forecasts at the Kamp catchment – experiences

with the flood events 2006 and 2007). Oesterr. Wasser Abfallwirtsch. 60 (3–4), a13–a18.
Blöschl, G., Sivapalan, M., Viglione, A., Wagener, T., Savenije, H., in press. Runoff prediction in ungauged basins Synthesis across processes, places and scales. Cambridge University Press.
Boughton, W., Droop, O., 2003. Continuous simulation for design flood estimation – a review. Environ. Modell. Softw. 18, 309–318.
Boughton, W.C., Hill, P., 1997. A Design Flood Estimation System using Data Generation and a Daily Water Balance Model. Report 97/8 in CRC for Catchment Hydrology, Monash University, Melbourne, Australia.
Boughton, W., Srikanthan, S., Weinmann, E., 2002. Benchmarking a new design flood estimation system. Aust. J. Water Resour. 6 (1), 45–52.
Brath, A., Montanari, A., Moretti, G., 2002. On the use of simulation techniques for the estimation of peak river flows. In: Proceedings of the International Conference on Flood Estimation, International Commission for the Hydrology of the Rhine Basin, Lelystad, Netherlands, pp. 587–599.
Calver, A., Stewart, E., Goodsell, G., 2009. Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation. J. Flood Risk Manage. 2, 24–31. http://dx.doi.org/10.1111/j.1753-318X.2009.01018.x.
Camici, S., Tarpanelli, A., Brocca, L., Melone, F., Moramarco, T., 2011. Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modelling. Water Resour. Res. 47, W05527. http://dx.doi.org/10.1029/2010WR009298.
DIN 18130-1, Deutsches Institut fuer Normung E.V., 1998. Baugrund – Untersuchung von Bodenproben; Bestimmung des Wasserdurchlässigkeitsbeiwerts – Teil 1: Laborversuche (Soil – investigation and testing; Determination of the coefficient of water permeability – Part 1: Laboratory tests). DIN 18130-1, German National Standard, Berlin, Germany.
DVWK, Deutscher Verband für Wasserwirtschaft und Kulturbau, 1996. Ermittlung der Verdunstung von Land – und Wasserflächen (Estimation of evapotranspiration from land and water surfaces). DVWK-Merkblätter, 238, Bonn, Germany.
DVWK, Deutscher Verband für Wasserwirtschaft und Kulturbau, 1999. Hochwasserabflüsse – I. Einsatz von Niederschlag- Abflussmodellen zur Ermittlung von Hochwasserabflüssen (Floods – On the use of rainfall runoff models for flood estimation). DVWK-Schriften, 124, Bonn, Germany.
Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resour. Res. 8 (4), 878–898. http://dx.doi.org/10.1029/WR008i004p00878.
Falkenmark, M., Chapman, T.C., 1989. Comparative Hydrology: An Ecological Approach to Land and Water Resources. UNESCO, Paris.
Faulkner, D., Wass, P., 2005. Flood estimation by continuous simulation in the Don catchment, South Yorkshire, UK. Water Environ. J. 19 (2), 78–84.
FEH, Flood Estimation Handbook, 1999. Institute of Hydrology, Wallingford, UK.
Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., Blöschl, G., in press. Flood time scales – understanding the interplay of climate and catchment processes through comparative hydrology. Water Resour. Res. http://dx.doi.org/10.1029/2011WR011509.
Gutknecht, D., Blöschl, G., Reszler, Ch., Heindl, H., 2006. Ein "Mehr-Standbeine" – Ansatz zur Ermittlung von Bemessungshochwässern kleiner Auftretenswahrscheinlichkeit (A "multi-pillar"-approach to the estimation of low probability design floods). Oesterr. Wasser Abfallwirtsch. 58 (3/4), 44–50.
Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Adv. Water Resour. 25 (8–12), 1287–1304. http://dx.doi.org/10.1016/S0309-1708(02)00056-8.
Klemeš, V., 1993. Probability of extreme hydrometeorological events – a different approach. In: Proceedings of the Yokohama Symposium, Extreme Hydrological Events: Precipitation, Floods and Droughts, Yokohama, Japan, vol. 213. IAHS Publ., pp. 167–176.
Kohl, B., 2011. Das Niederschlags-/Abflussmodell ZEMOKOST (The rainfall runoff model Zemokost). PhD, University of Innsbruck, Austria.
Kohl, B., Markart, G., 2002. Dependence of surface runoff on rain intensity – results of rain simulation experiments. In: Proceedings of the International Conference on Flood Estimation, International Commission for the Hydrology of the Rhine Basin (March 6–8), Bern, Switzerland, pp. 139–146.
Kohl, B., Klebinder, K., Markart, G., Perzl, F., Pirkl, H., Riedl, F., Stepanek, L., 2008. Analyse und Modellierung der Waldwirkung auf das Hochwasserereignis im Paznauntal vom August 2005 (Analysis and modelling of the effect of forest cover on the 2005 flood event in the Paznaun valley). Conf. Proc. 2, 505–516 (Interprävent, Dornbirn).
Komma, J., Blöschl, G., Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. J. Hydrol. 357, 228–242.
Koutsoyiannis, D., 1994. A stochastic disaggregation method for design storm and flood synthesis. J. Hydrol. 156, 193–225. http://dx.doi.org/10.1016/0022-1694(94)90078-7.
Kuichling, E., 1889. The relation between rainfall and the discharge in sewers in populous districts. Trans. Am. Soc. Civ. Eng. 20 (1), 1–56.
Laio, F., Di Baldassarre, G., Montanari, A., 2009. Model selection techniques for the frequency analysis of hydrological extremes. Water Resour. Res. 45, W07416. http://dx.doi.org/10.1029/2007WR006666.
Lorenz, P., Skoda, G., 2000. Bemessungsniederschläge kurzer Dauerstufen ($D \leqslant 12$ Stunden) mit inadäquaten Daten (Design storms of short durations ($D \leqslant 12$ hours) with inadequate data). Mitteilungsblatt des Hydrographischen Dienstes in Österreich, 80, Austria, pp.1–24.
Markart, G., Kohl, B., Sotier, B., Schauer, T., Bunza, G., Stern, R., 2004. Provisorische Geländeanleitung zur Abschätzung des Oberflächenabflussbeiwertes auf alpinen Boden-/Vegetationseinheiten bei konvektiven Starkregen (Version 1.0) (A Simple Code of Practice for Assessment of Surface Runoff Coefficients for Alpine Soil-/Vegetation Units in Torrential Rain (Version 1.0)). Bundesamt und Forschungszentrum für Wald, BFW Dokumentation, 3, Vienna, Austria.
McKerchar, A.I., Macky, G.H., 2001. Comparison of a regional method for estimating design floods with two rainfall-based methods. New Zeal Hydrol Soc – J. Hydrol. (NZ) 40 (2), 129–138.
Merz, R., Blöschl, G., 2008a. Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. Water Resour. Res. 44 (8), W08432. http://dx.doi.org/10.1029/2007WR006744.
Merz, R., Blöschl, G., 2008b. Flood frequency hydrology: 2. Combining data evidence. Water Resour. Res. 44 (8), W08433. http://dx.doi.org/10.1029/2007WR006744.
Nash, J.E., 1958. The Form of the Instantaneous Hydrograph. General Assembly of Toronto, 42 (3). IAHS Publ., pp. 114–118.
Ökostra-93, 1992. Österreichische koordinierte Starkniederschlagsregionalisierung und –auswertung, Heft 3: Eine optimierte Starkniederschlagsauswertung (Analysis and regionalisation of extreme rainfalls in Austria, Vol. 3: an optimised extreme rainfall analysis). Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW), BMLFUW Forschungsbericht, Vienna, Austria.
Packman, J.C., Kidd, C.H.R., 1980. A logical approach to the design storm concept. Water Resour. Res. 16 (6), 994–1000. http://dx.doi.org/10.1029/WR016i006p00994.
Pilgrim, D.H., Cordery, I., 1975. Rainfall temporal patterns for design floods. J. Hydraul. Div. 101 (1), 81–95.
Pirkl, H., Markart, G., Kohl, B., 2000. Von Fachkartierungen zu flächenhaften Prozessdarstellungen – Aggregierungsschritte als Weg (Aggregating process information from field surveys to the catchment scale). International Symposium Interpraevent (June 2000), Villach, Austria. Conf Proc 3, 259–270.
Rickenmann, D., 1996. Fliessgeschwindigkeit in Wildbächen und Gebirgsflüssen (Flow velocities in torrents and alpine rivers). Schweizerischer Wasserwirtschaftsverband, Wasser, Energie, Luft, 88(11/12), Switzerland, pp. 298–304.
Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., Merz, R., Blöschl, G., in press. Step changes in the flood frequency curve – process controls. Water Resour. Res. http://dx.doi.org/10.1029/2011WR011187.
Sivapalan, M., Blöschl, G., Merz, R., Gutknecht, D., 2005. Linking flood frequency to long-term water balance incorporating effects of seasonality. Water Resour. Res. 41 (6), W06012. http://dx.doi.org/10.1029/2004WR003439.
Viglione, A., Merz, R., Blöschl, G., 2009. On the role of the runoff coefficient in the mapping of rainfall to flood return periods. Hydrol. Earth Syst. Sci. 13 (5), 577–593.
Viglione, A., Castellarin, A., Rogger, M., Merz, R., Blöschl, G., 2012. Extreme rainstorms: comparing regional envelope curves to stochastically generated events. Water Resour. Res. 48, W01509. http://dx.doi.org/10.1029/2011WR010515.
Viglione, A., Merz, R., Blöschl, G., in preparation. Flood frequency hydrology 3. Bayesian framework. Water Resour. Res.
Weilguni, V., 2009. Bemessungsniederschläge in Österreich (Design storms in Austria). Wiener Mitteilungen: Hochwasser, Bemessung, Risikoanalyse und Vorhersage, Austria, vol. 216, pp.71–84.
Zeller, J., 1981. Starkniederschläge und ihr Einfluss auf Hochwasserereignisse (Extreme storms and their influence on floods). Eidgenössische Anstalt für das Forstliche Versuchswesen, Switzerland, Report 126, second ed.