Natural Hazards
and Earth System
Sciences

Event generation for probabilistic flood risk modelling: multi-site peak flow dependence model vs. weather-generator-based approach

Benjamin Winter[1,2], Klaus Schneeberger[1,2], Kristian Förster[3], and Sergiy Vorogushyn[4]

[1]Institute of Geography, University of Innsbruck, Innrain 52f, 6020 Innsbruck, Austria
[2]alpS, Grabenweg 68, 6020 Innsbruck, Austria
[3]Institute of Hydrology and Water Resources Management, Leibniz University Hannover,
Appelstr. 9A, 30167 Hannover, Germany
[4]GFZ German Research Centre for Geosciences, Hydrology Section, Telegrafenberg, 14473 Potsdam, Germany

Correspondence: Benjamin Winter (benjamin.winter@uibk.ac.at)

Abstract. Flood risk assessment is an important prerequisite for risk management decisions. To estimate the risk, i.e. the probability of damage, flood damage needs to be either systematically recorded over a long period or modelled for a series of synthetically generated flood events. Since damage records are typically rare, time series of plausible, spatially coherent event precipitation or peak discharges need to be generated to drive the chain of process models. In the present study, synthetic flood events are generated by two different approaches to modelling flood risk in a meso-scale alpine study area (Vorarlberg, Austria). The first approach is based on the semi-conditional multi-variate dependence model applied to discharge series. The second approach relies on the continuous hydrological modelling of synthetic meteorological fields generated by a multi-site weather generator and using an hourly disaggregation scheme. The results of the two approaches are compared in terms of simulated spatial patterns of peak discharges and overall flood risk estimates. It could be demonstrated that both methods are valid approaches for risk assessment with specific advantages and disadvantages. Both methods are superior to the traditional assumption of a uniform return period, where risk is computed by assuming a homogeneous return period (e.g. 100-year flood) across the entire study area.

1 Introduction

In recent decades several large flood events occurred across Europe resulting in direct damage exceeding EUR 1 billion (Kundzewicz et al., 2013). Growing flood damage due to socio-economic and land-use changes as well as a possible increase of flood hazards in a warmer climate (Hoegh-Guldberg et al., 2018) calls for robust flood risk assessment. A reliable estimation of flood damage is an essential prerequisite for profound decision making (de Moel et al., 2015). The most straightforward estimation of possible flood risk would be a statistical evaluation of documented flood damage across the area of interest. In practice, systematic damage records are rare and mostly not available for longer periods (Downton and Pilke, 2005), whereas the major interest, for example in the re-insurance industry, is on losses due to extreme events such as the 200-year return period to fulfil the Solvency II Directive regulations (European Union, 2009).

Following the European flood directive, flood risk is defined as "the combination of the probability of a flood event and of the potential adverse consequences [...]" (European Union, 2007). In other words, flood risk is defined by the probability of damage. Hence, for risk estimation, a flood event, including its probability of occurrence (hazard) on the one hand and the vulnerability of exposed values on the other hand, needs to be considered (Klijn et al., 2015). Since risk assessment is currently not feasible based on empirical data, modelling approaches based on synthetic flood scenarios are

often deployed (e.g. Lamb et al., 2010; Falter et al., 2015; Schneeberger et al., 2019).

In a traditional approach, the hydrological load is estimated by means of extreme-value statistics using river gauge data and transformed into corresponding inundated areas by hydrodynamic models (Teng et al., 2017). The monetary damage can then be assessed in combination with susceptibility functions which describe the relationship between one or more flood hazard characteristics (e.g. inundation depth and flow velocity) and damage for the elements at risk (Merz et al., 2010). This approach implies two strong assumptions. First, the return period of flood discharge is assumed to be equal to the return period of the resulting damage. Second, a uniform return period across the entire study area is considered and resulting damage estimates are accumulated. The first assumption can be relaxed by modelling a continuous series of synthetic flood events. As a result, a long series of damage values can be generated and used for analysing damage frequency distribution (Achleitner et al., 2016). The second assumption of homogeneous flood return periods may be valid for small areas (de Moel et al., 2015). With increasing scale, the assumption of a homogeneous return period becomes unlikely, as precipitation and flood footprints are inhomogeneous in space. This assumption can lead to an overestimation of risk for specific return periods in large river basins (Thieken et al., 2015; Vorogushyn et al., 2018; Metin et al., 2020). To overcome the second limitation, realistic spatially heterogeneous events need to be generated across the area of interest which fully represent the spatial variability of flooding (Schneeberger et al., 2019).

Generation of spatially heterogeneous flood events in terms of precipitation fields or discharges is of current scientific interest (Keef et al., 2013; Falter et al., 2015; Falter, 2016; de Moel et al., 2015; Speight et al., 2017; Diederen et al., 2019; Diederen and Liu, 2019; Schneeberger et al., 2019). There are different approaches to generating large event series of heterogeneous flood events. One possibility is the application of multi-variate statistical methods to discharge series, such as copula models (Jongman et al., 2014; Serinaldi and Kilsby, 2017; Brunner et al., 2019) or the semi-parametric conditional model proposed by Heffernan and Tawn (2004) (hereinafter referred to as "HT-model" or "HTm"). These models consider the pairwise dependence of peak discharges at multiple locations and generate synthetic series of multiple dependent flow peaks. The second possibility is based on the generation of spatially distributed meteorological fields by a weather generator, either station-based with subsequent interpolation (Falter et al., 2016; Falter, 2016; Breinl et al., 2017; Evin et al., 2018; Raynaud et al., 2019) or raster-based (Buishand and Brandsma, 2001; Peleg et al., 2017). Synthetic meteorological fields are subsequently used to drive hydrological simulations to generate streamflow values across the study area.

The two presented approaches estimate the hydrological load in the river network at multiple locations but are different in their nature. This leads to the key question of the present study: does it matter which approach is chosen in the context of flood risk modelling, and what are the advantages and disadvantages of the two? We answer this question by comparing the set of heterogeneous flood events from the HT-model with the one resulting from a weather generator and subsequent rainfall–runoff modelling. Both methods are embedded in a probabilistic flood risk model used to estimate the effect of chosen methods on flood losses. To the best of the authors' knowledge, there is no study to date in which the two approaches are directly compared. Additionally, the flood risk corresponding to homogeneous flood scenarios of certain return periods ("traditional" approach) is derived and compared to the other two approaches.

This paper is organised as follows: first, the study area is shortly described. In Sect. 2 the flood risk model is introduced and the two different approaches for heterogeneous event generation are presented in details. Section 3 presents the results of the comparison, which are discussed in the following section. Finally, conclusions summarise the major findings.

2 Study area

The flood risk model is applied in the westernmost province of Austria, Vorarlberg. The region is characterised by a strong altitudinal gradient between the Rhine River valley (≈ 400 m a.s.l.) and the high mountain ranges of the Alps (> 3000 m a.s.l.). As a result of the high relief energy, the rivers are characterised by a fast hydrological response with short concentration times. The mountainous landscape of in total 2600 km^2 is dominated by forest, meadows and pastures with only small percentage of settlement area (Sauter et al., 2019). Due to steep topography, asset values are concentrated in the lowlands of larger valley floors, especially alongside the Rhine and Ill rivers. Vorarlberg is characterised by one of the highest precipitation amounts in Austria, conditioned by predominantly westerly flows and strong orographic effects (BMLFUW, 2007). During the last decades, the province was affected by several severe flood events in 1999, 2002, 2005 and 2013. The most devastating recent flood event in August 2005 caused about EUR 180 million direct tangible losses for the private and public sector, including infrastructure (Habersack and Krapesch, 2006). Figure 1 provides an overview of the study area, including the river network, settlement areas and locations of river gauging stations as well as meteorological stations.

3 Methods and data

The probabilistic flood risk model (PRAMo) used in the presented work consists of three different modules: the hazard module comprising the generation of long time series of flood events; the vulnerability module used to evaluate pos-
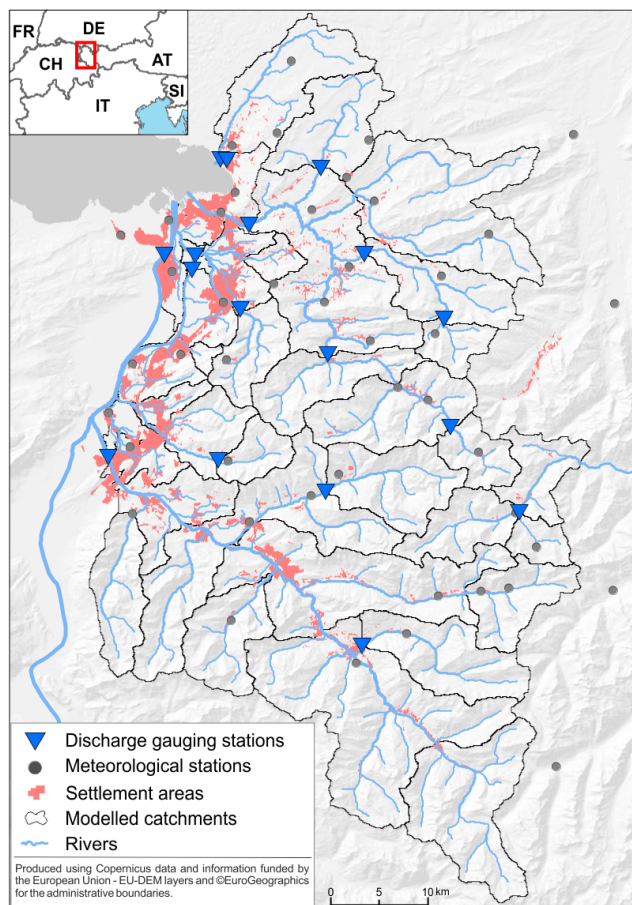
Figure 1. Study area and the location of meteorological and river gauging stations.

sible adverse consequences of flood events with a certain exceedance probability; and the risk assessment module, which combines the results of the hazard and vulnerability modules to estimate the loss per event and resulting risk (Schneeberger et al., 2019). The output of the flood risk model consists of expected annual damage and exceedance probability curves of damage. PRAMo was previously driven by the synthetic flood event series of coherent peak discharges generated by the HT-model (Schneeberger and Steinberger, 2018). A second event generation approach based on a multisite, multi-variate weather generator and continuous rainfall–runoff modelling was recently introduced by Winter et al. (2019) and is used for comparison with the HT-model-based approach and the assumption of homogeneous return periods. Figure 2 provides an overview of the modules and the simulation steps, which are described in more details in the following.

In this study, data of 17 gauging stations (1971–2013) are applied for the HT approach. The continuous simulation of the WeGen approach is based on daily time series from 1971 to 2013 for 45 meteorological stations (cf. Fig. 1). At hourly time steps data for only 23 sites starting from 2001 are avail-

able. Stations without hourly information were interpolated by an inverse distance-weighting (IDW) scheme (for details see Winter et al., 2019).

3.1 Hazard module I: HT-model

The hazard module generates time series of spatially distributed synthetic flood events. In the first approach, we apply the conditional extreme-value model (HT-model) proposed by Heffernan and Tawn (2004) to peak flows. In this approach, flood events are understood as a set of spatially consistent peak discharges at multiple locations of stream gauges. Spatial consistency is ensured by considering the correlation structure of peak flows from the past observation period. Discharge time series at 17 gauges across the study area are used to parameterise the HT-model. In the first step, the observed data are standardised by a marginal model to a Laplace distribution. In the second step, the dependency between the stations is modelled for the case of peak flow at one station being above a certain threshold. According to Lamb et al. (2010), the HT-model can be interpreted as a multi-site peak-over-threshold approach. Due to strong seasonality of streamflow in Vorarlberg, the HT-model is separately parameterised for winter and summer periods (Schneeberger and Steinberger, 2018).

For the set of synthetic flood peaks at each of the 17 gauge locations we estimate the return period based on the generalised extreme value (GEV). A flood event is characterised by exceedance of a certain streamflow at a single location or multiple locations with a defined time period. As a threshold for defining a widespread flood event, a return period of 30 years was selected in the present study. The output of the HT-model in terms of synthetic flood peaks is available at the locations of gauging stations. Hence, for the river segments without observations, the flows and their respective return periods need to be estimated. We apply the top-kriging approach (Skøien et al., 2006) for the spatial interpolation of model results to the entire river network. This method takes into account the nested structure of river catchments, which makes the results more robust compared to traditional regional regression-based approaches (Laaha et al., 2014; Archfield et al., 2013). A more detailed description of the HT-model is provided in Schneeberger and Steinberger (2018) and Schneeberger et al. (2019).

3.2 Hazard module II: WeGen

The second approach is based on a stochastic weather generator used to drive a hydrological model. Long-term daily precipitation and temperature series are generated, with a multi-site, multi-variate weather generator based on the autoregressive model (Hundecha et al., 2009). Daily precipitation amounts are generated from mixed gamma and generalised Pareto distributions fitted to individual weather stations. The mixed distribution is shown to better capture extreme pre-
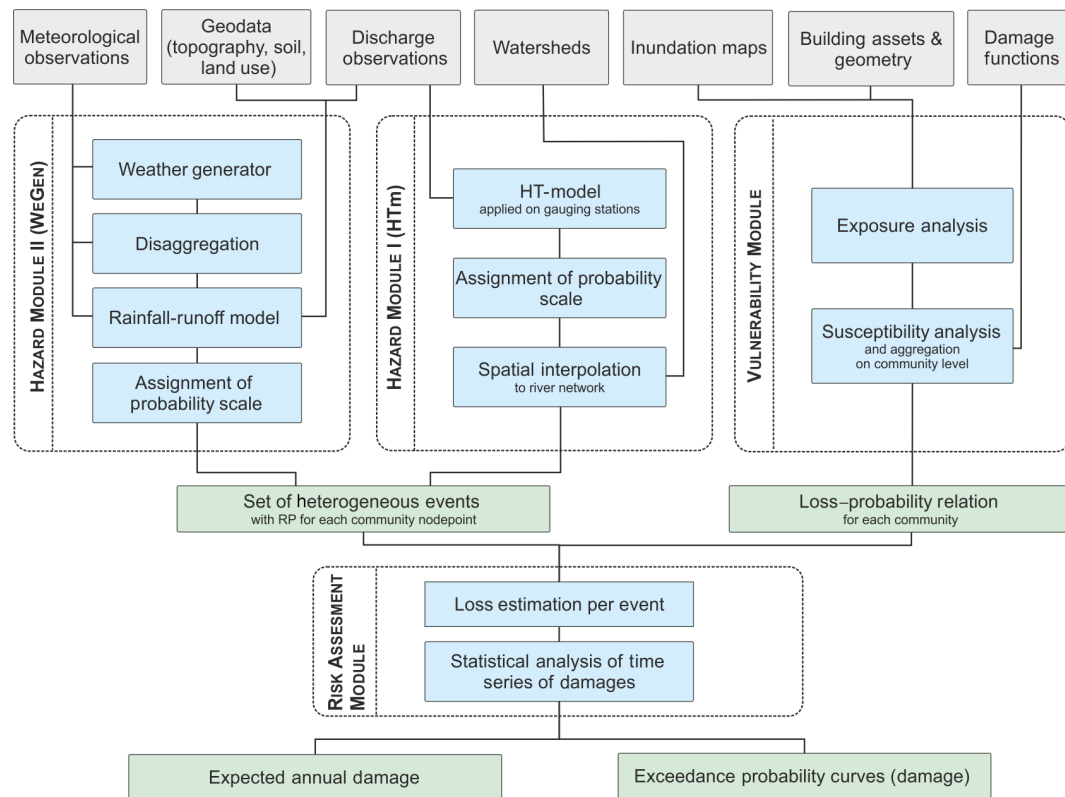
Figure 2. Flowchart of the PRAMo flood risk model including two different approaches for flood event generation.

cipitation while robustly modelling the bulk of precipitation amounts (Vrac and Naveau, 2007). With respect to seasonal patterns, the fitting is applied on a monthly basis. Occurrence and amount of precipitation are modelled considering the autocorrelation and inter-site correlation structure. The mean temperature is then modelled conditioned to the simulated precipitation (Hundecha and Merz, 2012). As the study area is characterised by mostly alpine topography with short catchment response times, the hydrological model needs to be driven by meteorological input at sub-daily resolution to estimate realistic peak flows (e.g. Dastorani et al., 2013). A non-parametric k-nearest-neighbour algorithm based on the method of fragments is applied to disaggregate the generated daily values to hourly time steps (Winter et al., 2019). For a day to disaggregate, the generated daily values of temperature and precipitation from the weather generator are compared against observed daily data at all stations. Subsequently, k-nearest neighbours in terms of lowest Euclidean distances between generated and observed daily values are selected. Next, one matching day is randomly sampled from the selected neighbours, and the corresponding relative temporal patterns from the match day are transferred to the input day (method of fragments). In contrast to the previous study (Winter et al., 2019), a centred moving window of 30 d is applied instead of the identical months in order to restrict the search of possible matching days. The modification increases

the variability between the disaggregated days and reduces the maximum search distance on a temporal scale, especially for days at the beginning and end of a month.

Following the generation of meteorological data at the locations of the weather stations, a spatial interpolation to continuous meteorological fields is necessary for the application of the rainfall–runoff model. Complex methods for spatial interpolation can be applied (e.g. Goovaerts, 2000; Plouffe et al., 2015); however, for the long-term simulation a computationally efficient approach is needed. The interpolation was carried out by a inverse distance-weighting scheme including a stepwise lapse rate to account for the complex topography (Bavay and Egger, 2014).

Finally, the semi-distributed conceptual rainfall–runoff model HQsim is applied to simulate streamflow across all catchments of the study area (Kleindienst, 1996). HQsim is forced by precipitation and temperature data and has previously been used in various studies in alpine catchment areas (e.g. Senfter et al., 2009; Achleitner et al., 2012; Dobler and Pappenberger, 2013; Bellinger, 2015; Winter et al., 2019). A simulated annealing algorithm is used for the model calibration against observed discharge data at the gauging stations (Andrieu et al., 2003). From a long synthetic discharge series, relevant flood events are identified and extracted. For this, a flood frequency analysis at all points of interest based on fitting the GEV distribution using the L moments is car-

ried out. Analogously to the HT-model approach, a threshold of a 30-year return period, at least at one site across the study area is applied to define relevant flood events. A more detailed description of the modelling chain, including the disaggregation procedure, is given in Winter et al. (2019).

3.3 Vulnerability module

While the hazard module computes the hydrological load, the vulnerability module assesses the possible negative consequences in terms of exposed objects and monetary damage. The module is based on the widely used approach of combining the exposure and susceptibility of elements at risk in the inundated areas (Koivumäki et al., 2010; Merz et al., 2010; Huttenlau and Stötter, 2011; Meyer et al., 2013; Cammerer et al., 2013; de Moel et al., 2015; Falter, 2016; Wagenaar et al., 2016). The module calculates losses for each community in the study area for a number of predefined return periods (or probabilities) (i.e. RP = 30, 50, 100, 200 and 300 years). The results of the vulnerability module are loss–probability relations for each community, describing the expected damage for the corresponding return periods. To derive a continuous relation, a linear interpolation between available data points (RP damage) is applied. The loss–probability relations are used as input in the risk assessment module and combined with the simulated return periods (hazard module) at each community to derive risk curves.

At the scale of a community (on average 28 km²), a homogeneous return period of hydrological load is assumed and associated with the total community loss. For the loss calculation we use "official" inundation maps. The inundation maps are based on 1-D hydrodynamic modelling in rural areas and 2-D modelling in urban areas (IAWG, 2010). The boundary conditions for the hydrodynamic simulation are taken from the Austrian flood risk zoning project HORA (Merz et al., 2008).

The estimation of monetary damage for the elements at risk is based on the relative damage functions combined with the total asset values. A damage function describes the relative loss of value as a function of water depth (Merz et al., 2010). If available, additional damage influencing parameters, such as flow velocity or contamination, can be considered for damage assessment (Merz et al., 2013). In accordance with Schneeberger et al. (2019), the one parametric damage model of Borter (1999) is applied in the present study. The damage model was derived for Switzerland, which is a direct neighbour to the Austrian province Vorarlberg with a similar topography and building structure. More precise site-specific damage functions are not available for the study region.

The damage estimation is conducted on a single-object basis for residential buildings only. To derive the flood losses, the available inundation maps are combined with the asset datasets and damage function. Subsequently, the object-based loss data are aggregated for each community. The ab-

solute building values indexed to 2013 according to the construction price index (Statistik Austria, 2019) are derived by calculating mean cubature values from local insurance data and transferred to the entire building stock of the study area (Huttenlau et al., 2015). Since derived values are based on insurance data, they are consequently defined as replacement values.

3.4 Risk assessment module

The risk assessment module brings together the results of the hazard and vulnerability modules to generate a time series of losses and calculates the resulting risk curve for the area of interest (Schneeberger et al., 2019). In order to combine the results, each spatial unit (community) is represented by a defined model node point in the river network. For each generated heterogeneous flood scenario, the recurrence intervals are derived for all model node points (hazard module) and combined with the respective loss–probability relation to compute losses (vulnerability module). By integrating the losses at all model node points, i.e. for each community, the total loss for every generated event can be calculated. By evaluating the overall modelled time series of events, a continuous time series of damage is generated. Finally, the time series of damage can be statistically analysed to derive the expected annual damage (EAD) and to construct risk curves (Schneeberger et al., 2019). More detailed information about the vulnerability and risk assessment module, including a schematic overview of the module interaction, is provided in Schneeberger et al. (2019).

## 3.5 Assessment of spatial coherence of generated events

A core element of the probabilistic flood risk model is the generation of plausible, spatially heterogeneous flood events. To investigate the spatial coherence of synthetic events generated by two different approaches, two spatial dependence measures proposed by Keef et al. (2009) are applied. The first measure $P_{i,j}(p)$ describes the probability that a dependent site i exceeds a certain threshold, given that a conditional site j is exceeding a threshold $q_p(Q_j)$ as well:

$$P_{i,j}(p) = \Pr\left(Q_i > q_p(Q_i) \mid Q_j > q_p(Q_j)\right), \tag{1}$$

where (p) is the level of extremeness (quantile) and Q_i and Q_j are the dependent and conditioned runoff series, respectively. The calculation of the thresholds is based on a 3 d block maximum, which was found to be appropriate in this region (Schneeberger and Steinberger, 2018). The second spatial dependence measure $N_j(p)$ is an overall summary metric and describes the average probability of all dependent sites i to be high, given that the conditional site j is high, defined as

$$N_j(p) = \frac{\sum\limits_{i \neq j} \Pr\left(Q_i > q_p(Q_i) \mid Q_j > q_p(Q_j)\right)}{n-1}. \tag{2}$$
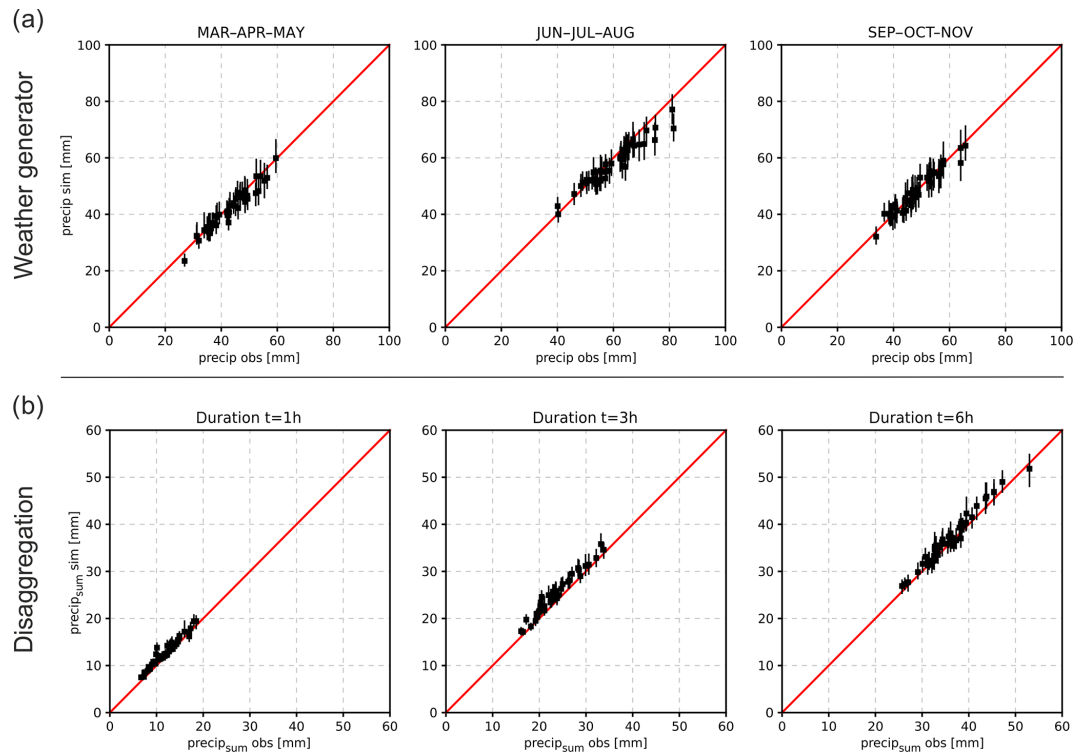
Figure 3. Validation results of the weather generator and the revised disaggregation procedure for all stations ($n = 45$). The bars represent the median and the 5–95 % quantile range of 100 realisations for the weather generator and disaggregation. **(a)** Weather generator: 99 % quantile of daily precipitation for generated data compared with observed data for spring, summer and autumn. **(b)** Disaggregation: 99.9 % quantile of 13 years of disaggregated data is compared to observed data, for the precipitation sum of 1, 3 and 6 h duration.

In the case of the WeGen approach the dependence matrices were computed for the peak discharges at the gauging station locations resulting from the combined simulations of the weather generator and rainfall–runoff model.
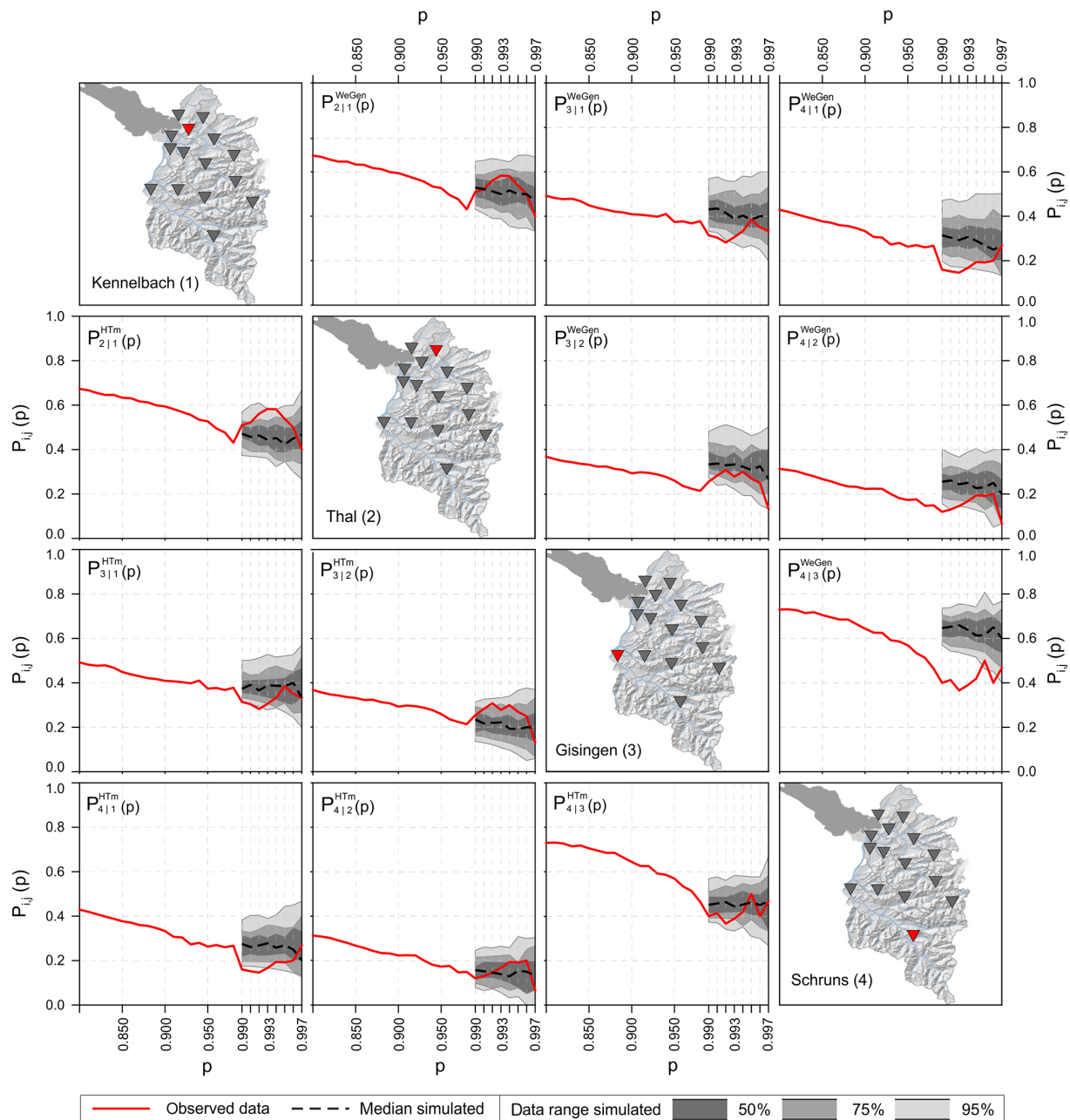
4 Results

4.1 Simulation results of the continuous modelling approach (WeGen)

To assess the performance of the continuous modelling approach, extreme precipitation of simulated data is compared to observed station data (daily: 1971–2013; hourly: 2001–2013) for the weather generator and disaggregation procedure. The median and the uncertainty range represented by the 5 and 95 % quantiles of 100 model realisations are compared to the observed data. Figure 3a shows the results for the 99 % quantile of daily precipitation (wet days) for all 45 station and spring (March–April–May), summer (June–July–August) and autumn (September–October–November). In general, the characteristics of the observed daily precipitation are well reproduced by the weather generator. A few stations, however, show a slight underestimation in summer (mainly June and August). The validation results for all

months separately, including maximum and minimum simulated daily temperatures, are provided by Winter et al. (2019). To validate the disaggregation procedure, the hourly data are first aggregated to daily data and subsequently disaggregated back to hourly time steps. For the comparison of disaggregated precipitation, 99, 99.9 and 99.95 % quantiles are calculated and compared to the observed values. The results for the 99.9 % quantile show a good agreement between observed and simulated precipitation intensities for the three analysed rainfall durations: 1, 3 and 6 h (Fig. 3b). Results for the 99 and 99.95 % quantile are shown in Winter et al. (2019).

The rainfall–runoff model is calibrated (2001–2007) and validated (2008–2013) in a classical split-sample approach (Klemeš, 1986) for all catchments of the study area against observed river gauging data. On average, a Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) of 0.68 and 0.67 and a Kling–Gupta efficiency (KGE; Kling et al., 2012) of 0.75 and 0.74 are achieved for the calibration and validation periods, respectively. Detailed results for the individual catchments, including a comparison of design flood estimates with a flood frequency analysis and a design storm approach, are given in Winter et al. (2019).

Figure 4. Comparison of observed (42 years) and simulated conditioned exceedance probability $P_{i,j}(p)$. The range of the simulated results is based on 42 years of simulation with 100 realisations. The plots in the lower triangle correspond to the HT model, whereas those in the upper triangle show the WeGen results.

4.2 Spatial patterns of generated flood events

For the analysis of spatial coherence, 100 simulations using each of the two event generation approaches (HT-model and WeGen) were carried out. Each simulation comprised 42 years of data corresponding to the length of the observed discharge series. Figure 4 illustrates exemplary results for four gauging stations and both methods. Each plot shows the dependence measure between the two stations depicted on the maps in the principal diagonal. The gauges Kennelbach and Gisingen are the two largest catchments of the study area (about 80 % of the total area). The examples Schruns and Thal are subcatchments of Gisingen and Kennelbach, respectively, and thus represent two strongly related gauge pairs. The measure is calculated for discharge values with exceedance probability between $p = 0.99$ and $p = 0.997$, above which the data are too few ($n < 15$) to calculate a meaningful $P_{i,j}$ value. Based on the empirical distribution
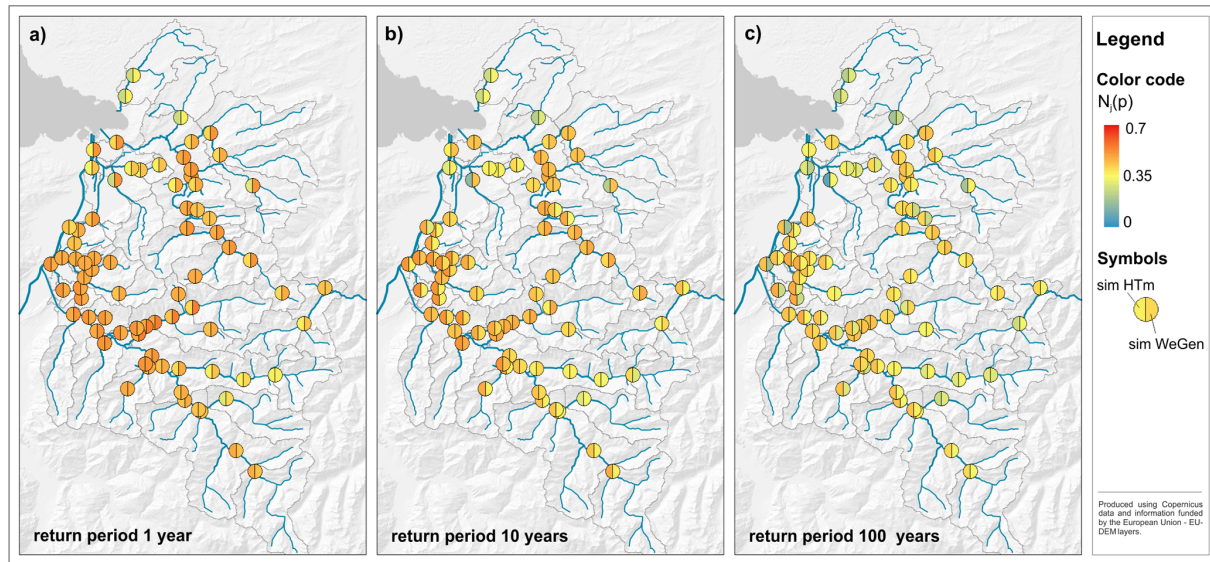
Figure 5. Spatial dependence measure $N_j(p)$ for the community node points at the river network and three different return periods. The results show the median for the HT-model and WeGen approach based on 30 realisations of 1000 years of simulation.

function of the 3 d block maxima series, a p value of 0.99 refers to a return period of approximately 1 year, and a p value of 0.997 refers to a return period of roughly 3 years.

In general, the spatial dependence declines with the level of extremeness. For more extreme runoff situations, the dependence structure is less stable and prone to a large variability. The HT-model results in the lower triangle reproduce the observed spatial patterns between the stations well. The observed measure is in $\approx 90\%$ of the cases inside the simulated data range (2.5–97.5 % quantile). The results of the WeGen approach follow the general observed patterns of lower dependence (e.g. $P_{i,j}(p) \approx 0.2$ for Thal (2) vs. Schruns (4)) and higher dependence (e.g. $P_{i,j}(p) \approx 0.5$ for Kennelbach (1) vs. Thal (2)). However, the results are biased towards a higher dependence, such that only half of the results correspond well to the observed data.

To analyse the dependence structure of high flows across the study area, the measure $N_j(p)$ is calculated for all node points corresponding to different communities. The measure is calculated for p values corresponding to the 1-, 10- and 100-year return period. As the simulation of the two approaches is not limited to the length of the observed data, the results are based on the median of 30 realisations of 1000 years of HT-model and WeGen simulations (Fig. 5). The length is chosen to be far above the highest return period of available homogeneous inundation data (RP300), and the number of 30 realisations is dictated by the computational limitations of the continuous simulation at an hourly time step. Both approaches (see Fig. 5a–c) show a decline of spatial dependence towards higher return periods. The general patterns of lower spatial dependence in the southern part of the study area and of the individual northern catchments

are visible. The node points downstream are characterised by a higher dependence. For a high return period of 100 years (Fig. 5c), the simulated spatial dependence is higher for the HT-model than for the WeGen results in contrast to the findings for the lower return periods. The results are regionally different. Whereas the dependence measure is higher for the HT-model in the western part of the study area, the northeastern catchments show a higher degree of dependence for the WeGen approach.

4.3 Comparison of risk curves

To compare the effect of the two approaches of synthetic event generation on the overall estimated loss, flood risk curves are calculated. Confidence intervals are derived based on 30 realisations of 1000-year simulations. Furthermore, the risk curve based on the assumption of homogeneous return period floods across all catchments is derived based on five inundation maps corresponding to the return periods between 30 and 300 years. The two synthetic event generators result in a comparable range of overall estimated flood risk (see Fig. 6). The WeGen approach systematically overestimates the risk computed by the HT-model. The relative difference between the estimated median values ((WeGen−HTm) / WeGen) is approximately 17.5 %. The uncertainty increases with increasing return period of damage alongside the extrapolation of the input time series. On average 172 damage events are generated per 1000 years of simulation in the WeGen approach, compared to about 167 for the HT-model. Both approaches show a significantly lower damage in comparison to the assumption of homogeneous scenarios for specific return periods. The estimated damage of
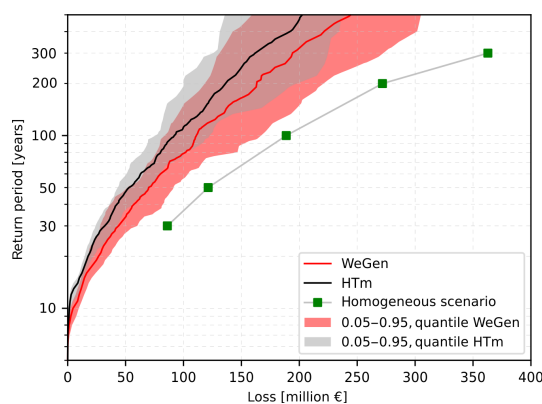
Figure 6. Risk curves for WeGen and HT-model approach in comparison to the results of a homogeneous scenario. The median and quantile confidence intervals are based on 30 realisations of 1000 years of simulation. Monetary values are normalised to the year 2013.

a homogeneous 100-year flood scenario is $\approx 50\,\%$ above the HT-model results and still $40\,\%$ above the WeGen approach.

The sets of generated heterogeneous flood events reflect a large variability of plausible spatial patterns. Hence the estimated flood risk is the result of a combination of these patterns. Figure 7 shows multiple examples of generated flood events corresponding to an estimated damage of EUR 100 ± 1 million for both model approaches. The general severity in terms of flood hazard (without consideration of flood risk) is given by the unit of flood hazard (UoFH). The measure UoFH is a simple proxy of hazard severity defined as the total number of sites at which the threshold of 30 years return period is exceeded (Schneeberger et al., 2019). Even though the selected severity of displayed flood event is rather high, some of the generated events are still spatially limited. The event with the lowest UoFH of 46 corresponds to $\approx 50\,\%$ of all sites exceeding the 30-year threshold. The most widespread event (UoFH $= 77$) corresponds to about $90\,\%$ of the sites exceeding the threshold. This result reflects the spatial distributions of elements at risk with a settlement concentration alongside the larger valley areas in the study area (cf. Fig. 1). Thus, the damage corresponding to an event is largely influenced by the region affected. If the overall comparison is conducted at hazard level only, the impact of widespread flood events may be overestimated, while the impact of spatially limited events in densely populated areas is underestimated.

5 Discussion

Both approaches, the HT-model and the WeGen approach, simulate complex, spatially heterogeneous patterns of synthetic flood events. In the present study, the HT-model outperforms the WeGen approach in terms of reproducing the observed dependence patterns of peak flows at the gaug-

ing stations. The HT-model makes use of the observed river gauging data and models their dependence structure directly. In contrast, the WeGen approach models the dependence structure only indirectly based on the meteorological input data.

The overall river network and especially small ungauged tributaries do however rely on the top-kriging interpolation in the case of the HT-model approach and are not able to react independently to the larger river system. This explains the higher dependence structure on the community node points, while at the river gauges the results do correspond well to the observed values. Nevertheless, in both cases the capability to capture spatial effects of a certain spatial scale in the end depends on the density of the measuring network and its data quality.

The WeGen approach seems to overestimate the overall spatial dependence in the study area in comparison to the observed values. This was also found in a previous study, comparing a different set of gauging stations (Winter et al., 2019). One possible reason could be that the spatial interpolation of the meteorological data by the rather simple IDW approach, without consideration of shading or other effects, and the rather short length of hourly input data for the disaggregation procedure might affect the spatial patterns towards a stronger dependence. More importantly, the WeGen model itself seems to overestimate the dependence between stations particularly for higher return period thresholds. This is in line with the results of the recent evaluation of the weather generator (Ullrich et al., 2019), which suggest an overestimation of correlation of extreme precipitation between individual stations leading to an overestimation of areal rainfall. The correlation structure of the weather generator is fitted on a monthly basis, independently of the rainfall intensities, and thus does mix low-intensity, large-scale rainfalls and small-scale convective events. The simulated stronger spatial dependence in certain areas with high damage potential also contributes to the higher flood risk estimate by the WeGen approach.

Only one possible combination of weather generator, disaggregation procedure and rainfall–runoff model was applied for the WeGen approach. Thus, by the application of an alternative weather generator with different assumptions about the spatial dependence or tail distribution, the resulting risk estimates may change. This counts as well for the application of a different rainfall–runoff model or alternative disaggregation procedure (e.g. Müller-Thomy et al., 2018). Thus, the result of a higher risk estimate for the WeGen approach in comparison to the HT approach can not be generalised to other model combinations.

The two approaches to synthetic event generation differ substantially in terms of estimated damage from the one assuming a uniform return period across the whole study area (Fig. 6). The flood losses for individual return periods above the 30-year threshold under the homogeneous assumption are largely overestimated. This result confirms the necessity to take heterogeneous spatial patterns into account. An event
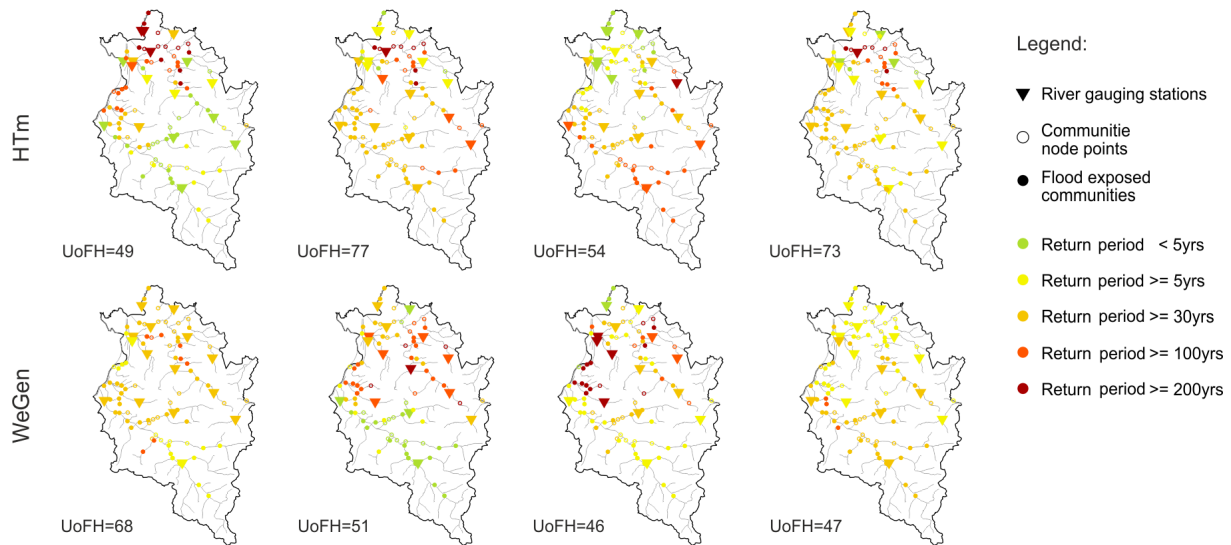
Figure 7. Examples of flood events with an estimated flood damage of EUR 100 ± 1 million flood damage for the HT-model and WeGen approach. The general severity of flood events is characterised by the unit of flood hazard (UoFH).

where every community in the study area is affected by discharges exceeding the 30-year return period during a single event is rare. Based on a total of 30 000 years of simulation, less then 10 % of the communities experience losses simultaneously in more than 50 % of events (Fig. 8). It can be expected that with increasing spatial scale the likelihood that a large number of communities will experience high return period discharges and losses in a single event will decrease (Metin et al., 2020). Therefore, generation of spatially consistent heterogeneous flood events is particularly important with increasing spatial scale. At the same time, considering dependence of meteorological and hydrological variables at multiple locations with increasing scale and an increasing number of dependent locations becomes more challenging.

A fundamental difference between the two approaches resides in the way of considering the hydrological processes. The HT-model takes a purely statistical approach by analysing the dependence of peak discharges above a certain threshold. It does not explicitly consider hydrological processes which generate extremes. For instance, the non-linearity of catchment response is not explicitly taken into account, but only so far it is imprinted in the previously observed peaks used for model parameterisation. The combination of the weather generator and rainfall–runoff modelling describes the hydrological processes in a spatially consistent and time-continuous way. Hence, the effect of soil moisture accumulation and pre-event catchment conditions are explicitly modelled. By the application of a fully distributed, physically based model, the hydrological process description could even be improved, for example, by solving full energy balance equations for snow melt or evapotranspiration (e.g. Förster et al., 2014, 2018). On the downside, a further increase in model complexity might compromise the model
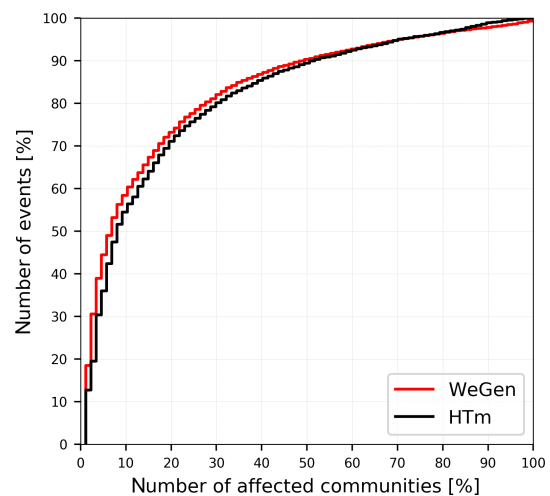
Figure 8. Relative number of flood events exceeding a 30-year flood threshold and corresponding relative number of affected communities. The results are based on 30 000 years of simulation.

parameter identifiability, increase calibration effort and computational burden, and increase input data demand (temperature, precipitation, radiation, humidity and wind speed).

In general, continuous hydrological modelling generates full hydrographs at all locations, which allows for direct coupling with hydraulic models as, for example, applied in Falter et al. (2015) and Falter (2016). The direct coupling of the We-Gen approach with a 1-D to 2-D hydrodynamic model would also allow consideration of hydrodynamic interactions in the river network and their possible effect on the risk estimates. This may, for example, be the reduction of risk downstream due to dike overtopping and failure upstream. In the case of

Table 1. Summary of advantages and disadvantages of the WeGen and HT-model approach to generating heterogeneous flood events.

Categories		HT-model		WeGen
Computational complexity	(+)	Low processing costs (local processing)	(−) (−)	Processing intensive (HPC necessary) Complex data interfaces between different models
Output	(−)	Return periods at all sites for modelled events only	(+)	Continuous hydrographs at all modelled sites
Hydraulic coupling	(−)	Event hydrographs need to be deducted to drive a hydraulic model	(+)	Continuous description of hydraulic boundary conditions allows unsteady hydraulic modelling
Processes	(−)	No information about individual hydrological processes	(+)	Continuous description of hydrological system and modelled processes
Hydrological changes	(−) (+)	No explicit modelling of scenarios (e.g. climate or land-use scenarios) possible Runoff trends can be integrated	(+)	Scenarios can be modelled explicitly (e.g. climate or land-use scenarios)
Transferability	(+)	Model is well transferable to other study areas	(−)	Model chain is transferable; however all components must be set up and calibrated for new study areas

the HT-model, only peak discharge of events is estimated, not the entire hydrograph. Hence, these results cannot be used directly as a boundary condition for unsteady hydraulic simulations. Assumptions on the shape of a hydrograph would be required.

In addition, the continuous modelling approach is capable of explicitly modelling scenarios of changing hydrological boundary conditions. For instance, changes in the climate system can be taken into account in the generation of meteorological fields by conditioning the rainfall and temperature probability distributions (e.g. Hundecha and Merz, 2012). Also possible changes in land use can be considered by parameterising hydrological models accordingly (Rogger et al., 2017). As the HT-model approach is based on observed streamflow only, change scenarios may be included in terms of trends. However, they cannot be modelled explicitly. A continuous simulation approach requires a vast amount of processed data, including multiple data interfaces between the different modelling steps and results in high computational costs. This is especially true if sub-daily simulations are applied that require an additional disaggregation scheme. In contrast, the purely statistical HT-model shows its merit with its efficient data processing, easily applicable on local computers. A further advantage of the HT-model is the transferability of the approach. While each of the modelling steps of the continuous approach, from the weather generator to the hydrological models, needs to be implemented, calibrated and validated for every new study area, the HT-model only needs to be fitted to new discharge time series which is less complex. Different advantages and disadvantages of both approaches are summarised in Table 1.

The presented approaches are subject to different uncertainties. The confidence intervals presented in Fig. 6 are, for example, based on the random processes generating heterogeneous flood events of each method (multiple realisations). However, there are other uncertainties which are not explicitly addressed, for example uncertainties related to the topological kriging of the HT-model results or uncertainties related to the hydrological model in the WeGen approach. Some uncertainties pertain to both methods, such as the choice and fitting of the extreme-value distributions. A comprehensive assessment by propagating the uncertainties of all sub-models throughout the model chain is currently precluded by computational constraints particularly relevant for the WeGen approach.

A further important point, currently not considered in both approaches, is dike failure scenarios. In the study area, for example, no inundation is considered for the Rhine River due to its high protection level. Nonetheless, the probability of a dike failure is non-zero and could have a devastating effect. In this sense, the consideration of flood volumes beside peak estimates could be another important extension to describe the severity of flood events (e.g. Dung et al., 2015; Lamb et al., 2016).

A traditional validation of the overall risk model in terms of a comparison of observed to simulated data is hardly possible as comprehensive databases of loss events are often not available (Thieken et al., 2015). In the present study, damage data based on an insurance portfolio were available for the 2005 event. The data are, however, only a subset of the overall elements at risk, and, due to rather low sublimits (maximum insurance payout), the full losses remain unknown. Finally, without a larger set of loss events it is not possible to

assign a meaningful return period to the 2005 event to val-
idate the risk outcome in a traditional way. Nonetheless, by
applying and comparing different methods, the plausibility
of the results can be checked (Molinari et al., 2019). Fur-
thermore, the uncertainties related to the choice of methods
for generating heterogeneous flood events seem to be lower
in comparison to other aspects of the probabilistic flood risk
model, such as the choice of the applied damage functions
(Winter et al., 2018).

6   Conclusions

The question of whether the choice of method for generating
heterogeneous flood events for flood risk modelling matters
can be answered in different ways. Both approaches, the HT-
model and continuous WeGen approach, were generally ca-
pable of modelling spatially plausible flood events across the
study area. By direct comparison to observed spatial patterns,
the HT-model approach performed better than the WeGen ap-
proach in our study area in terms of correctly representing
the observed dependence structure. A stronger modelled de-
pendence of extreme precipitation resulted in high areal rain-
fall in the WeGen approach and higher overall risk compared
to the HT-model. The median damage from 30 000 years of
simulation is about 17.5 % larger in the WeGen approach
than in the HT-model. The representation of the dependence
structure for simulation of extremes needs to be further im-
proved for the weather generator. Nevertheless, the choice
of method for generating heterogeneous flood events might
have a smaller impact than, for example, the choice of the
applied damage functions (Winter et al., 2018).

To conclude, both methods are valid approaches to over-
coming the simplified assumption of uniform return period
across a study area. Accordingly, when designing a flood risk
study, the choice of the approach should consider the specific
advantages and disadvantages of the two methods and data
availability. If computational efficiency and quick transfer-
ability are in focus, the HT-model approach might be a bet-
ter choice. In contrast, if unsteady hydraulic modelling is re-
quired for the targeted application, the continuous modelling
of generated meteorological fields is more appropriate.

Code and data availability. For Austria, daily meteorological and
river gauging data are available at https://ehyd.gv.at (last access:
28 May 2020) (BMLRT, 2020). The applied meteorological data
for the Deutscher Wetterdienst (DWD) stations are freely avail-
able at https://opendata.dwd.de (last access: 28 May 2020) (DWD,
2020). Underlying loss data are not publicly available. The Me-
teoIO is available at https://models.slf.ch/p/meteoio/ (last accss:
28 May 2020) (Bavay and Egger, 2014). The applied weather gener-
ator and the rainfall–runoff model HQsim are currently not publicly
available.

References

Achleitner, S., Schöber, J., Rinderer, M., Leonhardt, G., Schöberl,
    F., Kirnbauer, R., and Schönlaub, H.: Analyzing the op-
    erational performance of the hydrological models in an
    alpine flood forecasting system, J. Hydrol., 412–413, 90–100,
    https://doi.org/10.1016/j.jhydrol.2011.07.047, 2012.
Achleitner, S., Huttenlau, M., Winter, B., Reiss, J., Plörer, M., and
    Hofer, M.: Temporal development of flood risk considering set-
    tlement dynamics and local flood protection measures on catch-
    ment scale: An Austrian case study, Int. J. River Basin Manage.,
    14, 273–285, https://doi.org/10.1080/15715124.2016.1167061,
    2016.
Andrieu, C., Freitas, N., Doucet, A., and Jordan, M.: An Introduc-
    tion to MCMC for Machine Learning, Mach. Learn., 50, 5–43,
    https://doi.org/10.1023/A:1020281327116, 2003.
Archfield, S. A., Pugliese, A., Castellarin, A., Skøien, J. O., and
    Kiang, J. E.: Topological and canonical kriging for design flood
    prediction in ungauged catchments: An improvement over a
    traditional regional regression approach?, Hydrol. Earth Syst.

Sci., 17, 1575–1588, https://doi.org/10.5194/hess-17-1575-2013, 2013.

Bavay, M. and Egger, T.: MeteoIO 2.4.2: a preprocessing library for meteorological data, Geosci. Model Dev., 7, 3135–3151, https://doi.org/10.5194/gmd-7-3135-2014, 2014.

Bellinger, J.: Uncertainty Analysis of a hydrological model within the Flood Forecasting of the Tyrolean River Inn, Dissertation, University of Innsbruck, Innsbruck, 2015.

BMLFUW (Ed.): Hydrologischer Atlas Österreichs, 3. Lieferung, Wien, 2007.

BMLRT – Bundesministerium für Landwirtschaft, Regionen und Tourismus: eHYD, available at: https://ehyd.gv.at/, last access: 28 May 2020.

Borter, P.: Risikoanalyse bei gravitativen Naturgefahren: Fallbeispiele und Daten, Naturgefahren, Bern, 1999.

Breinl, K., Strasser, U., Bates, P., and Kienberger, S.: A joint modelling framework for daily extremes of river discharge and precipitation in urban areas, J. Flood Risk Manage., 10, 97–114, https://doi.org/10.1111/jfr3.12150, 2017.

Brunner, M. I., Furrer, R., and Favre, A.-C.: Modeling the spatial dependence of floods using the Fisher copula, Hydrol. Earth Syst. Sci., 23, 107–124, https://doi.org/10.5194/hess-23-107-2019, 2019.

Buishand, T. A. and Brandsma, T.: Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling, Water Resour. Res., 11, 2761–2776, https://doi.org/10.1029/2001WR000291, 2001.

Cammerer, H., Thieken, A. H., and Lammel, J.: Adaptability and transferability of flood loss functions in residential areas, Nat. Hazards Earth Syst. Sci., 13, 3063–3081, https://doi.org/10.5194/nhess-13-3063-2013, 2013.

Dastorani, M., Koochi, J., and Darani, H. S.: River instantaneous peak flow estimation using daily flow data and machine-learning-based models, J. Hydroinform., 15, 1089–1098, https://doi.org/10.2166/hydro.2013.245, 2013.

de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., and Ward, P. J.: Flood risk assessments at different spatial scales, Mitig. Adapt. Strat. Global Change, 20, 865–890, https://doi.org/10.1007/s11027-015-9654-z, 2015.

Diederen, D. and Liu, Y.: Dynamic spatio-temporal generation of large-scale synthetic gridded precipitation: with improved spatial coherence of extremes, Stoch. Environ. Res. Risk Assess., https://doi.org/10.1007/s00477-019-01724-9, in press, 2019.

Diederen, D., Liu, Y., Gouldby, B., Diermanse, F., and Vorogushyn, S.: Stochastic generation of spatially coherent river discharge peaks for continental event-based flood risk assessment, Nat. Hazards Earth Syst. Sci., 19, 1041–1053, https://doi.org/10.5194/nhess-19-1041-2019, 2019.

Dobler, C. and Pappenberger, F.: Global sensitivity analyses for a complex hydrological model applied in an Alpine watershed, Hydrol. Process., 27, 3922–3940, https://doi.org/10.1002/hyp.9520, 2013.

Downton, M. W. and Pilke, R. A.: How Accurate are Disaster Loss Data? The Case of U.S. Flood Damage, Nat. Hazards, 35, 211–228, https://doi.org/10.1007/s11069-004-4808-4, 2005.

Dung, N. V., Merz, B., Bárdossy, A., and Apel, H.: Handling uncertainty in bivariate quantile estimation – An application to flood hazard analysis in the Mekong Delta, J. Hydrol., 527, 704–717, https://doi.org/10.1016/j.jhydrol.2015.05.033, 2015.

DWD – Deutscher Wetterdienst: Open Data-Server, available at: https://opendata.dwd.de/, last access: 28 May 2020.

European Union: European Union on the assessment and management of flood risks: Directive 2007/60/EC of the European Parliament and the Council, Official Journal of the European Community, Luxembourg, 2007.

European Union: on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II): Directive 2009/138/EC of the European Parliament and the Council, Official Journal of the European Union, Luxembourg, 2009.

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

Falter, D.: A novel approach for large-scale flood risk assessments: continuous and long-term simulation of the full flood risk chain, Dissertation, University of Potsdam, Potsdam, 2016.

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, J. Hydrol., 524, 182–193, https://doi.org/10.1016/j.jhydrol.2015.02.021, 2015.

Falter, D., Dung, N. V., Vorogushyn, S., Schröter, K., Hundecha, Y., Kreibich, H., Apel, H., Theisselmann, F., and Merz, B.: Continuous, large-scale simulation model for flood risk assessments: proof-of-concept, J. Flood Risk Manage., 9, 3–21, https://doi.org/10.1111/jfr3.12105, 2016.

Förster, K., Meon, G., Marke, T., and Strasser, U.: Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany), Hydrol. Earth Syst. Sci., 18, 4703–4720, https://doi.org/10.5194/hess-18-4703-2014, 2014.

Förster, K., Garvelmann, J., Meißl, G., and Strasser, U.: Modelling forest snow processes with a new version of WaSiM, Hydrolog. Sci. J., 63, 1540–1557, https://doi.org/10.1080/02626667.2018.1518626, 2018.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113–129, https://doi.org/10.1016/S0022-1694(00)00144-X, 2000.

Habersack, H. and Krapesch, G.: Hochwasser 2005 – Ereignisdokumentation: der Bundeswasserbauverwaltung, des Forsttechnischen Dienstes für Wildbach- und Lawinenverbauung und des Hydrographischen Dienstes, BMLFUW, Wien, 2006.

Heffernan, J. E. and Tawn, J. A.: A conditional approach for multivariate extreme values (with discussion), J. Roy. Stat. Soc. B, 66, 497–546, https://doi.org/10.1111/j.1467-9868.2004.02050.x, 2004.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., and Zhou, G.: Impacts of 1.5 °C Global Warming on Natural and Human Systems, in: Global Warming of 1.5 °C. An IPCC Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, edited by: Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan,

C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield T., available at: https://www.ipcc.ch/site/assets/uploads/sites/2/2019/06/SR15_Chapter3_Low_Res.pdf (last access: June 2020), 2018.

Hundecha, Y. and Merz, B.: Exploring the relationship between changes in climate and floods using a model-based analysis, Water Resour. Res., 48, W04512, https://doi.org/10.1029/2011WR010527, 2012.

Hundecha, Y., Pahlow, M., and Schumann, A.: Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes, Water Resour. Res., 45, W12412, https://doi.org/10.1029/2008WR007453, 2009.

Huttenlau, M. and Stötter, J.: The structural vulnerability in the framework of natural hazard risk analyses and the exemplary application for storm loss modelling in Tyrol (Austria), Nat. Hazards, 58, 705–729, https://doi.org/10.1007/s11069-011-9768-x, 2011.

Huttenlau, M., Schneeberger, K., Winter, B., Reiss, J., and Stötter, J.: Analysis of loss probability relation on community level: a contribution to a comprehensive flood risk assessment, in: Disaster Management and Human Health Risk IV, edited by: Sener, S. M., Brebbia, C. A., and Ozcevik, O., WIT Press, Wessex, 171–182, 2015.

IAWG: HORA/Vorarlberg: Hydraulische Neuberechnung für Vorarlberg, unpublished, 2010.

Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., Bouwer, L. M., Pflug, G., Rojas, R., and Ward, P. J.: Increasing stress on disaster-risk finance due to large floods, Nat. Clim. Change, 4, 264–268, https://doi.org/10.1038/NCLIMATE2124, 2014.

Keef, C., Tawn, J., and Svensson, C.: Spatial risk assessment for extreme river flows, J. Roy. Stat. Soc. C, 58, 601–618, https://doi.org/10.1111/j.1467-9876.2009.00672.x, 2009.

Keef, C., Tawn, J. A., and Lamb, R.: Estimating the probability of widespread flood events, Environmetrics, 24, 13–21, https://doi.org/10.1002/env.2190, 2013.

Kleindienst, H.: Erweiterung und Erprobung eines anwendungsorientierten hydrologischen Modells zur Ganglinensimulation in kleinen Wildbacheinzugsgebie, Diplomarbeit, LMU, München, 1996.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrolog. Sci. J., 31, 13–24, https://doi.org/10.1080/02626668609491024, 1986.

Klijn, F., Merz, B., Penning-Rowsell, E. C., and Kundzewicz, Z. W.: Preface: climate change proof flood risk management, Mitig. Adapt. Strat. Global Change, 20, 837–843, https://doi.org/10.1007/s11027-015-9663-y, 2015.

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

Koivumäki, L., Alho, P., Lotsari, E., Käyhkö, J., Saari, A., and Hyyppä, H.: Uncertainties in flood risk mapping: a case study on estimating building damages for a river flood in Finland, J. Flood Risk Manage., 3, 166–183, https://doi.org/10.1111/j.1753-318X.2010.01064.x, 2010.

Kundzewicz, Z. W., Pińskwar, I., and Brakenridge, G. R.: Large floods in Europe, 1985–2009, Hydrolog. Sci. J., 58, 1–7, https://doi.org/10.1080/02626667.2012.745082, 2013.

Laaha, G., Skøien, J. O., and Blöschl, G.: Spatial prediction on river networks: Comparison of top-kriging with regional regression, Hydrol. Process., 28, 315–324, https://doi.org/10.1002/hyp.9578, 2014.

Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P., and Batstone, C.: A new method to assess the risk of local and widespread flooding on rivers and coasts, J. Flood Risk Manage., 3, 323–336, https://doi.org/10.1111/j.1753-318X.2010.01081.x, 2010.

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?, Hydrol. Process., 30, 2463–2481, https://doi.org/10.1002/hyp.10882, 2016.

Merz, B., Kreibich, H., and Apel, H.: Flood risk analysis: uncertainties and validation, Österreichische Wasser- und Abfallwirtschaft, 60, 89–94, https://doi.org/10.1007/s00506-008-0001-4, 2008.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Assessment of economic flood damage, Nat. Hazards Earth Syst. Sci., 10, 1697–1724, https://doi.org/10.5194/nhess-10-1697-2010, 2010.

Merz, B., Kreibich, H., and Lall, U.: Multi-variate flood damage assessment: a tree-based data-mining approach, Nat. Hazards Earth Syst. Sci., 13, 53–64, https://doi.org/10.5194/nhess-13-53-2013, 2013.

Metin, A. D., Dung, N. V., Schröter, K., Vorogushyn, S., Guse, B., Kreibich, H., and Merz, B.: The role of spatial dependence for large-scale flood risk estimation, Nat. Hazards Earth Syst. Sci., 20, 967–979, https://doi.org/10.5194/nhess-20-967-2020, 2020.

Meyer, V., Becker, N., Markantonis, V., Schwarze, R., van den Bergh, J. C. J. M., Bouwer, L. M., Bubeck, P., Ciavola, P., Genovese, E., Green, C., Hallegatte, S., Kreibich, H., Lequeux, Q., Logar, I., Papyrakis, E., Pfurtscheller, C., Poussin, J., Przyluski, V., Thieken, A. H., and Viavattene, C.: Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps, Nat. Hazards Earth Syst. Sci., 13, 1351–1373, https://doi.org/10.5194/nhess-13-1351-2013, 2013.

Molinari, D., de Bruijn, K. M., Castillo-Rodríguez, J. T., Aronica, G. T., and Bouwer, L. M.: Validation of flood risk models: Current practice and possible improvements, Int. J. Disast. Risk Reduct., 33, 441–448, https://doi.org/10.1016/j.ijdrr.2018.10.022, 2019.

Müller-Thomy, H., Wallner, M., and Förster, K.: Rainfall disaggregation for hydrological modeling: is there a need for spatial consistence?, Hydrol. Earth Syst. Sci., 22, 5259–5280, https://doi.org/10.5194/hess-22-5259-2018, 2018.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P.: An advanced stochastic weather generator for simulating 2-D high-resolution climate variables, J. Adv. Model. Earth Syst., 9, 1595–1627, https://doi.org/10.1002/2016MS000854, 2017.

Plouffe, C. C., Robertson, C., and Chandrapala, L.: Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A

case study of Sri Lanka, Environ. Model. Softw., 67, 57–71, https://doi.org/10.1016/j.envsoft.2015.01.011, 2015.

Raynaud, D., Hingray, B., Evin, G., Favre, A.-C., and Chardon, J.: Assessment of meteorological extremes using a synoptic weather generator and a downscaling model based on analogs, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-557, in review, 2019.

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J. C., Bodner, G., Borga, M., Chaplot, V., Gallart, F., Glatzel, G., Hall, J., Holden, J., Holko, L., Horn, R., Kiss, A., Kohnová, S., Leitinger, G., Lennartz, B., Parajka, J., Perdigão, R., Peth, S., Plavcová, L., Quinton, J. N., Robinson, M., Salinas, J. L., Santoro, A., Szolgay, J., Tron, S., van den Akker, J. J. H., Viglione, A., and Blöschl, G.: Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research, Water Resour. Res., 53, 5209–5219, https://doi.org/10.1002/2017WR020723, 2017.

Sauter, I., Kienast, F., Bolliger, J., Winter, B., and Pazur, R.: Changes in demand and supply of ecosystem services under scenarios of future land use in Vorarlberg, Austria, J. Mount. Sci., 12, 2793–2809, https://doi.org/10.1007/s11629-018-5124-x, 2019.

Schneeberger, K. and Steinberger, T.: Generation of Spatially Heterogeneous Flood Events in an Alpine Region – Adaptation and Application of a Multivariate Modelling Procedure, Hydrology, 5, 5, https://doi.org/10.3390/hydrology5010005, 2018.

Schneeberger, K., Huttenlau, M., Winter, B., Steinberger, T., Achleitner, S., and Stötter, J.: A Probabilistic Framework for Risk Analysis of Widespread Flood Events: A Proof-of-Concept Study, Risk Anal., 39, 125–139, https://doi.org/10.1111/risa.12863, 2019.

Senfter, S., Leonhardt, G., Oberparleiter, C., Asztalos, J., Kirnbauer, R., Schöberl, F., and Schönlaub, H.: Flood Forecasting for the River Inn, in: Sustainable Natural Hazard Management in Alpine Environments, edited by: Veulliet, E., Johann, S., and Weck-Hannemann, H., Springer, Berlin, Heidelberg, 35–67, https://doi.org/10.1007/978-3-642-03229-5_2, 2009.

Serinaldi, F. and Kilsby, C. G.: A Blueprint for Full Collective Flood Risk Estimation: Demonstration for European River Flooding, Risk Anal., 37, 1958–1976, https://doi.org/10.1111/risa.12747, 2017.

Skøien, J. O., Merz, R., and Blöschl, G.: Top-kriging – geostatistics on stream networks, Hydrol. Earth Syst. Sci., 10, 277–287, https://doi.org/10.5194/hess-10-277-2006, 2006.

Speight, L. J., Hall, J. W., and Kilsby, C. G.: A multiscale framework for flood risk analysis at spatially distributed locations, J. Flood Risk Manage., 10, 124–137, https://doi.org/10.1111/jfr3.12175, 2017.

Statistik Austria: Baupreisindex, available at: https://www.statistik.at/web_de/statistiken/wirtschaft/produktion_und_bauwesen/konjunkturdaten/baupreisindex/index.html, last access: 15 October 2019.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B., Dutta, D., and Kim, S.: Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, Environ. Model. Softw., 90, 201–216, https://doi.org/10.1016/j.envsoft.2017.01.006, 2017.

Thieken, A. H., Apel, H., and Merz, B.: Assessing the probability of large-scale flood loss events: a case study for the river Rhine, Germany, J. Flood Risk Manage., 8, 247–262, https://doi.org/10.1111/jfr3.12091, 2015.

Ullrich, S., Hegnauer, M., Dung, N. V., de Bruijn, K., Merz, B., Kwadijk, J., and Vorogushyn, S.: Comparative evaluation of two types of stochastic weather generators coupled to hydrological models for flood estimation, Oral, in: IUGG General Assembly, Montreal, 2019.

Vorogushyn, S., Bates, P. D., de Bruijn, K., Castellarin, A., Kreibich, H., Priest, S., Schröter, K., Bagli, S., Blöschl, G., Domeneghetti, A., Gouldby, B., Klijn, F., Lammersen, R., Neal, J. C., Ridder, N., Terink, W., Viavattene, C., Viglione, A., Zanardo, S., and Merz, B.: Evolutionary leap in large-scale flood risk assessment needed, Wiley Interdisciplin. Rev.: Water, 5, 1–7, https://doi.org/10.1002/wat2.1266, 2018.

Vrac, M. and Naveau, P.: Stochastic downscaling of precipitation: From dry events to heavy rainfalls, Water Resour. Res., 43, W07402, https://doi.org/10.1029/2006WR005308, 2007.

Wagenaar, D. J., de Bruijn, K. M., Bouwer, L. M., and de Moel, H.: Uncertainty in flood damage estimates and its potential effect on investment decisions, Nat. Hazards Earth Syst. Sci., 16, 1–14, https://doi.org/10.5194/nhess-16-1-2016, 2016.

Winter, B., Schneeberger, K., Huttenlau, M., and Stötter, J.: Sources of uncertainty in a probabilistic flood risk model, Nat. Hazards, 91, 431–446, https://doi.org/10.1007/s11069-017-3135-5, 2018.

Winter, B., Schneeberger, K., Dung, N. V., Huttenlau, M., Achleitner, S., Stötter, J., Merz, B., and Vorogushyn, S.: A continuous modelling approach for design flood estimation on sub-daily time scale, Hydrolog. Sci. J., 88, 1–16, https://doi.org/10.1080/02626667.2019.1593419, 2019.