Stochastic Environmental Research and Risk Assessment (2019) 33:1057–1074
https://doi.org/10.1007/s00477-019-01681-3

ORIGINAL PAPER

A probabilistic risk modelling chain for analysis of regional flood events

J. Oliver[1,2] · X. S. Qin[3] · H. Madsen[4] · P. Rautela[5] · G. C. Joshi[6] · G. Jorgensen[4]

Published online: 11 May 2019

Abstract
A probabilistic flood risk modelling chain is proposed for flood risk analysis with consideration of spatial spreading and temporal clustering of the flood events. The proposed method consists of (1) a continuous simulation of long-term climatic and hydrologic fields, (2) Monte-Carlo simulations of a 1D river and inundation model and (3) a probabilistic loss model. In the first stage, three multisite multivariate weather generators were tested and a K-nearest neighbor weather generator (KNN-CAD v4) was found most suitable to reproduce hydrological extremes. The methodology was then applied to identify variations in risk between spatially coherent and simplified event sets generated from 5000 years of synthetic data across four river basins in Uttarakhand, India. The results showed that preserving the spatial coherency of regional events lead to a spreading of the T-year local event distributions on both sides of a uniform regional T-year return period. The specificity of the exposure and the correlations between local events translated into minimal (4%) differences in losses at high return periods (> 50 year) but larger difference (20%) at low return periods (5 to 10-year). The redistribution of loss frequencies led to negligible differences in the Annual Average Losses between the different event sets. In this application, a simplified event set structure with uniform but appropriate T-year return periods proved to capture reasonably well the averaged risk metrics regionally. This study illustrates the complexity of estimating regional flood risk accumulation and exemplifies how a probabilistic risk model chain with continuous simulation can provide a more detailed picture of flood risks.

Keywords Flood risk analysis · Continuous simulation · Risk model chain · Probabilistic flood risk modelling

1 Introduction

Flood management practice is increasingly addressed from a risk perspective. The related approaches aim at quantifying the hazard, e.g. inundation depths and extents, with its consequences represented by the elements at risk and their susceptibility to flood (e.g. Kron 2005; Merz et al. 2010). Such a practice forms today the basis required by planners and decision makers to establish portfolios of cost-effective measures designed not only to avoid flooding but also to minimize the impact of floods on their economy and population (e.g. Luo et al. 2015). Flood risk analysis is

✉ X. S. Qin
xsqin@ntu.edu.sg

[1] Nanyang Environment and Water Research Institute (NEWRI), Interdisciplinary Graduate School, Nanyang Technological University, 1 Cleantech Loop, Singapore 637141, Singapore

[2] DHI Water & Environment (S) Pte. Ltd., 2 Venture Drive, #18-18 Vision Exchange, Singapore 608526, Singapore

[3] School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

[4] DHI, Agern Allé 5, 2970 Hørsholm, Denmark

[5] Disaster Mitigation and Management Centre (DMMC), Uttarakhand Secretariat Rajpur Rd, Karanpur, Dehradun, Uttarakhand 248001, India

[6] Project Implementation Unit (PIU), Uttarakhand Disaster Recovery Initiative 29, IIE (IT Park) SIDCUL Building, Sahastradhara Road, Dehradun, Uttarakhand 248001, India

traditionally carried out in a series of steps involving: (1) the definition of various driving scenarios, such as several (or one) T-year return period (RP) discharges or rainfall events developed from extreme value analysis of historical records within a catchment (Luo et al. 2018a); (2) the mapping of the T-year inundation scenarios using a T-year event applied uniformly at the reach or catchment scale; (3) the computation of the T-year damage associated with each T-year flood map using geospatial layers of exposed elements and associated damage functions. As an example, the European Flood Directive on the Assessment and Management of Flood Risks recommends its member states to perform this exercise for 3 (low, medium and high) RPs (European Commission 2007).

The assumption of a spatially uniform RP and the association of the same RP to describe the event, inundation scenario and damage frequency holds for small scale studies (e.g. Luo et al. 2018b) but not necessarily when larger scales are considered. Falter et al. (2015) demonstrated that spatial variations could exist in terms of hazard and losses for realistic event scenarios at the reach scale. Thieken et al. (2015) and Schneeberger et al. (2017) concluded that the use of uniform RPs over large-scale systems could lead to an overestimation of the flood risks. Different approaches should then be proposed to define regional scale risk management plans and policies, and account for the spatial spreading and temporal clustering of flooding across different administrative units, regions or countries to better describe risk cumuli (e.g. Kundzewicz et al. 2013; Jongman et al. 2014).

Recent research works have investigated new methodologies to capture the spatial heterogeneity of risks over large areas. Most methods make use of multivariate frequency analysis (MFA) to model spatial dependencies of hazards (or discharges) at multiple locations (Ghizzoni et al. 2012; Neal et al. 2013; Carreau and Bouvier 2016; Schneeberger et al. 2017) and between losses of correlated regions (Timonina et al. 2015; Jongman et al. 2014). However, these approaches failed at capturing the dependencies of temporal processes that drive flooding. To overcome these limitations, Serinaldi and Kilsby (2017) applied MFA with a continuous simulation approach to produce spatially and temporally correlated weekly discharge maxima. However, these approaches are still limited to the reproduction of instantaneous discharge peaks and do not include other flood hydrograph characteristics (e.g. volume) important for hazard and risk accumulation estimation.

While more complex and adding uncertainties in the modelling chain, using continuous rainfall and runoff simulation, instead of direct discharge observation and MFA methods, can present several benefits: (1) the event set quality can be enhanced as the simulation accounts for antecedent conditions (see Grimaldi et al. 2013; Haberlandt and Radtke 2014); (2) the spatial and temporal variation in peaks, volumes and hydrograph shapes are represented using the full description of the routing and accumulation processes; (3) the historical precipitation records are independent of anthropogenic influences that can alter the flow over time (e.g. dikes, dams or land-use); (4) the event set can be perturbated for climate change studies. The application of a continuous hydrological simulation will generate a realistic set of hazard scenarios spatially and temporally coherent at the regional scale. With this approach, no pre-defined assumption is made between the RP of the discharge peak locally and a possible RP of the event hazard at the regional scale. The continuous simulation concept can then be extended and applied to the last link of the risk modelling chain: by simulating regional losses for each regional hazard scenario, regional risk metrics can be established without any assumption made between the RPs of the loss and the hazard. Falter et al. (2015) introduced this approach and proposed a coupled model chain relying on a multi-site weather generator, a rainfall-runoff model, a river channel and floodplain model, and a loss model. The coupled model chain was applied to the meso-scale Mulde catchment in Germany and forced with 10,000 years of synthetic rainfall. Using this approach, they demonstrated that variability exists between peak discharge and loss RPs at the catchment and sub-catchment scales.

Better quantification of regional flood risks can help plan and allocate more adequately financial resources to allow flood prone and populated regions to economically grow. In India, the Uttarakhand Province, host of 10.1 million inhabitants, faces tremendous natural hazards caused by water (e.g. floods, flash floods, heavy precipitation, avalanches, landslides, Glacier Lake Outburst Flow). The recent calamity due to heavy downpour and subsequent devastating floods in various rivers on 16–17 June 2013 led to heavy losses of life and property (see Das 2013; Rana et al. 2013). While flash floods are predominant hazards in the mountain zones, the low-lying areas in the Terai and Dun valleys, which are more densely populated and economically active, could be affected more severely by fluvial flooding. To better understand the susceptibility of the State to natural disasters, it is desired to have thorough investigations of the spatial dynamics of flood hazards and risks at the outlet of the four main rivers draining the State.

To date, very limited studies have tried to describe adequately risk at the regional scale or understand the limitations of using a uniform T-year RP as currently practiced, merely none for the Uttarakhand region. Existing methods allow to capture the spatial heterogeneity of flood events but rarely reproduce the temporal dependencies of

flooding processes. Besides, no approaches have been proposed that can account for other possible variabilities in the hazard or loss models. In this study, we aim to investigate the differences between simplified event set structures (e.g. uniform T-year RP) and spatially coherent and continuous simulations on flood hazard and risks. For this purpose, we developed a continuous risk modelling chain similar to Falter et al. (2015) but enhanced with a probabilistic representation of the hazard and losses. The system relies on synthetic weather data produced with a multi-site multivariate weather generator. The generated weather fields are used to run continuously a rainfall-runoff model that provides inflow to a river routing and floodplain model. The spatially coherent inundation maps are then used to generate different event sets (with or without spatial variations). The different event catalogues are used to produce losses using a probabilistic loss model. The model chain is applied to four of the main tributaries of the river Ganges draining the State of Uttarakhand, India.

2 Flood risk modelling chain

Figure 1 shows the modelling chain for flood hazard and risk calculations. The chain relies on the generation of long-term continuous weather variables synthetically derived from a multi-site weather generator. The long-term time series are used to produce continuous discharge and inundation maps. Different event sets are produced from the continuous and spatially coherent data and different degrees of simplification of its underlying structure (e.g. spatial and discharge variability) are passed to investigate the significance of its different components. Catalogues of inundation maps corresponding to each event set are combined and used for damage modelling. The probabilistic flood risk modelling is carried out using existing software packages (i.e. ARC GIS, MIKE, and CAPRA). The overall hydrodynamic and Monte-Carlo simulations rely on the MIKE1D Application Programming Interface (API) (see Unphon 2009), implemented in C#. Three open-source multi-site weather generators have been tested including MulGETS (Chen et al. 2014) and TripleM (Breinl et al. 2013, 2015) implemented in Matlab and a modified version of KNN-CAD v4 (King et al. 2015) in Python (Mandal et al. 2017). Pre- and post-processing of data and results and extreme value analysis are performed with the statistical software R and 'extRemes' package.

2.1 Synthetic weather generation

In recent years numerous statistically based multi-site weather generators have been developed (e.g. Hundecha et al. 2009; Breinl et al. 2013; Evin et al. 2018; Chen et al. 2015; Sparks et al. 2018). They have in most cases been tested in temperate climate and in terms of their capacity to reproduce historical rainfall correlations and statistics. Besides, few direct comparisons of multi-site weather generator performance have been done. In their study, Vallam and Qin (2016) demonstrated that performances may varies with the specific characteristics of rainfall to be preserved (i.e. extreme values and spatial correlation). The application of multi-site weather generators over large scale and their capacity to reproduce hydrological extremes remains largely unexplored. Falter et al. (2015) demonstrated the ability of the multi-site rainfall generator of Hundecha and Merz (2012) to reproduce discharge extremes in the Mulde catchment (7400 km^2), Germany. Chen et al. (2015) demonstrated the ability of MulGETS to reproduce, although with some overestimation, the extreme flows characteristics of Lac-Saint-Jean watershed (73,800 km^2) in Quebec. Three multisite weather generators are applied to produce daily weather variables. The selected models are open-source and their underlying algorithms have been validated in numerous case studies.

2.1.1 k-nearest neighbor: KNN-CAD v4

The KNN-CAD v4 is a block-bootstrap non-parametric generator with perturbation based on the K-nearest neighbour (KNN) algorithm (King et al. 2015). KNN types of weather generators have been described and applied in many research studies, proving to produce adequately rainfall and temperature (Sharma and Lall 1999; Buishand and Brandsma 2001; Bannayan and Hoogenboom 2008; Lu and Qin 2014; Vallam and Qin 2016) and extreme flows (Leander and Buishand 2007; Wit and Buishand 2007). In this version, King et al. (2015) introduced two features to the KNN resampling procedure: (1) a perturbation of the resampled observed data to produce values outside the observed range while preserving daily autocorrelations, and (2) the use of block resampling to improve temporal correlations.

2.1.2 Multisite Markov model: TripleM

The Multisite Markov Model (TripleM) is a reduced complexity multi-site generator developed by Breinl et al. (2013). The model, semi-parametric in nature, relies on a Markov process to simulate rainfall occurrence and a parametric sampling and reshuffling of the precipitation amounts (see Clark et al. 2004). The overall process is explained in details by Breinl et al. (2013, 2015, 2017) and summarized as follows: (1) daily snapshot of precipitation occurrences are clustered according to their similarities (using binary or k-means methods); (2) a Markov process simulates synthetic time-series of clustered vectors as

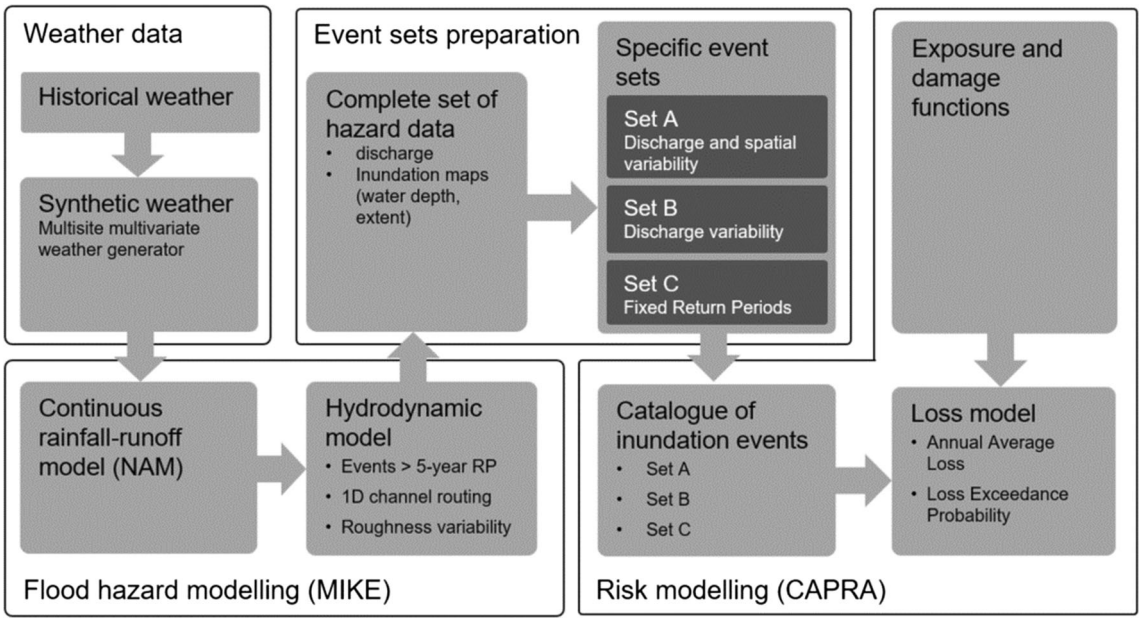

Fig. 1 Schematization of the main modelling components and event set preparation adopted for flood hazard and risk calculations

function of a number of previous days; (3) each clustered vector is replaced by a daily precipitation snapshot sampled from the same cluster; (4) synthetic rainfall values are generated at each location using a fitted distribution (e.g. Weibull) and optionally using a correlated Cholesky decomposition (Watkins 2010); (5) sampled rainfall amounts are reshuffled based on their original ranks in the vectors produced in step (3) to preserve correlation between stations. The TripleM generates precipitation field which can be complemented by temperature and evaporation fields simulated using KNN.

2.1.3 MulGETS

The MulGETS weather generator (Brissette et al. 2007; Chen et al. 2014) simulates daily precipitation and minimum/maximum temperature using temporally independent but spatially correlated random numbers. A spatially correlated random field of rainfall occurrence or amount is generated using the method of Iman and Conover (1982) to describe rainfall over the length of the simulated period. The random field is then treated as a single site weather generator to produce the spatially correlated rainfall occurrences or amounts. Occurrences are generated from a first-order two-state Markov chain considering wet and dry status of the 2 previous days. Amounts can be produced from a multi-exponential or multi-gamma distribution. The MulGETS generates precipitation and temperature fields which can be complemented by evaporation fields simulated using KNN.

2.2 Flood hazard modelling

Hydrological modelling is carried out using the NAM rainfall-runoff model developed at the Institute of Hydrodynamics and Hydraulic Engineering at the Technical University of Denmark (Nielsen and Hansen 1973) and widely used in hydrological studies (see for example: Rulin et al. 2008; Wang et al. 2016). The NAM model is a lumped, conceptual model simulating continuously rainfall-runoff processes as a function of weather variables (including precipitation, evaporation, temperature and solar radiations) and accounting for water content in four mutually interrelated storage compartments including snow, surface, root zone and groundwater. The continuous rainfall, temperature and potential evaporation fields are generated using the multisite weather generator and the hydrological model run continuously. The events above the 5-year RP are identified and used for hydrodynamic simulation.

The hydrodynamic modelling is performed with the one-dimensional engine MIKE1D (DHI 2017), which is based on the full St Venant equations. Where overland flow patterns are parallel to the main river, such treatment of flow routing in channels can be extended to simulate floodplain flows, producing fairly good approximations of the main flood characteristics (Horritt and Bates 2002). The 1D river channel model could also be coupled dynamically to 2D hydrodynamic models to represent more complex floodplain flow patterns.

2.3 Event sets preparation

Three main types of event sets are generated to identify the contribution of different elements of the event set structure: (1) Event Set A capturing all possible events occurring during the simulation period while preserving the spatially coherency at the regional scale, (2) Event Set B capturing all possible events but assuming full correlation between the RP of the local events (uniform T-year RP regionally), and (3) Event Set C, similar to Event Set B with uniform T-year RP at the regional scale, but for only a few events of selected RP.

Event Set A represents an exhaustive and realistic set of scenarios capturing both discharge and spatial variability. Continuous simulations are performed and discharge peaks above a 5-year RP identified at each location l where risk analysis is performed. An event is defined using a 5-day window around the date of a peak at each location. Each event is associated with the same frequency of occurrence equal to the inverse of the number of years simulated. In this scenario, we can expect to have one regional event composed by local events of different RPs.

$$Event\ set\ A \begin{cases} Event_i = \bigcup_{locl} Event_l(p) \\ F(Event_i) = 1/X \end{cases} \quad (1)$$

where Event Set A is composed by N Events (i = 1, 2, …, N); N is the total number of peak combinations considering the periods when local events occur at each location l; $locl = \{Loc1, Loc2, \ldots, LocL\}$ the L locations considered for risk analysis; $Event_l(p)$ is the event occurring at location l during period p; $F(Event_i)$ is the annual probability of occurrence of $Event_i$; X is the number of years of synthetic weather data simulated.

Event Set B assumes full correlation between RPs of local events. The regional event set is built from Event Set A. But instead of grouping the events by date of occurrence, the events are grouped by their rank in terms of discharge at each location:

$$Event\ set\ B \begin{cases} Event_j = \bigcup_{locl} Event_l(j) \\ F\left(Event_j\right) = 1/X \end{cases} \quad (2)$$

where Event Set B is composed by Y events (j = 1, 2, …, Y); Y a fixed number of peaks above a 5-year RP; $Event_l(j)$ is the event occurring at location l with rank j; $F\left(Event_j\right)$ is the annual probability of occurrence of $Event_j$.

Event set C is analogous to the event set structure typically used for risk analysis. Here, a few T-year RPs are used to map hazards and estimate risks. The event set is built by extracting single events by using their ranks from Event Set B, where the rank indexes should correspond to the T-year RPs used. With this formulation, Event Set C is constituted by a limited number of events representing an exceedance level in risk estimation:

$$Event\ set\ C \begin{cases} Event_k = \bigcup_{locl} Event_l(k) \\ F(Event_k) = k/X \end{cases} \quad (3)$$

where Event Set C is composed by k events (k = 1, 2, …, K); K is the fixed number of ranked events (or RPs) investigated; $Event_l(k)$ is the event occurring at location l with rank k; $F(Event_k)$ is the annual probability of occurrence of $Event_k$.

2.4 Risk modelling

The open-source "Comprehensive Approach to Probabilistic Risk Assessment" platform (CAPRA) (ERN-AL 2009a) is used for risk computation. The platform design allows to perform probabilistic hazards (e.g. earthquakes, flooding, and hurricanes) and risk analysis. The platform has been frequently used in risk studies (Marulanda et al. 2013; Salgado-Gálvez et al. 2014; Velásquez et al. 2012; Oliver et al. 2018). CAPRA-GIS (ERN-AL 2011) computes risk metrics such as Average Annual Loss (AAL) and Loss Exceedance Probability (LEP) from catalogues of hazard footprints and assets (exposure) associated with their damage functions. The LEP curve establishes the frequencies of occurrence of losses originating from flooding events and exposed assets across an event set of all plausible scenarios and accounting for the uncertainties in water depth and the damage potential. The Average Annual Loss (AAL) is calculated as the integral of the LEP curve. Formulation details are provided in ERN-AL (2009b) and Oliver et al. (2018).

It is common practice, in flood risk analysis, to produce an event set limited to only a few selected and unique exceedance probabilities (e.g. Event Set C). In this case the AAL can be approximated using the trapezoidal rule (Meyer et al. 2009):

$$AAL = \sum_{i=1}^{N-1} \frac{L(Event_i) + L(Event_{i+1})}{2} [F(Event_{i+1}) - F(Event_i)] \quad (4)$$

where $L(Event_i)$ is the loss associated with $Event_i$ of annual probability $F(Event_i)$ and N corresponds to a number of single and unique probabilities of losses in the event set.

3 Application to the Uttarakhand province

3.1 Study area

The study domain covers four catchments in Uttarakhand Province, India: the Yamuna (YAM), Sarda (SAR), Kosi

(KOS) and Upper Ganges (GAN) (Fig. 2). Physiographically, the Uttarakkhand Province can be divided into three zones: Himalayas, the Shivaliks and the Tarai regions. Elevation ranges from 200 to 7818 m (Mount Nanda Devi) with approximately 50% of the area characterized by steep slopes and above 2000 m altitude. Uttarakhand presents two main climatic regions. The climate in the northern and largest part of Uttarakhand is typically Himalayan with cold winters and snowfall, intense rainfall during the monsoon, and mild summers. The area is bestowed with a relatively high average annual rainfall (around 1200 mm) reaching up to 2400 mm in some areas with 82% received during the monsoon season (Kumar 2017). The mountain range itself exerts an appreciable extent of influence on monsoon and rainfall patterns depending on altitude and orientation. The southern slope receives high rainfall with longer sunshine duration and thus the area normally remains dry due to high rate of evapotranspiration. On the contrary, the northern slope receives lesser rainfall, though remains moist due to short sunshine duration. Data used for the flood hazard model were collected with the support of the Project Implementation Unit (PIU) and originated from the Central Water Commission (CWC), the India Meteorological Department (IMD), the Uttarakhand Irrigation Department, the National Remote Sensing Center (NRSC), publicly available datasets (e.g. GPM, SRTM), and field surveys. The exposure and vulnerability models were based on DHI et al. (2017) and published socio-economic statistics and studies, publicly available dataset (e.g. OpenStreetMap, Wikimapia) and site surveys.

3.2 Historical weather and discharge data

Due to the complex and heterogeneous terrain, weather systems are expected to present large variations in the study domain. In view of the limited number of ground monitoring stations, particularly in the high mountains, a variety of weather data products have been used:

- The long-term (from 1901 to 2015) gridded (0.25°) daily rainfall field produced by the Indian Meteorological Department (IMD) (Pai et al. 2014). The product is based on archived records of 7000 rain-gauge stations and interpolated in space and time using an Inverse

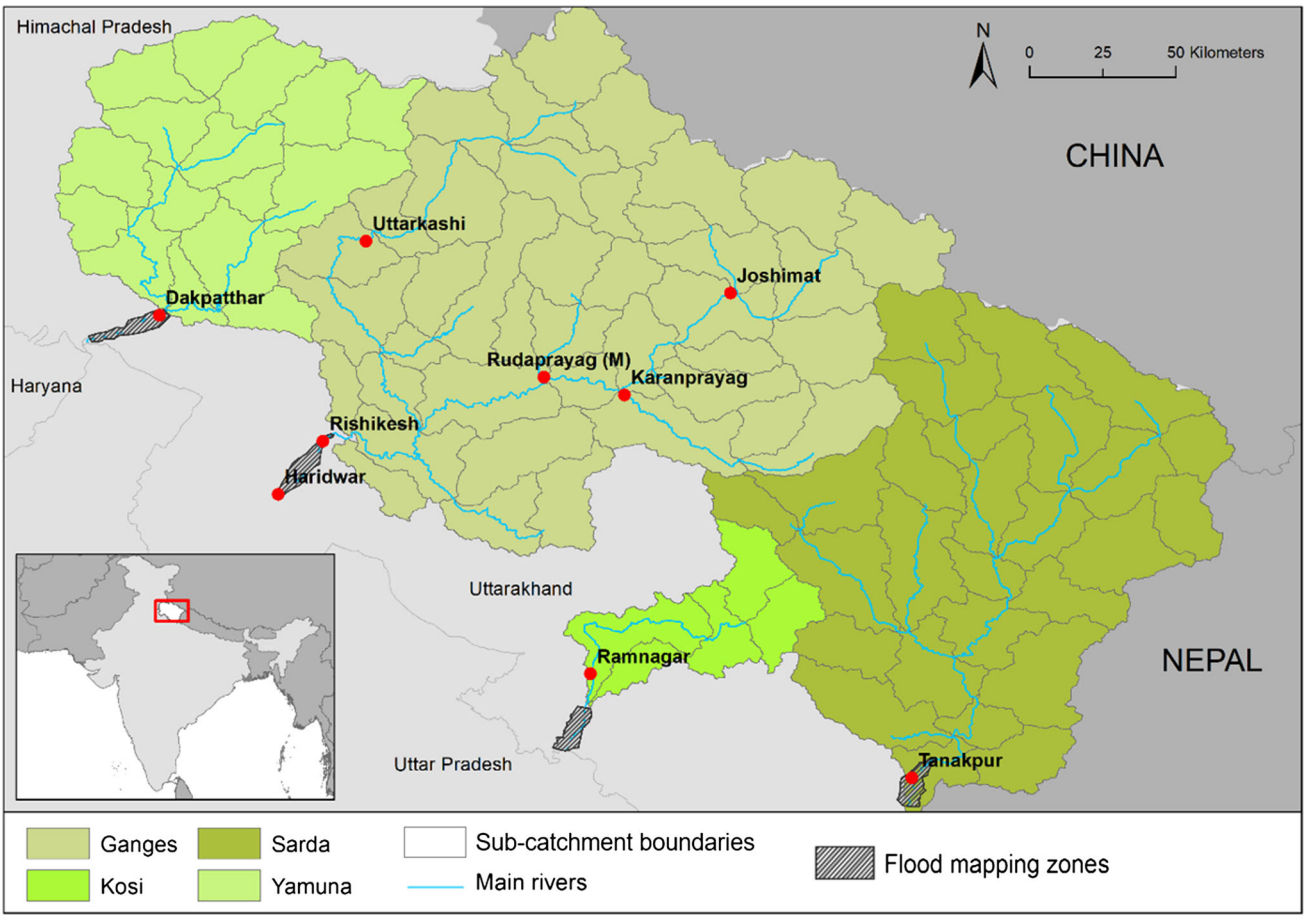

Fig. 2 Study domain including the four major river catchments in Uttarakhand with hydrological and hydrodynamic model components (sub-catchments and river network) and flood mapping locations

Weighted Distance (IWD) method (Shepard 1968). The IMD rainfall dataset has been applied in numerous hydrological studies (Prakash et al. 2016; Parida et al. 2017).

- The hourly gridded (0.1°) but short-term (2014–2015) GPM calibrated V03D data (Hou et al. 2014). The data is considered capable of describing rainfall in the high mountains poorly covered by the synoptic network. The dataset is used to identify correction ratios of low and medium rainfall to be applied to the IMD dataset using the mean yearly rainfall accumulation between GPM and IMD products for catchments with mean elevation above 3500 m. The correction factors suggest a spatial redistribution of rainfall with an accumulated depth reduction in the high mountains and an increase in Himalayan foreland, which is consistent with the rainfall behavior regime expected in the Himalayas (Dhar and Nandargi 2004; Bookhagen and Burbank 2006).
- The daily gridded (0.25°) Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) rainfall product (Funk et al. 2015). It is applied to the areas covering Nepal.
- The gridded (0.25°) daily data from the European Centre for Medium Range Weather Forecasts ECMWF ERA-Interim daily data (Dee et al. 2011). They are used for estimating evaporation and temperature.

Discharge data is provided by the Central Water Commission (CWC) as 10 days averaged values covering the monsoon period (June–October) for the period 2010–2014 and daily averaged Annual Maximum Series (1980–2015) for selected locations along the rivers Bagirathi, Alaknanda and Mandakini (directly downstream the highest mountain zones where rainfall data quality is uncertain). The Annual Maximum Series (1980–2015) of the Ganges discharge at the Bimghoda Dam near Haridwar was also made available by the Indian Irrigation Department.

3.3 Model set-up

3.3.1 Synthetic weather data

Three multisite weather generators are used to produce daily weather variables over a 500-year simulation period. KNN-CAD has been configured with a fixed temporal window of 20 days translating to a 'k' value of 24 according to the formulation of Yates et al. (2003). A perturbation of 10% and 20% and a block size of 1 and 10 are used to simulate rainfall, temperature and evaporation data. As recommended by Mandal et al. (2017), a log-normal perturbation is applied to the rainfall and evaporation fields while normal perturbation is applied to temperature. The TripleM weather generator has been tested using Weibull and a mixture of Weibull-GPD distributions. The Weibull-GPD distributions are combined using a Peak-Over-Threshold (POT) approach with the threshold established using the 'parameter stability plot' method (see Coles 2001). Finally, the MulGETS weather generator has been tested using both multi-exponential and multi-gamma distributions, with a first order Markov chain and monthly clustering. After performance evaluation, a weather generator is selected and 5000 years of synthetic data are produced to adequately represent event variability in the 5 to 500-year RP critical for loss estimation.

3.3.2 Runoff simulation

The NAM model domain is divided in 89 sub-catchments (see Fig. 2 and Table 1) of spatial resolution similar to the weather data (600 km^2) and each associated with time series of climate variables. Within each sub-catchment, orographic effects on rainfall are accounted for using correction factors and regional regression established between GPM and elevation. Lumped conceptual rainfall-runoff models such as NAM typically require continuous discharge observations and rainfall data over several years to derive accurately the catchment model parameters (Madsen 2000). Regionalization techniques can also be applied to derive model parameters in un-gauged locations (Hundecha et al. 2002; Bárdossy 2007).

Considering the limitation of the hydrometric datasets, a manual calibration of the NAM model parameters (CQOF; CK1,2) is done based on expert judgement and supported by the analysis of the geomorphic characteristics of the catchments. In the NAM model, the parameters with largest influence on discharge peaks are the overland flow runoff coefficient (CQOF) and routing time constant (CK1,2) (see Madsen 2000) which are analogous to the runoff curve number, developed in 1954 by the SCS, and the lag time of the Unit Hydrograph Method (UHM), respectively. Those parameters are derived here to guide the manual calibration process of the NAM model. The Curve Number is derived using an approach similar to Hong and Adler (2008) based on reclassified land cover from the European Space Agency (ESA)—Climate Change Initiative Land Cover project (CCI LC 2012) and soil types from the Harmonized World Soil Database (USDA-NRCS) reclassified in 4 Hydrologic Soil Groups HSG following Rawls and Brakensiek (1989). The lag time is computed according to the Soil Conservation Service (SCS) formula (ASCE 1996):

$$T_{lag} = \frac{(3.28 \times 10^3 \cdot L)^{0.8} \left(\frac{1000}{CN} - 9\right)^{0.7}}{1900 \cdot Y^{0.5}} \quad (5)$$

Table 1 Summary of model domains and hydrological distributions

River	Abbreviation	Catchment size (km^2)	Sub-catchments	Total rivers length (km)
Ganges	GAN	21,100	38	1206
Sarda	SAR	14,980	26	766
Yamuna	YAM	11,600	20	368
Kosi	KOS	1860	5	159

where L is the hydraulic length (km), Y the average slope (%), and CN the areal weighted Curve Number (–).

The CQOF and CK1,2 parameters are first assumed identical to the CN and lag-time values and then adjusted uniformly across all sub-catchments to achieve the best fit of simulated annual peak flows with the observed. The remaining parameters controlling interflow, baseflow and snow storage and melt are established manually to fit the rising and receding limbs of the 10 days average monsoon flow data and annual water balances.

3.3.3 Hydrodynamic model and inundation mapping

The approach, which assumes that the overland flow patterns are parallel to the main river, is judged acceptable for the study area presenting relatively steep river gradients and well-defined U-Shaped floodplains. The approach is also computationally efficient to allow for the required large number of simulations. The river network is extracted from the 10 m (LSE20: 8 m) Cartosat Digital Surface Model (DSM) produced by the National Remote Sensing Centre (NRSC). The DSM is hydrologically corrected using the carving technique proposed by Soille et al. (2003) and buildings removed using a block minimum smoothening filtering procedure relying on manually digitized building polygons. Following river delineation, cross-sections are extracted from the DTM every 3 km and the NAM simulated runoff is distributed along the river network. The main dams and reservoirs are simulated using a simple broad-crest weir formulation and considered at full reservoir level.

At the four locations where flood mapping is intended, cross-sections are extracted every 500 m to 2 km from the DTM. Low flow conveyance, not captured in the original DSM, is considered minimal compared to conveyance of the meandering river bed and floodplain and no correction is applied. The water retention structures are assumed fully open during the flood event and able to accommodate the peak flow. The Upper Ganges (GAN) river section features a 100-year flood protection embankment. Planar surface interpolation of water level between computation nodes and overlaid with the DTM is then used to produce inundation maps.

3.3.4 Exposure and vulnerability

Exposure and vulnerability functions were developed by DHI et al. (2017) from reports, field surveys or extracted from the literature (Fig. 3). The exposure considers two categories of assets most susceptible to flooding: agriculture and residential. Damage curves associated with fluvial flooding are defined as a function of the household contents/assets (e.g. furnishings, electric equipment) susceptible to be damaged for single, double and multiple stories houses. The curves were derived from detailed surveys of more than 1250 households across the Uttarakhand State and validated against other household contents damage curves published in the relevant literature (e.g. Reese and Ramsay 2010). A single composite damage curve capturing the main crops present in the area (paddy, sugarcane, dry crops, beans and potato) and based on Huizinga et al. (2017) is used for agriculture. Assets valuation has been established from surveys and published production data.

3.4 Event sets and subsets

From the 3 main event sets proposed, sub-scenarios including variabilities in different parameters are investigated (Table 2). The results of the fully continuous simulation performed are used in both Events sets A and B with A preserving the spatial structure of the events. A and B implicitly capture the variability of discharge by accounting for all potential events that occur within the 5000 years and without making assumption between loss and hazard RPs. We tested how the probabilistic risk estimates compare against those based on combinations of different flood hazard maps. By considering a few number of uniform RPs, Event sets C do not account for discharge variability around the RPs selected and assume identical RPs between hazard and losses. C1 considers 8 RPs as recommended by the PIU for analysis in Uttarakhand. C2 considers the same RPs below the 100-year threshold. C3 and C4 follow the European Flood Directive (2007/60/EC) and uses three hazard maps of low, medium and high probabilities. The 100 and 1000 years are typical design criteria for hydraulic infrastructure in Uttarakhand and are used for medium and low probabilities. The low RP is 5 years for C3 and 10 years for C4 to investigate the impact of choice on the lowest RP. Vulnerability uncertainty is accounted for in A

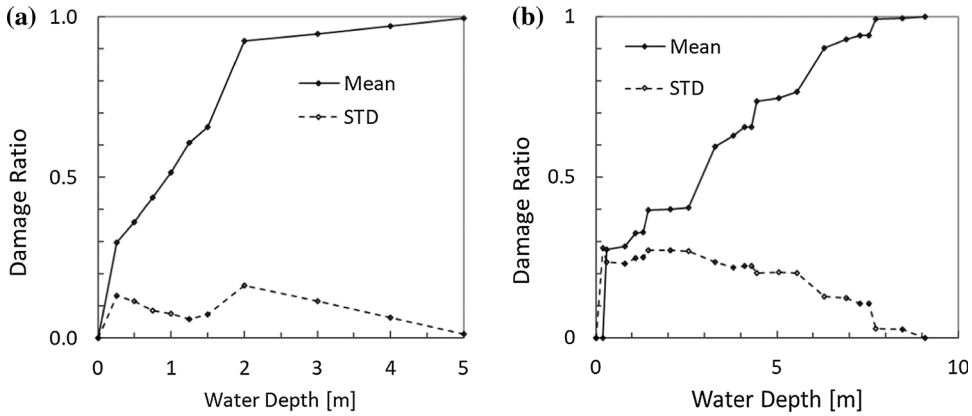

Fig. 3 Examples of vulnerability curves **a** composite agriculture and **b** residential multi-storey. (*Data source*: DHI et al. 2017)

Table 2 Summary of event sets and scenarios considered with variabilities included in different parameters

Event set	Loss estimation approach/fixed probability levels considered (year)	Variabilities considered			
		Hazard			Vulnerability
		Spatial	Discharge	Roughness	
A	Fully continuous	x	x	x	x
B1	Fully continuous		x	x	x
B2	Fully continuous		x	x	
C1	5, 10, 25, 50, 100, 250, 500, 1000			x	
C2	5, 10, 25, 50, 100			x	
C3	5, 100, 1000			x	
C4	10, 100, 1000			x	

and B1. Only the mean of the damage functions is used in B2 and C.

4 Results

4.1 Flood hazard and risk model performance and limitations

A manual calibration of the hydrological NAM model has been performed by comparing simulated and observed annual peaks for the 1980–2015 period. After quality control, 3 years, presenting excessive (> 60%) deviations between rainfall and discharge observations, indicating considerable errors in the distributed rainfall fields, were removed from the historical dataset to avoid propagating input data error in the synthetic event sets. Of the remaining years, an average positive bias of 15–30% in simulated peaks is considered acceptable to account for the differences between the daily averaged observations and the instantaneous simulations (Fig. 4a). The regional calibration approach is validated at the catchment outlet in Haridwar with a mean peak error of 1.9% (Fig. 4b). The calibration of the GAN basin is judged satisfactory considering the data limitation, and the NAM parameter estimation approach is applied to the 3 remaining basins (KOS, SAR, and YAM).

Inundation depths and extents are principally driven by discharge peaks, topography and roughness. With no adequate calibration data for flood levels and extents, roughness (Manning's n) is treated pragmatically as an uncertain and spatially uniform variable. It is resampled at each location from a truncated lognormal distribution (see Ghosn et al. 2003) with parameters ($\mu = 0.045$; $\sigma = 0.025$ μ; min = 0.035, max = 0.07) derived from land-use and literature guidelines (Chow 1959).

4.2 Synthetic discharge performance evaluation

The performance of the weather generators was evaluated by comparing the flood frequency curves derived from the hydrodynamic model simulations using generated and observed weather data. 500 years of synthetic climate data are first produced using the different WG. The flood frequency curves for the different configurations were estimated by resampling. 300 random samples of length 32

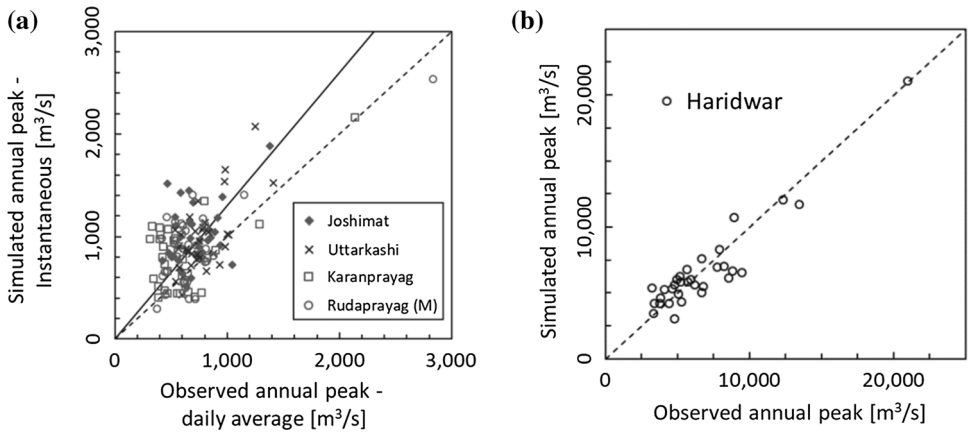

Fig. 4 Simulated versus historical annual peak discharges for **a** calibration at 4 locations in the upstream Ganges catchment, and **b** validation of the regional calibration approach in Haridwar

were taken with replacement from each of the 500 years of annual maxima. Each sample was fitted to the Generalised Extreme Value (GEV) with an estimation of the parameters based on the method of L-moments (Hosking and Wallis 2005). The median was derived from the 300 flood frequency curves in each configuration and compared to the Weibull plotting positions of the annual maxima based on the simulation with observed weather data.

Figure 5a shows the range of variations in flood quantiles generated by the different WG configurations. MULGETS severely underestimates the discharge extremes with both distributions (Gamma or exponential). TripleM performs similarly to the MULGETS with a single rainfall distribution (Weibull) but shows improvement when using a mixture of distributions (Weibull and GPD) to describe low-moderate and high extreme rainfalls. The KNN with a perturbation factor of 0.9 leads to a better preservation of the extremes when using a block resampling of 10 days, which allows to capture better the longer time scale of routing and accumulation processes taking place over the full catchment. The increased perturbation factor of 0.8 (20% of the rainfall amount perturbed) leads to the best representation of the extreme value distribution. Figure 5b depicts the change in cross-correlation between each of the 56 stations and between the historical data and synthetic data. Mean correlation changes vary between − 2 and + 5% with the KNN with block size of 10 and perturbation factor of 0.8 showing the largest increase in correlation. The spatial correlation of the different weather generators is judged well-preserved and suitable for subsequent analysis.

The KNN-CAD weather generator with block size 10 and perturbation factor 0.8 was selected to produce 5000 years of synthetic climate data. To validate the model chain, flood frequency curves were derived at the 4 locations where flood mapping is performed, using the same resampling approach but with 1000 random samples. The median and the 50% and 95% confidence intervals are estimated from the 1000 fitted frequency curves. The resulting frequency curves (Fig. 6) show a very good agreement between the derived flood quantiles and the historical peaks and confirm the suitability of the event model chain configuration selected.

4.3 Impacts on frequencies of discharge events

From the 5000 synthetic years, the largest 1000 peaks—capturing events above the 5-year RP—from each individual location were used for inundation mapping. Event Set A was constructed using the peak occurrence periods and this led to 1866 unique and spatially coherent regional events from which 54% occurred simultaneously at all 4 locations. At each location, discharge RPs were established by rank (from the 1000 largest local peaks) or set to 1 if no event occurred (events with RP below 5 year). For each regional event in Event Set A, local event RPs are plotted between locations (Fig. 7). Event Sets B and C implicitly assume full correlation (uniform T-year) of event RPs locally. The results illustrate the diversity of local event magnitudes that can be expected at the regional scale from the spatially coherent Event Set A.

Table 3 summarizes the regional frequency of occurrences, as they occur in Event Set A, of local event RPs conditional to the occurrence of the local event RP in at least one or all the locations simultaneously. Again, the uniform threshold RP can be considered representative of the Event Sets B or C. The combinations of locations vary from 2 to 4 for threshold RPs ranging from 5 to 500 years. The results show a consistent increase (all locations above the threshold RP) and decrease (at least one location above the threshold RP) when the number of locations under consideration increases. For example, a 100-year event triggered in any of the locations occurs with a 62-year RP frequency across the regional event set while the same

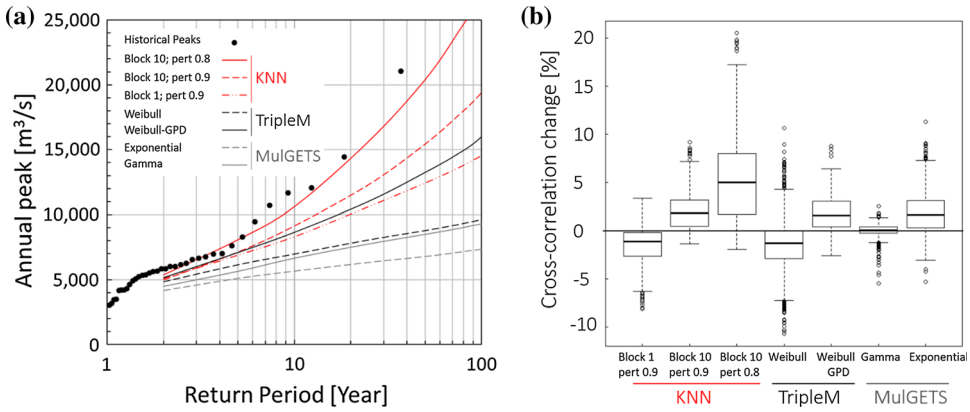

Fig. 5 **a** Comparison of flood frequency curves based on generated and observed weather data for Haridwar and **b** cross-correlations changes of all 56 rainfall time-series between historical and 500 years of synthetic rainfall

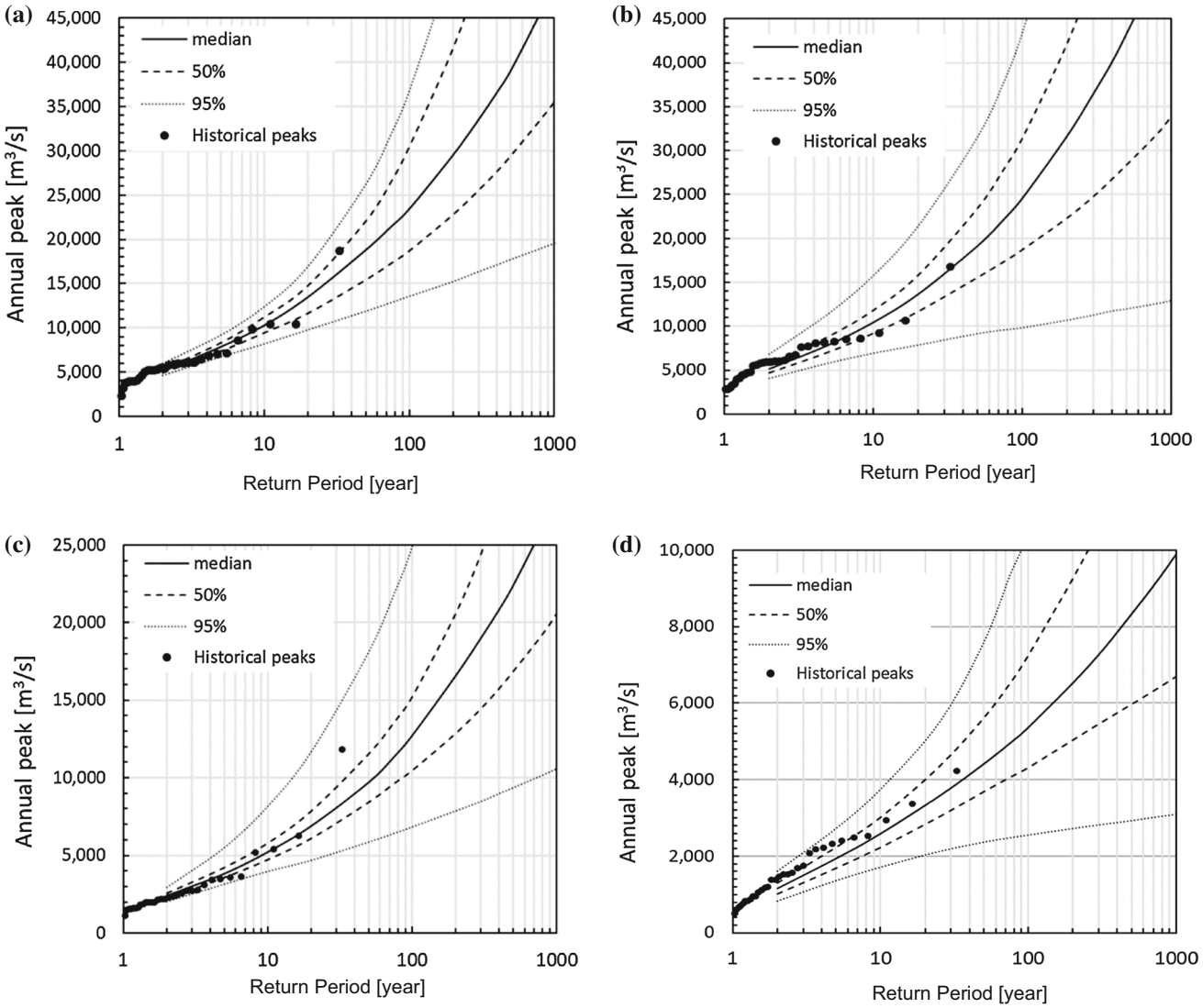

Fig. 6 Comparison of flood frequency curves based on generated and observed weather data: **a** GAN; **b** SAR; **c** YAM; and **d** KOS

100-year triggering threshold occurs with a 278-year RP. A second observation can also be made: the decorrelated local events lead to increase in frequencies of local T-year events at the regional scale (e.g. the 50-year local event would occur every 33 years at the regional scale) which can have implication in the loss estimation.

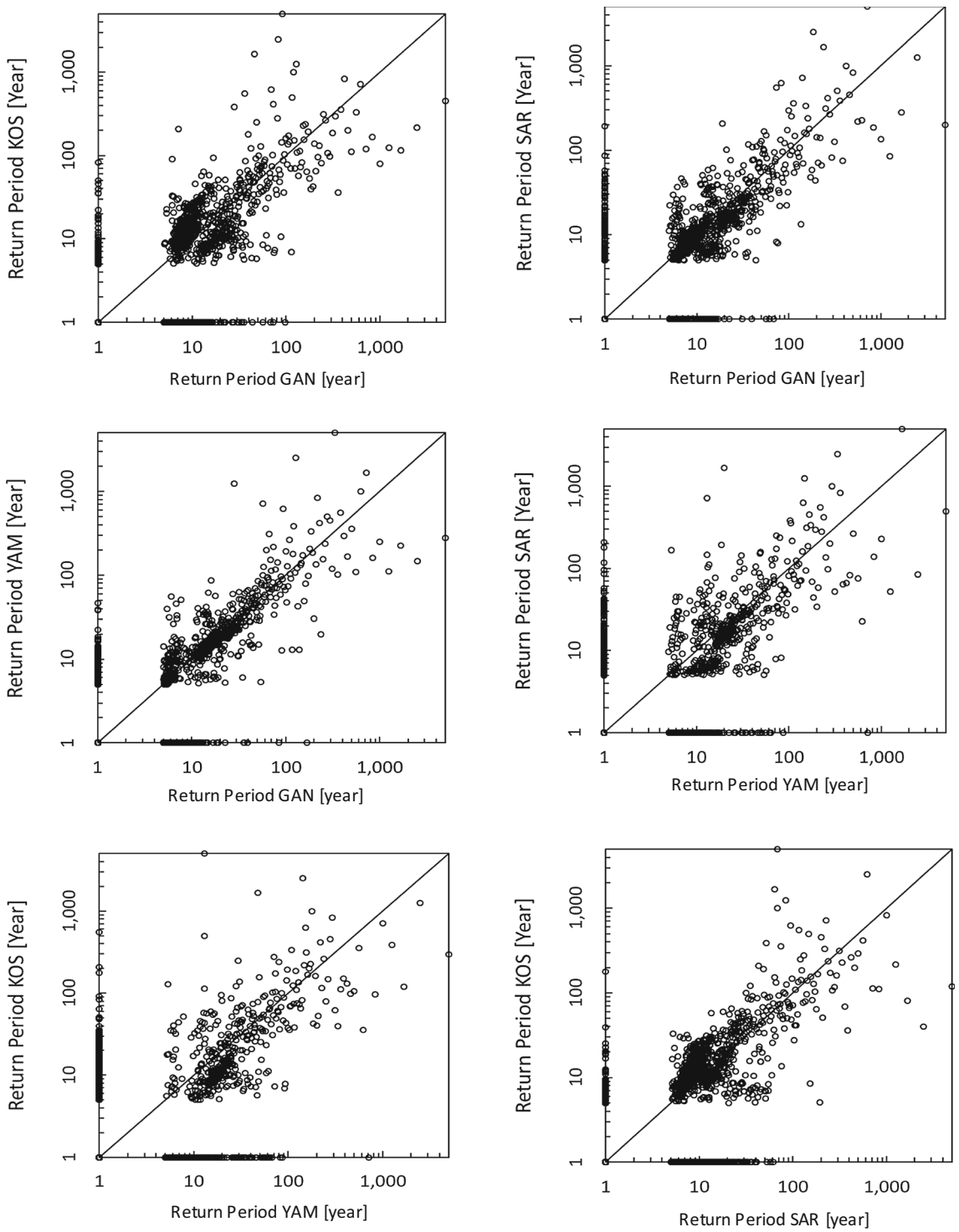

Fig. 7 Correlation plots of peak discharge RP between the different catchments for the spatially coherent Event Set A

4.4 Impacts on risk metrics

Figure 8 shows the exposure distribution, LEP curves and variation of loss frequencies for the different event sets. The exposure exceedance curves (Fig. 8a) illustrate an uneven distribution of exposed assets between locations. Residential exposure is dominated (85%) by the GAN location above the 100-year RP, which corresponds to the protection level at this location. The residential exposure in SAR and YAM increases quickly but stabilizes after the 50-year RP. As a consequence, little variation is seen between exposure exceedance level of Event Sets A and B.

Table 3 Regional events probabilities in the spatial coherent Event Set A for different RP and across different spatial scale

Local event simultaneously above RP in	RP (year)	Mean regional probability (year) of local events above RP considering		
		4 locations	3 locations	2 locations
1 or more catchments	5	3	3	4
All catchments		14	11	9
1 or more catchments	10	6	7	8
All catchments		21	18	15
1 or more catchments	25	15	16	19
All catchments		48	44	37
1 or more catchments	50	33	35	39
All catchments		102	88	71
1 or more catchments	100	62	67	76
All catchments		278	214	159
1 or more catchments	500	179	222	306
All catchments		–	5000	3333

Event Sets B and C can be represented by the spatially uniform RP (column 2)

Despite the spatial variation in local RPs, events of 100-year RP and above in the GAN location are most frequently associated with events above the 25-year RP in the remaining locations, leading to a very similar range of exposure despite the variations in discharge. The agriculture exposure distribution is dominated by the KOS location (> 60%) but more variations exist between locations. The main variations between locations are within the 5 to 30-year RP range leading also to more variation between Event Sets A and B within this range.

The LEP curves (Fig. 8b) show a behavior similar as the exposure exceedance curves. Interestingly, losses start to occur, in Event Set A and for both classes of assets, with a higher level of probability than in other events sets. Event Set A also clearly displays, for agriculture assets, a decrease of the low RPs losses (approximately − 50% and − 10% for the 5 and the 10-year RP respectively). After the 20-year RP the losses fluctuates between ± 4% depending on the scenarios. The differences between Event Sets A and B are minimal for residential assets dominated by losses at a single location (GAN). There, the influence of vulnerability uncertainty (A or B1 against B2) is more important for the higher RPs (> 300-year).

Variations in probabilities for identical loss levels between Event Sets (Fig. 8c) also allow to clearly identify the main zones where differences occur, and the significance of the variabilities considered. The 3 to 5-year RP corresponds in all cases to a decrease in loss probability as Event Set A can generate losses below 5-year RP (Table 3) and no losses are expected to occur below a uniform 5-year RP (Event Set B and C). Loss probabilities decrease between the 5 and 25-year RPs. The 25 to 500-year RP zone shows more stable results between event sets with higher perturbations (negative and positive) around the 100-year mark for residential assets being protected at this RP in the highest exposure location (GAN). After the 500-year RP, losses are overestimated when considering a spatially uniform RP (Event Set B1) and underestimated when not considering vulnerability uncertainty (Event Set B2 and C1).

Table 4 shows the variation in AAL between event sets for the different asset classes. The difference in AAL only due to the spatial variations is negligible (Scenario A vs. B1 and B2). In this application, the increase and decrease in loss frequencies experienced in the LEP curves cancel each other. Still larger differences are experienced in the standard deviation of the individual event losses. The standard deviation of agriculture losses is more than twice the standard deviation produced by the other events sets. This result is a direct consequence of the spatial redistribution of local event intensities in Event Set A. More surprisingly the standard deviation of Event Set A is slightly below (− 25%) for residential assets. This result is attributed to the extreme losses occurring in GAN for a limited number of events and in all event sets (50 events above the 100-year design protection occurring over the 5000 year) and the much large number of low losses with limited variations produced by Event Set A (1840 events) compared to Event Set B (1000 events). The 8 RPs used for AAL estimation (Scenario C1) (less than 1%) present a very good approximation of the risks. For residential assets, the 22% of error for Event Set C1 can be reduced to less than 5% if an additional RP just below the 100-year design

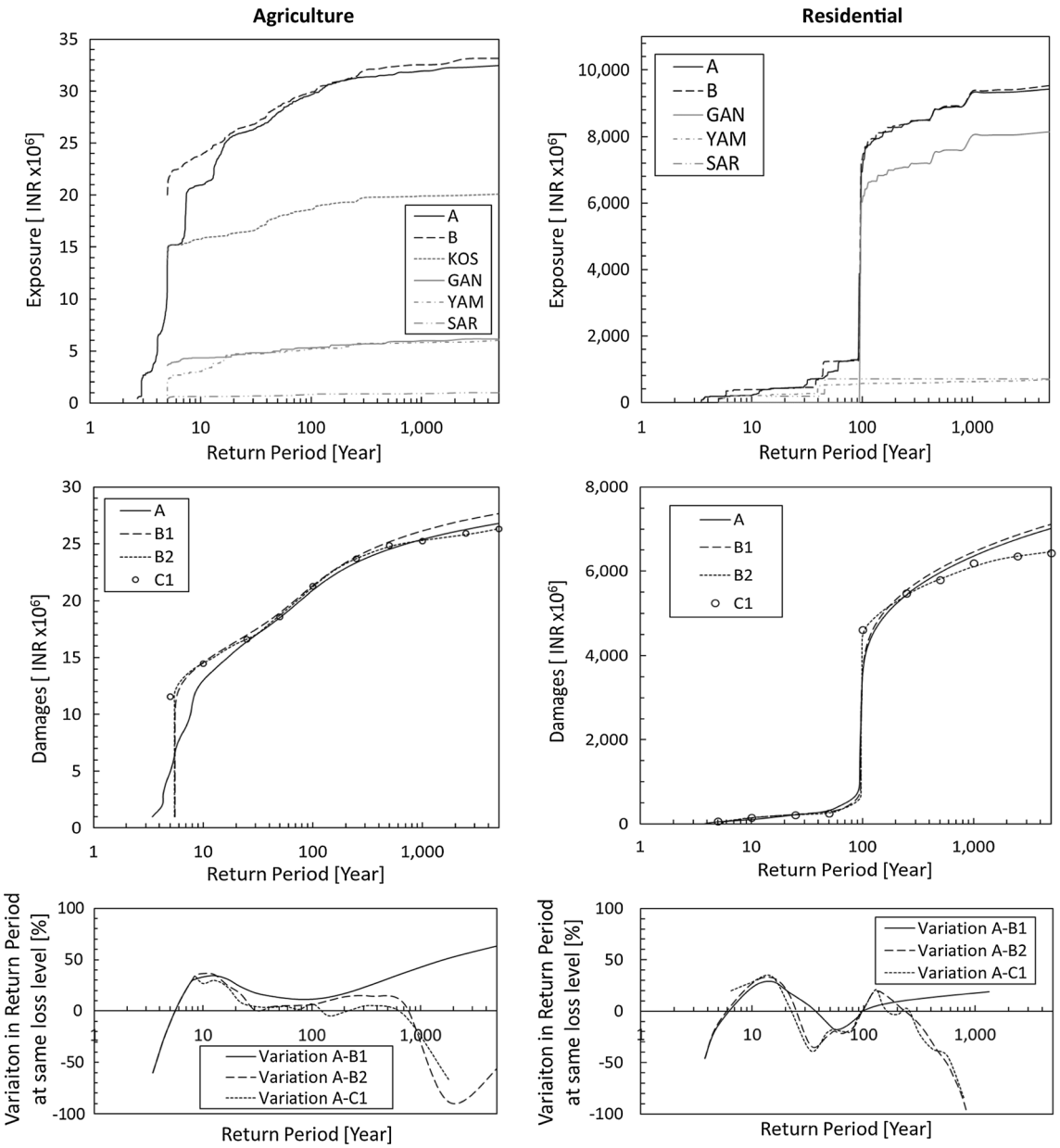

Fig. 8 **a** Exposure exceedance curves, **b** loss exceedance curves and **c** variation in frequencies at the same loss level for agriculture and residential assets and Event Sets A, B.1, B.2 and C.1

Table 4 Average annual losses, standard deviation of losses and percentage variations for agriculture and residential sectors considering the different event sets and sub-sets

Scenario	Agriculture				Residential			
	AAL	σ	AAL variation (%)		AAL	σ	AAL variation (%)	
	(INR × 0^6)		From A	From C.1	(INR × 10^6)		From A	From C.1
A	3.06	6.22	–	–	84.98	848	–	–
B.1	3.06	2.62	0.0	–	84.91	1149	0.0	–
B.2	3.06	2.62	0.0	–	84.59	1148	0.0	–
C.1	3.02	4.94	− 1.0	–	103.82	2805	22.3	–
C.2	2.91	3.32	− 5.0	− 4.0	74.82	1778	− 12.5	− 28.4
C.3	3.34	5.74	9.1	10.1	496.64	2598	484.9	378.4
C.4	1.83	4.38	− 40.0	− 39.4	266.37	2591	213.7	156.6

protection level is used in the integration procedure to account for the sudden loss increase.

5 Discussion

This study presented an efficient probabilistic flood hazard and risk modelling chain relying on 5000 years of continuous climate simulations to produce exhaustive and spatially coherent discharge events at the regional scale. The model was applied to four river basins in Uttarakhand, India to investigate agriculture and residential economic risks. The different elements of the model chain have been developed using expert judgement and the best available datasets, complemented by site surveys. However, in Uttarakhand, as in many regions of the world, there is a lack of consistent and comprehensive data for flood hazard and risk model validation. The river systems in Uttarakhand can be complex with man-made structures (levees) and reservoirs (e.g. Tehri dam) able to assert control on part of the high flow. For clarity, these structures have been simplified in the present analysis. With this in mind, the risk results in absolute terms should be used with care but the relative variations between event sets are considered valid and relevant for the study purpose. Also, the flood hazard mapping relies on the assumption of unidirectional flow, which is considered valid in the study context. Nevertheless, the proposed modelling chain is generally applicable in other areas of the world, but it may be revised in more complex floodplain systems with an efficient 2-dimensional floodplain model.

Overall, the parametric multi-site weather generators tested performed similarly or better than the KNN weather generator in preserving spatial correlation with the added advantage of preserving local extreme distribution. However, in the study context, the KNN type of weather generator using block resampling and perturbations, worked best to reproduce the extreme historical discharge patterns at the catchment scale. The MulGETS and TripleM underestimated, in some case considerably, the discharge extremes. The difference of performance of the MulGETS when compared to the results of Chen et al. (2015), can be attributed to the larger number and more dense rainfall data applied here (28 locations for a catchment area of 21,100 km^2 against 15 locations over an area of 73,800 km^2) and the large difference in rainfall pattern between the Himalayan monsoon and Quebec. It also illustrates the limitation of the Markov chain process to deal with spatial variability (e.g. Wilks 1998; Brissette et al. 2007; Khalili et al. 2009). Parametric models might need further attention to be able to reproduce the geospatial and temporal pattern of heavy and extreme rainfalls and the extreme discharge accumulation if those models are to be used in flood risk studies.

The continuous climatic and hydrologic simulations were used to investigate risk metrics derived from the continuous and spatially coherent event set and from the application of T-year return periods events uniformly at the regional scale. From a discharge event perspective, preserving the spatial coherency at the regional scale led to a spreading of local T-year event distributions (and their possible losses) on both sides of a regionally uniform T-year RP (i.e. a 100-year event will be exceeded regionally with a 278-year RP but local 100-year events can occur with a 62-year RP regionally).

The specificity of the exposure distributions and the correlations between local events magnitude translated into smaller differences (2–4%) in losses at high RPs (above 50-year) but larger differences (20%) at low RPs (5 to 10-year). These results are consistent to the results of Theiken et al. (2015) and Schneeberger et al. (2017) although the differences of the spatially varied event set are much less marked for the high RPs due to the high discharge correlation, exposure concentration and limited number of locations. Most importantly, this study illustrated another phenomenon that differentiates spatially coherent and uniform RP event sets: the redistribution of losses in the spatially coherent event set led to a shift in the minimum RP generating losses from 5-year (assuming a uniform T-year event set) to 3-year corresponding to the de-correlated events occurring at individual locations. These variations (both positive and negative) led to negligible differences between the average annual losses derived from the spatially coherent and continuous event set or appropriate uniform T-year events.

Here the number and choice of the T-year RPs should be made with care. The use of too few or not adequate RPs could lead to large errors ($-$ 40 to $+$ 200%) in the regional averaged risk metrics. These results, aligned with the work of Ward et al. (2011), illustrate the importance of selecting the right RPs to describe more accurately the LEP curve: the RPs should be (1) as low as possible to effectively capture where losses start (e.g. in this study a 1 or 2-year RP threshold could have provided more accurate results), and (2) adequately resolve sudden changes in losses caused by protection design criteria used for critical assets groups (e.g. 100-year flood protection of residential assets in the Upper Ganges basin).

Finally, the spatially coherent event set produced higher standard deviation (above 200% increase) compared with uniform T-year RP event sets. These differences could adversely affect financial plans such as contingency funds or risk transfer schemes designed using such risk metric.

6 Conclusion

In summary, the spatial variability of flood events, exposure and the critical frequencies at which changes occur in the risk metrics and across different spatial scales need to be better understood and captured. The present study investigated the impact of spatial variability in the risk modelling chain and risk cumuli. It identified a possible modelling approach and the different tools available with their limitations. Some challenges remain, for example, proper representation of extreme hydrological events at the regional scale—using weather generators or joint probability approaches, and efficient simulations of a large number of inundation events. Many studies address these challenges individually, but more studies are required to develop, investigate and promote the integration of these different parts in the risk modelling chain.

Acknowledgements The study was funded by the Economic Development Board (EDB) of Singapore (IPP Scholarship) and the DHI-NTU Research Centre and Education Hub. The work is closely affiliated with the Disaster Risk Assessment of Uttarakhand, an initiative of the Uttarakhand Disaster Recovery Programme (UDRP) funded by the World Bank. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the World Bank or the UDRP. The authors acknowledge the support of the UDRDP leadership team. Special thanks are directed to Mr. Amit Singh Negi, Program Director, and Dr. Girish Chandra Joshi of the UDRP for their support and kind permission to share the outcomes of our work in the public domain. This work directly complements their extensive risk mitigation programme in Uttarakhand.

References

American Society of Civil Engineers (1996) Hydrology handbook, 2nd edn. ASCE, Reston

Bannayan M, Hoogenboom G (2008) Weather analogue: a tool for real-time prediction of daily weather data realizations based on a modified k-nearest neighbor approach. Environ Model Softw 23(6):703–713

Bárdossy A (2007) Calibration of hydrological model parameters for ungauged catchments. Hydrol Earth Syst Sci Discuss 11(2):703–710

Bookhagen B, Burbank DW (2006) Topography, relief, and TRMM-derived rainfall variations along the Himalaya. Geophys Res Lett 33(8):L08405

Breinl K, Turkington T, Stowasser M (2013) Stochastic generation of multi-site daily precipitation for applications in risk management. J Hydrol 498:23–35

Breinl K, Turkington T, Stowasser M (2015) Simulating daily precipitation and temperature: a weather generation framework for assessing hydrometeorological hazards. Meteorol Appl 22(3):334–347

Breinl K, Di Baldassarre G, Lopez MG, Hagenlocher M, Vico G, Rutgersson A (2017) Can weather generation capture precipitation patterns across different climates, spatial scales and under data scarcity? Sci Rep 7(1):5449

Brissette FP, Khalili M, Leconte R (2007) Efficient stochastic generation of multi-site synthetic precipitation data. J Hydrol 345(3–4):121–133

Buishand TA, Brandsma T (2001) Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. Water Resour Res 37(11):2761–2776

Carreau J, Bouvier C (2016) Multivariate density model comparison for multi-site flood-risk rainfall in the French Mediterranean area. Stoch Environ Res Risk Assess 30(6):1591–1612

Chen J, Brissette FP, Zhang XJ (2014) A multi-site stochastic weather generator for daily precipitation and temperature. Trans ASABE 57(5):1375–1391

Chen J, Brissette FP, Zhang XJ (2015) Hydrological modeling using a multisite stochastic weather generator. J Hydrol Eng 21(2):04015060

Chow V (1959) Open channel hydraulics. McGraw-Hill, New York

Clark MP, Gangopadhyay S, Brandon D, Werner K, Hay L, Rajagopalan B, Yates D (2004) A resampling procedure for generating conditioned daily weather sequences. Water Resour Res 40(4):W04304

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London

Das PK (2013) The Himalayan Tsunami"–Cloudburst, flash flood & death toll: a geographical postmortem. IOSR J Environ Sci Toxicol Food Technol 7:33–45

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597

Dhar ON, Nandargi S (2004) Rainfall distribution over the Arunachal Pradesh Himalayas. R Meteorol Soc Weather 59(6):155–157

DHI (2017) M11, a modelling system for rivers and channels. Reference manual. DHI Water & Environment, Hørsholm

DHI, ERN, AIT Joint Venture (2017) Typologies, vulnerability functions and costs applied in the DRA. Technical Project Report (Component 2, Output 1). Uttarakhand Disaster Recovery Project, India

European Commission (2007) Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. Official Journal of the European Union

Evaluación de Riesgos Naturales América Latina-ERN-AL (2009a) Vulnerabilidad de edificaciones e infraestructura. Tomo I—Metodologia de Modelacion Probabilistica de Riesgos Naturales. Informe Técnico ERN-CAPRA-T1-5. https://ecapra.org. Accessed 31 Oct 2018

Evaluación de Riesgos Naturales América Latina-ERN-AL (2009b) Metodo de Analisis Probabilista del Riesgo. Tomo I—Metodologia de Modelacion Probabilistica de Riesgos Naturales. Informe Técnico ERN-CAPRA-T1-6. https://ecapra.org. Accessed 31 Oct 2018

Evaluación de Riesgos Naturales América Latina-ERN-AL (2011) CAPRA-GIS v2.0. Program for the probabilistic risk assessment. https://ecapra.org. Accessed 31 Oct 2018

Evin G, Favre AC, Hingray B (2018) Stochastic generation of multi-site daily precipitation focusing on extreme events. Hydrol Earth Syst Sci 22(1):655

Falter D, Schröter K, Dung NV, Vorogushyn S, Kreibich H, Hundecha Y, Merz B (2015) Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. J Hydrol 524:182–193

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S et al (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2:150066

Ghizzoni T, Roth G, Rudari R (2012) Multisite flooding hazard assessment in the Upper Mississippi River. J Hydrol 412:101–113

Ghosn M, Moses F, Wang J (2003) NCHRP report: design of highway bridges for extreme events. Transportation Research Board of the National Academics, Washington

Grimaldi S, Petroselli A, Arcangeletti E, Nardi F (2013) Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. J Hydrol 487:39–47

Haberlandt U, Radtke I (2014) Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows. Hydrol Earth Syst Sci 18(1):353–365

Hong Y, Adler RF (2008) Estimation of global SCS curve numbers using satellite remote sensing and geospatial data. Int J Remote Sens 29(2):471–477

Horritt MS, Bates PD (2002) Evaluation of 1D and 2D numerical models for predicting river flood inundation. J Hydrol 268(1):87–99

Hosking JRM, Wallis JR (2005) Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge

Hou AY, Kakar RK, Neeck S, Azarbarzin AA, Kummerow CD, Kojima M et al (2014) The global precipitation measurement mission. Bull Am Meteorol Soc 95(5):701–722

Huizinga J, Moel H de, Szewczyk W (2017) Global flood depth-damage functions. Methodology and the database with guidelines. JRC Working Papers JRC105688. Joint Research Centre, Seville

Hundecha Y, Merz B (2012) Exploring the relationship between changes in climate and floods using a model-based analysis. Water Resour Res 48(4):W04512

Hundecha Y, Zehe E, Bárdossy A (2002) Regional parameter estimation from catchment properties for the prediction of ungauged basins. In: Proceedings of the PUB Kick-off meeting, pp 20–22

Hundecha Y, Pahlow M, Schumann A (2009) Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. Water Resour Res 45(12):W12412

Iman RL, Conover WJ (1982) A distribution-free approach to inducing rank correlation among input variables. Commun Stat Simul Comput 11(3):311–334

Jongman B, Hochrainer-Stigler S, Feyen L, Aerts JC, Mechler R, Botzen WW et al (2014) Increasing stress on disaster-risk finance due to large floods. Nat Clim Change 4(4):264

Khalili M, Brissette F, Leconte R (2009) Stochastic multi-site generation of daily weather data. Stoch Environ Res Risk Assess 23(6):837–849

King LM, McLeod AI, Simonovic SP (2015) Improved weather generator algorithm for multisite simulation of precipitation and temperature. JAWRA J Am Water Resour Assoc 51(5):1305–1320

Kron W (2005) Flood risk = hazard• values• vulnerability. Water Int 30(1):58–68

Kumar V (2017) Statistical distribution of rainfall in Uttarakhand, India. Appl Water Sci 7(8):4765–4776

Kundzewicz ZW, Pińskwar I, Brakenridge GR (2013) Large floods in Europe, 1985–2009. Hydrol Sci J 58(1):1–7

Leander R, Buishand TA (2007) Resampling of regional climate model output for the simulation of extreme river flows. J Hydrol 332(3–4):487–496

Lu Y, Qin XS (2014) A coupled K-nearest neighbour and Bayesian neural network model for daily rainfall downscaling. Int J Climatol 34(11):3221–3236

Luo P, He B, Takara K, Xiong Y, Nover D, Duan W, Fukushi K (2015) Historical assessment of Chinese and Japanese flood management policies and implications for managing future floods. Environ Sci Policy 48:265–277

Luo P, Apip B He, Duan W, Takara K, Nover D (2018a) Impact assessment of rainfall scenarios and land-use change on hydrologic response using synthetic Area IDF curves. J Flood Risk Manag 11(S1):84–97

Luo P, Mu D, Xue H, Ngo-Duc T, Dang-Dinh K, Takara K, Nover D, Schladow G (2018b) Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions. Sci Rep 8(1):12623

Madsen H (2000) Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. J Hydrol 235(3–4):276–288

Mandal S, Breach P, Gaur A, Simonovic S (2017) Tools for downscaling climate variables: a technical manual. Water resources research report no. 097. Department of Civil and Environmental Engineering, Western University, Canada

Marulanda MC, Carreno ML, Cardona OD, Ordaz MG, Barbat AH (2013) Probabilistic earthquake risk assessment using CAPRA: application to the city of Barcelona, Spain. Nat Haz 69(1):59–84

Merz B, Hall J, Disse M, Schumann A (2010) Fluvial flood risk management in a changing world. Nat Hazard Earth Syst 10(3):509

Meyer V, Haase D, Scheuer S (2009) Flood risk assessment in European river basins—concept, methods, and challenges exemplified at the Mulde river. Integr Environ Assess Manag 5(1):17–26

Neal J, Keef C, Bates P, Beven K, Leedal D (2013) Probabilistic flood risk mapping including spatial dependence. Hydrol Process 27(9):1349–1363

Nielsen SA, Hansen E (1973) Numerical simulation of the rainfall runoff process on a daily basis. Nordic Hydrol 4:171–190

Oliver J, Qin XS, Larsen O, Meadows M, Fielding M (2018) Probabilistic flood risk analysis considering morphological dynamics and dike failure. Nat Hazards 91(1):287–307

Pai DS, Sridhar L, Rajeevan M, Sreejith OP, Satbhai NS, Mukhopadhyay B (2014) Development of a new high spatial resolution (0.25×0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. Mausam 65(1):1–18

Parida BR, Behera SN, Bakimchandra O, Pandey AC, Singh N (2017) Evaluation of satellite-derived rainfall estimates for an extreme rainfall event over Uttarakhand, Western Himalayas. Hydrology 4(2):22

Prakash S, Mitra AK, Pai DS, Kouchak A (2016) From TRMM to GPM: how well can heavy rainfall be detected from space? Adv Water Resour 88:1–7

Rana N, Singh S, Sundriyal YP, Juyal N (2013) Recent and past floods in the Alaknanda valley: causes and consequences. Curr Sci 105(9):1209–1212

Rawls WJ, Brakensiek DL (1989) Estimation of soil water retention and hydraulic properties. Unsaturated flow in hydrologic modeling. Springer, Dordrecht, pp 275–300

Reese S, Ramsay D (2010) RiskScape: flood fragility methodology. National Institute of Water and Atmospheric Research, Wellington, p 42

Rulin O, Liliang R, Weiming C, Zhongbo Y (2008) Application of hydrological models in a snowmelt region of Aksu River Basin. Water Sci Eng 1(4):1–13

Salgado-Gálvez MA, Zuloaga-Romero D, Bernal GA, Mora MG, Cardona OD (2014) Fully probabilistic seismic risk assessment considering local site effects for the portfolio of buildings in Medellín, Colombia. Bull Earthq Eng 12(2):671–695

Schneeberger K, Huttenlau M, Winter B, Steinberger T, Achleitner S, Stötter J (2017) A probabilistic framework for risk analysis of widespread flood events: a proof-of-concept study. Risk Anal 39(1):125–139

Serinaldi F, Kilsby CG (2017) A blueprint for full collective flood risk estimation: demonstration for European river flooding. Risk Anal 37(10):1958–1976

Sharma A, Lall U (1999) A nonparametric approach for daily rainfall simulation. Math Comput Simul 48(4–6):361–371

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM national conference. ACM, pp 517–524

Soille P, Vogt J, Colombo R (2003) Carving and adaptive drainage enforcement of grid digital elevation models. Water Resour Res 39(12):1366

Sparks NJ, Hardwick SR, Schmid M, Toumi R (2018) IMAGE: a multivariate multi-site stochastic weather generator for European weather and climate. Stoch Environ Res Risk Assess 32(3):771–784

Thieken AH, Apel H, Merz B (2015) Assessing the probability of large-scale flood loss events: a case study for the river Rhine, Germany. J Flood Risk Manag 8(3):247–262

Timonina A, Hochrainer-Stigler S, Pflug G, Jongman B, Rojas R (2015) Structured coupling of probability loss distributions: assessing joint flood risk in multiple river basins. Risk Anal 35(11):2102–2119

Unphon H (2009) Re-engineering for evolvability. Dissertation, IT University of Copenhagen

Vallam P, Qin XS (2016) Multi-site rainfall simulation at tropical regions: a comparison of three types of generators. Meteorol Appl 23(3):425–437

Velásquez CA, Cardona OD, Mora MG, Yamin LE, Carreño ML, Barbat AH (2012) Hybrid loss exceedance curve (HLEC) for disaster risk assessment. Nat Haz 72(2):455–479

Wang L, Liu HL, Bao AM, Pan XL, Chen X (2016) Estimating the sensitivity of runoff to climate change in an alpine-valley watershed of Xinjiang, China. Hydrol Sci J 61(6):1069–1079

Ward PJ, De Moel H, Aerts JCJH (2011) How are flood risk estimates affected by the choice of return-periods? Nat Hazards Earth Syst 11(12):3181

Watkins DS (2010) Fundamentals of matrix computations, vol 64. Wiley, New York

Wilks DS (1998) Multisite generalization of a daily stochastic precipitation generation model. J Hydrol 210(1–4):178–191

Wit MJM, Buishand TA (2007) Generator of rainfall and discharge extremes (GRADE) for the Rhine and Meuse basins. Rijkswaterstaat RIZA report 2007.027/KNMI publication 218, Lelystad, The Netherlands

Yates D, Gangopadhyay S, Rajagopalan B, Strzepek K (2003) A technique for generating regional climate scenarios using a nearest-neighbor algorithm. Wat Resou Res 39(7):WR001769