

A vine copula-based ensemble projection of precipitation intensity–duration–frequency curves at sub-daily to multi-day time scales

Boen Zhang¹, Shuo Wang^{1,2,*}, Hamid Moradkhani³, Louise Slater⁴, and Jiafeng Liu⁵

¹Department of Land Surveying and Geo-Informatics and Research Institute for Land and Space,
The Hong Kong Polytechnic University, Hong Kong, China

²Shenzhen Research Institute, The Hong Kong Polytechnic University, Shenzhen, China

³Department of Civil, Construction, and Environmental Engineering, Center for Complex
Hydrosystems Research, The University of Alabama, Tuscaloosa, AL, USA

⁴School of Geography and the Environment, University of Oxford, Oxford, UK

⁵China Aero Geophysical Survey and Remote Sensing Center, Beijing, China

*Corresponding author: Phone: (852) 3400-3896; Email: shuo.s.wang@polyu.edu.hk

Key Points:

- Stochastic spatiotemporal downscaling of precipitation is achieved to derive IDF curves at sub-daily to multi-day time scales.
- A vine copula multi-model ensemble approach is proposed to improve the accuracy and reliability of climate-model-based IDF curves.
- Extreme precipitation is projected to increase by up to 30% in intensity and nearly two times in frequency over 196 Chinese cities.

Abstract

Precipitation intensity–duration–frequency (IDF) curves play a crucial role in the design and planning of urban infrastructure to reduce the risk of urban flooding and rainfall-triggered landslides. However, changing rainfall characteristics in a warming climate render conventional IDF curves inappropriate due to the statistical assumption of stationarity. In this study, we develop vine copula-based projections of future IDF curves at sub-daily to multi-day time scales with a multi-model ensemble of five regional climate simulations over China. Stochastic spatiotemporal downscaling of precipitation is achieved to generate extreme precipitation simulations at a high spatial (0.1°) and temporal (3 hourly) resolution. These downscaled simulations are combined by the vine copula to improve the reliability and accuracy of climate-model-based IDF curves relative to historical observations. Our findings reveal that climate-model-based stochastic downscaling of precipitation reproduces the IDF curves well based on historical observations in China. The vine copula multi-model ensemble approach outperforms Bayesian model averaging (BMA) by generating more accurate and reliable IDF curves. The urban areas of 196 Chinese cities are projected to experience an increase in extreme precipitation of up to 30% in intensity and nearly two times the frequency of historical events under a high emission scenario (RCP8.5). The current urban infrastructure of more than half of the 196 cities would thus be inadequate to prevent losses caused by rainfall-triggered hazards if designed solely based on historical precipitation observations. Compared to the climate-model-based IDF curves, we find that statistical IDF curves based on a nonstationary time covariate (i.e., extrapolating historical trends) are likely to underestimate the risk of urban infrastructure failures under a warming climate. This work highlights that urban infrastructure design guidelines in China should be upgraded to adapt existing IDF curves to the changing climate.

43 **Keywords:** IDF curves; Climate projection; Copula; Stochastic downscaling

1. Introduction

The past decade has witnessed a number of rare heavy rainfall events in many countries leading to catastrophic flooding or landslides, as exemplified by recent events in Zhengzhou, China (Yin et al., 2022), New York, USA (The Washington Post, 2021), western Germany (Reuters, 2021), Milan Malpensa Airport, Italy (Wanted in Milan, 2021), and Singapore (FloodList, 2021). Zhengzhou was hit by massive flooding on 20 July 2021 due to an unprecedented 622.7 mm of rain that fell in 24 hours, which led to 292 known deaths and affected more than 11 million people, despite having invested over \$80 million since 2016 in upgrading urban infrastructure (Li & Wang, 2022; Yin et al., 2022). The intensity and frequency of heavy rainfall events are expected to increase under a warming climate, as highlighted in recent studies and the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Qing et al., 2022; Seneviratne et al., 2021; Zhu et al., 2018). This raises the question of whether the conventional design of urban infrastructure is still adequate in the context of climate change (Mondal & Daniel, 2019; Switzman et al., 2017; Wright et al., 2021; You & Wang, 2022; Zhang et al., 2022).

Current infrastructure engineering design relies on precipitation intensity-duration-frequency (IDF) curves to estimate rainfall intensities that the structures will be designed to withstand during their lifetime (Hassanzadeh et al., 2014, 2019; Ombadi et al., 2018; Ouarda et al., 2019; Ritschel et al., 2017; Q. Xu et al., 2018). IDF curves graphically represent estimates of probable precipitation intensity associated with different durations and return periods, the estimation of which involves fitting a representative probability distribution (e.g., the generalized extreme value (GEV) distribution) to annual maximum precipitation time series (Butcher et al., 2021; Marra & Morin, 2015; Ritschel et al., 2017; So et al., 2017; Srivastav et al., 2014; Zhou et

al., 2019). As extreme precipitation intensifies with warming at a rate consistent with the increase in atmospheric moisture, the assumption of stationarity used in the traditional development of IDF curves has been challenged (Boukhelifa et al., 2018; Courty et al., 2019; Innocenti et al., 2017).

To address the challenge of developing nonstationary IDF curves, time-varying covariates such as time and mean temperature have been employed (Cheng & Aghakouchak, 2014; Ouarda et al., 2019; Sarhadi & Soulis, 2017; Slater, Anderson, et al., 2021; Slater, Villarini, et al., 2021; Srivastava et al., 2021). Such IDF estimates rely on the statistical assumption that the nonstationary model adequately represents the changing properties of rainfall extremes. Such an assumption, however, may be invalid for the future period, and simple extrapolation of historical trends may lead to erroneous spatiotemporal projections of future rainfall extremes (Ganguli & Coulibaly, 2017; Luke et al., 2017; Serinaldi & Kilsby, 2015). In contrast, a second approach for estimating future IDF curves, i.e., climate-model-based IDF curves, directly uses climate simulations under present and future climates to address the nonstationary behavior of extreme precipitation (Gaur et al., 2018, 2020; Mailhot et al., 2007; Ragno et al., 2018; Simonovic et al., 2016). Such a method provides a physics-based avenue for projection and has been used recently as a promising alternative to purely statistical projections based on temporal extrapolation of underlying trends (Bezak et al., 2016; Hosseinzadehtalaei et al., 2018, 2020; Lima et al., 2018).

Although the aforementioned methodological advances have been achieved in the development of IDF curves, little effort has been devoted to evaluating urban IDF curves on sub-daily scales and projecting their response to the changing climate, especially on a national scale (Harpold & Kohler, 2017; Jalowska & Spero, 2019; Le et al., 2018; Sarhadi & Soulis, 2017). A major reason is the limited availability of long rainfall records with a reasonable level of

accuracy and fine time scales (i.e., sub-daily). The resolutions of widely used general circulation models (GCMs) in the Coupled Model Intercomparison Project Phase 5 and 6 (CMIP5 and CMIP6) and RCMs in the Coordinated Regional Downscaling Experiment (CORDEX) are not fine enough to assess future changes of urban IDF curves on sub-daily time scales (Chandra et al., 2015; Ganguli & Coulibaly, 2019; Ragno et al., 2018). It is thus necessary to downscale the outputs from GCMs or RCMs to the desired spatiotemporal resolutions (Agilan & Umamahesh, 2016; J. Li et al., 2017; Yan et al., 2021). In addition, the substantial spread and bias of climate models are not well constrained in previous studies (Hosseinzadehtalaei et al., 2018; Ragno et al., 2018; Srivastava et al., 2021). A universal, robust method for developing a multi-model projection ensemble is unavailable, even though a variety of model weighting strategies have been proposed, such as Bayesian model averaging (BMA) (Raftery et al., 2005). In BMA, the conditional probability density function (PDF) of each model is commonly assumed to follow a normal or gamma distribution, and such an assumption may be invalid for rainfall extremes (Madadgar & Moradkhani, 2014). Another limitation of BMA is that data transformation is usually needed to transform raw model predictions to the space of posterior distribution. To cope with these limitations, Madadgar and Moradkhani (2014) proposed an approach that integrates the copula functions with BMA to estimate the posterior distribution of model predictions without a need to transform the model predictions and assuming the form of the posterior distribution. Similarly, the vine copula approach enables a flexible assessment of complex interactions among multiple rainfall extremes without any assumption regarding the type of marginal probability distributions (Bevacqua et al., 2017; Hatami & Nazemi, 2022; Sun et al., 2021; Zaerpour et al., 2021). Therefore, it remains to be seen whether a vine copula-based multi-

model ensemble approach may improve the reliability of climate-model-based IDF curves, relative to more traditional ensemble approaches.

In this work, we develop a vine copula-based ensemble projection of IDF curves at sub-daily to multi-day time scales. Specifically, stochastic spatiotemporal downscaling of daily precipitation, generated from five regional climate simulations in the CORDEX experiment, is developed for deriving urban IDF curves over China from 3-hour to 7-day time scales. A vine copula multi-model ensemble approach is also proposed to more accurately downscale rainfall extremes at different time scales, thereby improving the accuracy and reliability of climate-model-based IDF curves. We assess future changes of the IDF curves averaged over the urban areas of 196 Chinese cities under a high-emission scenario (RCP8.5). Furthermore, we also compare the differences between the climate-model-based and covariate-based methods in terms of projecting future changes in the urban IDF curves.

The paper is organized as follows. Section 2 describes the spatiotemporal downscaling method, the vine copula multi-model ensemble IDF curves, performance metrics for evaluating IDF curves, and data sets used in this study. Section 3 presents the evaluation of multi-model IDF curves and Section 4 discusses projected future urban IDF curves for 196 cities in China. Section 5 discusses the results and presents a comparison between the climate-model-based and time covariate-based IDF curves. Finally, Section 6 summarizes our study and highlights the main findings.

2. Methods and Data Sources

2.1 Spatiotemporal downscaling of precipitation

To investigate climate change impacts on urban IDF curves, a spatiotemporal downscaling method is developed to generate reliable precipitation at a fine spatiotemporal scale based on regional climate simulations. The spatiotemporal downscaling consists of a bias-correction spatial disaggregation (BCSD) method for spatial downscaling and a precise temporal disaggregation method for breaking down daily precipitation into sub-daily time series.

To perform spatial downscaling of precipitation using the BCSD method, the fine-scale precipitation observations are regridded to match the coarse-resolution precipitation simulations, and then quantile mapping is constructed between the observed and simulated precipitation for each grid cell and each day of the year (DOY) with a ± 15 -day window centered over the respective day (Wood et al., 2002). This well-established mapping approach is then used to correct the regridded past and future precipitation simulations to the same distributions as the observed precipitation, and bilinearly interpolate the simulations to the same grid as the precipitation observations. To preserve spatial details of the fine-scale precipitation, a scaling factor (i.e., the ratio of the simulated precipitation average to the observed precipitation average) is calculated for each DOY and then multiplied by the daily interpolated precipitation simulations for past and future climates, thereby generating precipitation simulations at a fine spatial scale.

The spatially downscaled daily precipitation requires temporal disaggregation to be used for developing IDF curves at sub-daily scales. To this end, the sub-daily (e.g., 3-hour) precipitation process is assumed as a cyclostationary process with monthly stationarity. Let $\{R_c(t) | t \in T, c = 1, \dots, N\}$ represent the cyclostationary precipitation process, in which T denotes an indexed set

and c is cyclically varying index over t . Under such an assumption, the sub-daily precipitation process can be decomposed into $N = 12$ stationary stochastic processes (Papalexiou, 2018). Each of the twelve processes describes the precipitation process for each month and is represented by reproducing its marginal distribution function $F_R(r)$ and the autocorrelation structure (ACS) $\rho_R(\tau)$. In addition, each of the twelve processes has a parent-Gaussian process $\{X(t)\}$ with a standard Gaussian marginal $\Phi_X(x)$ and an ACS $\rho_X(\tau)$. Since it is easy to perform random simulations of Gaussian processes, the basic idea of the stochastic temporal downscaling method is to identify two functions f and g such that $R(t) = f(X(t))$ and $\rho_X(\tau) = g(\rho_R(\tau))$, thereby using the two functions to transform a Gaussian random variable into sub-daily precipitation. These two functions enable transformations from the marginal distribution and the linear ACS of the Gaussian process $\{X(t)\}$ to those of the process $\{R(t)\}$, respectively. The function f is easily identified because there is a well-known data transformation approach, namely Gaussian anamorphosis, which can be used to transform a Gaussian random variable X into the sub-daily precipitation R with an arbitrary distribution (e.g., the generalized gamma distribution in this study) (Wang et al., 2018):

$$R_c = G_{R_c}^{-1}(\Phi_X(X)) \quad (1)$$

where G_{R_c} is the theoretical cumulative distribution function (CDF) of sub-daily precipitation for month c . Therefore, the stochastic temporal downscaling method only requires a well-defined function g to transform the linear ACS.

Let $R(t)$ and $R(\tau) := R(t-\tau)$ be a pair of $\{R(t)\}$ random precipitation processes at an arbitrary time point t , lagged by τ , and having a correlation coefficient $\rho_R(\tau)$ given by

$$\rho_R(\tau) = \frac{E(R(t)R(\tau)) - \mu_R^2}{\sigma_R^2} \quad (2)$$

where μ_R and σ_R represent the mean and standard deviation of R , respectively. $E(R(t)R(\tau))$ represents the expectation and can be estimated using the fundamental probability theorem (Feller, 1971). According to the theorem, given a real function g such that $R = g(X)$, the expectation of R can be expressed as

$$E(R(t)) = E(g(X(t))) = \int_{-\infty}^{\infty} g(x) f_R(r) dy \quad (3)$$

where $f_R(r)$ represents the known PDF of R . Since equation 2 can be applied in two dimensions, $E(R(t)R(\tau))$ in equation 1 can be expressed as

$$\begin{aligned} E(R(t)R(\tau)) &= E\left(Q_R\left(\Phi_X(X(t))\right)Q_R\left(\Phi_X(X(\tau))\right)\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_R\left(\Phi_X(X(t))\right)Q_R\left(\Phi_X(X(\tau))\right) \\ &\quad \times \varphi_{X(t)X(\tau)}(x(t), x(\tau); \rho_X(\tau)) dx(t) dx(\tau) \end{aligned} \quad (4)$$

where $\Phi_X(X(t)) = (1 - \text{erf}(-x/\sqrt{2}))/2$. Equation 4 is inserted into equation 2, leading to an autocorrelation transformation function (ACTF) that enables the calculation of $\rho_R(\tau)$ from $\rho_X(\tau)$. However, the stochastic temporal downscaling of precipitation seeks to transform the parent-Gaussian process $\{X(t)\}$ into the desired process $\{R(t)\}$. It is thus necessary to quantify the inverse relationship for transforming the autocorrelation value of sub-daily precipitation ρ_R into the autocorrelation value of parent-Gaussian random variables $\rho_X(\tau)$. For simplicity and numerical stability, the ACTF can be expressed as a simple parametric function, i.e.,

$$\rho_X = \frac{(1 + a\rho_R)^{1-b} - 1}{(1 + a)^{1-b} - 1} \quad (5)$$

To estimate the unknown parameters a and b , a set of (ρ_R, ρ_X) pairs can be generated by making numerical estimates of equation 4 along with equation 2 for a fixed number of ρ_X values (Papalexiou, 2018).

The ACSs $\rho_R(\tau)$ of the process $\{R(t)\}$ is assessed using the observed sub-daily precipitation time series. For instance, the Pareto II ACS is used in this study:

$$p_R(\tau, a, b) = \exp(-(\tau/b)^c), \quad b > 0, \quad c > 0 \quad (6)$$

The fitted equation 6 is inserted into equation 2 and then is used to perform the numerical estimation along with equation 4, thereby rendering it possible to fit equation 5. The fitted equation 5 is used to estimate the ACS of the corresponding parent-Gaussian process, which is then simulated using autoregressive models $AR_c(p)$ of sufficiently large order p to reproduce the $p_{Xc}(\tau)$ ACSs of the parent. The generated random Gaussian time series for each month are transformed into the sub-daily precipitation time series by using the Gaussian anamorphosis (equation 1).

While the generated sub-daily random precipitation time series reproduces the marginal distribution and the linear ACS, it cannot be exactly the same as the observed sub-daily precipitation time series, and errors will inevitably exist for specific points in time. Such errors can be reduced by repeating the aforementioned simulation process until a predetermined error goal is fulfilled. However, our intention is to disaggregate the daily precipitation time series generated from climate simulations into the sub-daily scale, rather than reproduce precipitation observations. More importantly, climate change impacts on precipitation regimes are not represented in the generated sub-daily precipitation time series. Therefore, the generated sub-daily precipitation time series is aggregated by day to obtain daily precipitation, and then a scaling factor (i.e., the ratio of daily precipitation generated from climate simulations to the aggregated daily precipitation) is calculated for each calendar day. The scaling factor is multiplied by the generated sub-daily precipitation time series. Such a process not only combines the added value of the spatial downscaling and the temporal disaggregation in reproducing sub-

daily precipitation at a fine spatial scale, but also takes into account climate change impacts on the future sub-daily precipitation intensity.

2.2 Vine copula multi-model ensemble IDF curves

The spatiotemporal downscaling of precipitation from a single RCM is inadequate for reproducing the observation-based IDF curves well. It is thus necessary to combine the time series of extreme precipitation extracted from multiple downscaled model outputs to address the uncertainty in model structures and improve the reliability of climate-model-based IDF curves. To this end, a vine copula multi-model ensemble approach is developed based on regular vine copulas. This approach constructs a joint multivariate probability distribution of rainfall extremes from multiple downscaled precipitation model outputs and observations. The basic idea of this approach is to implement repeated conditional sampling from the well-calibrated vine copula given the extremes extracted from downscaled precipitation, thus generating multi-model ensemble simulations of extreme precipitation (Zhang et al., 2021). Assume that $\mathbf{r} = (r_1, \dots, r_{n-1}, r_n)$ signifies rainfall extremes generated from $n-1$ downscaled precipitation outputs and the corresponding observation. The joint PDF between simulations (r_1, \dots, r_{n-1}) and the observation r_n can be expressed as

$$p(r_1, \dots, r_{n-1}, r_n) = p_1(r_1) \cdot \dots \cdot p_{n-1}(r_{n-1}) \cdot p_n(r_n) \cdot c(u_1, \dots, u_{n-1}, u_n) \quad (7)$$

where $p_i(r_i)$ represents the marginal PDF and u_i represents the marginal cumulative probability, $i = 1, \dots, n$; c represents the copula density.

Since copulas are inflexible in high dimensions and high-dimensional copula families are limited, the vine copula, also known as pair-copula construction (PCC), has been proposed to graphically represent the high-dimensional dependence structure (i.e., equation 7) as vines

comprising a nested set of trees with nodes that are joined by edges (Bedford & Cooke, 2002). The canonical vine (C-vine) and the drawable vine (D-vine) are two special regular vines and are also the most widely used decompositions. Equation 7 can be decomposed through the C-vine and D-vine copulas as equations 8 and 9, respectively.

$$p(r_1, \dots, r_{n-1}, r_n) = \prod_{k=1}^n p_k(r_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1, \dots, j-1} \{P(r_j | r_1, \dots, r_{j-1}), P(r_{j+i} | r_1, \dots, r_{j-1})\} \quad (8)$$

$$p(r_1, \dots, r_{n-1}, r_n) = \prod_{k=1}^n p_k(r_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1, \dots, i+j-1} \{P(r_i | r_{i+1}, \dots, r_{i+j-1}), P(r_{i+j} | r_{i+1}, \dots, r_{i+j-1})\} \quad (9)$$

where $P(\cdot|\cdot)$ represents the conditional cumulative probability. If rainfall extremes from two models are used, the joint density between the simulations (r_1, r_2) and the observation (r_o) can be decomposed through a 3-dimensional vine copula as

$$p(r_1, r_2, r_o) = p(r_1) \cdot p(r_2) \cdot p(r_o) \cdot c(u_1, u_2, \theta_{1,2}) \cdot c(u_1, u_o, \theta_{1,o}) \cdot c(h(u_2, u_1, \theta_{1,2}), h(u_o, u_1, \theta_{1,o}), \theta_{2,o|1}) \quad (10)$$

where $\theta_{1,2}$, $\theta_{1,o}$, and $\theta_{2,o|1}$ represent the parameters of bivariate copulas; the h -function is the conditional distribution function. For example, $h(u_2, u_1, \theta_{1,2})$ is expressed as

$$h(u_2, u_1, \theta_{1,2}) = F(u_2 | u_1) = \frac{\partial C_{2,1} \{F_2(r_2), F_1(r_1), \theta_{1,2}\}}{\partial F_1(r_1)} \quad (11)$$

where $F_1(r_1) = u_1$ and $F_2(r_2) = u_2$ represent marginal cumulative distribution functions (CDFs).

Since the number of possible vine copula structures grows exponentially with the number of variables, a feasible solution is to use the sequential maximal spanning tree algorithm along with the Akaike information criterion (AIC) or the Bayesian Information Criterion (BIC) to identify an appropriate structure (Dißmann et al., 2013). After determining the vine structure, a conditional CDF of r_o can be constructed by recursively applying the h -function:

$$P(r_o | r_1, r_2) = \frac{\partial C_{o,2|1} (P(r_o | r_1), P(r_2 | r_1))}{\partial P(r_2 | r_1)} = h[h(u_o, u_1, \alpha_{1,o}), h(u_2, u_1, \alpha_{1,2}), \alpha_{2,o|1}] \quad (12)$$

Thus, the inverse form of equation 12 can be used to combine multi-model rainfall extremes for probabilistic ensemble projections:

$$\hat{r}_o = f(r_1, r_2, \tau) = P_o^{-1} \left\{ h^{-1} \left(\tau \left| h \left(P_2(r_2) \mid P_1(r_1), \alpha_{1,2} \right), \alpha_{2,o|1} \right) \mid P_1(r_1), \alpha_{1,o} \right) \right\}, \tau \in (0,1) \quad (13)$$

where τ represents random probability levels (e.g., $\tau = 0.01, 0.1, \dots, 0.99$); P represents marginal CDFs. To achieve reliable model results, MC simulations are used to generate multiple (e.g., 500) samples of τ from the uniform distribution $U(0, 1)$, leading to multiple realizations of r_o .

Each realization of r_0 represents a stochastic simulation of rainfall extremes and is fitted to a generalized extreme value (GEV) distribution, which is a commonly used distribution for deriving IDF curves (Ragno et al., 2018; Srivastava et al., 2021). The CDF of the GEV distribution is

$$G(r \mid \mu, \sigma, \xi) = \exp \left\{ - \left(1 + \xi \cdot (r - \mu) / \sigma \right)^{-1/\xi} \right\}, \quad \xi \cdot (r - \mu) / \sigma > 0 \quad (14)$$

where μ , σ , and ξ represent the location, scale, and shape parameters of the distribution, respectively. These parameters were calculated using the L-moments estimation method, and the parameter uncertainty was estimated by performing a parametric bootstrap (Gilleland & Katz, 2016; Marra et al., 2017). A GEV distribution is fitted to the annual maxima precipitation for different durations (e.g., 3-, 6-, 9-, 12-, 18-, 24-, 48-, 72-, 96-, 120-, 144-, and 168-hours in this study) to estimate the precipitation intensities corresponding to various design return periods (e.g., 2-, 5-, 10-, 25-, 50-, and 100-year in this study). Thus, the relationship between the intensity and frequency (i.e., return period) of extreme precipitation for different durations can be characterized using a series of curves along with confidence intervals (Figure 1). Such IDF curves are generated from multiple realizations of r_0 based on the vine copula and thus consider both the uncertainty in the climate model structure and the inter-model dependence structure.

The uncertainty information in the IDF curves can be used to estimate the statistical significance of the difference in precipitation intensity of IDF curves between past and future climates. For example, the vine copula-based ensemble approach generates 500 realizations of extreme precipitation time series for a given duration (e.g., 3-hour). Each of the 500 time series is used to fit the GEV distribution and the underlying uncertainty is quantified using a parametric bootstrap, in which random sampling with replacement is performed 500 times, leading to 500 estimated precipitation intensities for a given design return period (e.g., 100-year). This process of uncertainty quantitation generates a total of 500×500 estimated design values for both past and future climates. The final uncertainty in the IDF curves is defined as the 2.5th and 97.5th percentile interval of the 500×500 design value estimates. The design values under the future climate are significantly ($P < 0.05$) higher than those under the past climate only if at least 95% of the 500×500 design value estimates under the future climate exceed the 97.5th percentile under the past climate.

To improve climate-model-based IDF curves in China, the vine copula multi-model ensemble approach was used to combine annual maximum precipitation time series for different durations obtained from the spatiotemporal downscaling (see Section 2.1) of RCM-simulated precipitation in the CORDEX experiment for the past (1979–2005) and future (2058–2099) periods. Since the vine copula builds a multidimensional relationship between the rainfall extremes from five downscaled RCM model outputs and the observed extremes in a probability space, it is necessary to select appropriate univariate CDFs to transform the five model simulations and the observation into uniform variables on the interval (0, 1). A total of 9 probability distributions were examined, including gamma, log-normal, log-logistic, Gumbel, generalized extreme value (GEV), generalized Pareto (GP), log-gamma, generalized gamma, and

four-parameter kappa distributions. The optimal marginal distribution was identified using the Kolmogorov-Smirnov (K-S) test, while the corresponding distribution parameters were estimated by the maximum likelihood method. The bivariate copula families considered in this study include Gaussian, Student t, Clayton, Gumbel, and Frank. To assess the robustness of the proposed approach, it is necessary to perform a quantitative comparison with previous approaches. BMA is one of the most widely used multi-model combination techniques in the hydroclimate community and is thus selected as the comparative benchmark (Raftery et al., 2005; Zhang & Wang, 2021). The BMA approach was used to combine multiple time series of annual maximum precipitation for different durations derived from the spatiotemporal downscaling (see Section 2.1) of RCM-simulated precipitation, and the BMA-based annual maximum precipitation was then used to derive IDF curves.

2.3 Performance metrics for evaluating IDF curves

To evaluate the IDF curves derived from the spatiotemporal downscaling of climate simulations, three performance metrics are used, including accuracy, reliability, and sharpness. Accuracy is evaluated using the mean absolute percentage error (MAPE):

$$\text{MAPE}(\%) = \left(\frac{\text{IDF}_{\text{simulation}} - \text{IDF}_{\text{observation}}}{\text{IDF}_{\text{observation}}} \right) \times 100\% \quad (15)$$

MAPE is normalized and insensitive to absolute values of rainfall, which enables the performance of IDF curves to be evaluated across China regardless of variations in climate regimes. Since IDF curves are subject to various sources of uncertainty, reliability and sharpness are also examined in this study to verify probabilistic IDF curves. The reliability represents the percentage of simulation-based IDF estimates captured within the confidence intervals of observation-based IDF curves. The sharpness is defined as the average width of confidence

intervals. For example, Figure 1 presents a simple illustration of IDF curve evaluation. Method 1 achieves higher reliability (100%) compared to Method 2 since all IDF estimates generated from Method 1 fall within the 95% confidence interval of the observation-based IDF curve. Nevertheless, Method 2 achieves lower sharpness of IDF curves compared to Method 1 since Method 2 leads to a narrower confidence interval (3.5 mm) than Method 1 (6 mm). It should be noted that the desired performance of IDF curves is characterized by high accuracy (i.e., low MAPE), high reliability, and low sharpness.

2.4 Data sources

The MSWEP V2 (Multi-source weighted-Ensemble Precipitation, version 2) 0.1° 3-hour precipitation product was used to derive gridded IDF curves in China during the period of 1979–2020 since it has been shown to outperform other precipitation datasets in terms of accuracy and spatiotemporal resolution (Beck et al., 2019). An ensemble of five regional climate model simulations from the CORDEX East Asia experiment was used to perform stochastic spatiotemporal downscaling of precipitation and thus investigate climate change impacts on IDF curves. Specifically, the RCM REMO2015 was used to dynamically downscale three GCMs (MOHC-HadGEM2-ES, MPI-M-MPI-ESM-LR, and NCC-NorESM1-M) that participated in the Coupled Model Intercomparison Project Phase 5 (CMIP5). The RCM RegCM4 was used to dynamically downscale two (MOHC-HadGEM2-ES and NCC-NorESM1-M) GCMs in the CMIP5. All five simulations have the same horizontal resolution of about $0.22^{\circ} \times 0.22^{\circ}$ (~25 km) and cover the historical period (1970–2005) and a future period (2006–2099). Future simulations used in this study were forced with a high-emission scenario (RCP8.5) during the period of 2058–2099.

To derive IDF curves over urban areas, urban boundaries were identified based on the artificial impervious area mapped from 30-m resolution Landsat images in 2018. Specifically, the inner non-urban area of each city was firstly filled through a kernel density estimation method and cellular-automata simulations, thereby generating initial urban clusters. Since there can be quite a few small and isolated urban clusters for a single city, morphological dilation and erosion operations were performed to improve the delineation of urban boundaries around urban fringe areas (X. Li et al., 2020). In addition, a buffer zone was drawn around each cluster, with the area of the buffer equal to one-quarter of the area of the cluster (Angel et al., 2016). Such an operation merges all close but disconnected urban clusters whose buffers intersect one another, thereby generating the urban boundary of each city. Since the spatial resolution (0.1°) of downscaled precipitation was still coarse compared to a city scale, we selected the cities where the urban area was covered by at least two 0.1° grids, leading to 196 cities for analyzing urban IDF curves. More details on the delineation of urban boundaries can be found in X. Li et al. (2020) and Xu et al. (2020).

3. Evaluation of IDF curves

We performed stochastic spatiotemporal downscaling of precipitation using the approach described in Section 2.1 and evaluated its performance in reproducing urban IDF curves. Figure 2 presents the bias (indicated by the MAPE), reliability, and sharpness of precipitation IDF curves averaged over the urban areas of 196 Chinese cities, given return periods of 10 and 50 years, generated from multi-model ensemble mean simulations. The corresponding results for return periods of 2, 5, 25, and 100 years are presented in Figure S1 of the supplementary material. Results show that the stochastic downscaling method leads to a reasonable level of accuracy

(MAPE < 20%) for the 1-in-10 year IDF curves for more than 80% of cities (Figure 2a). The reliability of 10-year and 50-year IDF curves exceeds 50% for more than 80% of cities (Figures 2b and 2e), indicating that at least six out of twelve return levels of extreme precipitation in the IDF curves are captured within the uncertainty range of multi-model ensemble simulations. This indicates that the stochastic spatiotemporal downscaling method reproduces the extreme precipitation regimes well at sub-daily to multi-day time scales. In addition, the sharpness of IDF curves increases (from 8 mm on average to 22 mm) from the return period of 10 years to 50 years (Figures 2c and 2f).

To evaluate the urban IDF curves given different return periods, we present PDF plots of the MAPE, reliability, and sharpness across the 196 cities (Figure 3). The IDF curves given different return periods have a comparable level of accuracy, with the MAPE generally ranging from 0 to 40% and increasing with the return period (Figure 3a). Nevertheless, the IDF curves for long return periods generally have relatively high reliability (Figure 3b). For example, the IDF curves for the 100-year return period have a mean reliability of 95% across 196 cities, which is significantly ($P < 0.05$) higher than the mean reliability (92%) for the return period of 50 years. Figure 3c shows that the sharpness generally increases with return periods (lower values are better), which is not surprising since the sample of extreme precipitation decreases with the return period. We also find that the bias in the return levels of simulated extreme precipitation first increases and then decreases from the 3-hour duration to the 168-hour duration (Figure 4). For example, the simulated 3-hour extreme precipitation given the return period of 10 years has a mean relative bias of 1%, and the corresponding biases are approximately -21% and -8% for the 24-hour and 7-day (i.e., 168-hour) extreme precipitation, respectively. Such a sensitivity of the

bias to the time intervals may result from the combined effects of the temporal disaggregation and spatial downscaling of precipitation.

Since the multi-model ensemble mean climate simulation along with the spatiotemporal downscaled precipitation show biases in the urban IDF curves, we employ the vine copula ensemble approach to combine the time series of rainfall extremes generated from multi-model ensemble downscaling. The vine copula-based ensemble rainfall extremes are then used to construct the IDF curves. We compared the accuracy, reliability, and sharpness of IDF curves given the 10-year return period generated from different multi-model combination approaches (Figures 5a–o). Results show that the ensemble mean simulation generally leads to a MAPE of below 50% (Figure 5a), a reliability of over 50% (Figure 5f), and a sharpness of below 10 mm (Figure 5k) over China. Northwest China shows relatively large MAPE of up to 50% (Figure 5a), which may result from the bias of RCMs over arid areas (Zhu et al., 2018). The BMA methods do not show improvements in the accuracy and reliability of IDF curves upon the ensemble mean simulation (Figures 5b–d and 5g–i), despite their relatively low sharpness (Figures 5l–n). In comparison, the IDF curves generated from the vine copula-based multi-model ensemble approach have a MAPE of below 10% (Figure 5e) and a reliability of over 90% (Figure 5j) for more than 90% of land area in China. The vine copula-based multi-model ensemble approach leads to a sharpness of below 5 mm for more than 80% of land area in China, which is comparable to the sharpness generated from the ensemble mean simulation (Figure 5o).

In addition, summaries of bias, reliability, and sharpness for IDF curves across return periods are presented in Figures 5p–r. Results show that the vine copula method leads to a generally lower bias (MAPE) and higher reliability than the ensemble mean (ENS) and BMA methods (Figures 5p–q). The mean MAPE of IDF curves generated from the ENS and BMA

methods is generally higher than 8% across return periods, but the mean MAPE generated from the vine copula method is generally lower than 8% (Figure 5p). The vine copula method also leads to a mean reliability of over 90% across return periods, while other methods generally lead to relatively low reliability (Figure 5q). Figure 5r shows that the vine copula method generates similar distributions of sharpness to the ENS method. Therefore, the vine copula method improves both the accuracy and reliability of multi-model IDF curves relative to the ENS and BMA methods. It should be noted that the BMA methods lead to relatively low (< 4 mm) sharpness (i.e., lower uncertainty) of IDF curves (Figure 5r). Nevertheless, the reliability and accuracy of IDF curves generated from BMA methods are lower than the ensemble mean and vine copula methods. Such IDF curves are not desirable because the first priority in hydroclimate applications is accuracy and reliability.

To further compare the performance of multi-model approaches in reproducing IDF curves, we selected the urban area of Hong Kong and Zhengzhou to evaluate the precipitation IDF curves given the 10-year return period during the period of 1979–2005. Since Hong Kong and Zhengzhou have both suffered from serious consequences of heavy rainfall, climate-model-based IDF curves with a reasonable level of accuracy play a crucial role in infrastructure design and risk assessment. We show that the IDF curves for the 10-year return period, generated from the ensemble mean simulations for Hong Kong and Zhengzhou, underestimate the return level of extreme precipitation compared to those generated from observations, with a MAPE of 20% and 17%, respectively (Figures 6a and 6c). In comparison, the vine copula-based multi-model ensemble method substantially reduces the MAPE to 1.1% and 1.2% for the IDF curves over Hong Kong and Zhengzhou, respectively (Figures 6b and 6d).

4. Projected changes in extreme precipitation and IDF curves

To project future changes in the IDF curves, the daily precipitation generated from climate simulations in the CORDEX experiment was spatially and temporally downscaled to the 3-hour time scale and the 0.1° spatial scale. The vine copula-based multi-model ensemble approach was then used to generate probabilistic projections of future changes in extreme precipitation intensity for different durations and different return periods. Figure 7 presents the cumulative probability plots of the relative difference in extreme precipitation between past and future climates. We find that climate change is generally intensifying extreme precipitation with an increase in all return periods regardless of the event duration. For example, the intensity of 3-hour heavy rainfall events given a return period of 10 years shows an increase of more than 50% over approximately 10% of land area (i.e., land grid cells) in China and a decrease of 10% on average over 36% of land area (Figure 7a). We also find that the fraction of land area with a projected increase in the intensity of extreme precipitation events expands with return periods. This may suggest a higher risk of failure for urban infrastructure in the case of more severe events with longer return periods. Such a finding has important implications for the design of China's urban infrastructure because most Chinese urban infrastructure is designed with a return period of 5 years at present (Qi et al., 2021).

We investigated future changes in the 3-hour extreme precipitation averaged over the urban areas of 196 Chinese cities given different return periods (Figure 8). Results show that the 3-hour extreme precipitation is expected to increase for over half of the 196 cities regardless of the return periods. The fraction of cities with increased intensity of 3-hour extreme precipitation also increases with the return period. For example, 56% of cities show increases in the intensity of 3-hour extreme precipitation given a return period of 2 years (Figure 8a), and such increases would

occur over 80% of cities given return periods of 50 and 100 years. More importantly, more than half of the 196 cities show significant ($P < 0.05$) increases in the 3-hour extreme precipitation intensity given return periods of 10, 25, 50, and 100 years, and the average percentage increases are 30%, 44%, 56%, 70%, respectively. We also show that the Chinese megacities, such as Beijing, Shanghai, and Guangzhou, do not show increases in the 3-hour extreme precipitation intensity, but the small and medium cities are expected to experience substantial increases. For example, future changes in the 3-hour extreme precipitation in Beijing given a 10 year return period are projected to vary from -18% to 18% . In comparison, Zhenjiang, a comparatively small city located in the Yangtze River Delta, is projected to experience $64\text{--}148\%$ and $227\text{--}400\%$ increases in the magnitude of 3-hour heavy rainfall given return periods of 10 and 100 years, respectively.

To further investigate future changes in short-duration extreme precipitation over urban areas, we present spatial patterns of the relative difference in 3-hour and 24-hour heavy rainfall between past and future climates in four selected urban areas, given the return period of 50 years (Figure 9). The four urban areas include Beijing and Tianjin (BTH), the Yangtze River Delta (YRD), the Pearl River Delta (PRD), and Zhengzhou. The central urban area of Beijing shows little change in the 3-hour and 24-hour heavy rainfall intensity, but the urban fringe area shows at least 40% increases (Figures 9a and 9b). Most urban areas of Shanghai show little change in the 3-hour heavy rainfall intensity but significant increases in the 24-hour heavy rainfall intensity (Figures 9c and 9d). In comparison, the 3-hour and 24-hour heavy rainfall intensity show little change in most of the PRD except eastern Shenzhen (Figures 9e and 9f). To put these changes into perspective, a frequency analysis of observed heavy rainfall in Zhengzhou indicated that the 622.7 mm of 24-hour rainfall that occurred on 20 July 2021 had a return period of greater than

1,000 years (Ryan Woo & Stella Qiu, 2021). Here we show that the 1-in-50 year magnitude of 24-hour heavy rainfall is projected to increase by at least 20% for most urban areas in Zhengzhou (Figure 9h).

To show climate change impacts on precipitation IDF curves over urban areas, we selected 12 cities to present the IDF curves under historical and future climates given return periods of 10, 50, and 100 years (Figures 10–12). For most cities the IDF curves under the future climate (RCP8.5) are significantly different from those under the past climate for specific time intervals (i.e., duration), but they do not present consistent upward or downward shifts. For example, the extreme precipitation intensity in Shanghai shows significant ($P < 0.05$) increases on the 3-hour to 2-day time scales between past and future climates, but such increases are not exhibited for time scales of more than 2 days (Figures 10h, 11h, and 12h). In comparison, Shenzhen shows a marked upward shift in the IDF curves between past and future climates (Figures 10k, 11k, and 12k). Urumqi, the capital city of Xinjiang province in the arid area of northwestern China, in contrast, shows significant ($P < 0.05$) decreases in the intensity of long-duration (> 24 hours) extreme precipitation between past and future climates (Figures 10a, 11a, and 12a). In addition, the intensity of short-duration (< 12 hours) heavy rainfall does not show significant ($P < 0.05$) decreases in all the 12 cities but increases for multiple cities, such as Shenyang (Figure 10d), Hangzhou (Figure 10l), and Xi'an (Figure 12e).

After investigating the change in extreme event intensity for a fixed return period, we explored changes in the frequency of 3-hour and 24-hour heavy rainfall events for given magnitudes. Specifically, we chose the intensity of 1-in-50 and 1-in-100 year events under the past climate to estimate their expected occurrence under the future climate (RCP8.5). Figure 13 shows the return periods expected in the future (2058–2099) of the corresponding historical

events. Events that are currently (1979–2020) expected to occur every 50 or 100 years are becoming more frequent in more than half of Chinese cities. For example, the 3-hour heavy rainfall events, which currently occur every 50 years on average, are projected to occur more frequently for 78% of cities, with a mean return period of 27 years across the 196 cities under the future climate (RCP8.5) (Figure 13a). Similarly, the frequency of 24-hour heavy rainfall with a return period of 100 years is projected to more than double over 42% of cities (i.e., occur with an updated return period of shorter than 50 years) (Figure 13d). Such an increased frequency of rare heavy rainfall events may raise public safety concerns in urban environments, especially in the densely populated Chinese coastal cities.

5. Discussion

Covariate-based and climate-model-based IDF curves are two commonly used methods to address changes in the frequency and intensity of rainfall events under a warming climate. Covariate-based nonstationary IDF curves can be generated by introducing time-varying covariates (such as time) and constructing a link function (such as a simple linear model) between the covariates and the parameters of an Extreme Value distribution such as the GEV. In contrast, climate-model-based IDF curves are generated directly by employing extreme precipitation simulations under present and future climates. However, these two methods each have their own limitations and there is currently no well-accepted method for updating IDF curves that considers future climate change impacts (Benestad et al., 2021; J. Li et al., 2017).

Here we present a comparison of the two methods used to project future changes in the IDF curves over the urban area of 196 Chinese cities (Figure 14). We developed models with either one or both of the location (μ) and scale (σ) parameters of the GEV distribution as a function of

time, and selected the best fitting model with the smallest BIC (Slater, Villarini, et al., 2021). The use of a time covariate and a linear function may not always provide the best representation of nonstationary extreme precipitation, but can avoid overfitting of the observation period and unrealistic prediction of future design values (Fadhel et al., 2017; Ragno et al., 2018). We find that 25 out of 196 cities show significantly ($P < 0.05$) decreasing trends of 1% on average per year in the annual maximum 3-hour rainfall during the period of 1979–2020. Approximately 36% of cities show an increasing trend, but these increases are significant ($P < 0.05$) only in 6 cities (Figure 14a). In comparison, over 84% of cities show no significant trends in the annual maximum 3-hour rainfall. To project future IDF curves with the covariate-based approach, we assume these historical trends are likely to continue in the future climate and employ the time covariate to project future changes in the distribution of extreme precipitation (Figure 14b). We find that the covariate-based approach leads to increases in heavy rainfall intensity over 44% of cities, while approximately 80% of cities show increases in the climate-model-based projections (Figures 14c and 14d). There are only 63 out of 196 cities showing increases in the 3-hour heavy rainfall intensity generated from both the covariate-based and climate-model-based methods. Of those 63 cities, 37 show greater design values estimated by the covariate-based IDF curves than by climate-model-based IDF curves.

The discrepancies between the two methods do not necessarily imply that the covariate-based model is unreliable; the difference suggests that the two methods are capturing different effects. The covariate-based method is constructed based on historical trends of extreme precipitation: it considers all the different factors that have affected precipitation in the past (i.e., not only climate), and assumes these are likely to continue into the future. In contrast, the climate-model approach employs scenario-based projections, which consider only the effect of

future changes in the climate. The climate-model-based method provides a physics-based projection of IDF curves which leads to greater design values of heavy rainfall than the covariate-based method for 76% of China's cities.

To further compare future changes in the precipitation IDF curves generated by the climate-model-based and covariate-based methods, we summarized the fraction of 196 cities showing increases in the design values of IDF curves given different return periods and different durations (Figures 15a and 15c). The covariate-based method leads to a similar fraction of increases for an individual duration (Figure 15a) since we do not consider the statistical uncertainty in the nonstationary model here. The climate-model-based method generally shows a greater fraction of cities with increased IDF design values (Figure 15c) than the covariate-based method (Figure 15a). More importantly, the design values from the climate-model-based IDF curves generally show greater increases than those from the covariate-based IDF curves between past and future climates (Figure 15d). In comparison, the design values in the covariate-based IDF curves show little increase ($< 5\%$) between past and future climates for durations longer than 12 hours (Figure 15b). Such results may partially result from the low significance of the trends in extreme precipitation (see Figure S2 in the supplementary material) and the limited ability of a nonstationary statistical model using only a time-based covariate to capture future climate change signals. Such nonstationary statistical models can be improved by selecting appropriate covariates associated with extreme precipitation (such as mean temperature), but there is no well-accepted covariate accounting for the physical process of extreme precipitation (Yan et al., 2021). Therefore, although both the nonstationary statistical approach and the climate-model-based approach may be wrong in terms of the sign of projected change, climate-model-based IDF curves should be preferred to reduce possible risks of hydrologic failure since they provide a

relatively worst-case scenario to update IDF curves in light of future climate change (Medeiros de Saboia et al., 2020).

The future projections of IDF curves are subject to substantial uncertainties including the spatiotemporal downscaling, vine copula simulations, and the GEV parameters. The contribution of uncertainty sources was evaluated using the ANOVA-based variance decomposition method (Bosshard et al., 2013). Results show that the GEV parameters and vine copula-based ensemble simulations are the major uncertainty source, accounting for on average 35% and 32% of the total uncertainty, respectively, in the projection of IDF curves (see Figure S3 in the supplementary material). Such results are consistent with those in previous studies which recognized the substantial uncertainty contribution of climate model structures in the projection of IDF curves (Switzman et al., 2017).

This study has several limitations. The intense and widespread heavy rainfall events were not considered in the future projections of IDF curves. Previous studies have shown that the spatial dependence of precipitation could increase under a warming climate, leading to more intense and widespread heavy rainfall events (Rastogi et al., 2020). It is thus desired to consider possible changes in the spatial dependence between IDF curves in future studies (see Figure S4 in the supplementary material). The BCSD-based statistical downscaling has its disadvantages, such as correcting biases through quantile-mapping (QM) and downscaling through interpolation, but it can well represent the spatial pattern of extreme precipitation (see Figure S5 in the supplementary material). It is thus necessary to use multiple downscaling approaches to enhance the credibility of the future projections of IDF curves in future studies.

The GEV distribution may not always be the optimal choice to simulate precipitation extremes, but it is appropriate for more than 99% of grid cells in this study based on the

Kolmogorov-Smirnov (K-S) test (see Figure S6 in the supplementary material). To better represent precipitation extremes, the nonparametric kernel density estimation can be used in future studies since it relaxes the assumption of prior distribution. In addition, the vine copula-based ensemble simulations are subject to the curse of dimensionality, but we used only five climate simulations in this study and thus this issue did not affect our approach. It will be more appropriate and practical to use the nonparametric pair-copula construction if more climate simulation datasets are used (Nagler et al., 2017).

6. Summary and Conclusions

In this study, we present a vine copula-based ensemble projection of future changes in precipitation IDF curves at sub-daily to multi-day time scales. Specifically, we develop stochastic spatiotemporal downscaling of daily precipitation in China, based on five regional climate simulations, to produce IDF curves for past (1979–2020) and future (2058–2099) climates under RCP8.5. A vine copula multi-model ensemble approach is proposed to enhance the accuracy and reliability of multi-model ensemble simulations of extreme precipitation and thus the climate-model-based IDF curves. The vine copula approach is compared with existing multi-model combination approaches in terms of accuracy, reliability, and sharpness of IDF curves. Furthermore, we investigate future changes in the intensity and frequency of extreme precipitation as well as the IDF curves averaged over 196 Chinese cities. The climate-model-based IDF curves are also compared with those generated from a nonstationary time covariate-based method.

Our findings reveal that the spatiotemporal downscaling method reproduces the extreme precipitation regimes well at sub-daily to multi-day time scales, albeit with potential

underestimation of extreme precipitation intensity by no more than 20%. The vine copula multi-model ensemble approach achieves higher reliability and accuracy in reproducing the climate-model-based precipitation IDF curves compared to the ensemble mean and BMA techniques. More than half of the 196 Chinese cities are expected to experience extreme precipitation increases of up to 30% in intensity and nearly twice the frequency compared to historical records, especially for return periods of over 10 years. Such increases lead to a significant upward shift of precipitation IDF curves, especially for short-duration (< 12 hours) heavy rainfall. This suggests that the current urban infrastructure in more than half of China's cities would be inadequate to prevent losses and damages caused by rainfall-triggered hazards under a high emission scenario (RCP8.5) if designed solely based on historical precipitation observations. Our findings also reveal that the climate-model-based precipitation IDF curves show higher design values of heavy rainfall under the future climate than the covariate-based IDF curves. The covariate-based nonstationary IDF curves may underestimate the risk of failure for urban infrastructure, especially when the nonstationary behavior of extreme precipitation is not characterized well by the covariates embedded in the time-varying distribution of extreme precipitation.

Although we perform a comprehensive assessment of future changes in Chinese precipitation IDF curves, our findings come with a major limitation related to the inability of regional climate models to represent land surface processes affecting local extreme precipitation well. For example, the intensification of urbanization-induced extreme precipitation has been reported in recent studies but is not considered in RCMs (Y. Li et al., 2020; Wu et al., 2019; Yu et al., 2022). This may lead to an underestimation of future extreme precipitation intensity in urban areas and the IDF curves. It is thus necessary to further improve the accuracy of

precipitation simulations by accounting for urban dynamics through high-resolution climate simulations in future studies.

Although the Chinese Sponge City Program (SCP) initiated in 2013 requires a 1-in-30 year standard for urban infrastructure (Qi et al., 2021), the effort to build a “sponge city” might be compromised by the increasing intensity and frequency of extreme events with long return periods (e.g., 1-in-50 and 1-in-100 year events). Our work reveals that the design guidelines of Chinese urban infrastructure should be updated to consider possible changes in the IDF curves induced by the changing climate. Such updates would help reduce the risk of catastrophic hazards like the 2021 deadly flooding in Zhengzhou. To facilitate the use of and open access to China’s urban IDF curves, we established an interactive web map (<https://boenzhang.github.io/urban-IDF-curves/>) to visualize urban IDF curves across China’s 196 cities under past and future climates. This web map may help inform urban planning policies to reduce vulnerability to future rainfall-triggered hazards and to achieve climate-resilient urban development.

Our estimates should not be used for the design and planning of specific urban infrastructure since the spatiotemporal downscaling of precipitation is still insufficient to capture the spatial variability of urban precipitation extremes. Rather, we provide evidence that the precipitation IDF curves could shift upward and thus increase the risk of failure for urban infrastructure under a warming climate. This finding is important because many previous studies of IDF curves either overlooked future changes in extreme precipitation or failed to represent the nonstationary behavior of extreme precipitation adequately under the present climate.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 51809223), the Hong Kong Research Grants Council Early Career Scheme (Grant No. 25222319), and the Environment and Conservation Fund (Grant No. ECF 119/2021). The MSWEP V2 precipitation product can be downloaded at <http://www.gloh2o.org/mswep/>. We acknowledge the World Climate Research Program's Working Group on Coupled Modeling, which is responsible for CORDEX, and we thank the climate modeling groups for producing and making their model outputs available at <https://esg-dn1.nsc.liu.se/projects/cordex/>. The precipitation IDF curves averaged over 196 Chinese cities under past and future climates are freely accessible at <https://doi.org/10.5281/zenodo.6409196>. The urban boundaries of 196 Chinese cities are freely accessible at <https://doi.org/10.5281/zenodo.6409606>.

References

- Agilan, V., & Umamahesh, N. V. (2016). Is the covariate based non-stationary rainfall IDF curve capable of encompassing future rainfall changes? *Journal of Hydrology*, 541, 1441–1455. <https://doi.org/10.1016/j.jhydrol.2016.08.052>
- Angel, S., Blei, A. M., Parent, J., Lamson-Hall, P., Galarza-Sanchez, N., Civco, D. L., & Thom, K. (2016). Atlas of urban expansion—2016 edition. *Areas and Densities*. Cambridge, MA: NYU Urban Expansion Program at New York University, UN-Habitat, and the Lincoln Institute of Land Policy.
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., et al. (2019). MSWep v2 Global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100(3), 473–500. <https://doi.org/10.1175/BAMS-D-17-0138.1>
- Bedford, T., & Cooke, R. M. (2002). Vines - A new graphical model for dependent random variables. *Annals of Statistics*, 30(4), 1031–1068. <https://doi.org/10.1214/aos/1031689016>
- Benestad, R. E., Lutz, J., Dyrddal, A. V., Haugen, J. E., Parding, K. M., & Dobler, A. (2021). Testing a simple formula for calculating approximate intensity-duration-frequency curves. *Environmental Research Letters*, 16(4). <https://doi.org/10.1088/1748-9326/abd4ab>
- Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: Analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*, 21(6), 2701–2723. <https://doi.org/10.5194/hess-21-2701-2017>

- Bezak, N., Šraj, M., & Mikoš, M. (2016). Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. *Journal of Hydrology*, 541, 272–284. <https://doi.org/10.1016/j.jhydrol.2016.02.058>
- Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., & Schär, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, 49(3), 1523–1536. <https://doi.org/10.1029/2011WR011533>
- Boukhelifa, M., Meddi, M., & Gaume, E. (2018). Integrated Bayesian estimation of intensity-duration-frequency curves: Consolidation and extensive testing of a method. *Water Resources Research*, 54(10), 7459–7477. <https://doi.org/10.1029/2018WR023366>
- Butcher, J. B., Zi, T., Pickard, B. R., Job, S. C., Johnson, T. E., & Groza, B. A. (2021). Efficient statistical approach to develop intensity-duration-frequency curves for precipitation and runoff under future climate. *Climatic Change*, 164(1–2). <https://doi.org/10.1007/s10584-021-02963-y>
- Chandra, R., Saha, U., & Mujumdar, P. P. (2015). Model and parameter uncertainty in IDF relationships under climate change. *Advances in Water Resources*, 79, 127–139. <https://doi.org/10.1016/j.advwatres.2015.02.011>
- Cheng, L., & Aghakouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, 4, 1–6. <https://doi.org/10.1038/srep07093>
- Courty, L. G., Wilby, R. L., Hillier, J. K., & Slater, L. J. (2019). Intensity-duration-frequency curves at the global scale. *Environmental Research Letters*, 14(8). <https://doi.org/10.1088/1748-9326/ab370a>

- 728 Dißmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating
729 regular vine copulae and application to financial returns. *Computational Statistics and Data
730 Analysis*, 59(1), 52–69. <https://doi.org/10.1016/j.csda.2012.08.010>
- 731 Fadhel, S., Rico-Ramirez, M. A., & Han, D. (2017). Uncertainty of Intensity–Duration–
732 Frequency (IDF) curves due to varied climate baseline periods. *Journal of Hydrology*, 547,
733 600–612. <https://doi.org/10.1016/j.jhydrol.2017.02.013>
- 734 Feller, W. (1971). *An introduction to probability theory and its applications*. New York: Wiley.
- 735 FloodList. (2021). Singapore – roads swamped, cars stranded after 114mm of rain. Retrieved
736 September 25, 2021, from <https://floodlist.com/asia/singapore-floods-august-2021>
- 737 Ganguli, P., & Coulibaly, P. (2017). Does nonstationarity in rainfall require nonstationary
738 intensity-duration-frequency curves? *Hydrology and Earth System Sciences*, 21(12), 6461–
739 6483. <https://doi.org/10.5194/hess-21-6461-2017>
- 740 Ganguli, P., & Coulibaly, P. (2019). Assessment of future changes in intensity-duration-
741 frequency curves for Southern Ontario using North American (NA)-CORDEX models with
742 nonstationary methods. *Journal of Hydrology: Regional Studies*, 22.
743 <https://doi.org/10.1016/j.ejrh.2018.12.007>
- 744 Gaur, A., Schardong, A., & Simonovic, S. (2018). Effects of global warming on precipitation
745 extremes: Dependence on storm characteristics. *Water Resources Management*, 32(8),
746 2639–2648. <https://doi.org/10.1007/s11269-018-1949-x>
- 747 Gaur, A., Schardong, A., & Simonovic, S. P. (2020). Gridded extreme precipitation intensity–
748 duration–frequency estimates for the Canadian landmass. *Journal of Hydrologic
749 Engineering*, 25(6), 5020006. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001924](https://doi.org/10.1061/(asce)he.1943-5584.0001924)

- Gilleland, E., & Katz, R. W. (2016). extRemes 2.0: An extreme value analysis package in R. *Journal of Statistical Software*, 72(8 SE-Articles), 1–39. <https://doi.org/10.18637/jss.v072.i08>
- Harpold, A. A., & Kohler, M. (2017). Potential for changing extreme snowmelt and rainfall events in the mountains of the Western United States. *Journal of Geophysical Research: Atmospheres*, 122(24), 13,219–13,228. <https://doi.org/10.1002/2017JD027704>
- Hassanzadeh, E., Nazemi, A., & Elshorbagy, A. (2014). Quantile-based downscaling of precipitation using genetic programming: Application to IDF curves in Saskatoon. *Journal of Hydrologic Engineering*, 19(5), 943–955. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000854](https://doi.org/10.1061/(asce)he.1943-5584.0000854)
- Hassanzadeh, E., Nazemi, A., Adamowski, J., Nguyen, T. H., & Van-Nguyen, V. T. (2019). Quantile-based downscaling of rainfall extremes: Notes on methodological functionality, associated uncertainty and application in practice. *Advances in Water Resources*, 131, 103371. <https://doi.org/10.1016/j.advwatres.2019.07.001>
- Hatami, S., & Nazemi, A. (2022). Compound changes in temperature and snow depth lead to asymmetric and nonlinear responses in landscape freeze–thaw. *Scientific Reports*, 12(1), 1–13. <https://doi.org/10.1038/s41598-022-06320-6>
- Hosseinzadehtalaei, P., Tabari, H., & Willems, P. (2018). Precipitation intensity–duration–frequency curves for central Belgium with an ensemble of EURO-CORDEX simulations, and associated uncertainties. *Atmospheric Research*, 200(September 2017), 1–12. <https://doi.org/10.1016/j.atmosres.2017.09.015>

- 771 Hosseinzadehtalaei, P., Tabari, H., & Willems, P. (2020). Climate change impact on short-
 772 duration extreme precipitation and intensity–duration–frequency curves over Europe.
 773 *Journal of Hydrology*, 590(March), 125249. <https://doi.org/10.1016/j.jhydrol.2020.125249>
- 774 Innocenti, S., Mailhot, A., & Frigon, A. (2017). Simple scaling of extreme precipitation in North
 775 America. *Hydrology and Earth System Sciences*, 21(11), 5823–5846.
 776 <https://doi.org/10.5194/hess-21-5823-2017>
- 777 Jalowska, A. M., & Spero, T. L. (2019). Developing PIDF curves from dynamically downscaled
 778 WRF model fields to examine extreme precipitation events in three Eastern U.S.
 779 metropolitan areas. *Journal of Geophysical Research: Atmospheres*, 124(24), 13895–13913.
 780 <https://doi.org/10.1029/2019JD031584>
- 781 Kendall, M. G. (1975). *Rank correlation methods (4th edn.)* Charles Griffin. San Francisco, CA.
- 782 Le, P. D., Leonard, M., & Westra, S. (2018). Modeling spatial dependence of rainfall extremes
 783 across multiple durations. *Water Resources Research*, 54(3), 2233–2248.
 784 <https://doi.org/10.1002/2017WR022231>
- 785 Li, J., Johnson, F., Evans, J., & Sharma, A. (2017). A comparison of methods to estimate future
 786 sub-daily design rainfall. *Advances in Water Resources*, 110(February), 215–227.
 787 <https://doi.org/10.1016/j.advwatres.2017.10.020>
- 788 Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., et al. (2020). Mapping global urban
 789 boundaries from the global artificial impervious area (GAIA) data. *Environmental Research*
 790 *Letters*, 15(9). <https://doi.org/10.1088/1748-9326/ab9be3>
- 791 Li, X., & Wang, S. (2022). Recent increase in the occurrence of snow droughts followed by
 792 extreme heatwaves in a warmer world. *Geophysical Research Letters*, 49(13),
 793 e2022GL099925. <https://doi.org/10.1029/2022gl099925>

- 794 Li, Y., Fowler, H. J., Argüeso, D., Blenkinsop, S., Evans, J. P., Lenderink, G., et al. (2020).
795 Strong intensification of hourly rainfall extremes by urbanization. *Geophysical Research*
796 *Letters*, 47(14), e2020GL088758. [https://doi.org/https://doi.org/10.1029/2020GL088758](https://doi.org/10.1029/2020GL088758)
- 797 Lima, C. H. R., Kwon, H. H., & Kim, Y. T. (2018). A local-regional scaling-invariant Bayesian
798 GEV model for estimating rainfall IDF curves in a future climate. *Journal of Hydrology*,
799 566(September), 73–88. <https://doi.org/10.1016/j.jhydrol.2018.08.075>
- 800 Luke, A., Vrugt, J. A., AghaKouchak, A., Matthew, R., & Sanders, B. F. (2017). Predicting
801 nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the
802 United States. *Water Resources Research*, 53(7), 5469–5494.
803 <https://doi.org/10.1002/2016WR019676>
- 804 Madadgar, S., & Moradkhani, H. (2014). Improved Bayesian multimodeling: Integration of
805 copulas and Bayesian model averaging. *Water Resources Research*, 50(12), 9586–9603.
806 <https://doi.org/10.1002/2014WR015965>
- 807 Mailhot, A., Duchesne, S., Caya, D., & Talbot, G. (2007). Assessment of future change in
808 intensity-duration-frequency (IDF) curves for Southern Quebec using the Canadian
809 Regional Climate Model (CRCM). *Journal of Hydrology*, 347(1–2), 197–210.
810 <https://doi.org/10.1016/j.jhydrol.2007.09.019>
- 811 Marra, F., & Morin, E. (2015). Use of radar QPE for the derivation of intensity-duration-
812 frequency curves in a range of climatic regimes. *Journal of Hydrology*, 531, 427–440.
813 <https://doi.org/10.1016/j.jhydrol.2015.08.064>
- 814 Marra, F., Morin, E., Peleg, N., Mei, Y., & Anagnostou, E. N. (2017). Intensity-duration-
815 frequency curves from remote sensing rainfall estimates: Comparing satellite and weather

radar over the eastern Mediterranean. *Hydrology and Earth System Sciences*, 21(5), 2389–
2404. <https://doi.org/10.5194/hess-21-2389-2017>

Medeiros de Saboia, M. A., de Souza Filho, F. de A., Helfer, F., & Rolim, L. Z. R. (2020).
Robust strategy for assessing the costs of urban drainage system designs under climate
change scenarios. *Journal of Water Resources Planning and Management*, 146(11),
05020022. [https://doi.org/10.1061/\(asce\)wr.1943-5452.0001281](https://doi.org/10.1061/(asce)wr.1943-5452.0001281)

Mondal, A., & Daniel, D. (2019). Return levels under nonstationarity: The need to update
infrastructure design strategies. *Journal of Hydrologic Engineering*, 24(1), 04018060.
[https://doi.org/10.1061/\(asce\)he.1943-5584.0001738](https://doi.org/10.1061/(asce)he.1943-5584.0001738)

Nagler, T., Schellhase, C., & Czado, C. (2017). Nonparametric estimation of simplified vine
copula models: comparison of methods. *Dependence Modeling*, 5(1), 99-120.
<https://doi.org/10.1515/demo-2017-0007>

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. lin. (2018). Developing intensity-duration-
frequency (IDF) curves from satellite-based precipitation: Methodology and evaluation.
Water Resources Research, 54(10), 7752–7766. <https://doi.org/10.1029/2018WR022929>

Ouarda, T. B. M. J., Yousef, L. A., & Charron, C. (2019). Non-stationary intensity-duration-
frequency curves integrating information concerning teleconnections and climate change.
International Journal of Climatology, 39(4), 2306–2323. <https://doi.org/10.1002/joc.5953>

Papalexiou, S. M. (2018). Unified theory for stochastic modelling of hydroclimatic processes:
Preserving marginal distributions, correlation structures, and intermittency. *Advances in*
Water Resources, 115, 234–252. <https://doi.org/10.1016/j.advwatres.2018.02.013>

Qi, Y., Chan, F. K. S., O'Donnell, E. C., Feng, M., Sang, Y., Thorne, C. R., et al. (2021).
Exploring the development of the Sponge City Program (SCP): The case of Gui'an new

- 839 district, Southwest China. *Frontiers in Water*, 3, 41.
840 <https://doi.org/10.3389/frwa.2021.676965>
- 841 Qing, Y., Wang, S., Ancell, B. C., & Yang, Z.-L. (2022). Accelerating flash droughts induced by
842 the joint influence of soil moisture depletion and atmospheric aridity. *Nature*
843 *Communications*, 13, 1139. <https://doi.org/10.1038/s41467-022-28752-4>
- 844 Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). Using Bayesian model
845 averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5), 1155–1174.
846 <https://doi.org/10.1175/MWR2906.1>
- 847 Ragno, E., AghaKouchak, A., Love, C. A., Cheng, L., Vahedifard, F., & Lima, C. H. R. (2018).
848 Quantifying changes in future intensity-duration-frequency curves using multimodel
849 ensemble simulations. *Water Resources Research*, 54(3), 1751–1764.
850 <https://doi.org/10.1002/2017WR021975>
- 851 Rastogi, D., Touma, D., Evans, K. J., & Ashfaq, M. (2020). Shift toward intense and widespread
852 precipitation events over the United States by mid-21st century. *Geophysical Research*
853 *Letters*, 47, e2020GL089899. <https://doi.org/10.1029/2020GL089899>
- 854 Reuters. (2021). Further flooding feared in western Germany with death toll above 80. Retrieved
855 October 17, 2021, from [https://www.reuters.com/world/europe/floods-germany-claim-81-](https://www.reuters.com/world/europe/floods-germany-claim-81-victims-more-than-1000-missing-2021-07-16/)
856 [victims-more-than-1000-missing-2021-07-16/](https://www.reuters.com/world/europe/floods-germany-claim-81-victims-more-than-1000-missing-2021-07-16/)
- 857 Ritschel, C., Ulbrich, U., N  vir, P., & Rust, H. W. (2017). Precipitation extremes on multiple
858 timescales - Bartlett-Lewis rectangular pulse model and intensity-duration-frequency
859 curves. *Hydrology and Earth System Sciences*, 21(12), 6501–6517.
860 <https://doi.org/10.5194/hess-21-6501-2017>

- Ryan Woo, & Stella Qiu. (2021). At least 25 dead as rains deluge central China's Henan province | Reuters. Retrieved March 29, 2022, from <https://www.reuters.com/world/china/heavy-rainfall-kills-12-central-chinas-henan-provincial-capital-xinhua-2021-07-20/>
- Sarhadi, A., & Soulis, E. D. (2017). Time-varying extreme rainfall intensity-duration-frequency curves in a changing climate. *Geophysical Research Letters*, 44(5), 2454–2463. <https://doi.org/10.1002/2016GL072201>
- Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's Tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. <https://doi.org/10.1080/01621459.1968.10480934>
- Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. Di, et al. (2021). Weather and climate extreme events in a changing climate. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (In Press). Cambridge University Press.
- Serinaldi, F., & Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources*, 77, 17–36. <https://doi.org/10.1016/j.advwatres.2014.12.013>
- Simonovic, S. P., Schardong, A., Sandink, D., & Srivastav, R. (2016). A web-based tool for the development of Intensity Duration Frequency curves under changing climate. *Environmental Modelling and Software*, 81, 136–153. <https://doi.org/10.1016/j.envsoft.2016.03.016>

- 882 Slater, L., Villarini, G., Archfield, S., Faulkner, D., Lamb, R., Khouakhi, A., & Yin, J. (2021).
 883 Global changes in 20-year, 50-year, and 100-year river floods. *Geophysical Research*
 884 *Letters*, 48(6). <https://doi.org/10.1029/2020GL091824>
- 885 Slater, L., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., et al. (2021).
 886 Nonstationary weather and water extremes: A review of methods for their detection,
 887 attribution, and management. *Hydrology and Earth System Sciences*, 25(7), 3897–3935.
 888 <https://doi.org/10.5194/hess-25-3897-2021>
- 889 So, B. J., Kim, J. Y., Kwon, H. H., & Lima, C. H. R. (2017). Stochastic extreme downscaling
 890 model for an assessment of changes in rainfall intensity-duration-frequency curves over
 891 South Korea using multiple regional climate models. *Journal of Hydrology*, 553, 321–337.
 892 <https://doi.org/10.1016/j.jhydrol.2017.07.061>
- 893 Srivastav, R.K., Schardong, A. & Simonovic, S.P. (2014). Equidistance quantile matching
 894 method for updating IDF curves under climate change. *Water Resources Management*, 28,
 895 2539–2562. <https://doi.org/10.1007/s11269-014-0626-y>
- 896 Srivastava, A. K., Grotjahn, R., Ullrich, P. A., & Sadegh, M. (2021). Pooling data improves
 897 multimodel IDF estimates over median-based IDF estimates: Analysis over the
 898 Susquehanna and Florida. *Journal of Hydrometeorology*, 22(4), 971–995.
 899 <https://doi.org/10.1175/jhm-d-20-0180.1>
- 900 Sun, C., Huang, G., Fan, Y., Zhou, X., Lu, C., & Wang, X. (2021). Vine copula ensemble
 901 downscaling for precipitation projection over the Loess Plateau based on high-resolution
 902 multi-RCM outputs. *Water Resources Research*, 57(1).
 903 <https://doi.org/10.1029/2020WR027698>

- Switzman, H., Razavi, T., Traore, S., Coulibaly, P., Burn, D. H., Henderson, J., et al. (2017). Variability of future extreme rainfall statistics: Comparison of multiple IDF projections. *Journal of Hydrologic Engineering*, 22(10), 04017046. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001561](https://doi.org/10.1061/(asce)he.1943-5584.0001561)
- The Washington Post. (2021). New York braces for new round of flooding after summer of wild rains. Retrieved April 11, 2022, from <https://www.washingtonpost.com/weather/2021/09/23/flood-watch-new-york-city/>
- Wang, S., Ancell, B. C., Huang, G. H., & Baetz, B. W. (2018). Improving robustness of hydrologic ensemble predictions through probabilistic pre- and post-processing in sequential data assimilation. *Water Resources Research*, 54(3), 2129–2151. <https://doi.org/10.1002/2018WR022546>
- Wanted in Milan. (2021). Malpensa: Milan airport under water in flash floods. Retrieved September 25, 2021, from <https://www.wantedinmilan.com/news/malpensa-milan-airport-under-water-in-flash-floods.html>
- Wood, A. W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research: Atmospheres*, 107(D20), 4429. <https://doi.org/10.1029/2001JD000659>
- Wright, D. B., Samaras, C., & Lopez-Cantu, T. (2021). Resilience to extreme rainfall starts with science. *Bulletin of the American Meteorological Society*, 102(4), E808–E813. <https://doi.org/10.1175/BAMS-D-20-0267.1>
- Wu, M., Luo, Y., Chen, F., & Wong, W. K. (2019). Observed link of extreme hourly precipitation changes to urbanization over coastal South China. *Journal of Applied*

- 926 *Meteorology and Climatology*, 58(8), 1799–1819. [https://doi.org/10.1175/JAMC-D-18-](https://doi.org/10.1175/JAMC-D-18-0284.1)
- 927 0284.1
- 928 Xu, G., Zhou, Z., Jiao, L., & Zhao, R. (2020). Compact urban form and expansion pattern slow
- 929 down the decline in urban densities: A global perspective. *Land Use Policy*, 94.
- 930 <https://doi.org/10.1016/j.landusepol.2020.104563>
- 931 Xu, Q., Chen, J., Peart, M. R., Ng, C. N., Hau, B. C. H., & Law, W. W. Y. (2018). Exploration
- 932 of severities of rainfall and runoff extremes in ungauged catchments: A case study of Lai
- 933 Chi Wo in Hong Kong, China. *Science of the Total Environment*, 634, 640–649.
- 934 <https://doi.org/10.1016/j.scitotenv.2018.04.024>
- 935 Yan, L., Xiong, L., Jiang, C., Zhang, M., Wang, D., & Xu, C. Y. (2021). Updating intensity–
- 936 duration–frequency curves for urban infrastructure design under a changing environment.
- 937 *Wiley Interdisciplinary Reviews: Water*, 8(3), 1–24. <https://doi.org/10.1002/wat2.1519>
- 938 Yin, J., Gu, H., Liang, X., Yu, M., Sun, J., Xie, Y., et al. (2022). A possible dynamic mechanism
- 939 for rapid production of the extreme hourly rainfall in Zhengzhou City on 20 July 2021.
- 940 *Journal of Meteorological Research*, 36(1), 6–25. [https://doi.org/10.1007/s13351-022-](https://doi.org/10.1007/s13351-022-1166-7)
- 941 1166-7
- 942 You, J., & Wang, S. (2021). Higher probability of occurrence of hotter and shorter heat waves
- 943 followed by heavy rainfall. *Geophysical Research Letters*, 48(17), e2021GL094831.
- 944 <https://doi.org/10.1029/2021GL094831>
- 945 Yu, X., Gu, X., Kong, D., Zhang, Q., Cao, Q., Slater, L. J., et al. (2022). Asymmetrical shift
- 946 toward less light and more heavy precipitation in an urban agglomeration of East China:
- 947 Intensification by urbanization. *Geophysical Research Letters*, 49(4), e2021GL097046.
- 948 <https://doi.org/10.1029/2021GL097046>

- Zaerpour, M., Papalexiou, S. M., & Nazemi, A. (2021). Informing stochastic streamflow generation by large-scale climate indices at single and multiple sites. *Advances in Water Resources*, 156, 104037. <https://doi.org/10.1016/j.advwatres.2021.104037>
- Zhang, B., & Wang, S. (2021). Probabilistic characterization of extreme storm surges induced by tropical cyclones. *Journal of Geophysical Research: Atmospheres*, 126(3), e2020JD033557. <https://doi.org/10.1029/2020JD033557>
- Zhang, B., Wang, S., & Zhu, J. (2021). A weighted ensemble of regional climate projections for exploring the spatiotemporal evolution of multidimensional drought risks in a changing climate. *Climate Dynamics*, 1–20. <https://doi.org/10.1007/s00382-021-05889-4>
- Zhang, B., Wang, S., Qing, Y., Zhu, J., Wang, D., & Liu, J. (2022). A vine copula-based polynomial chaos framework for improving multi-model hydroclimatic projections at a multi-decadal convection-permitting scale. *Water Resources Research*, 58(6), e2022WR031954. <https://doi.org/10.1029/2022wr031954>
- Zhou, Z., Smith, J. A., Wright, D. B., Baeck, M. L., Yang, L., & Liu, S. (2019). Storm catalog-based analysis of rainfall heterogeneity and frequency in a complex terrain. *Water Resources Research*, 55(3), 1871–1889. <https://doi.org/10.1029/2018WR023567>
- Zhu, J., Huang, G., Wang, X., Cheng, G., & Wu, Y. (2018). High-resolution projections of mean and extreme precipitations over China through PRECIS under RCPs. *Climate Dynamics*, 50(11–12), 4037–4060. <https://doi.org/10.1007/s00382-017-3860-1>

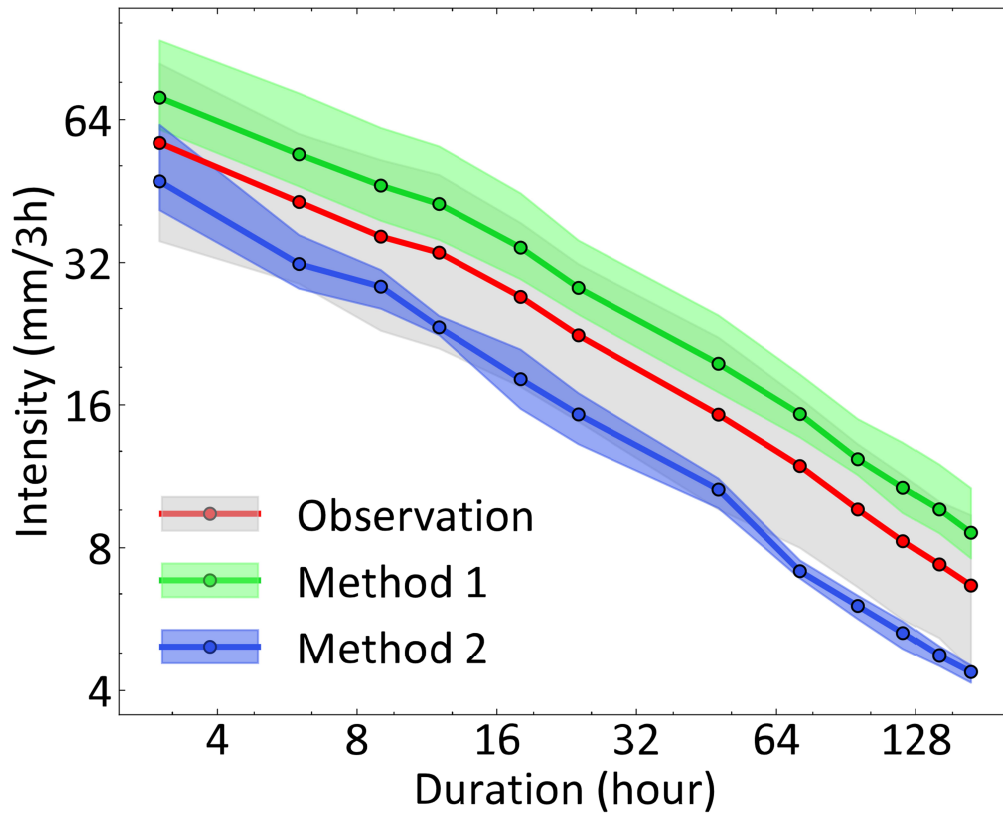
969 **Figures**

Figure 1. An illustrative example of evaluating IDF curves. The grey ribbon represents the 95% confidence interval owing to the uncertainty in the parameters of the GEV distribution. The green and blue ribbons represent the 95% confidence interval owing to the uncertainty in the parameters of the GEV distribution and the structural uncertainty of climate models.

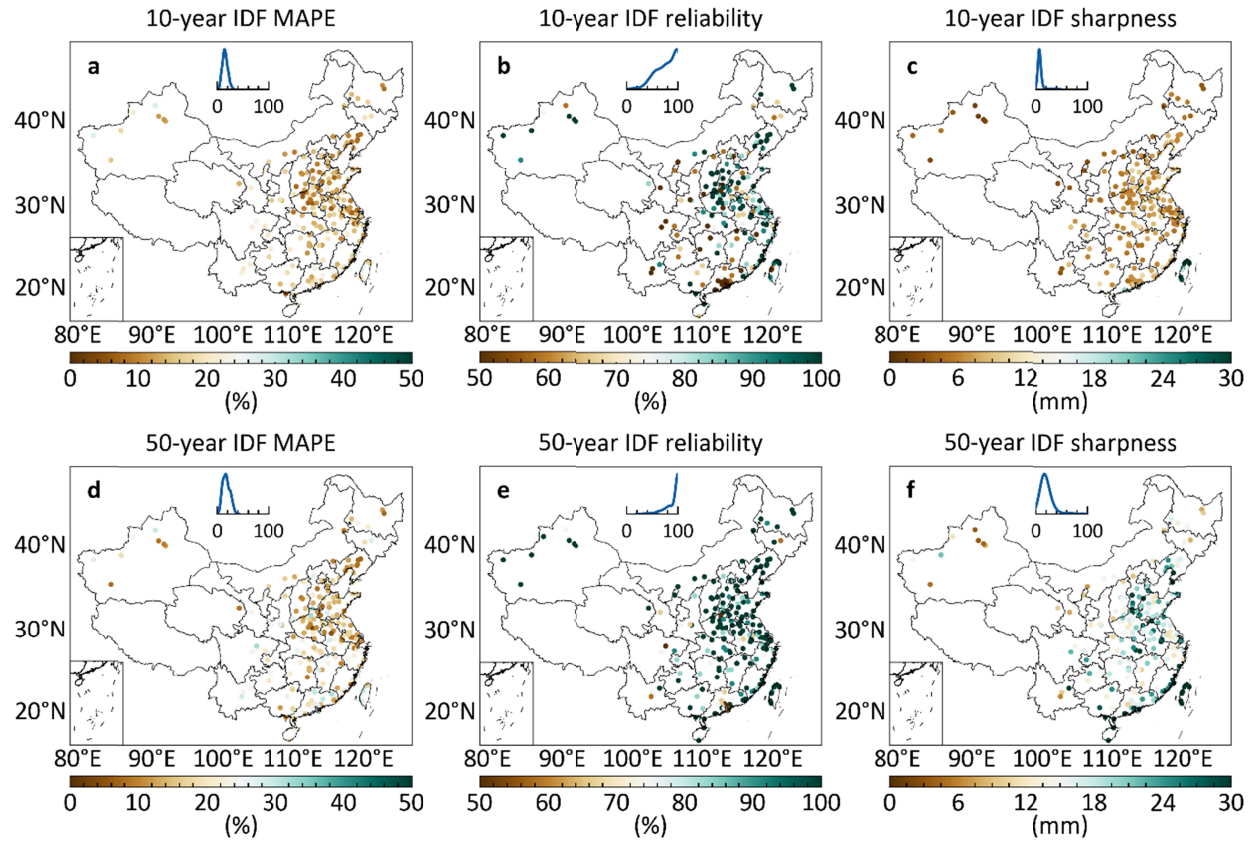


Figure 2. The MAPE, reliability, and sharpness of IDF curves generated based on the multi-model ensemble stochastic spatiotemporal downscaling of precipitation in China, given the return periods of 10 and 50 years. MAPE = mean absolute percentage error. The density plot in each panel corresponds to the probability density function (PDF) of values across 196 Chinese cities.

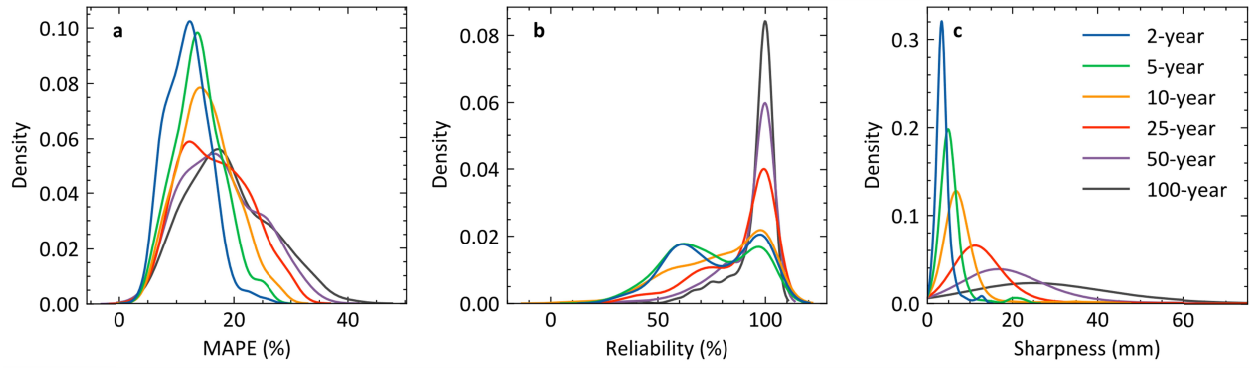


Figure 3. The MAPE, reliability, and sharpness of IDF curves generated based on the multi-model ensemble stochastic spatiotemporal downscaling of precipitation across 196 Chinese cities given the return periods of 2, 5, 10, 25, 50, and 100 years. The legend in **c** is also applicable to **a** and **b**.

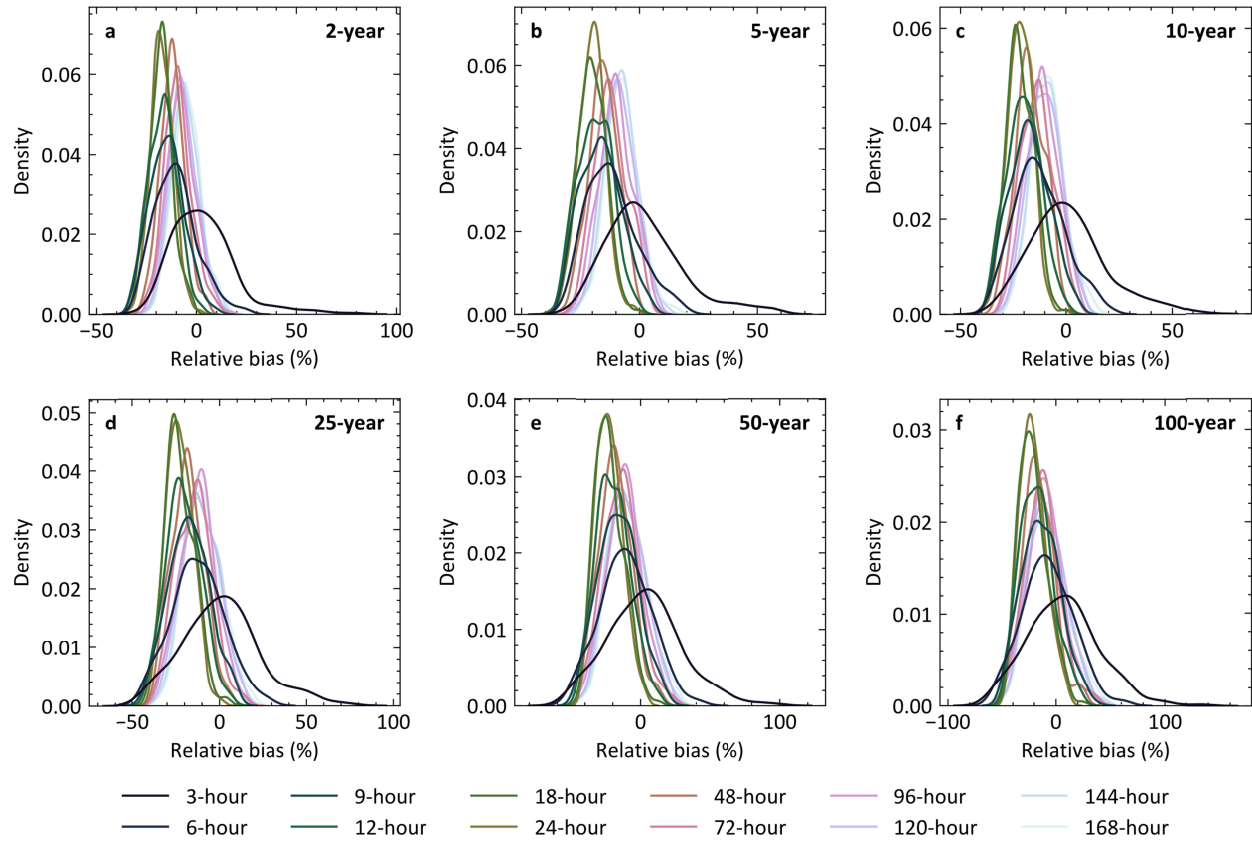


Figure 4. The PDF plots of the relative bias for the multi-model ensemble mean return levels of extreme precipitation across 196 cities, given different durations (from 3 hours to 168 hours) and return periods (2, 5, 10, 25, 50, and 100 years).

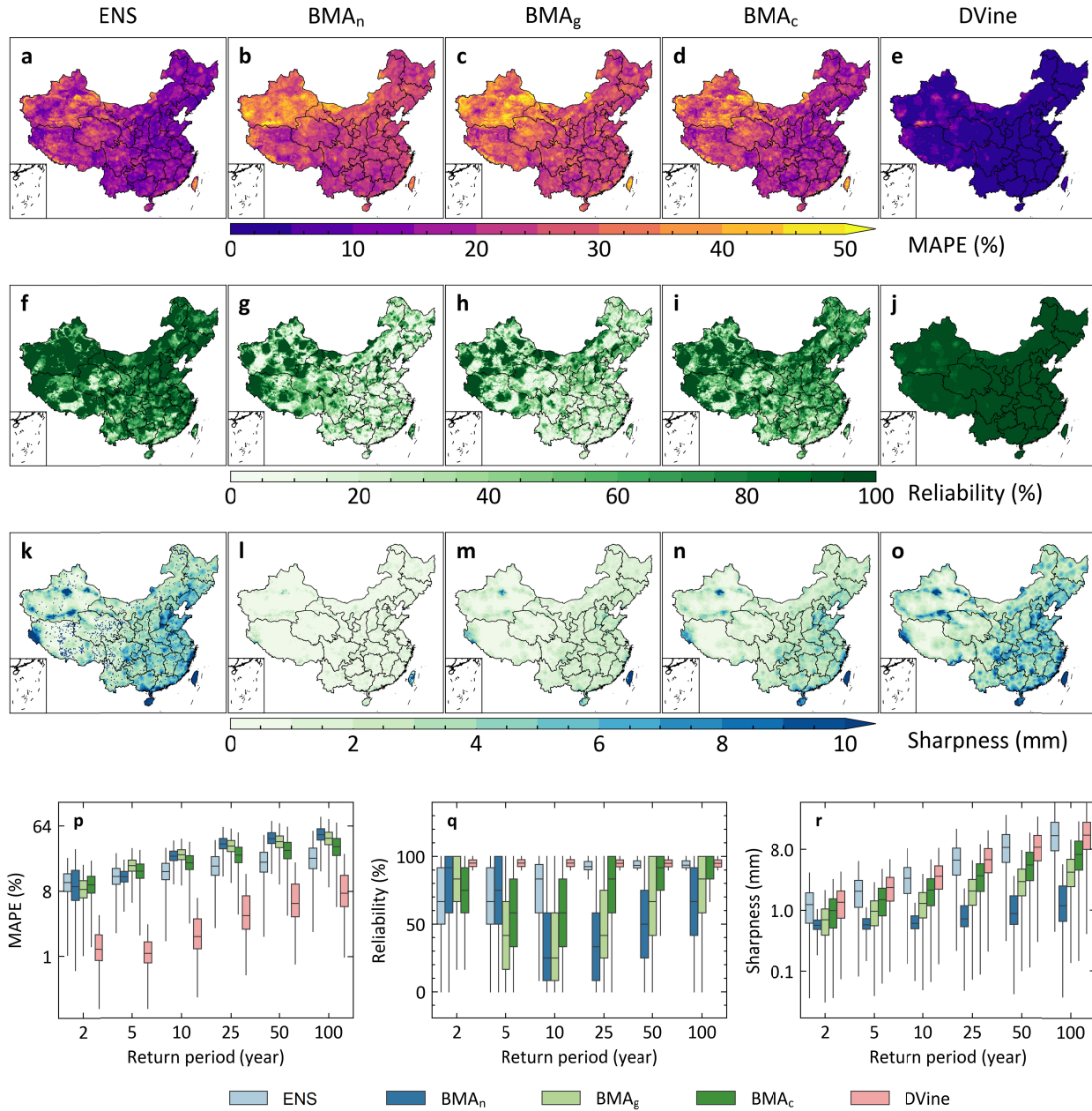


Figure 5. a–o The MAPE, reliability, and sharpness of IDF curves generated based on the stochastic spatiotemporal downscaling of precipitation using different multi-model combination approaches, given a return period of 10 years. p–r The MAPE, reliability, and sharpness of IDF curves given different return periods. MAPE = mean absolute percentage error. ENS = ensemble mean. DVine = drawable vine copula. BMA_n, BMA_g, and BMA_c represent Bayesian model

1002 averaging in which the conditional PDF is assumed to follow a normal, gamma, and copula-
1003 based distribution, respectively. The boxes in **p-r** are drawn from the 25th percentile (Q_1) to the
1004 75th percentile (Q_3) with a horizontal line in the middle denoting the median. The upper and
1005 lower whiskers represent the values of $Q_3 + 1.5 \times \text{IQR}$ and $Q_1 - 1.5 \times \text{IQR}$, respectively, where
1006 IQR denotes the interquartile range equal to $Q_3 - Q_1$.

1007

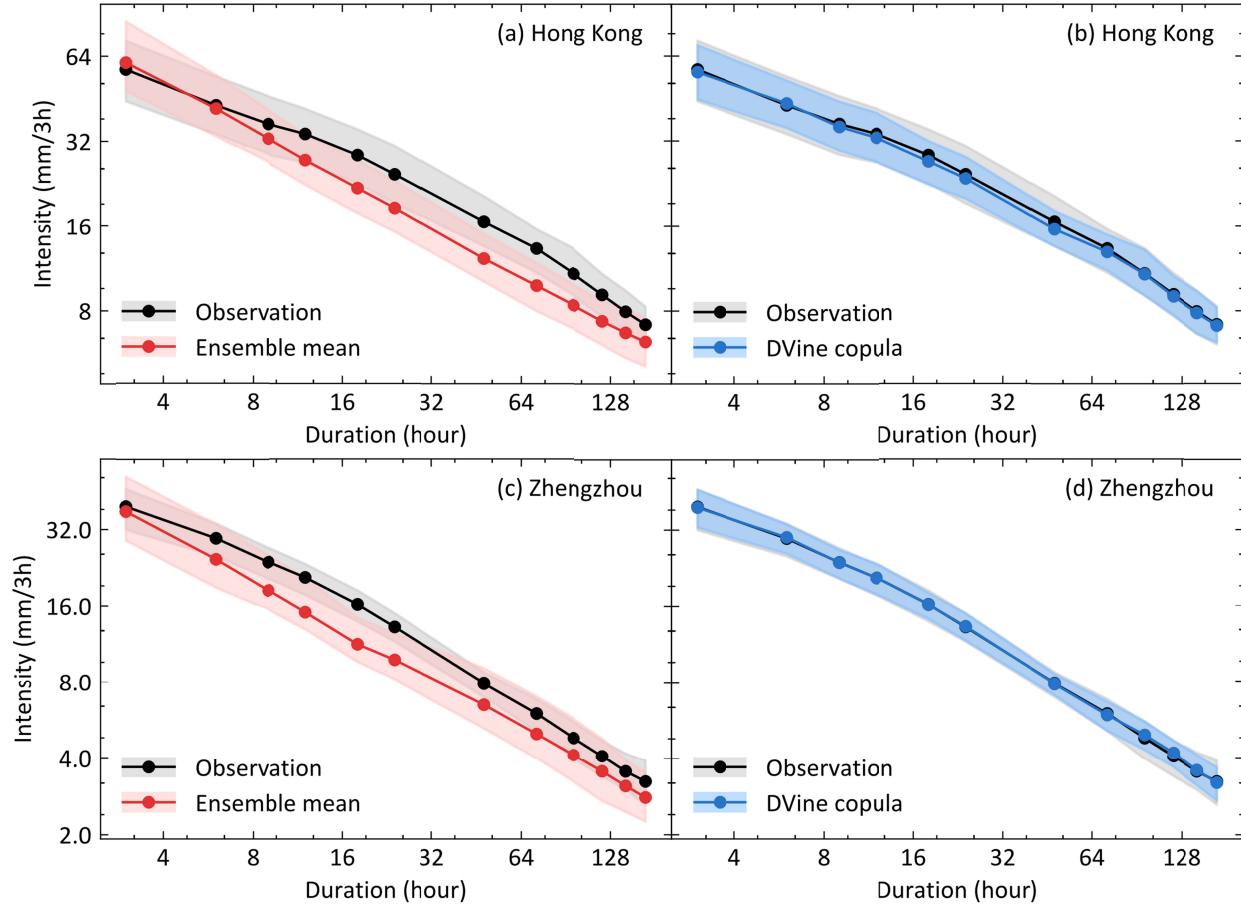


Figure 6. Comparison of the ensemble mean method and the DVine copula method used to reproduce precipitation IDF curves over the urban area in **a, b** Hong Kong and **c, d** Zhengzhou during the period of 1979–2005, given a return period of 10 years. The grey ribbon represents the bootstrapped 95% confidence interval for design value estimates. The red ribbon represents the 2.5th and 97.5th percentile interval of 5×500 design value estimates generated from 500 parametric bootstrap samples for five downscaled RCM model outputs. The blue ribbon represents the 2.5th and 97.5th percentile interval of 500×500 design value estimates generated from 500 parametric bootstrap samples for 500 vine copula-based ensemble simulation outputs.

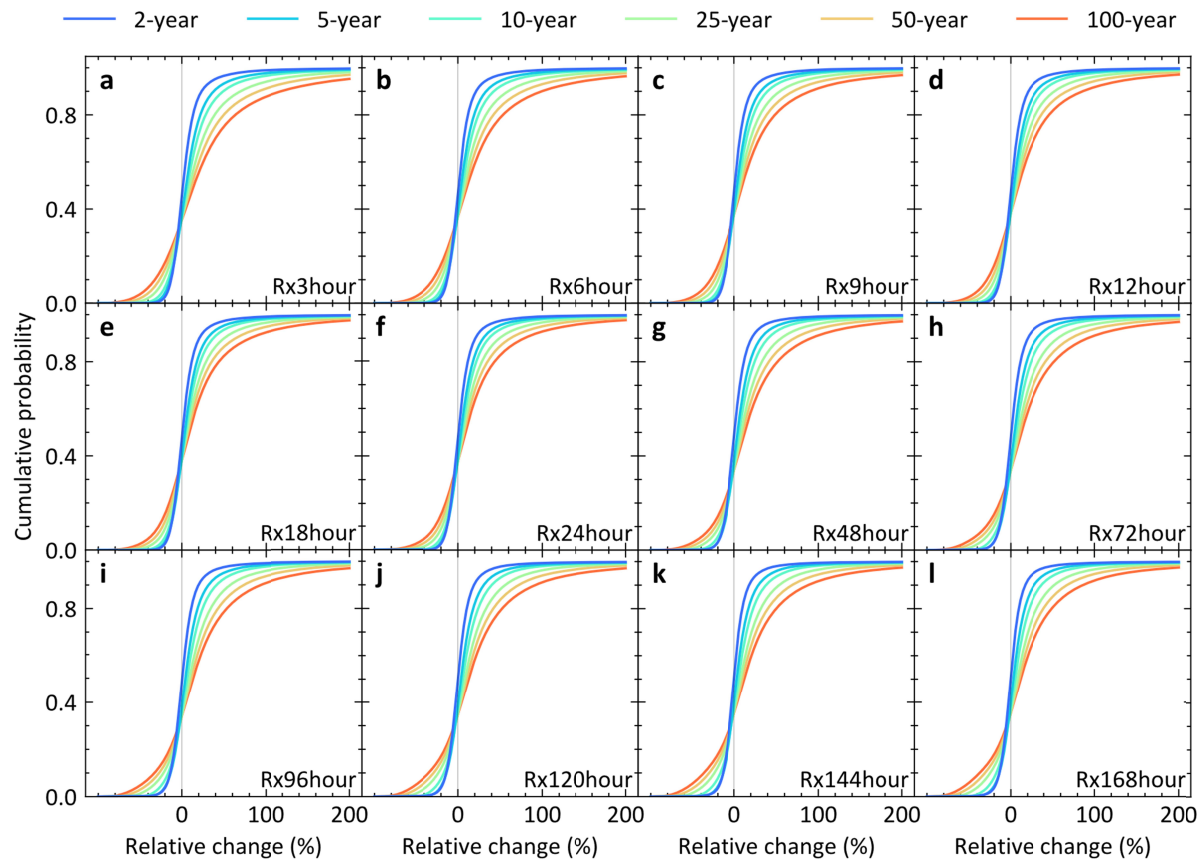


Figure 7. Cumulative probability of the relative difference in extreme precipitation across all the 0.1° land grids of China between past (1979–2020) and future (2058–2099) climates in terms of different return periods and different event durations.

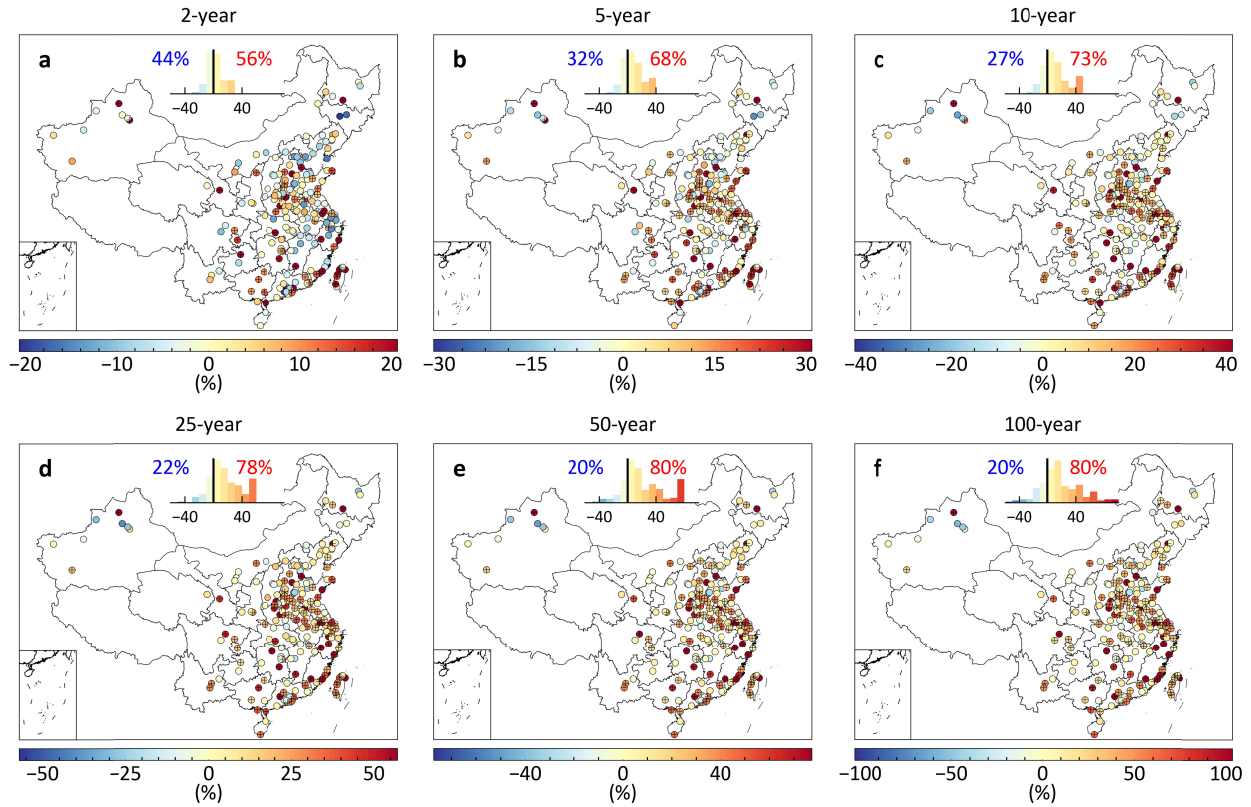


Figure 8. Spatial patterns of the relative difference in 3-hour heavy rainfall between past and future climates across 196 cities given different return periods. The circles with crosses represent cities where the 3-hour heavy rainfall intensity under the future climate is significantly ($P < 0.05$) higher than under the present climate. The statistical significance is estimated using the method described in section 2.2. The inset bar chart in each panel indicates the proportion of cities with decreasing (blue) or increasing (red) rainfall intensity.

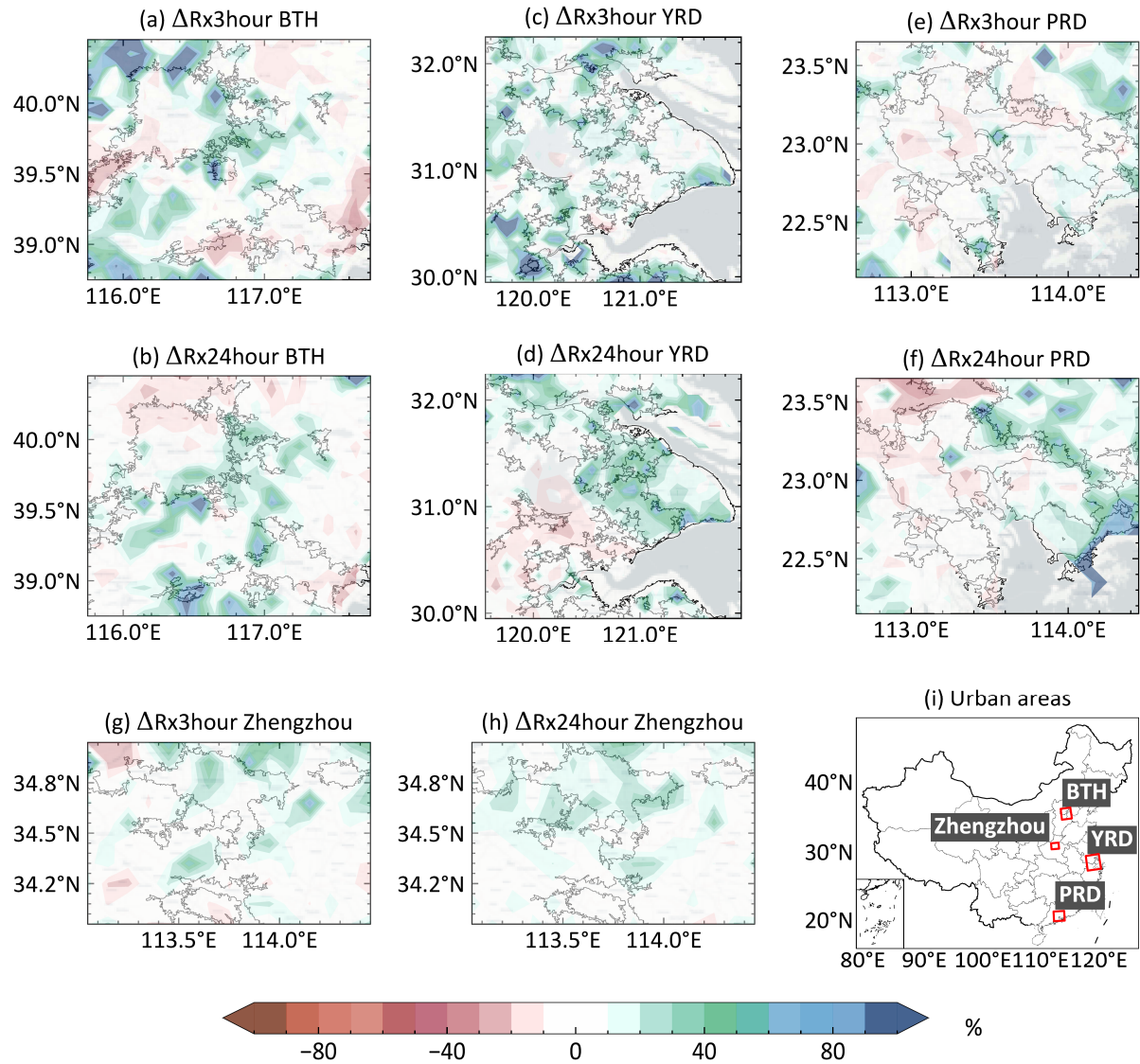


Figure 9. Spatial patterns of the relative difference in 3-hour and 24-hour heavy rainfall between past and future climates over four selected urban areas, given a return period of 50 years. BTH represents the Beijing-Tianjin-Hebei metropolitan area. YRD represents the Yangtze River Delta and PRD represents the Pearl River Delta.

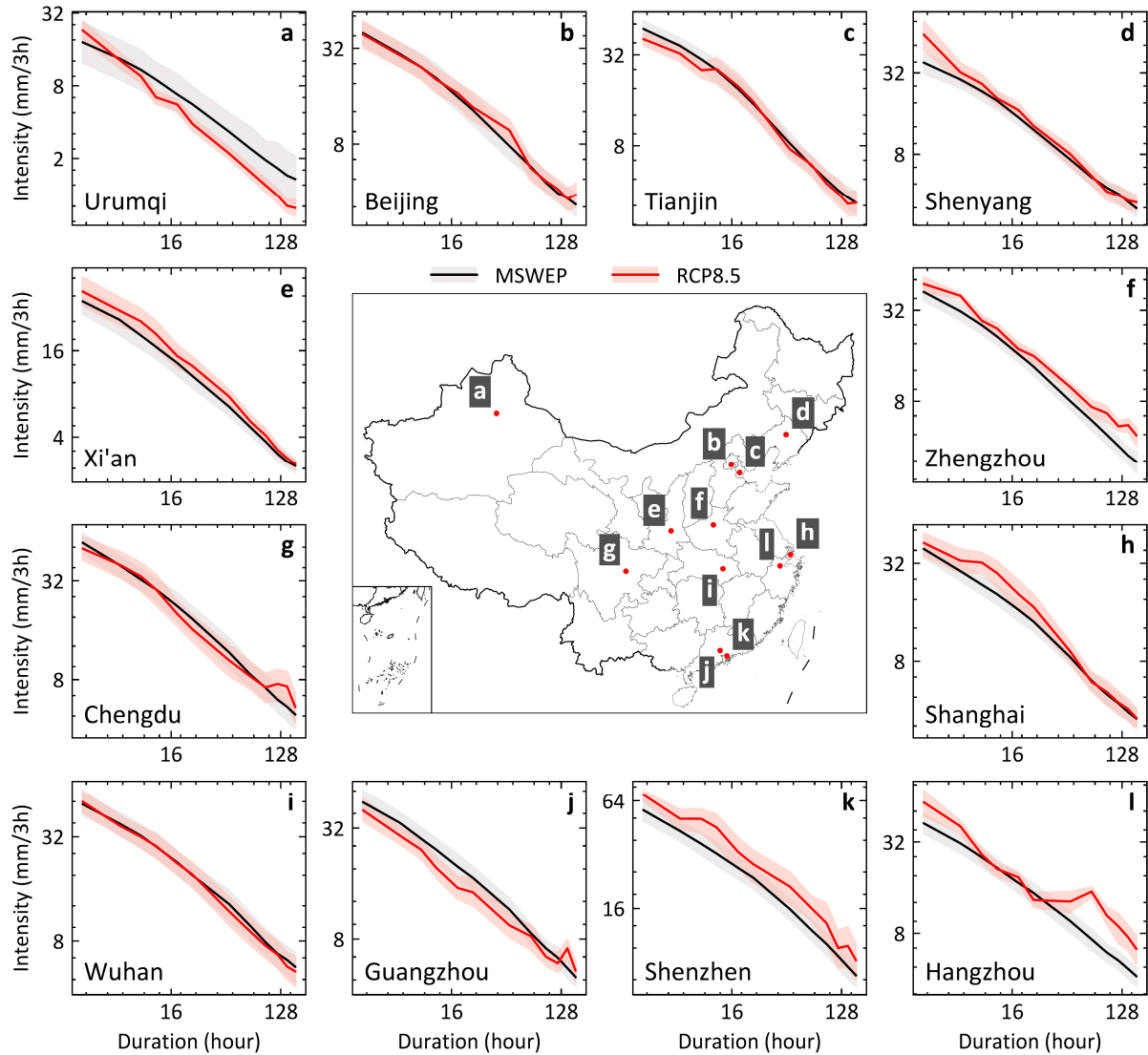


Figure 10. IDF curves with 95% confidence intervals between past (MSWEP) and future (RCP8.5) climates, given a return period of 10 years. The grey ribbon represents the bootstrapped 95% confidence interval for the design value estimates. The red ribbon represents the 2.5th and 97.5th percentile interval of 500×500 design value estimates generated from 500 parametric bootstrap samples for 500 vine copula-based ensemble simulation outputs.

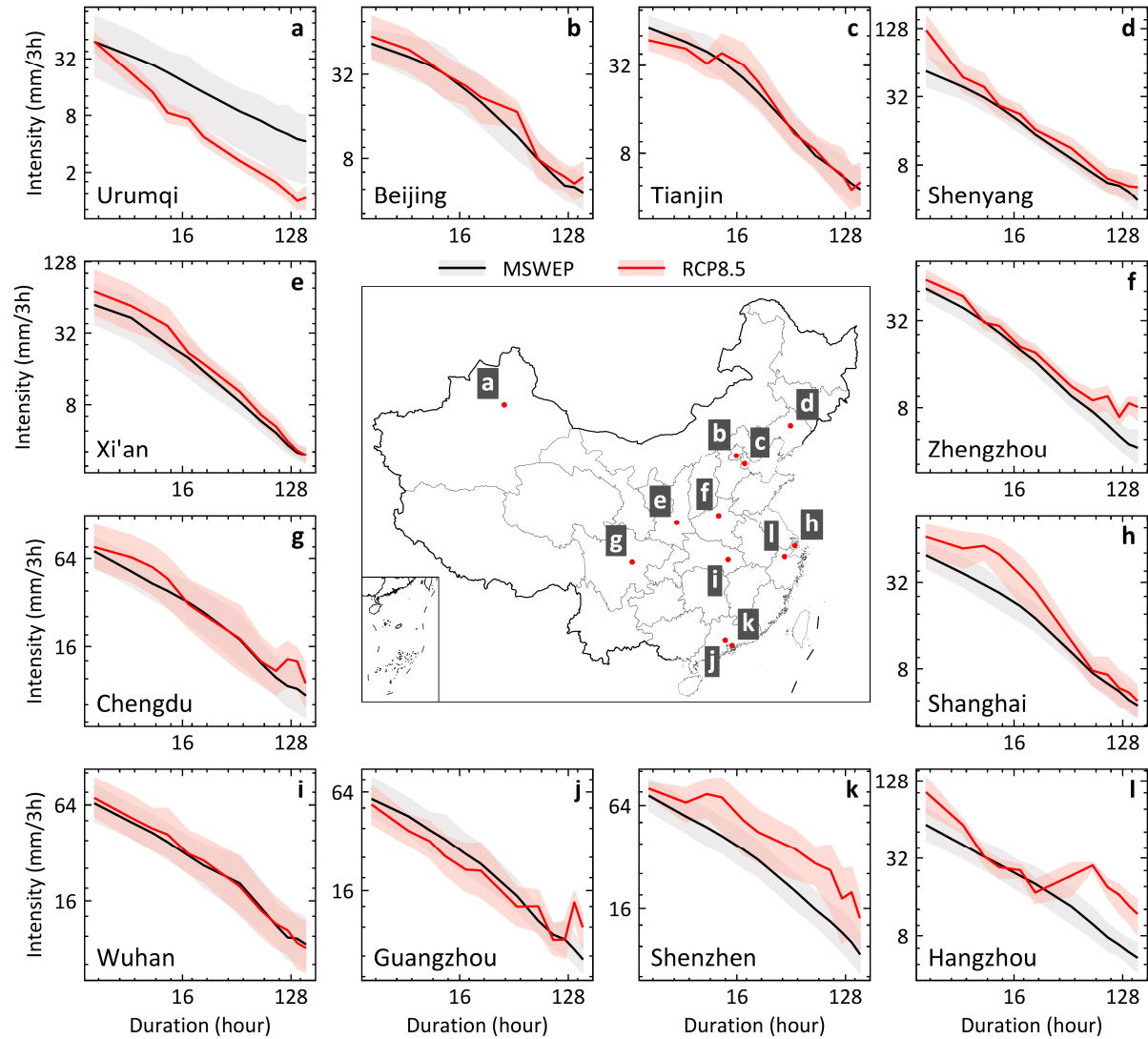


Figure 11. IDF curves with 95% confidence intervals between past (MSWEP) and future (RCP8.5) climates, given a return period of 50 years. The setting of the ribbon plot is the same as Figure 10.

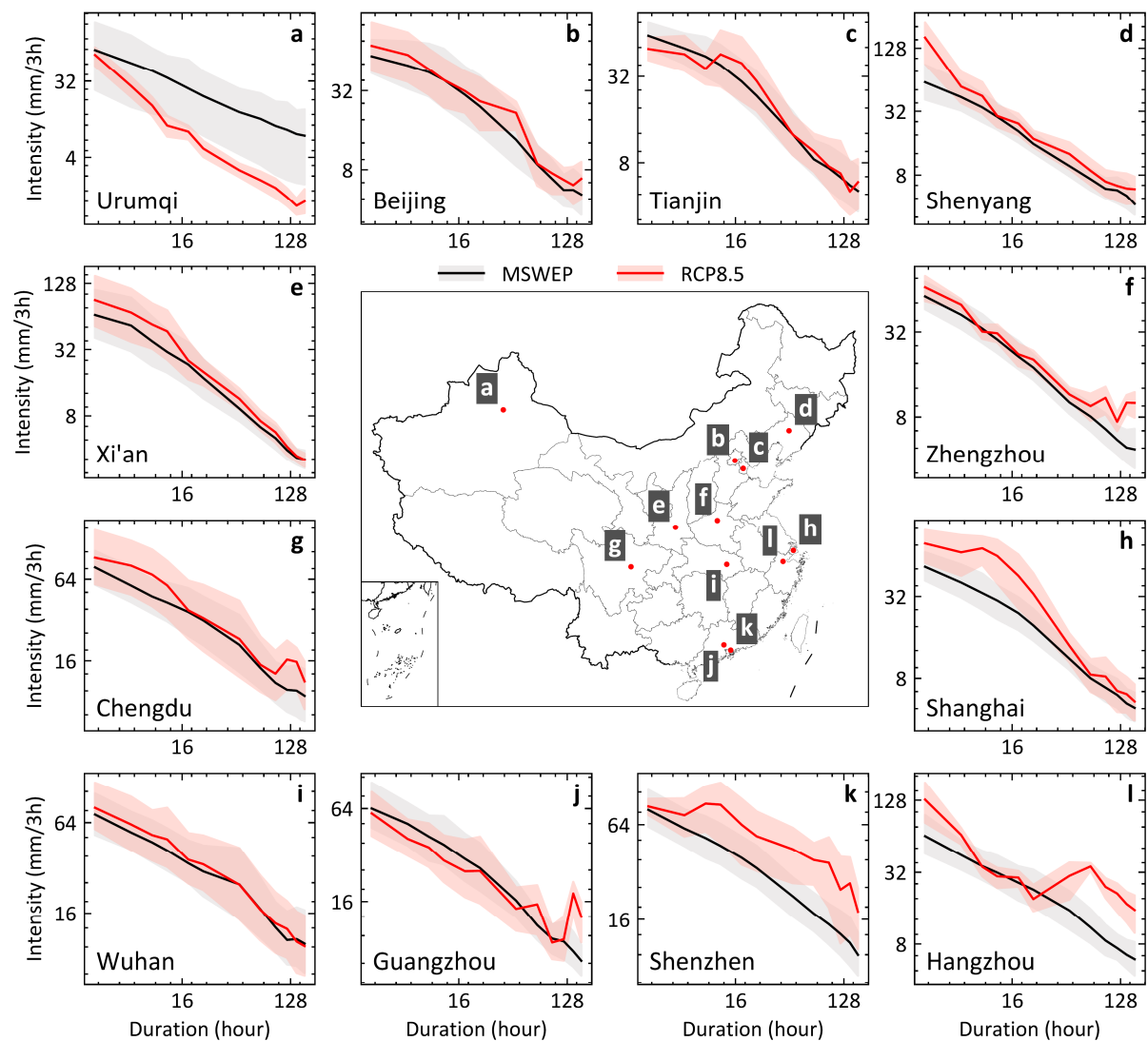


Figure 12. IDF curves with 95% confidence intervals between past (MSWEP) and future (RCP8.5) climates, given a return period of 100 years. The setting of the ribbon plot is the same as Figure 10.

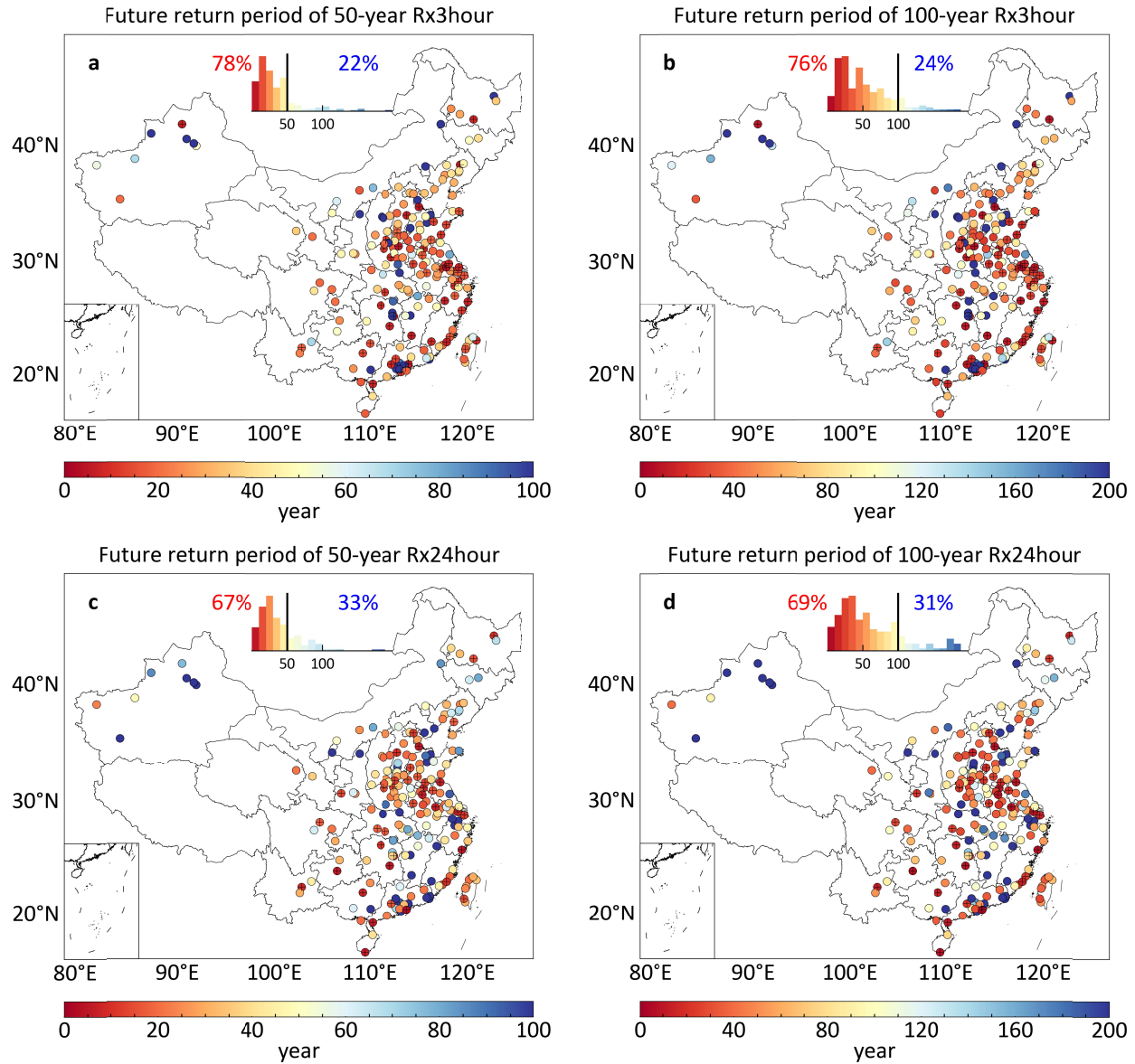


Figure 13. Future (2058–2099) return periods of 3-hour (Rx3hour) and 24-hour (Rx24hour) heavy rainfall currently associated with return periods of 50 and 100 years (1979–2005) over urban areas of 196 cities. Circles with crosses represent cities with a significant ($P < 0.05$) difference in the return period of heavy rainfall between past and future climates. The statistical significance is estimated using the method described in section 2.2. The inset bar chart in each panel indicates the proportion of cities with more (red) or less (blue) frequent heavy rainfall events under the future climate.

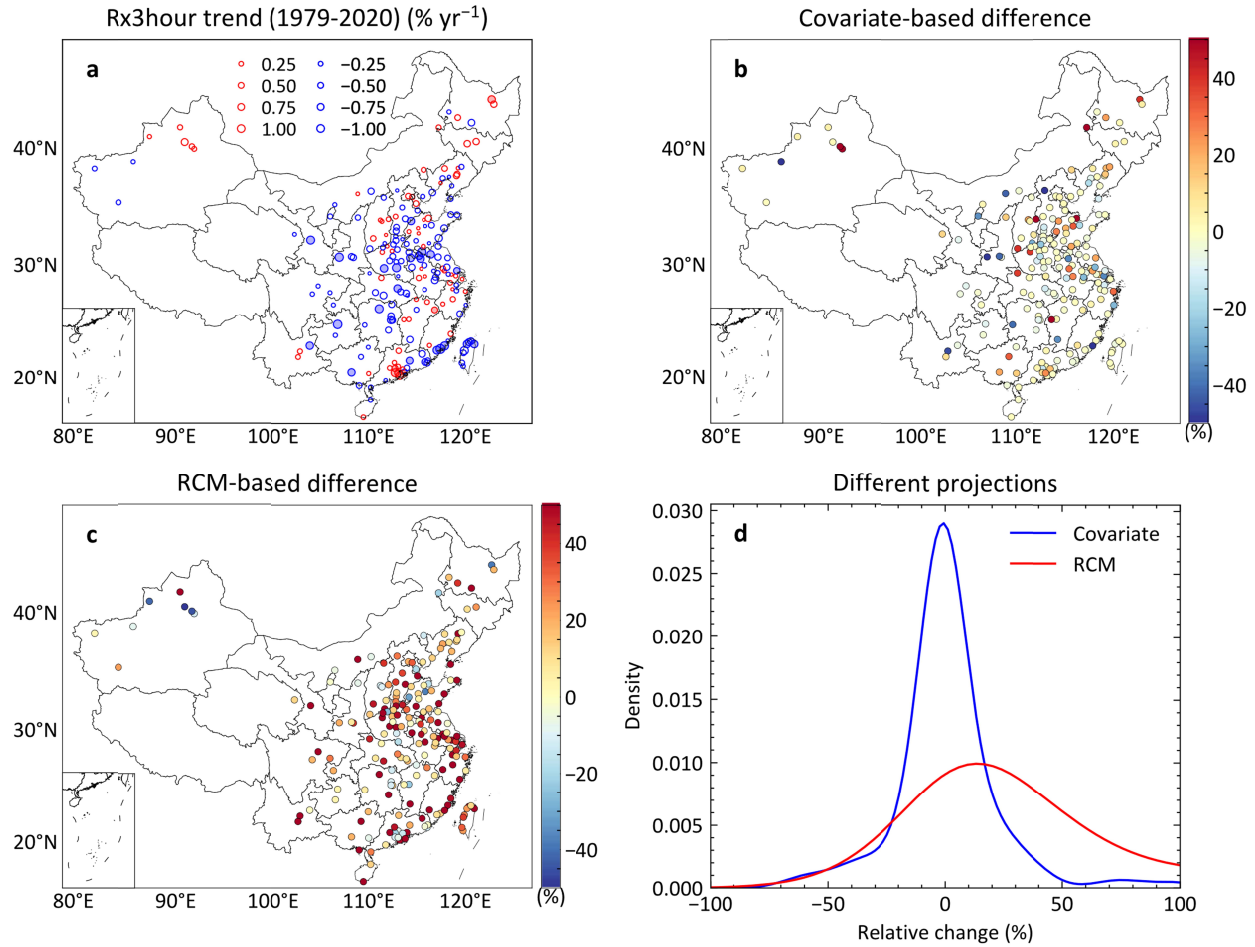


Figure 14. **a** Linear trends of annual maximum 3-hour rainfall estimated by Sen's slope estimator averaged over the urban areas of 196 Chinese cities from 1979 to 2020 (Sen, 1968); solid circles represents significant ($P < 0.05$) trends and the significance is estimated using Mann-Kendall's test (Kendall, 1975). **b, c** projected relative changes in 3-hour heavy rainfall associated with a 100-year return period generated using the covariate-based and climate-model-based methods between past (1979–2020) and future (2058–2099) climates. **d** PDF plots of relative changes in the 3-hour heavy rainfall associated with a 100-year return period across 196 cities.

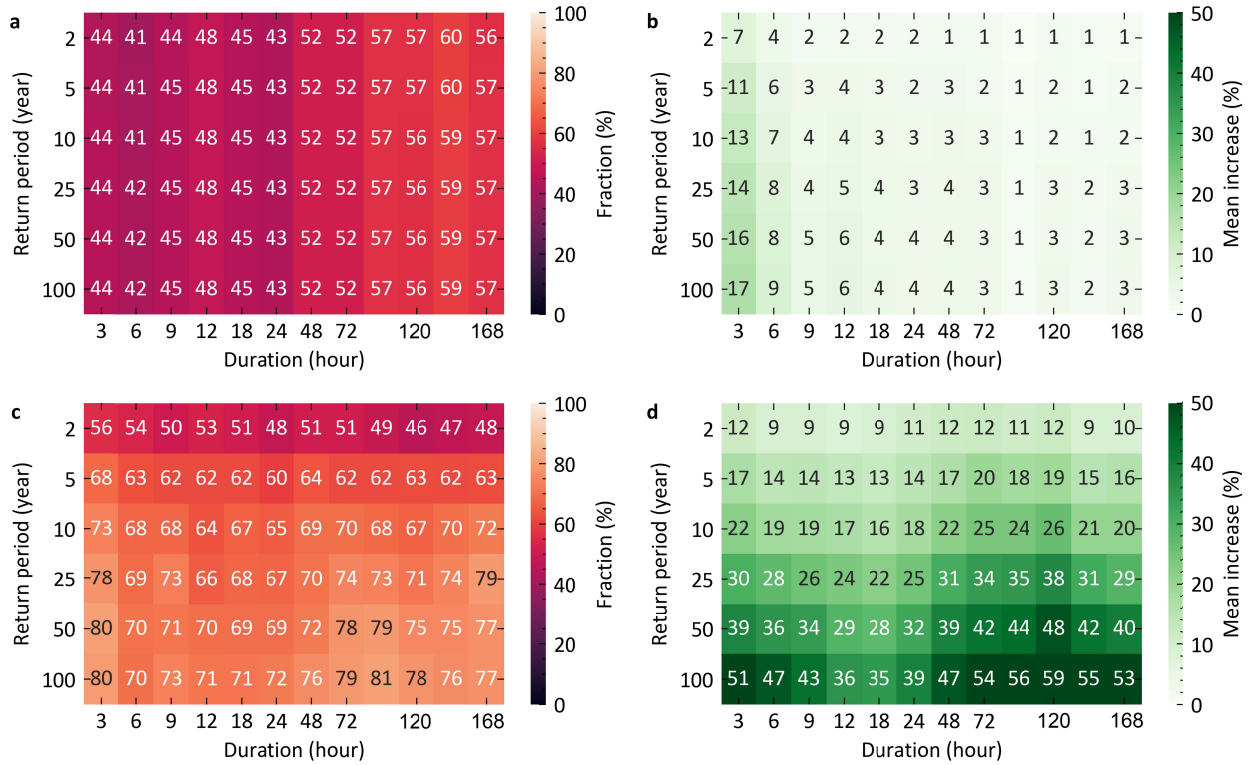


Figure 15. Projected changes in the design values of precipitation IDF curves generated from **a**, **b** the covariate-based nonstationary precipitation IDF curves and **c**, **d** the climate-model-based precipitation IDF curves given different time intervals (i.e., durations) and different return periods. **a**, **c** The fraction of 196 cities showing increases in extreme precipitation intensity between past (1979–2020) and future (2058–2099) climates. **b**, **d** The average increases of extreme precipitation intensity for the cities showing increased intensity between past and future climates.