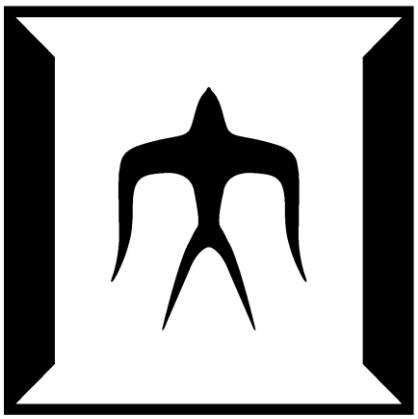
論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Quantification of the climate change impacts on extreme rainfall and the associated risk in Southeast Asia
著者(和文)	Zhao Wenpeng
Author(English)	Wenpeng Zhao
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12226号, 授与年月日:2022年9月22日, 学位の種別:課程博士, 審査員:木内 豪,神田 学,中村 恭志,髙木 泰士,VARQUEZ ALVIN CHRIST
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12226号, Conferred date:2022/9/22, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Quantification of the climate change impacts on extreme rainfall and the associated risk in Southeast Asia



Wenpeng Zhao

A dissertation submitted for the Degree of Doctor of Engineering

Global Engineering for Development, Environment and Society (GEDES)

School of Environment and Society

Tokyo Institute of Technology

July 2022

Contents

1. Introduction

1.1 Background

Human-induced global warming is widely recognized, with an estimated increase in global mean temperature by approximately 1.08±0.13 ℃ in 2021 above the pre-industrial level (WMO, 2021). Although 197 participating countries in the recently held 26th Conference of the Parties (COP26) have broadly agreed to limit the warming to 1.5 ℃ at the end of the 21st century by introducing new controlling measures of ground rules and emission reporting, the collective political will and policy frameworks may not be adequate to achieve the goal owing to the intrinsic contradictions in the agreement (O'Grady, 2021). One major influence of global warming is that heavy rainfall may intensify as the atmospheric water vapors increase at a rate of ~7% per degree warming based on the thermodynamic Clausius–Clapeyron (CC) (for abbreviations, see Table 1.1) relationship (Pendergrass, 2018; Trenberth et al., 2003). This intensification of heavy rainfall, induced by the natural and anthropogenic climate variability, can further increase the risks of flooding (Tabari, 2020) under the joint influence of other factors such as population growth, urbanization, and land-use change, which can be quite detrimental to a multitude of interdependent sectors including agriculture, infrastructure, socioeconomics in the affected regions (Kundzewicz et al., 2014; Shah et al., 2022). Therefore, quantifying the rainfall intensification and the resulting flooding risks is imperative to pave the way for strengthening the political will to achieve the 1.5℃ goal as well as for policymakers to decide mitigation and adaptation measures, especially for the low-income developing countries in Southeast Asia (SEA), which are vulnerable to heavy rainfall-induced floods due to the poor infrastructure (Hijioka et al., 2014).

The design of various flood-prevention infrastructures is affected by the likelihood of the heavy rainfall events mathematically represented as the intensity–duration–frequency (IDF) curves at various spatiotemporal scales (Nissen, and Ulbrich, 2017; Hosseinzadehtalaei et al., 2018, 2020). IDF curves, representing the relationship between the rainfall intensity, rainfall duration, and frequency of occurrence, are commonly used for infrastructure design. An example of the application is shown in Fig. 1.1; where IDF curves with different return periods can be transformed through the Chicago equation (Willems, 2013) into storm hyetographs, which subsequently are used for different

engineering designs (e.g., 2-year (10-year) for small-scale (large-scale) urban drainage system). Development of the IDF curves and the subsequent hydraulic studies and management strategies for the hydroclimate risk for any specific site rely upon the rainfall records based on the rain gauge stations. However, this task remains a challenge for the data-scarce regions, especially in Southeast Asia where many traditional rain gauge stations are designed to record only daily rainfall (https://portal.mrcmekong.org/home). Additionally, global warming requires us to update the IDF curves based on observations by using climate change predictions that can provide future rainfall datasets (Srivastav et al., 2014; Volosciuk et al., 2017). However, climate change predictions usually provide us with limited (in terms of coarse temporal and spatial scales) future rainfall information due to their high demand in computational power (Berg et al., 2019; Tabari et al., 2016), which is not sufficient enough for some infrastructure design (e.g., urban drainage system). Moreover, climate change predictions are subject to the model uncertainty arising from the unresolved physical processes, parameterization schemes, and coarse spatial resolution (usually 25-50 km) (Volosciuk et al., 2017). The uncertainty is found to be larger for projecting the future rainfall extremes (Ning et al., 2015; Ning and Bradley, 2016). Therefore, reducing the model uncertainty is essential in the development of IDF curves to help identify and act upon suitable adaptation measures in the face of global warming.

Table 1.1 List of abbreviations used in this study

Full name	Abbreviations	Full name	Abbreviations
Annual maximum rainfall time series	AMS	Intensity-duration-frequency	IDF
Artificial neural network	ANN	Markov chain monte carlo	MCMC
Areal reduction factor	ARF	Maximum likelihood method	MLE
Bias correction	BC	Monthly maximum rainfall time series	MMS
Clausius–Clapeyron	CC	Normalized quantile mapping	NQM
Climate change signal	CCS	Parameter estimation method	PEM
Cumulative distribution function	CDF	Phnom Penh City	PPC

Conditional artificial neural network	CANN	Quantile delta mapping	QDM
Can Tho City	CTC	Quantile mapping	QM
Coordinated regional climate downscaling experiment	CORDEX	Regional cliamte model	RCM
Extreme value analysis	EVA	Representative concentration pathway	RCP
General circulation model	GCM	Southeast Asia	SEA
Generalized extreme value	GEV	Zero-inflated generalized extreme value function	ZIGEV
Ho Chi Minh City	HCMC		

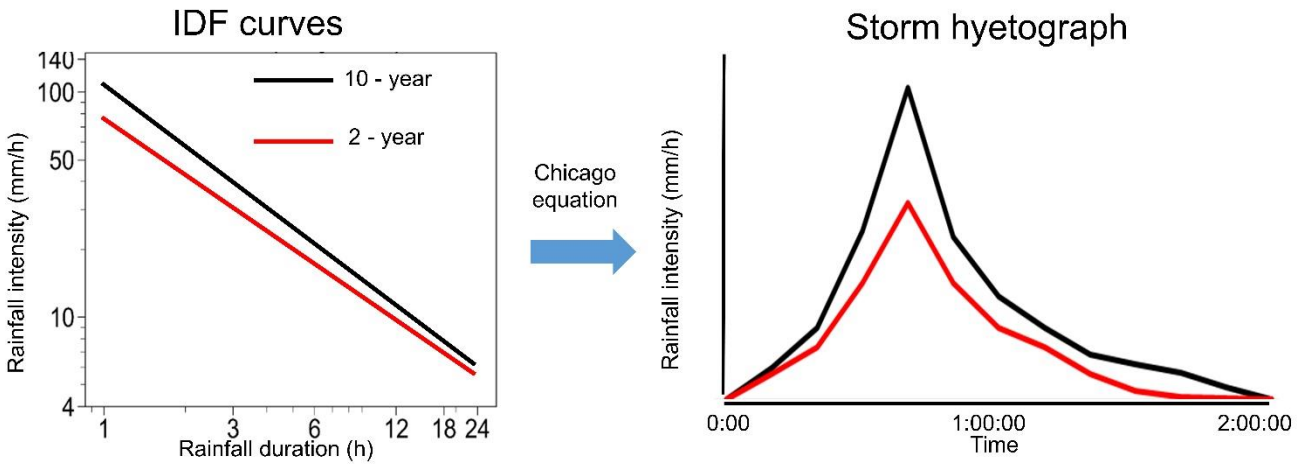


Fig. 1.1. An example of application of IDF curves to engineering design.

1.2 Purpose and structure of this study

In view of the abovementioned limitations, our study's objectives are three-fold.

1) To provide some cost-effective frameworks to develop and update the IDF curves in data-sufficient and -scarce regions (based on an assumption of homogeneity of hydrologic region).
2) To propose some novel bias correction (used for spatial downscaling) (Hereinafter, bias referring to the difference between climate model results and observed results) and temporal disaggregation methods for us to obtain future rainfall information at fine temporal and spatial scales.
3) To quantify different uncertainty sources in the development of IDF curves, followed by discussion on the adaptive measure under deep uncertainty.

The remainder of this study is presented as follows. Chapter 2 introduces a framework applied to

establish historical and future IDF curves in the case study cities (including two data-sufficient cities and one data-scarce city) based on latest climate models by using new temporal disaggregation, bias correction methods and spatial interpolation method. Chapter 3 extends the framework to a flood-prone and data-scarce country in SEA by using some auxiliary datasets, followed by the quantification of the governing physical mechanisms for the heavy rainfall (owing to various hydro-climate conditions in the study area, heavy rainfall is defined as the rainfall amount that deviates from the normal rainfall amount (e.g., annual maxima), instead of the exceedances over high thresholds, (e.g., rainfall amount > 100 mm/h) change in the study area. Thereafter, a new flood risk change index is applied to develop a province-level flood risk change map. In Chapter 4, we quantify the uncertainty in developing historical and future IDF curves and discuss on sustainable adaptation strategies that can be updated in the face of deep uncertainty. At last, conclusions and future studies are summarized in Chapter 5.

2. Establishment of IDF curves in data-sufficient and -scarce cities

2.1 Background

A recent Intergovernmental Panel on Climate Change special report has estimated that human activities have warmed the globe by 1 °C since the pre-industrial period, and at the current rate of temperature increase, global warming is expected to rise further to 1.5 °C between 2030 and 2052 (Hoegh-Guldberg et al., 2018). This rise is likely to occur earlier than foreseen by the Paris Agreement because almost all G20 economies (representing 75% of total 2010 greenhouse gas emissions) are not ambitious enough to implement the agreement or achieve their pledged contributions to the implemented policies (Roelfsema et al., 2020). Previous work has shown that extreme precipitation is expected to increase with global warming at a rate of approximately 7% per degree based on the thermodynamic CC relationship (Trenberth et al., 2003). However, this rate varies across different latitudes and is subject to the water availability (i.e., water supply and demand) of that region (Tabari, 2020; Tabari et al., 2019), which requires a regional climate change study focusing on extreme precipitation. This requirement is crucial for policymakers to decide mitigation and adaptation measures for reducing the negative impacts of global warming, and is imperative in the Southeast Asia (SEA) region, which is vulnerable to extreme precipitation-induced floods due to the poor infrastructure (e.g., urban drainage systems) of many developing countries (Hijioka et al., 2014).

Global climate models (GCMs), as the term suggests, provide information on future extreme precipitation at global scales. However, the coarse spatial resolution of GCMs (usually greater than 100 km) precludes their direct use at local scales (He et al., 2016; Maraun et al., 2010). To obtain finer spatial resolution information, there are two main approaches: dynamical downscaling and statistical downscaling. Dynamical downscaling denotes the use of higher resolution climate models (i.e., regional climate models (RCMs)), where RCMs use GCMs and reanalysis data as boundary conditions and improve physical processes to reproduce the regional climate conditions (e.g., rainfall) (Hewitson and Crane, 1996; Xu et al., 2019). For example, the RCMs (RegCM4, see in Tangang et al., 2019), with the Massachusetts Institute of Technology (MIT)-Emanuel convective scheme, can better simulate the temperature and precipitation in the regional scales (~ 25-50 KM) than GCMs (which

consist of a simplified convective scheme). Therefore, compared with GCMs, RCMs through dynamical downscaling involve the use of regional numerical models that include full sets of physics. On the other hand, statistical downscaling predicts the regional scale variables by developing the statistical relationships between the observed climate data and the output of GCMs for the same historical period without fully considering the physics (Chen et al., 2013a). Thereafter, after the validation of such statistical relationships in the historical period, these relationships can be used to develop regional-scale variables for the future period. However, dynamical downscaling is computationally expensive and time-consuming, which often exceeds the capability of institutions in developing countries (Trzaska and Schnarr, 2014). To solve this problem, the World Climate Research Programme has established the Coordinated Regional Climate Downscaling Experiment (CORDEX) (Giorgi et al., 2009; Gutowski and Giorgi, 2020). The Coordinated Regional Climate Downscaling Experiment Southeast Asia (CORDEX-SEA) was initiated in 2013 and represents the most comprehensive set of high-resolution (25 km) regional simulations over SEA (Tangang et al., 2020). Based on these CORDEX-SEA simulations, several papers researching future rainfall have been published. Tangang et al. (2020) evaluated the performance of CORDEX-SEA in simulating mean rainfall over SEA in the historical and future periods. Ge et al. (2019) used six ensemble members of CORDEX-SEA simulations to estimate the changes in extreme precipitation over SEA under the 1.5 °C and 2 °C global warming levels. Based on multi-model simulations of CORDEX-SEA, more intense rainfall events are expected to occur over most of Indochina during the middle of the 21st century (Tangang et al., 2018) and at the end of the 21st century (Supari et al., 2020).

Although these studies have used CORDEX-SEA simulations to project changes in daily rainfall in the future, there are no studies on changes in sub-daily (e.g., hourly) rainfall. More specifically, the changes in rainfall IDF curves, representing the relationship between the rainfall intensity, rainfall duration, and frequency of occurrence, based on CORDEX-SEA simulations are not well understood. IDF curves can be estimated by obtaining rainfall quantiles at different durations separately (Hosseinzadehtalaei et al., 2020) or through a general formula that includes rainfall intensity, duration, and frequency of occurrence in one analytical equation (Koutsoyiannis et al., 1998). Traditionally, this estimation of IDF curves is based on the observed rain gauge dataset for a single location (Liew et al., 2014), and for a large region based on the dataset provided by satellite (Endreny and Imbeah, 2009;

Marra et al., 2017), weather radar (Marra et al., 2017; Marra and Morin, 2015; Peleg et al., 2018) and reanalysis products (Cannon and Innocenti, 2019; Hosseinzadehtalaei et al., 2020). However, IDF curves developed by these historical datasets can misrepresent future conditions due to the non-stationarity in rainfall (Srivastav et al., 2014). Therefore, updating IDF curves based on the datasets of projected rainfall in the future period is important for the urban drainage system design in the developing countries in SEA. Regarding the projected rainfall datasets, while the RCMs with a high resolution are suggested in many studies (Poschlod et al., 2021; Tabari et al., 2016; Tangang et al., 2019), their availability is still limited in SEA. Therefore, the outputs from state-of-the-art RCMs provided by CORDEX-SEA simulations were explored in this chapter.

In addition to the lack of study on IDF curves based on CORDEX-SEA simulations, most climate impact studies used the raw outputs without a bias correction. Based on the evaluation of CORDEX-SEA in Tangang et al. (2019), a large discrepancy was found between the statistical properties of observed rainfall and those of modeled rainfall. Tangang et al. (2020) suggested that future studies should focus on removing the model bias provided by CORDEX-SEA simulations due to the parameterization schemes, unresolved process in these climate models, and the systematic model bias that arises from the large-scale atmospheric circulation (Volosciuk et al., 2017). Bias correction (BC) method is a standard practice in removing the systematic model bias because it increases the agreement of modeled variables with observed variables and hence decreases the ensemble spread of modeled variables (Chen et al., 2013a; Wang et al., 2016). Generally, BC method cannot be used in datasets with a large spatial gap (Maraun, 2013), but is reliable for (spatially) downscaling the daily data from RCMs to the local rain gauges with respect to the local distributional and time series characteristics (Ivanov and Kotlarski, 2017). For the climate change impact studies based on CORDEX-SEA simulations, previous studies used BC with quantile mapping (QM) to remove the systematic model bias (Ngai et al., 2020; Trinh-Tuan et al., 2019). The main task of QM is to adjust the cumulative distribution function (CDF) of the modeled variables (e.g., rainfall) to resemble the CDF of observed variables (Gudmundsson et al., 2012). In this way, the statistical properties of modeled variables will resemble those of observed variables. However, QM has been criticized for changing the raw modeled climate change signal (CCS, denoting the changes in the variables of interest inforamed by the climate models), due to its stationarity assumption (Maraun, 2013; Themeßl et al., 2012) and difficulty of

removing the bias in temporal structure (e.g., inter-annual variability and temporal dependence of different variables) (Chen et al., 2013b; Maraun, 2013; Volosciuk et al., 2017). Quantile delta mapping (QDM) (Cannon et al., 2015) and scaled distribution mapping (Switanek et al., 2017) were developed without a stationarity assumption. The nesting BC methods for correcting the bias at multiple temporal scales have shown success in correcting the inter-annual variability (Johnson and Sharma, 2012). Mehrotra and Sharma (2015, 2016) extended the nesting BC methods to simultaneously remove the bias across multiple variables. The results show that considering dependence in multiple variables removes the bias not only in individual variables but also in the correlation between these variables. Cannon (2018) developed the multivariate bias correction algorithm by transferring the statistical properties of multivariate distribution of observed variables to the corresponding multivariate distribution of modeled variables, which showed better accuracy in reproducing the spatiotemporal autocorrelation of rainfall fields. The multivariate bias correction algorithm has also been used for removing the bias in statistical properties and spatial dependence in multiple rain-gauge stations (Su et al., 2020). However, Hnilica et al. (2017) found that the multivariate bias correction algorithm generated unrealistic results when it was applied to variables other than those used for the development of the multivariate bias correction algorithm. Overall, BC methods have been changing from single time scale, variable, and station to multiple time scales, variables, and stations. Even though immense progress has been made in the development of BC methods, different types of assumptions behind these BC methods require us to select suitable BC methods before application (Chen et al., 2013a). This selection is important because the uncertainty associated with the choice of different BC methods is comparable to that associated with the choice of different climate models (Wang et al., 2016). A recent study also suggested that updating IDF curves based on CORDEX simulations should consider the uncertainty of different BC methods (Hosseinzadehtalaei et al., 2020).

Regarding the temporal resolution, few climate models of CORDEX-SEA simulations have provided hourly rainfall products, but most models' temporal resolution is not sufficiently fine to update the IDF curves. Moreover, sub-daily rainfall extremes generated mainly by convective events are not well reproduced by most climate models, such as the models with parameterized convection (Cannon and Innocenti, 2019). Thus, temporal disaggregation tools are needed to obtain rainfall datasets at finer time scales (e.g., hourly rainfall) based on the rainfall datasets at long time scales (e.g.,

daily rainfall) provided by CORDEX-SEA simulations. Previous studies have often used stochastic models, which generate statistically consistent rainfall events based on stochastic and probabilistic techniques, such as Hyetos (Kossieris et al., 2018; Koutsoyiannis, 2003; Koutsoyiannis and Onof, 2001), Cascade (Müller and Haberlandt, 2018), and MuDrain (Koutsoyiannis, 2003; Lu and Qin, 2014). However, the accuracy of the stochastic models is closely related to the prior assumption of rainfall generating mechanisms (Kossieris et al., 2018), trade-off between the computational complexity and incorporation of rainfall statistics into the models (Sharma and Mehrotra, 2010). Another category of temporal disaggregation models is based on non-parametric data-driven approaches that quantify the underlying mathematical relation between long-duration and short-duration rainfall without explicit prior physical assumption (Rajagopalan et al., 2010). The method of fragments makes use of the short-duration rainfall information at neighboring stations to disaggregate long-duration rainfall into short-duration fragments at the station of interest (Sharma and Srikanthan, 2006; Poschlod et al., 2018). Burian et al. (2000) and Burian and Durrans (2002) proved that an artificial neural network (ANN) can be a viable way to disaggregate hourly rainfall into sub-hourly rainfall. The data-driven model based on the K-nearest neighbor algorithm was found to perform better than Hyetos in reproducing the hourly rainfall series based on daily rainfall series (Lu and Qin, 2014). Nourani and Farboudfam (2019) proposed several data-driven models, which are especially suitable for rainfall disaggregation in mountainous regions. Mirhosseini et al. (2013, 2014) found that the ANN-based model was better than the stochastic model in reproducing short-duration rainfall extremes, and the stochastic model tended to underestimate the rainfall extremes. Data-driven models such as the ANN-based model are becoming increasingly popular in the field of hydrologic science, owing to more efficient algorithms, higher computing power, and ever-increasing amounts of data (He et al., 2016; Shen et al., 2018). At last, the development and update of IDF curves in the data-scarce areas (with only daily rain-gauge rainfall datasets available) is difficult owing to a lack of observed sub-daily that is not only needed to develop the historical IDF curves but to remove the bias in climate model simulations.

The objective of this chapter, therefore, is to project the IDF curves in the future period for designing urban drainage systems in tropical data-sufficient and -scarce cities based on a temporal disaggregation of bias-corrected climate models provided by CORDEX-SEA simulations. Through this projection, three important aspects are included. First, we propose a new BC method, called

normalized quantile mapping (NQM). We evaluate NQM, QM, and QDM by their capabilities of removing the systematic model bias in the historical period and preserving the raw modeled CCS in the future period. Based on this evaluation, we select a suitable BC method for each dataset provided by CORDEX-SEA. Second, for temporal disaggregation, we introduce an ANN-based model to obtain the future (sub-daily) monthly maximum series (MMS) based on the daily datasets provided by CORDEX-SEA. Third, for the establishment of IDF curves in data-scarce city, a spatial interpolation method is introduced. In this chapter a case study is conducted for Ho Chi Minh City (HCMC), Can Tho City (CTC), and Phnom Penh City (PPC). The remainder of this chapter is organized as follows. We introduce the research area and present our data in Section 2.2. The methodology of NQM, ANN-based model, and the method for constructing IDF curves are explained in Section 2.3. In Section 2.4, different downscaling and temporal disaggregation methods are compared, final IDF curves are presented, and future works are described; uncertainties in the IDF curves are also discussed. Finally, conclusions are presented in Section 2.5.

2.2 Study Area and Data

2.2.1 Study Area

In general, floods have been regarded as the dominant natural disaster in Southeast Asia, where the associated economic loss reached approximately 312.1 million USD per year (UNISDR and World Bank, 2010). To investigate the climate change impacts on floods, we selected three cities in flood-prone Southeast Asia, namely, Phnom Penh City (PPC) (the data-scarce city), Can Tho City (CTC), and Ho Chi Minh City (HCMC). PPC, the capital of Cambodia, is situated on the floodplain of the Mekong River (Fig. 2.1). The monsoon here brings approximately 96% of rainfall in the rainy season (April to November) every year. Therefore, the surface (pluvial) floods caused by the monsoon thunderstorms as well as river floods from monsoons and snowmelt-induced streamflow in the upper Mekong river are flooding PPC almost every day during the rainy season (Doyle, 2012; Flower and Fortnam, 2015). According to the questionnaire conducted in PPC, flash floods (one of the pluvial floods) in the urban area usually last within three hours (Lyna, 2019), which can cause severe damage to the prevailing socioeconomic activities (Fowler et al., 2021). The capacity of the drainage system tends to be exceeded during heavy rainfall events due to insufficient maintenance and the rapid

transformation of urban wetlands into residential lands driven by the socioeconomic and population growth in PPC (Lyna, 2019). Similar to PPC, CTC, as the biggest city located on the Mekong River delta of Vietnam, is also affected by the combined effects of pluvial floods and river floods. Moreover, due to the low topography of CTC (60-80 cm above mean sea level) and proximity to the sea (only about 84 km to the South China Sea), the river floods are likely to be exacerbated due to the global warming-induced future sea-level rise (Huong and Pathirana, 2013; Takagi et al., 2015). The third study site, i.e., HCMC, a coastal megacity, is also facing the risk of coastal floods caused by the high tide levels in addition to the pluvial floods and river floods (Vachaud et al., 2019). To reduce the flooding risks, Japan International Cooperation Agency (JICA) has been working with HCMC on urban flooding control, but the flooding problem still persists (Tran Ngoc et al., 2016). Further details of the location and elevation of the three study cities are shown in Fig. 2.1.

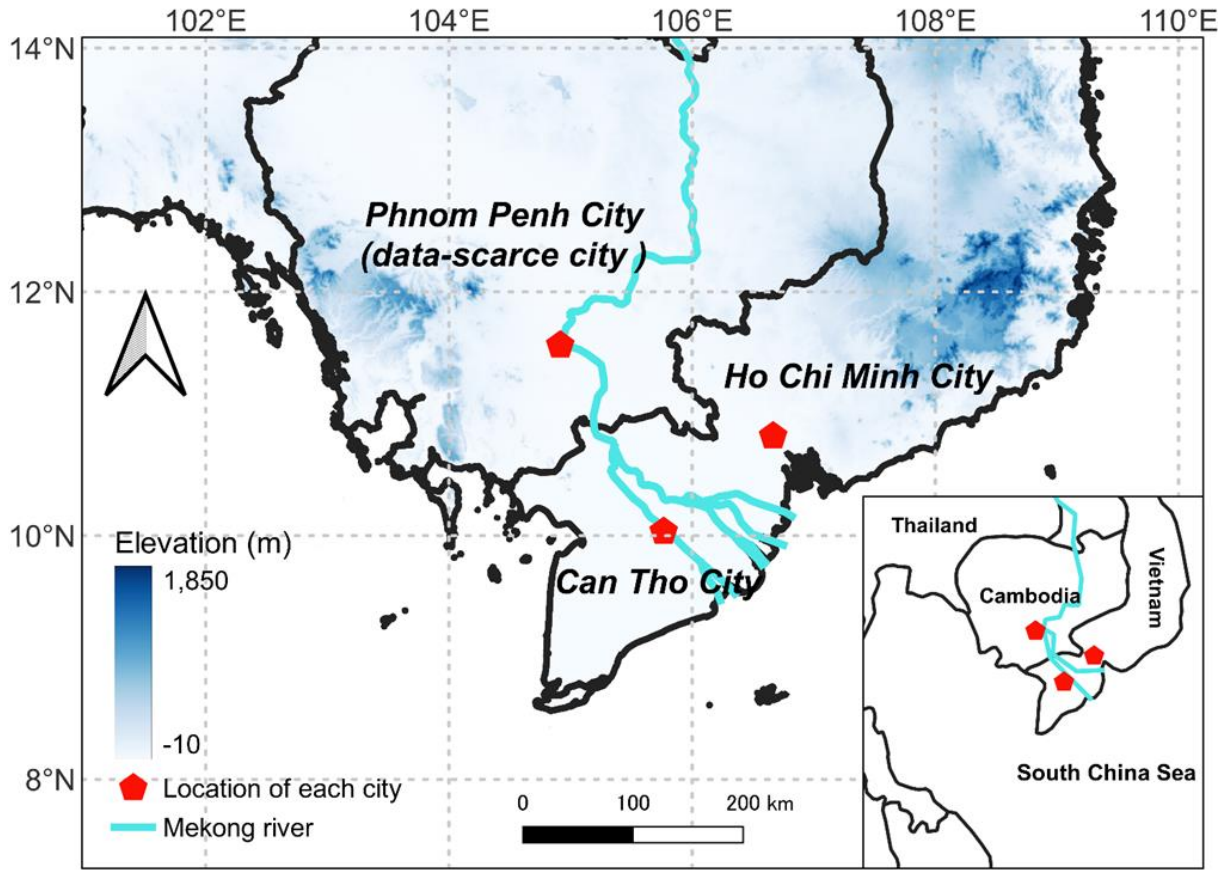


Fig. 2.1. Location and elevation (above mean sea level) of PPC, HCMC, and CTC.

2.2.2 Data

For HCMC and CTC, we used two types of rainfall datasets, as summarized in Table 2.1. Hourly observed rainfall at a single gauging station in each city was provided by the Southern Regional Hydro-Meteorological Center in Vietnam. Given that each city is located in flat terrain, the rainfall observed at a single station can be considered, to a certain extent, valid to represent the statistical rainfall characteristics in each city for the application in urban drainage design. In addition, there is almost no missing value (less than 0.02%) in the observed rainfall series. Daily rainfall for the future periods was provided by all the latest (as of June 2020) CORDEX-SEA simulations (https://esg-dn1.nsc.liu.se/projects/cordex/), except for the one based on the model named IPSL-CM5A-LR_RegCM4-3, which gave an unreasonable underestimation of rainfall in our study area as well as in SEA (Li et al., 2019). The data extracted from the nearest neighboring grid cell center was used for further analysis at the station of interest. For all CORDEX-SEA simulations, we used a high emission scenario (i.e., RCP8.5), which is viewed as the worst future case in the 21st century (Hausfather and Peters, 2020). Due to the limited record length and the setting of RCP scenarios (the period of projection under RCP scenarios starts from 2006) in CORDEX simulations, we constrained the historical period to 1986-2005. Future periods considered in the current study consist of the near future period (2026-2045) and far future period (2066-2085).

Table 2.1. Datasets used in HCMC and CTC. The index of CORDEX-SEA is listed in the first column.

No.	Dataset (Developer)	Resolution	Time
1	EC-EARTH_RegCM4-3 (ICTP[1])	0.22°(daily)	1986–2005 (Historical period) 2026–2045 (Near future period) 2066–2085 (Far future period)
2	NorESM1-M_REMO2015 (GERICS[2])		
3	MPI-ESM-MR_RegCM4-3 (ICTP)		
4	HadGEM2-ES_REMO2015 (GERICS)		
5	MPI-ESM-LR_REMO2015 (GERICS)		

6	HadGEM2-ES_RCA4 (SMHI[3])		
7	GFDL-ESM2M_RegCM4-3 (ICTP)		
-	Observations (SRHMC[4])	Point (hourly)	1986–2005 (Historical period)

1 The Abdus Salam International Centre for Theoretical Physics (ICTP)

2 Climate Service Center Germany (GERICS)

3 Swedish Meteorological and Hydrological Institute (SMHI).

4 Southern Regional Hydro-Meteorological Center (SRHMC)

For PPC, the long-term observed rainfall data is available only in daily duration at the Pochentong gauge station. Almost no missing values (<0.02%) were found in these observed datasets. The future rainfall datasets are listed in Table 2.2.

Table 2.2. List of RCMs from CORDEX simulations for RCP 4.5 and 8.5 used in PPC.

Driving model	**RCM version**	**Developer**	**Resolution**	**Time**
EC-EARTH	RegCM4-3	ICTP	0.22° (3h)	1986–2005 (Historical period) 2026–2045 (Near future period) 2066–2085 (Far future period)
GFDL-ESM2M				
MPI-ESM-MR				
HadGEM2-ES				
	RCA4	SMHI		

2.3 Methods

2.3.1 Procedure

Fig. 2.2 demonstrates the procedure for constructing the IDF curves in data-sufficient and -scaece cities. For the IDF curves established in HCMC and CTC, 1: Historical IDF curves (1986-2005) were constructed based on the rainfall quantiles (at different durations) of observed rainfall; 2: a new BC

method, NQM, was proposed and compared with QDM and QM. Based on this comparison, we selected a suitable BC method for each dataset provided by the CORDEX-SEA simulations; 3: selected BC methods were applied to downscale the future daily rainfall datasets projected for the periods of 2026–2045 and 2066–2085; 4: we compared two temporal disaggregation models based on how well they reproduce the observed MMS; 5: we generated sub-daily MMS from the downscaled daily rainfall in step 3 based on the superior disaggregation method in step 4; 6: IDF curves for the future periods were constructed based on the 'change factor' operation mentioned in detail in Section 2.3.5. The uncertainties in the projected IDF curves were also discussed. The development of IDF curves in PPC is based on the combined rainfall information in PPC, HCMC, and CTC. Firstly, daily rainfall provided by the RCMs was firstly spatial downscaled. Then the observed daily rainfall for the historical period and downscaled daily rainfall for the future period were used to estimate the historical and future return rainfall, respectively, based on the Bayesian generalized extreme value (Bayesian GEV) parameters estimated in HCMC and CTC. Thereafter, the estimated return rainfall was applied to construct the IDF curves in the historical and future periods.

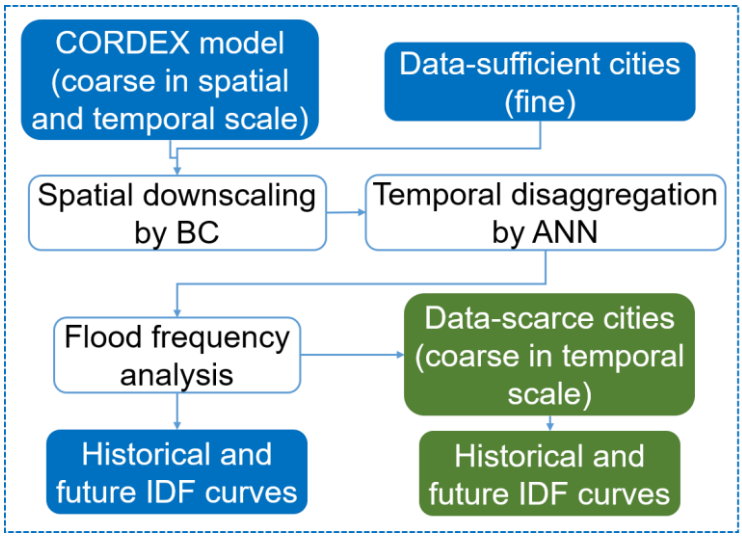


Fig. 2.2. Research framework for developing IDF curves in HCMC, CTC and PPC

2.3.2 Normalized Quantile Mapping

In this section, a new BC method, NQM, was proposed and compared with QM and QDM. Fig. 2.

3 illustrates how to remove the bias at a given empirical cumulative distribution function (ECDF) (quantile of order 0.9 in this figure) by QM, QDM, and NQM. In QM, the error correction value is the ratio of the raw historical model to the observations at ECDF ≈ 0.93. This error correction value remains the same between the raw future model and future downscaled data (that is called 'BC by QM', which represents the bias-corrected or downscaled future data by QM) at ECDF = 0.9. Thus, QM uses the error correction values established at a higher ECDF to remove the bias at a lower ECDF, which could alter the raw modeled CCS (Switanek et al., 2017). In QDM, the error correction value is the ratio of the raw future model data to the raw historical model data at ECDF = 0.9. This error correction value multiplied by observations is the 'BC by QDM' ('BC by QDM' = observations × raw future model / raw historical model).

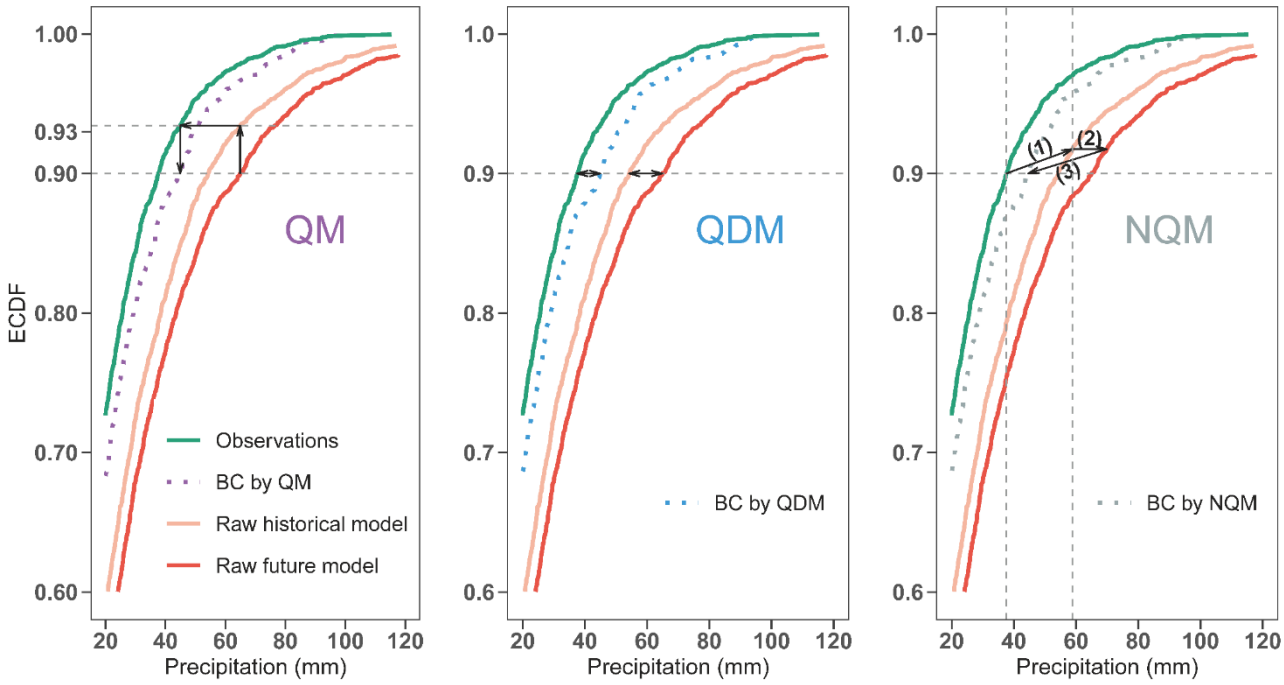


Fig. 2.3. Difference between the QM, QDM, and NQM methodologies. BC by QM (QDM, NQM) represents the bias-corrected (or downscaled) future projection by QM (QDM, NQM). Raw historical and future models represent climate model outputs (uncorrected outputs) for the historical and future periods, respectively. ECDF represents the empirical cumulative distribution function. '(1)', '(2)', and '(3)' in NQM correspond to the steps (1), (2), and (3) in Section 2.3.2.

QM and QDM assume that biases between the observations and the raw historical model exist at the same ECDF. In other words, the largest (second-largest,…, smallest) observed and modeled quantiles correspond to the same ECDF, which is unverifiable (Switanek et al., 2017). In contrast, in NQM, we determine the location of biases (the way to connect raw historical model and observations) based on a data normalization in step (1). This data normalization is used to change the values to a common scale, which is frequently applied in machine learning.

The methodology for applying NQM is explained as follows. Steps (1), (2), and (3) were used to remove the biases of the wet-day rainfall amount and correspond to the three arrow lines, in turn, in Fig. 2.3 (NQM). We removed the biases of the wet-day frequency in step (4). Step (5) was used to obtain the bias-corrected datasets.

Step (1): set a threshold as the daily minimum observed rainfall amount. Herein, we used 0.1 mm as the threshold, below which all raw modeled rainfall values were set to 0 mm. Next, observations (wet-day) and raw historical model variable (wet-day) were normalized by:

$$\hat{P}_{OBS} = \frac{P_{OBS} - \bar{P}_{OBS}}{\sigma(P_{OBS})};\ \hat{P}_{M,H} = \frac{P_{M,H} - \bar{P}_{M,H}}{\sigma(P_{M,H})} \tag{2.1}$$

where the subscripts OBS and M,H denote the observed and modeled variables (rainfall) in the historical period, respectively. P_{OBS} represents the observations (wet-day) and $P_{M,H}$ represents the raw historical model data (wet-day). $\bar{P}$ and $\sigma(P)$ are the mean and standard deviation values of the respective datasets (i.e., observations and raw historical model data). Thus, $\hat{P}_{OBS}$ and $\hat{P}_{M,H}$ are the normalized variables. Next, $\hat{P}_{OBS}$ is mapped into the range of $\hat{P}_{M,H}$ by:

$$K = \frac{max(\hat{P}_{M,H}) - min(\hat{P}_{M,H})}{max(\hat{P}_{OBS}) - min(\hat{P}_{OBS})}$$

$$\Delta = K \times \left(\hat{P}_{OBS} - max(\hat{P}_{OBS})\right) + max(\hat{P}_{M,H}) \tag{2.2}$$

where max (min) is the maximum (minimum) value of the respective datasets. K is the ratio of $\hat{P}_{M,H}$'s range to $\hat{P}_{OBS}$'s range. Δ is the location of $\hat{P}_{OBS}$ in the range of $\hat{P}_{M,H}$. Then, we linked the observations to the raw historical model by:

$$P^*_{OBS} = \Delta \times \sigma(P_{M,H}) + \bar{P}_{M,H} \tag{2.3}$$

where P^*_{OBS} is the location of P_{OBS} in the range of $P_{M,H}$. In other words, Eq. 2.3 gives the location

of biases between observations and raw historical model data.

Step (2): remove the biases by error correction values, which are the ratio of the future model data to the historical model data at the same ECDF:

$$\hat{P}_{BC} = F_{M,F}^{-1}\left(F_{M,H}(P^*_{OBS})\right) \tag{2.4}$$

where the subscripts BC and M,F denote the bias-corrected and uncorrected modeled variables (rainfall) in the future period, respectively. $\hat{P}_{BC}$ is the bias-corrected future data with a scaling factor, which is explained in Eq. 2.5. $F_{M,H}$ is the ECDF of climate model outputs for the historical period. $F_{M,F}^{-1}$ is the inverse ECDF of the climate model outputs for the future period.

Step (3): obtain bias-corrected future data P_{BC} by removing the scaling factor.

$$P_{BC} = \left(\frac{\frac{\hat{P}_{BC}-\bar{P}_{M,H}}{\sigma(P_{M,H})} - max(\hat{P}_{M,H})}{K} + max(\hat{P}_{OBS})\right) \times \sigma(P_{OBS}) + \bar{P}_{OBS} \tag{2.5}$$

Eq. 2.5 is the inverse operation of Eqs. 2.1, 2.2, and 2.3.

Step (4): remove the biases of the wet-day frequency based on Switanek et al. (2017).

$$WD_{BC} = WD_{M,F} \times \frac{WD_{OBS}/TD_{OBS}}{WD_{M,H}/TD_{M,H}} \tag{2.6}$$

where WD and TD are the number of wet days and the total number of days including dry days (days without rainfall), respectively. For example, let's assume that the total number of days was 1840 for the observed data, raw historical, and future model data, while those of wet days were 1339, 1714, and 1706, respectively. Based on step (4), WD_{BC} was found to be 1706 × (1339/1840) / (1714/1840) = 1333 days (1333 is the nearest integer value).

Step (5): place the bias-corrected future wet-day series back into the raw future model wet-day series in the correct temporal locations.

BC methods adjust the CDF of the modeled variables (e.g., rainfall) to resemble the CDF of observed variables (Gudmundsson, 2012). However, CDF only considers the datasets in ascending or descending order and does not consider the temporal structure (e.g., inter-annual variability) of the datasets. Therefore, in step (5), we sorted the bias-corrected datasets in the same order (or rank) as the raw future model datasets. In the previous example, based on steps (1), (2), and (3), we obtained a bias-corrected future wet-day series with a length of 1339. This length was different from 1333 in step (4). Therefore, first, the smallest 6 (1339-1333 = 6) rainfall values in this wet-day series were all set

to 0 mm. Second, the maximum daily rainfall in the raw future model was found to occur on September 13, 2040. Thus, the maximum bias-corrected rainfall was reinserted on that day. Third, a similar operation was performed for the rest of the 1332 (1333-1 = 1332) bias-corrected rainfall values. When the number of wet days from step (4) (1240) was larger than that from steps (1), (2), and (3) (1232), we selected 1240 from step (4) as the wet-day frequency. In this case, the largest 8 (1240-1232) rainfall values of the smallest 598 (1840-1232) raw future model rainfall values were placed in the correct temporal locations. Herein, we performed bias correction on a seasonal bias by grouping the datasets into different seasons (seasonal moving window), as suggested by Räty et al. (2014) because this can consider seasonality, or a monthly moving window if a longer duration of the observed rainfall is available.

2.3.3 Comparison of NQM, QDM, and QM

Two steps were followed to compare NQM, QDM, and QM. Firstly, we compared the reproducibility of the upper 95th percentile of the wet-day rainfall series (Q95th) by each method based on the observed datasets. Secondly, we compared how each method preserves the raw modeled CCS of CORDEX-SEA in Q95th in the future period. Q95th is a commonly used index for climate change impact studies focusing on rainfall extremes (Chen et al., 2013a). The quantile is calculated based on the default method in *quantile function* in R language. The details of these steps are described as follows.

In the first step, we chose 1986–1995 as the calibration period and 1996–2005 as the validation period. Each BC method was calibrated using the data for 1986–1995 and the systematic model bias was removed considering the data for 1996–2005. Then, we compared the bias-corrected model values with the observed values obtained for 1996–2005. Specifically, the comparison was represented by the relative error through $100\% * (Q95th_{BC} - Q95th_{OBS})/Q95th_{OBS}$. $Q95th_{BC}$ and $Q95th_{OBS}$ represent the Q95th of bias-corrected wet-day rainfall series and observed wet-day rainfall series, respectively.

In the second step, we checked how each BC method preserves raw modeled CCS in Q95th in future periods (2026–2045 and 2066–2085). For example, the raw future model showed a 20% increase in Q95th from the raw historical model. If the bias-corrected dataset showed a 10% increase in Q95th

from the observed dataset, the relative change of raw modeled CCS was 100% × (10%-20%) / 20% = -50%. These two steps were similar to Cannon et al. (2015) and Chen et al. (2013a). For each BC method, we set a threshold=0.1 mm, below which all modeled values were set to zero. QDM and QM were implemented by R packages *MBC* and *qmap* respectively, that are available online (https://cran.r-project.org/web/packages/). For each bias-corrected model, we recorded the mean values of relative error and relative change.

To select a suitable BC method for each dataset from CORDEX-SEA, results based on the previous two steps were combined. First, we rescaled the relative error series and relative change series as follows:

$$X^* = 100\% \times \frac{|X| - |X_{min}|}{|X_{max}| - |X_{min}|} \tag{2.7}$$

where $|X|$, $|X_{max}|$, and $|X_{min}|$ denote the absolute value of the data series (that is the relative error series or relative change series), maximum of this series, and minimum of this series, respectively. X^* denotes the mapped data series. Second, we took the average of the mapped relative error series and the mapped relative change series. This average is called the total error and is used for selecting a suitable BC method for each dataset from CORDEX-SEA in each station. This selection is to improve the accuracy of each ensemble member by considering the influence of bias correction and CCS preservation through Eq. 2.7, and thus reduce uncertainty in the ensemble mean and spread. These three BC methods implicitly conduct spatial downscaling, which has been applied in several studies for climate change impacts (Chen et al., 2014, 2018; Ivanov and Kotlarski, 2017; Li et al., 2017; Li and Babovic, 2019; Lima et al., 2018; Mullan et al., 2016).

2.3.4 Temporal Disaggregation Model

For temporal disaggregation, we introduced an ANN-based model that is used to obtain future sub-daily MMS based on daily rainfall datasets provided by CORDEX-SEA simulations. This ANN-based model is similar to the temporal disaggregation model used in Mirhosseini et al. (2014), which is used to reproduce MMS instead of the whole sub-daily rainfall series; However, there are two important advancements: (1) Instead of using all rainfall data information, we only extracted the extreme rainfall information for training, which is more computationally efficient, (2) to improve accuracy, we used the most accurate outputs simulated in previous steps (for 12 hourly rainfall) as the

inputs for later steps (for 6, 18 and 24 hourly rainfall). For comparison, we applied a widely used stochastic disaggregation model called Hyetos (Koutsoyiannis, 2003). Hyetos was improved by Kossieris et al. (2018), as it allowed for a temporal disaggregation from daily scale down into 1-min scale. This model is called Hyetosminute, which was implemented in this chapter using the R package (http://www.itia.ntua.gr/en/softinfo/3/).

More details on how to apply Hyetosminute and ANN-based model are described as follows. Hyetosminute was used for generating hourly rainfall series based on daily rainfall series. Hyetosminute, as a stochastic model, generates statistically consistent hourly rainfall events, thus needing suitable distribution functions to simulate the rainfall duration and intensity. The cell (rainfall events divided into different cells) intensity was simulated by Weibull distribution, which was also suggested by Kossieris et al. (2018). A preliminary analysis showed that a model considering dependence between cell duration and intensity performed better, which was similar to Kossieris et al. (2018). For the ANN-based model, we used a three-layered (input, hidden, and output layer) back-propagation network. The input layer sends the input variable (shown in Table 2.3) to the hidden layer (that produces the output variable based on nonlinear transformations), and the output layer gives the output variable (Burian et al., 2000). The objective of back-propagation is to minimize the error function (the mean squared error in our study) by adjusting the nonlinear transformations (Burian et al., 2000). We applied 20 hidden neurons with a dropout (rate = 0.5) between the hidden layer and output layer based on trial and error. Dropout is a technique where selected neurons are ignored randomly in the training to avoid overfitting. Such settings are problem-dependent and are based on a trial rule (Macukow, 2016; Mendenhall and Meiler, 2016). The input and output variables are listed in Table 2.3. Monthly maximum 1, 3, 6, 12, 18, 24 hourly rainfall series are signified by MMR1, MMR3, MMR6, MMR12, MMR18, MMR24, respectively. The optimum input variable was selected based on the minimum Akaike Information Criterion (AIC), which is widely used for input variable selection (Mirhosseini et al., 2014; Panchal et al., 2010). To set up the model, 80% of the datasets were used for training (80%) and validation for minimizing over-fitting (20%), and the remaining 20% for testing (Mirhosseini et al., 2014). For example, among 20-year observed hourly rainfall series, we used 154 (12 months × 20 years × 80% × 80%) observed MMR1 as the output variable for training. Based on the observed MMR1, we knew the exact time (i.e., month (m), day (t), the previous day ($t-1$), the

next day ($t+1$)) when MMR1 occurred and the amount (P) of rainfall accumulated during these times. Then, we had 154 sets of $(P_{t-1}, P_t, P_{t+1}, P_m)$ as the input variable. We also found that the ANN-based model simulated MMR6, MMR18, and MMR24 better if we used simulated MMR12 (sim_MMR12) as the input variable. Other combinations of input variables, such as $(P_{t-2}, P_{t-1}, P_t, P_{t+1}, P_m, sim\_MMR24)$, performed worse than those in Table 2.3 and are not shown here. Overall, we established 12 ANN-based models for six durations at two stations.

Table 2.3 shows that if we use the ANN-based model, we need to know the exact time when sub-daily rainfall extremes (e.g., MMR1) occurred. This information is unavailable in the future daily rainfall series provided by CORDEX-SEA. Therefore, firstly, we regarded daily rainfall as P_t and obtained 7305 (7305 days during 2026–2045 or 2066-2085) sets of $(P_{t-1}, P_t, P_{t+1}, P_m)$. Secondly, $(P_{t-1}, P_t, P_{t+1}, P_m)$ was input into ANN-based model to generate 7305 output variables. Next, we exacted MMS from these output variables. This procedure was also tested in the historical period (Section 2.4.2).

Table 2.3 Selecting the optimum input variables for ANN-based models based on Akaike Information Criterion (AIC). Output variables MMR1, MMR3, MMR6, MMR12, MMR18, and MMR24 represent monthly maximum 1, 3, 6, 12, 18, and 24 hourly rainfall series, respectively. P is the total rainfall in that particular day (t) and month (m). P_{t-1} and P_{t+1} are the daily rainfall centered by t. sim_MMR12 is the disaggregated MMR12 in that month. The best input variables are selected for each output variable based on the minimum AIC value (shown by the bold number).

Station	Input variable	Output variable					
		MMR1	**MMR3**	**MMR6**	**MMR12**	**MMR18**	**MMR24**
HCMC	$(P_{t-1}, P_t, P_{t+1}, P_m, sim\_MMR12$	4.69	4.47	4.36	None	4.42	4.64
	$(P_{t-1}, P_t, P_{t+1}, sim\_MMR12)$	4.75	4.44	**4.25**	None	**4.29**	**4.57**
	$(P_{t-1}, P_t, P_{t+1}, P_m)$	**4.42**	4.11	4.47	3.90	4.50	4.68
	(P_{t-1}, P_t, P_{t+1})	4.54	**4.09**	4.29	**3.62**	4.37	4.65
CTC	$(P_{t-1}, P_t, P_{t+1}, P_m, sim\_MMR12$	3.85	4.48	4.64	None	3.57	3.16

$(P_{t-1}, P_t, P_{t+1}, sim\_MMR12)$	4.12	4.36	**4.61**	None	**3.53**	**3.05**
$(P_{t-1}, P_t, P_{t+1}, P_m)$	**3.65**	4.33	4.66	3.46	4.04	3.74
(P_{t-1}, P_t, P_{t+1})	4.10	**4.28**	4.74	**3.41**	4.01	3.61

2.3.5 Constructing IDF Curves

For the historical period, extreme value analysis (EVA) was conducted using the annual maximum series (AMS) given from the observed rainfall series. For the future period, EVA was applied to AMS extracted from MMS, which is the product of downscaling and temporal disaggregation explained in Sections 2.3.3 and 2.3.4, to determine the return rainfalls (representing the rainfall intensity with a given return period) at different durations that were later used to construct the IDF curves. Compared with the peak over threshold method, the AMS method can ensure the independence of the data series and is suitable for the EVA in highly variable climatic conditions (Marra et al., 2017). For EVA, we applied a Bayesian generalized extreme value (GEV) distribution model proposed by Lima et al. (2018) based on the scaling-invariant property of the rainfall process (Blanchet et al., 2016; Koutsoyiannis et al., 1998). Lima and coworkers applied this model for estimating GEV parameters at several stations in a homogeneous hydrologic region simultaneously, while we used it for HCMC and CTC, separately. The scaling-invariant property implies that the statistical properties of rainfall with different rainfall durations are linked to each other by an operator consisting of the scaling ratio and scaling exponent (Blanchet et al., 2016). GEV has been widely used to fit the annual maxima in the field of hydrology (Coles et al, 2001; Papalexiou et al., 2013). The standard cumulative distributing function of GEV is defined by :

$$G(y) = exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, \text{for,} \left\{y: 1+\xi\left(\frac{y-\mu}{\sigma}\right) > 0\right\} \quad (2.8)$$

where $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$. μ, σ, and ξ are the location, scale, and shape parameters of GEV, respectively. y denotes the rainfall intensity data.

There are mainly two methods for parameter estimation based on the likelihood function: the maximum likelihood method and the Bayesian method (Muller et al., 2008). The Bayesian method is desirable since it can reasonably assess the uncertainty in estimated parameters (Mélèse et al., 2018). Moreover, to reduce the uncertainty arising from the scarcity of data, Bayesian method can include

other sources of information through the prior distributions of parameters (Muller et al., 2008). The joint posterior distribution with the complete parameter set Θ of the Bayesian GEV model is given by (Lima et al., 2018):

$$Pr\left(\frac{\Theta}{y}\right) \propto \prod_{i=1}^{N}\prod_{j=1}^{T}\prod_{d=1}^{D} GEV\left(y_{d,j,i}|\mu_{d,i},\sigma_{d,i},\xi_{d,i}\right)$$
$$\times N\left(log(\mu_{d,i})\middle|\alpha_i - H_i \cdot log\left(\frac{d+\theta_i}{24+\theta_i}\right),\tau_\mu^2\right) \times N\left(log(\mu_{d,i})|\beta_i - H_i \cdot log\left(\frac{d+\theta_i}{24+\theta_i}\right),\tau_\sigma^2\right)$$
$$\times N\left(\xi_{d,i}|\overline{\xi},\tau_\xi^2\right) \times N(\alpha_i|0,10^6) \times N(\beta_i|0,10^6)$$
$$\times \mathrm{B}(H_i|1,1,0,1) \times U(\theta_i|0,10) \times N\left(\overline{\xi}\middle|0,0.3^2\right)$$
$$\times Inv-Gamma\left(\tau_\mu^2|0.01,0.01\right) \times Inv-Gamma(\tau_\sigma^2|0.01,0.01)$$
$$\times Inv-Gamma\left(\tau_\xi^2|0.01,0.01\right) \quad (2.9)$$

where, y and Θ denote the rainfall intensity value and complete set of Bayesian GEV parameters, respectively, T and N represent the number of years and gauge stations, respectively. D is the set of rainfall durations considered in the construction of IDF curves. α and β are $log(\mu_{24})$ and $log(\sigma_{24})$ associated with the rainfall duration of 24h, respectively. $\overline{\xi}$ denotes the average shape parameter of the considered durations. τ_μ^2, τ_σ^2, and,τ_ξ^2 denote the variance in the respective scaling processes. $N(\cdot,\cdot)$ and $U(\cdot,\cdot)$ represent the normal distribution and uniform distribution, respectively. $B(\cdot,\cdot,a,b)$ denotes the beta distribution associated with the lower bound of a and upper bound of b. The inverse gamma distribution is signified by $Inv-Gamma(\cdot,\cdot)$. The priors for these parameters are explained in detail in Lima et al. (2018) and hence not repeated here. Owing to the lack of sub-daily datasets in PPC, Eq. 2.9 can only be used in HCMC and CTC. Subsequently, posterior distributions of $\overline{\xi}$ and τ_ξ^2 can be obtained and then used as the prior distributions for the scaling process in PPC,

$$Pr(\mu_{24},\sigma_{24},\xi_{24}|y) \propto \prod_{j=1}^{T} GEV\left(y_j|\mu_{24},\sigma_{24},\xi_{24}\right) \cdot N\left(\xi_{24}|\overline{\xi},\tau_\xi^2\right) \quad (2.10)$$

where $\mu_{24},\sigma_{24},$ and ξ_{24} denote the GEV parameters of 24h annual maximum series in PPC. In this case, to remove the difference between daily annual maximum series and 24h annual maximum series, the correction factor was used, as suggested by Young and McEnroe (2003).

Historical return rainfalls at different durations were used to construct the historical IDF curves. However, future return rainfalls, which were based on future disaggregated MMS by ANN, could distort the true MMS. Fig. 2.5 displays a similar distortion (e.g., light-tail problems in shorter durations) in the historical period. When constructing future IDF curves, this distortion needs to be considered.

Therefore, first, we generated ANN-based historical and future return rainfalls. Secondly, the 'change factor' (climate response that indicates the ratio of future rainfall intensities to historical ones) (Wilby et al., 2009) was computed by the ratio of ANN-based future return rainfalls to ANN-based historical ones. Finally, we constructed the IDF curves for future periods by multiplying the 'change factor' and observation-based historical return rainfalls (future IDF curves = observation-based IDF curves × future ANN-based IDF curves / historical ANN-based IDF curves). This procedure is similar to the development of IDF curves by using the ratio of model-based future rainfall intensity and model-based historical rainfall intensity in Liew et al. (2014) and Tabari et al. (2016).

Regarding IDF curves in PPC, we spatially downscaled the daily rainfall datasets by the BC method and constructed the IDF curves by Bayesian GEV-based interpolation (Eq. 2.10). The GEV parameters for 24h estimated from Eq. 2.10 and the scaling parameters previously estimated from Eq. 2.9 for HCMC and CTC were used to obtain the sub-daily GEV parameters for PPC. Subsequently, these GEV parameters were used to compute the return rainfall that was later used to construct IDF curves for PPC.

2.4 Results and Discussion

2.4.1 Comparison of NQM, QM, and QDM

Fig. 2.4 shows the total error of each dataset from CORDEX-SEA generated by the three BC methods. For each climate model dataset of CORDEX-SEA simulations, the most suitable BC method (either of QM, QDM, and NQM) was selected to minimize the total error integrating the influence of bias correction and CCS preservation. On the whole, we selected NQM, QDM, and QM as the most suitable BC method for downscaling 15, 11, and 2 rainfall datasets, respectively (this number corresponds to the number of black bars for each BC method shown in Fig. 2.4). This difference indicates that NQM and QDM performed better than QM. QM performed poorly mainly because QM uses the error correction values established at a higher ECDF to remove the bias at a lower ECDF, which could alter the raw modeled CCS (Switanek et al., 2017). The best BC methods remained almost the same in the near and far future periods and were similar in HCMC and CTC (Fig. 2.4). This similarity in selecting suitable BC methods for HCMC and CTC makes sense given the geographical location of these two cities.

In general, NQM and QDM performed better than QM. For some CORDEX-SEA products (e.g., far future case of Model 7 for HCMC), NQM gave a comparably small total error, although NQM was not selected as the best BC method. Given that NQM performed better than the other two BC methods with regard to certain datasets (e.g., No. 4 in Table 2.1), climate change impact studies in other cities are suggested to use NQM. In addition, since datasets downscaled by selected BC methods performed better than the datasets downscaled by one specific BC method, future study should consider selecting a suitable BC method for each climate model.

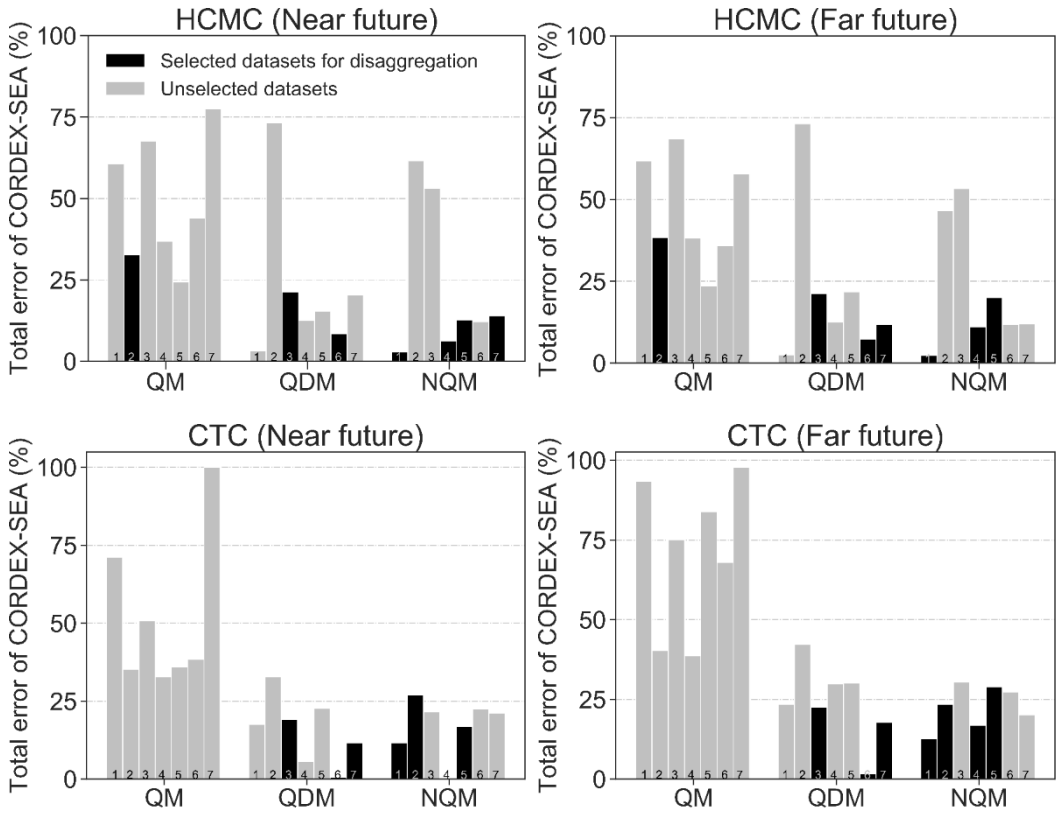


Fig. 2.4. Total error for datasets of CORDEX-SEA generated by three BC methods for HCMC (top) and CTC (bottom) in the near future and far future periods. The black bar indicates the minimum error among the three BC methods (QM, QDM, and NQM) for each dataset. The number at the bottom of each bar, 1 to 7, corresponds to the dataset number in Table 2.1.

2.4.2 Disaggregated MMS

Table 2.3 shows that the optimum input variables for ANN-based models are the same between HCMC and CTC. Comparing AIC between the input variable with sim_MMR12 and without sim_MMR12, such as $(P_{t-1}, P_t, P_{t+1}, sim\_MMR12)$ and (P_{t-1}, P_t, P_{t+1}), we found that sim_MMR12 improved accuracy for long durations (i.e., MMR6, MMR18, MMR24) while reduced it for short durations (i.e., MMR1, MMR3). This finding is probably because that compared with long-duration rainfall extreme, short-duration rainfall extreme is more related to the small-scale convective events (Fowler et al., 2021). Thus the long-duration rainfall extreme doesn't help a lot in improving accuracy for the short durations. The best input variables (shown by the bold number) were then compared with Hyetosminute based on how well they reproduce the observed MMS (Figs. 2.5 and 2.6) and used to obtain future MMS for different durations. The evaluation of ANN-based models is listed in Table 2.4, where the models performed similarly on the training datasets and test datasets, which indicates that there are no overfitting and underfitting problems.

Figs. 2.5 and 2.6 compare the quantile-quantile plot of the observed and simulated (disaggregated) MMS. Both ANN-based model and Hyetosminute gave satisfactory results in long durations (i.e., 12, 18, 24 h), while ANN-based model performed much better than Hyetosminute in short durations (i.e., 1, 3, 6 h). However, both ANN-based model and Hyetosminute tended to distort from (mainly underestimate) the observed MMS in the tails, especially in short durations. These results were in agreement with the results in Mirhosseini et al. (2014). Distortion problems are unavoidable, especially for short durations (Arnbjerg-Nielsen et al., 2013; Burian et al., 2000; Burian and Durrans, 2002; Kossieris et al., 2018; Mirhosseini et al., 2014). This is maybe because short-duration rainfalls are mainly derived from small-scale convective events (which usually last less than one day) that are difficult to represent by the daily rainfall datasets (Llasat, 2001).

Although we improved the ANN-based model by using the most accurate outputs (i.e., sim_MMR12) in previous steps as the inputs in later steps, as shown in Table 2.3, we further need to reduce uncertainty in future studies. First, we can consider other climate factors (such as temperature) as input variables to improve the accuracy of results, especially in short durations. Mirhosseini et al. (2014) showed that ANN-based model has better accuracy when using temperature as the input. Second, data for short durations, provided by high-resolution convection-permitting climate models, can be used directly (Cannon and Innocenti, 2019; Kendon et al., 2017).

In general, ANN-based model showed superior performance over Hyetosminute in short durations, and future work should consider using more input variables and convection-permitting climate models to reduce uncertainty.

Table 2.3. Selecting the optimum input variables for ANN-based models based on Akaike Information Criterion (AIC). Output variables MMR1, MMR3, MMR6, MMR12, MMR18, and MMR24 represent monthly maximum 1, 3, 6, 12, 18, and 24 hourly rainfall series, respectively. P is the total rainfall in that particular day (t) and month (m). P_{t-1} and P_{t+1} are the daily rainfall centered by t. sim_MMR12 is the disaggregated MMR12 in that month. The best input variables are selected for each output variable based on the minimum AIC value (shown by the bold number).

Station	Input variable	Output variable					
		MMR1	**MMR3**	**MMR6**	**MMR12**	**MMR18**	**MMR24**
HCMC	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, \mathrm{P}_m, sim\_MMR12)$	4.69	4.47	4.36	None	4.42	4.64
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, sim\_MMR12)$	4.75	4.44	**4.25**	None	**4.29**	**4.57**
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, \mathrm{P}_m)$	**4.42**	4.11	4.47	3.90	4.50	4.68
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1})$	4.54	**4.09**	4.29	**3.62**	4.37	4.65
CTC	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, \mathrm{P}_m, sim\_MMR12)$	3.85	4.48	4.64	None	3.57	3.16
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, sim\_MMR12)$	4.12	4.36	**4.61**	None	**3.53**	**3.05**
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1}, \mathrm{P}_m)$	**3.65**	4.33	4.66	3.46	4.04	3.74
	$(\mathrm{P}_{t-1}, \mathrm{P}_t, \mathrm{P}_{t+1})$	4.10	**4.28**	4.74	**3.41**	4.01	3.61

Table 2.4. Performance of ANN-based models based on Pearson coefficient r.

Station	Type	Duration					
		1 h	**3 h**	**6 h**	**12 h**	**18 h**	**24 h**
HCMC	Train	0.88	0.96	0.97	0.98	0.99	0.99

	Test	0.93	0.96	0.98	0.99	0.98	0.98
	Train	0.90	0.96	0.97	0.99	0.98	0.99
CTC							
	Test	0.92	0.93	0.94	0.99	0.99	0.99

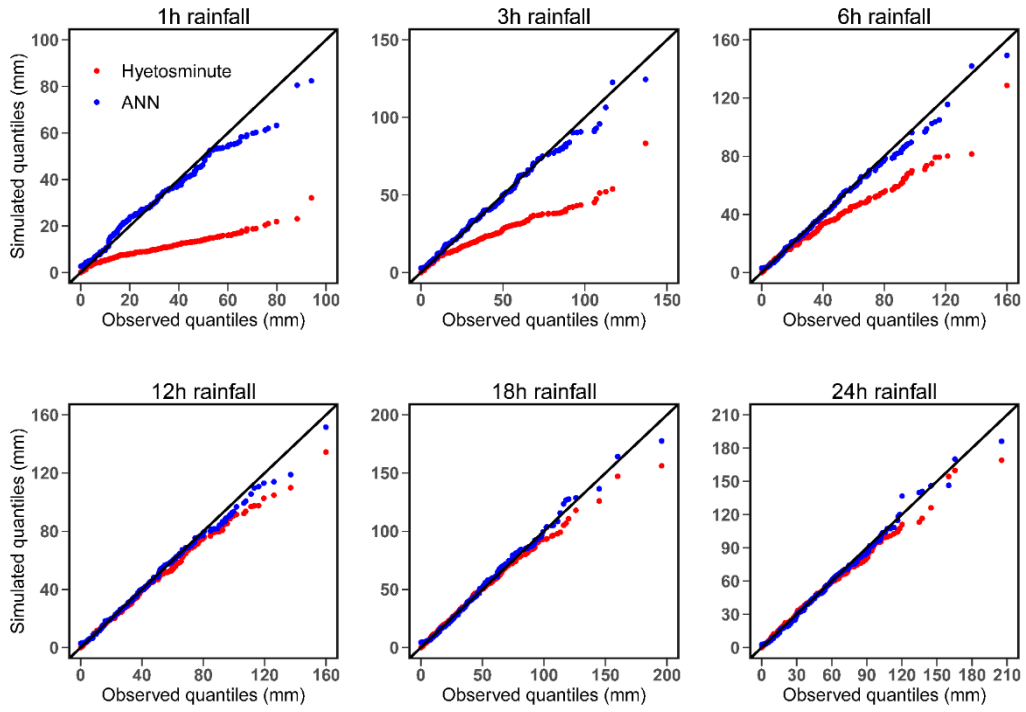
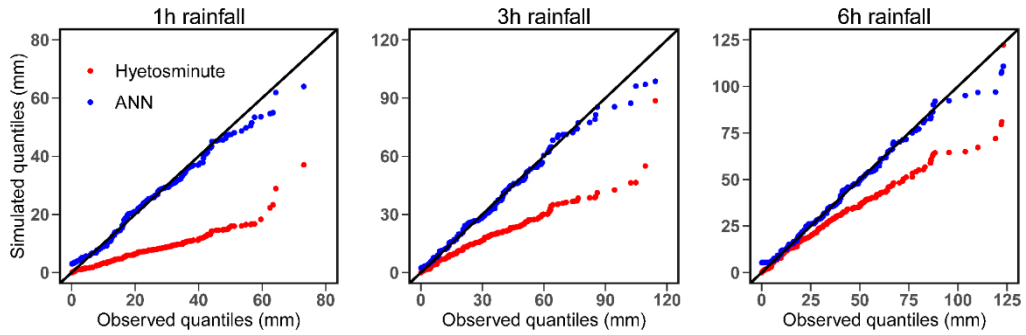


Fig. 2.5 Comparison of the quantile–quantile plot of the observed and disaggregated (wet-day) monthly maximum series in HCMC.



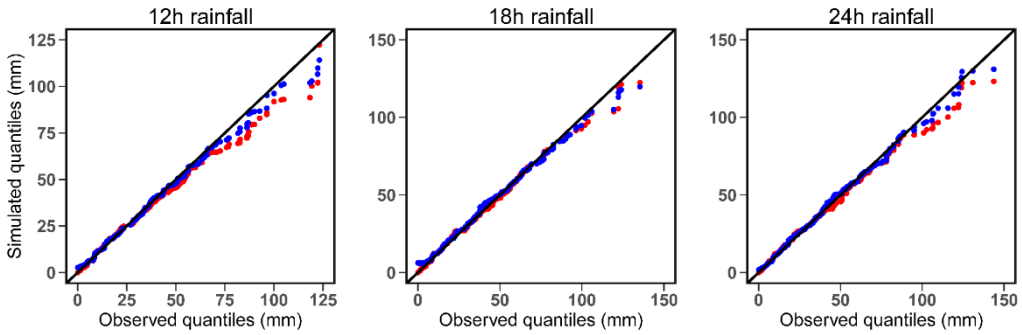


Fig. 2.6. Same as Fig. 2. 5, but for CTC.

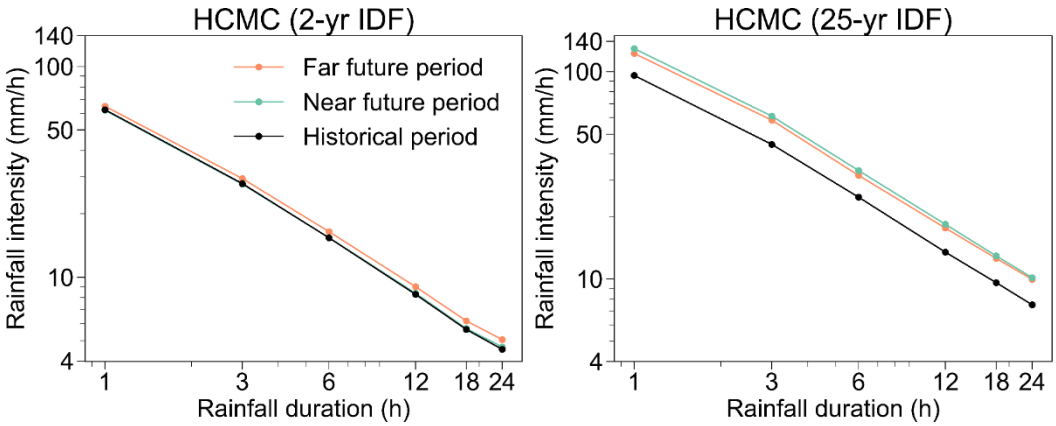
2.4.3 Historical and Future IDF curves

Figs. 2.7 and 2.8 display the IDF curves in historical and future periods. Comparing the 25-yr IDF curves for the near and far future periods, HCMC shows a higher rainfall intensity in the near future while CTC shows a lower rainfall intensity in this period. This observation can be attributed to the discrepancies among the climate models in simulating the changes in rainfall extremes. Based on a preliminary analysis (not shown), most raw climate models indicate that rainfall extremes (average of annual maximum daily rainfall) in the far future are smaller than those in the near future around HCMC, while giving an opposite result around CTC. This contradiction between climate modes provided by CORDEX-SEA simulations has also been identified by Supari et al. (2020), Tangang et al. (2020, 2019, 2018), and Trinh-Tuan et al. (2019). Owing to the complex climate and terrain in SEA, simulating the local rainfalls by regional climate models is still challenging (Tangang et al., 2020). With further development of CORDEX-SEA simulations, the uncertainties resulting from these simulations could be reduced in future studies. Despite the uncertainty, some common conclusions with regard to both HCMC and CTC can be drawn; these are mentioned as follows.

In the near future, for 2-yr IDF curves, which comprise the median values of annual maximum rainfall for different rainfall durations, HCMC showed an average change of 0.6% and CTC of -5.8% compared to the rainfall intensities during the historical period. The changes are in general agreement with the changes in annual maximum daily rainfall in this region reported by Tangang et al. (2018). In the far future, based on the 2-yr IDF curves, the rainfall intensity was expected to increase by 7.7% in HCMC and 1.8% in CTC, which is similar to changes in annual maximum daily rainfall in Supari et

al. (2020). Furthermore, based on the 25-yr IDF curves, the rainfall intensity was expected to increase by 34.5–36.8% (depending on the rainfall duration) in HCMC and by 15.7–30.4% in CTC in the near future, and 27.4–32.6% in HCMC, 55.4–72.8% in CTC in the far future. Thus, compared with IDF curves with a lower return period, IDF curves with a higher return period increased more from the historical period. This difference implies that the rainfall intensity for rarer rainfall events will intensify more in the future, which is similar to the results in Truong Ha, (2018), and in other regions like Europe and North America (Cannon and Innocenti, 2019; Kharin et al., 2018; Myhre et al., 2019). In addition to this observation, generally, more increase is found in longer-duration rainfall intensity in HCMC and CTC (Fig. 2.9). Despite the general increase, for some cases (e.g., 25-year in HCMC), longer-duration rainfall intensity increased less than that of shorter-duration one. This general increase compares well to a previous research on IDF curves in HCMC (Truong Ha, 2018). As extreme short-duration rainfall is generally responsible for flash flooding (Fowler et al., 2021), the less increase in extreme short-duration rainfall in our study region may be helpful in the timely implementation of efficient preventive measures for flash flooding.

Overall, rainfall intensity with higher return periods, especially for longer durations, is expected to increase more in both cities and future periods. These expected change in IDF curves urges us to update the existing drainage systems for higher return periods in HCMC and CTC.



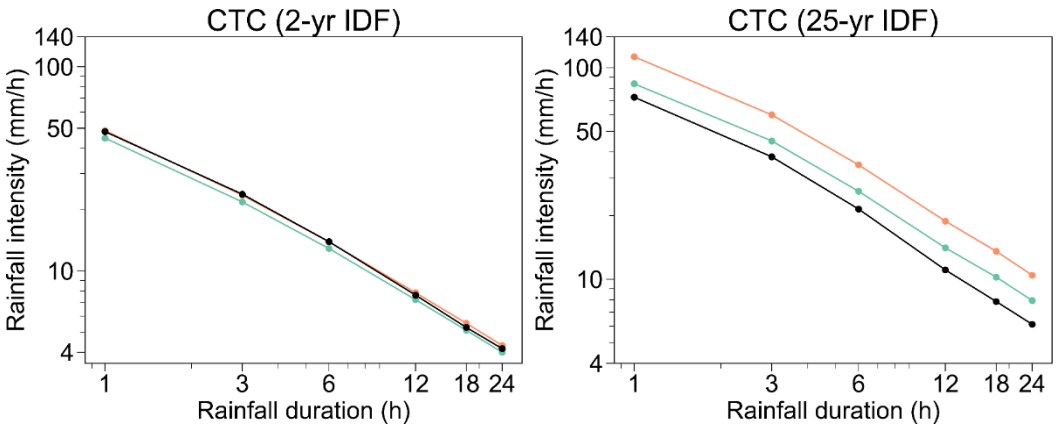


Fig. 2.7. 2 and 25-yr IDF curves for historical and future periods in HCMC (top) and CTC (bottom) based on observed and bias-corrected model datasets. IDF curves for future periods are the mean of the future projections based on the datasets (downscaled by selected BC methods) shown by black bars (after temporal disaggregation) in Fig. 2.4.

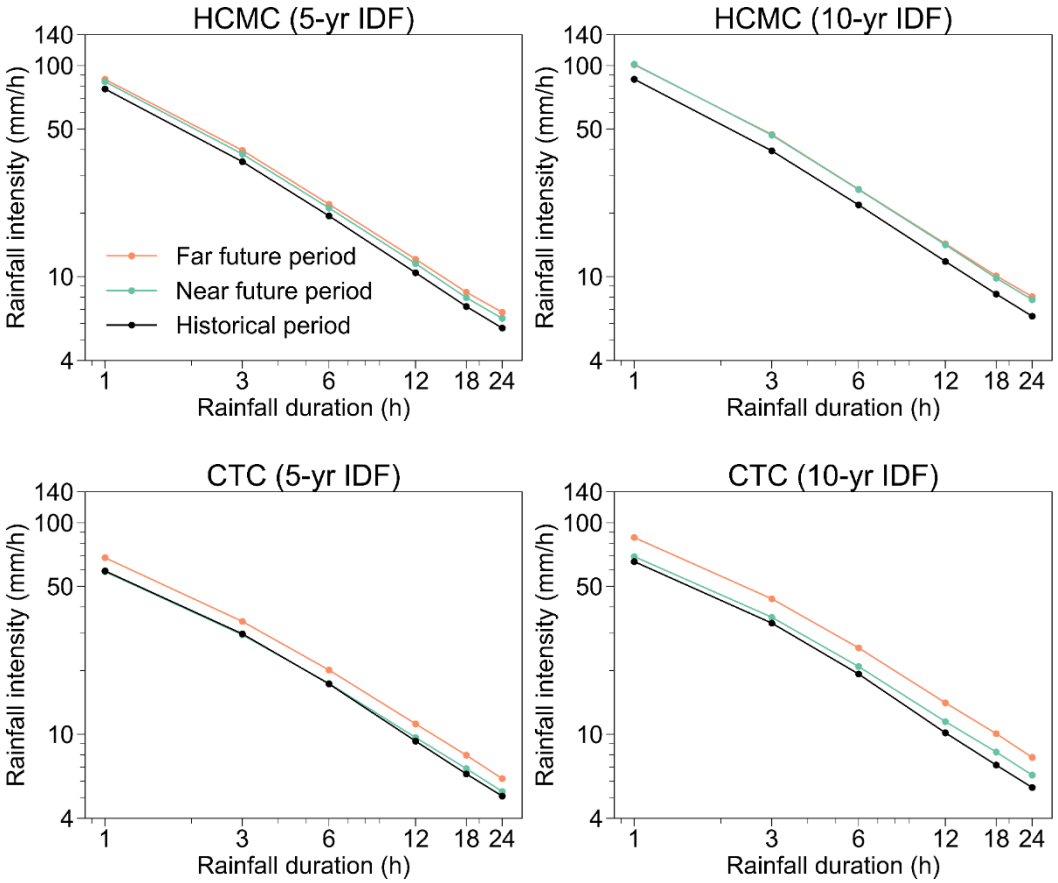


Fig. 2.8. Same as Fig. 2.7, but for 5 and 10-yr IDF curves.

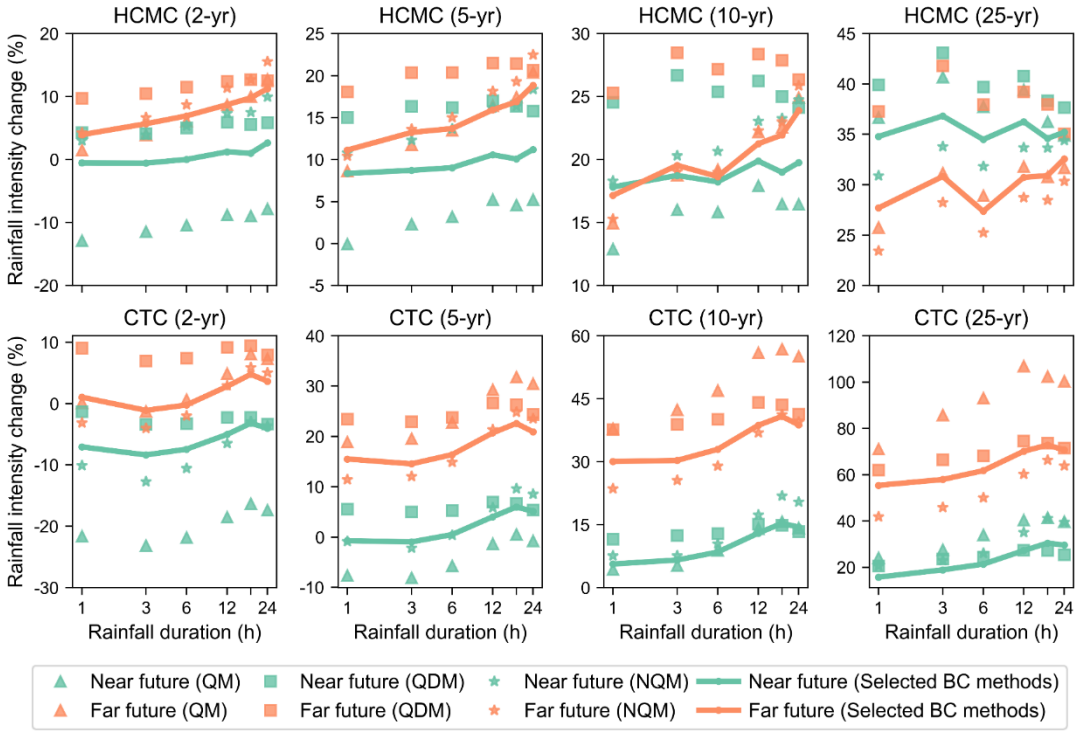


Fig. 2.9. Change in rainfall intensity in the future periods. Rainfall intensity for future periods is the mean of future projections obtained from seven models based on QM, QDM, NQM, and selected BC methods, respectively.

IDF curves for both the historical period (based on observations) and the future periods (based on the RCMs with BC) in PPC are shown in Fig. 2.10. 20-year rainfall intensity is projected to increase by 12.5–13.7% (depending on the rainfall duration) under RCP4.5 in the far future, which is much smaller than 16.4–26.2% in CTC and 31.4–49% in HCMC, during the same period. This difference is partly because we developed the historical and future IDF curves in this data-scarce city by integrating rainfall information at different levels (i.e., daily extreme rainfall in PPC and scaling properties for HCMC and CTC). This integration assumes that these cities are located within the homogenous hydrologic regions, which are difficult to check due to the limited number of existing gauge stations in the region. Nevertheless, this assumption is, to some extent, reasonable given that three cities are located in very flat terrain and relatively close proximity (Fig. 2.1), and Bayesian GEV model (Eqs.

2.9 and 2.10) through Markov Chain Monte Carlo (MCMC) wouldn't be possible to converge if these cities are located in the non-homogenous hydrologic regions. In the future study, further investigation on this assumption based on a delineation of homogenous regions in terms of scaling property is required (Ouali and Cannon, 2018). The use of clustering method (e.g., K-means clustering) to select the homogenous regions based on more rain-gauge observations can be recommended and is the future effort.

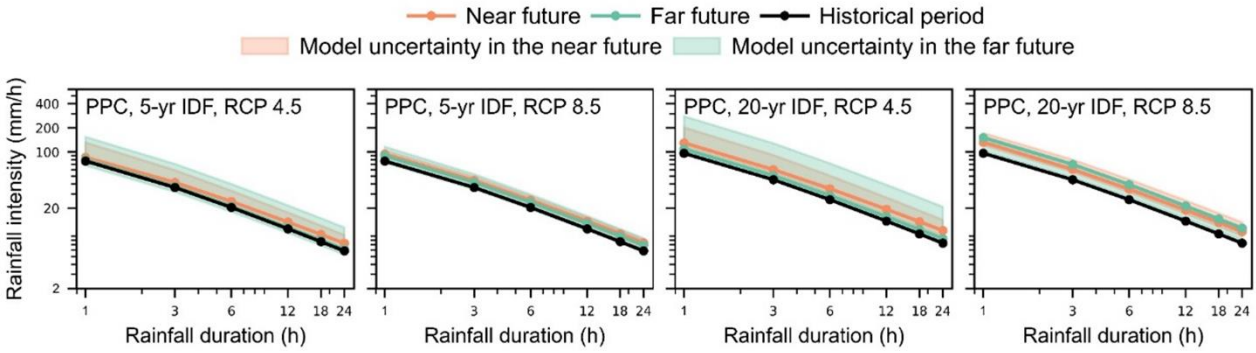


Fig. 2.10. 5- and 20-yr IDF curves in the historical period and future periods in PPC. The rainfall intensity in the historical period was based on observations, and that in the future periods was calculated based on the ensemble median of the RCMs (with BC) by using Bayesian GEV through MCMC. The model uncertainty was based on the 90% of the ensemble spread.

2.4.4 Uncertainty in IDF Curves

The IDF curves projected for the future period include different kinds of uncertainty due to the observed and modeled datasets, and the methods used to estimate the IDF curves. The observed datasets collected by the rain gauges are affected by systematic errors such as the underestimation due to the wind effects, and evaporation (Hosseinzadehtalaei et al., 2020), which is difficult to verify for the observed data we used. In addition, the observations with a record length of 20 years are easily affected by the natural variability of the climate system (Aalbers et al., 2018), thus causing uncertainty in the estimated return rainfall. This limited record length of data is mainly due to the limited observation period and the setting of the historical period in CORDEX-SEA simulations, which ends

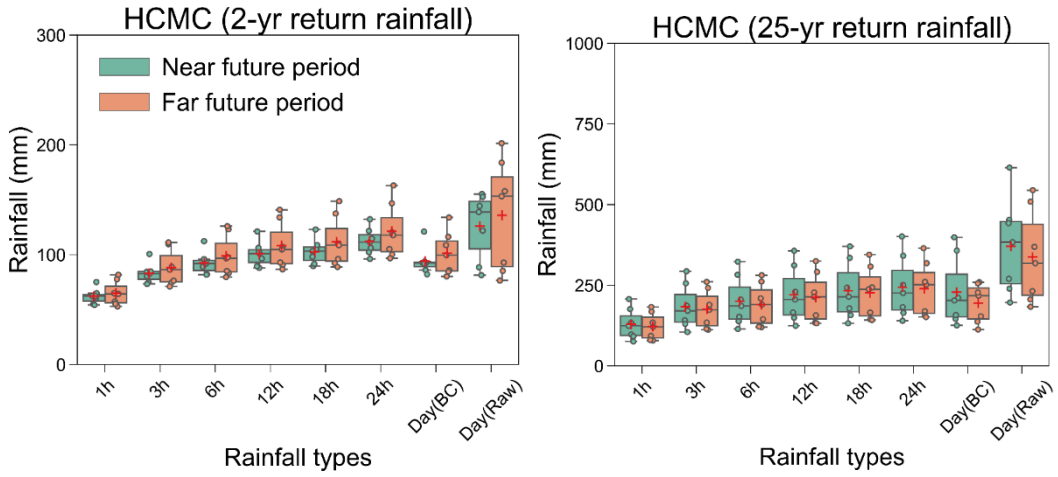
in 2005. Since the period of projection under RCP scenarios starts from 2006, we constrain the time span of observed rainfall examined in this chapter to 1986-2005. Nevertheless, the applied Bayesian GEV model uses more information to estimate the return rainfall, leading to a reduction in uncertainty caused by the natural variability. The uncertainty in the climate models can be attributed to various factors in RCMs and GCMs (e.g., physical parameterization schemes, coarse spatial resolutions, and boundary forcing) (Tangang et al., 2019, 2020). In RCMs, modeling the convective rainfall poses certain errors in the estimation of rainfall. For example, the convective rainfall scheme used in RegCM4-3 (Table 2.1), i.e., MIT-Emanuel cumulus parameterization scheme, was found to generally overestimate daily rainfall amount (Tangang et al., 2020) and daily rainfall extremes (e.g., R99pTOT) over SEA (Ge et al., 2019), and some models (e.g., No.3 in Table 2.1) even projected an opposite future change compared with the forcing GCMs (Tangang et al., 2020). Regarding the temperature, RCMs simulated the daily temperature better than GCMs in our research area, but still projected a different temperature from the forcing GCMs in the future period (Nguyen-Thi et al., 2021). Generally, RCMs affect the return rainfall estimates (Fig. 2.11) much more than the forcing GCMs (Berg et al. 2019). In addition, the coarse spatial resolution of the RCMs is difficult to provide us the high-resolution local rainfall fields in an urban environment, especially in complex terrain (Poschlod et al., 2018). Obtaining local rainfall fields from these RCMs based on BC methods could also cause uncertainty. All these limitations in climate models would cause significant uncertainty in return rainfall estimates (Poschlod et al., 2021). As SEA could be more affected by the enhanced convectional process in the future (Ge et al., 2019), RCMs with explicit convection schemes (focusing on the spatial resolution of 3-4 km) are needed to improve the results (Tabari et al., 2016). In addition to the shortcomings of these climate models, the natural variability of the climate system also limits the predictability of the CCS indicated by the climate models (Aalbers et al., 2018). These effects of natural variability on rainfall were found more obvious at smaller spatial scales and shorter durations (Hawkins and Suton, 2009; Aalbers et al., 2018). As these effects are difficult to estimate from observations with a limited record length, Poschlod et al. (2021) quantified their effects by using 50 ensemble members with the same emission scenario, model structure, and physics. Results show that the convectional process is more affected by natural variability. Therefore, it is also important to quantify the impact of natural variability on the estimated IDF curves, which is subject to the future work.

The uncertainty due to the choice of BC methods gives a substantial influence on the estimation of sub-daily rainfall extremes and resulting IDF curves. Figs. 2.11 and 2.12 display the uncertainty of projected return rainfall in the future periods. Day(BC) and Day(Raw) represent the daily return rainfall based on bias-corrected models and those based on raw climate models, respectively. To better compare these two kinds of return rainfall, we applied an areal reduction factor of 0.8 to Day(Raw) (Willems, 2013). However, the area reduction factor is usually site-dependent, which should not be seen as a reliable transformation of Day (Raw) (Poschlod et al., 2021). The range of the return rainfall (ensemble spread) is usually used as a measure of the uncertainty and ensemble median or mean (the former being more outlier-robust and the latter making full use of the obtained information) is advocated as a representation of the future (Kundzewicz et al., 2018). Maurer and Pierce (2014) found that after bias correction, the ensemble mean better preserved the raw modeled CCS than the ensemble median. This better performance based on the ensemble mean is also shown in Fig. 2.11, where the ensemble mean of Day(BC) showed a similar CCS (change from near future to far future) to that of Day(Raw), while ensemble median Day(BC) sometimes (e.g., 2-yr return rainfall in CTC) showed an opposite CCS to that of Day(Raw). Even though Fig. 2.11 displays some outliers in the certain boxplots (e.g., 1h, 25-year return rainfall in CTC), these outliers (defined by the boxplot) also represent possible and plausible futures (Kundzewicz et al., 2018). Therefore, if we apply BC methods to remove the systematic model bias, using the ensemble mean instead of the ensemble median is suggested.

Compared with Day(Raw), Day(BC) decreased the ensemble spread of daily return rainfalls by 47.1% in HCMC and 61.4% in CTC. This decrease indicates that the BC methods used in our study (especially NQM and QDM) reduced the uncertainty caused by the systematic model bias effectively owing to their approach to capture the CCS, although BC methods cannot be used to solve the fundamental problems of RCMs (e.g., parameterization schemes) (Maraun, 2016). Furthermore, ensemble means of Day(Raw) generally gave a larger return rainfall than those of Day(BC). This overestimation in rainfall extremes was also reported by Supari et al. (2020). A preliminary analysis (not shown) shows that some models project extremely large rainfalls, which cannot be seen as outliers without a compelling reason. The bias in these larger rainfalls is difficult to remove by almost all BC methods that calculate quantiles in a non-parametric way (Maraun et al., 2019). Future work could use a parametric method (such as Generalized Pareto distribution) to fit these extremes in the bias

correction (Switanek et al., 2017).

Some durations, such as 12, 18, and 24 h in CTC (Fig. 2.11, 25-yr return rainfall, near future), gave a wider ensemble spread than that of Day(BC). This wider spread is mainly due to the uncertainty arising from the temporal disaggregation model. This uncertainty from temporal disaggregation models cannot be eliminated (Burian et al., 2000; Burian and Durrans, 2002; Kossieris et al., 2018; Mirhosseini et al., 2014) but can be reduced by the solutions (i.e., using more input variables and convection-permitting climate models) mentioned in Section 2.4.2. Moreover, the temporal disaggregation model trained by the historical daily rainfall datasets was used in the future period based on the so-called stationarity assumption, which could introduce additional uncertainty. Although using the projected daily rainfall datasets for future periods reduces the uncertainty, the rainfall generating systems for the sub-daily rainfall are usually different from the daily rainfall (Cannon and Innocenti, 2019). Quantifying the uncertainty caused by the stationarity assumption is difficult due to the complexity of the rainfall generating systems and the interaction between thermodynamic and dynamic factors (Moustakis et al., 2021). Though the stationarity assumption is commonly accepted in climate change impact studies (Hosseinzadehtalaei et al., 2020, 2021; Lombardo et al., 2017), this assumption needs to be checked in the future study when climate models with higher accuracy are available (Berg et al., 2019).



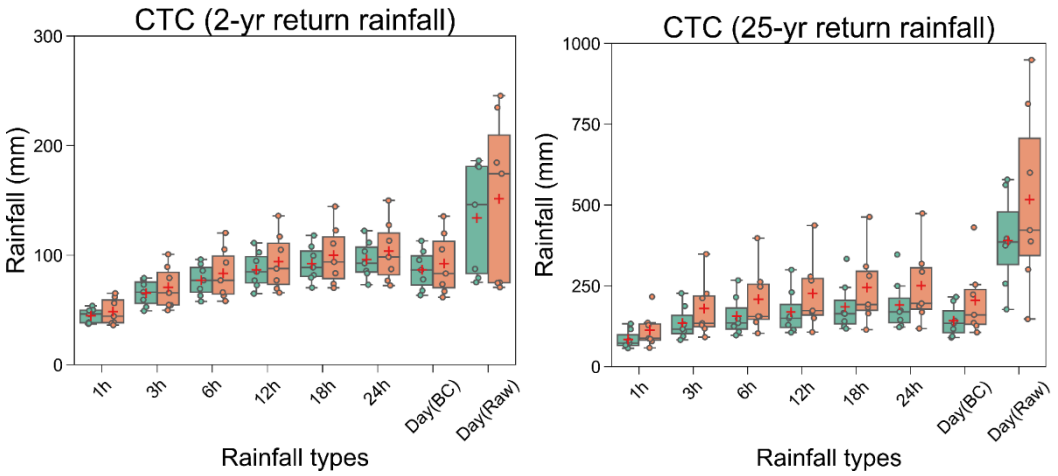
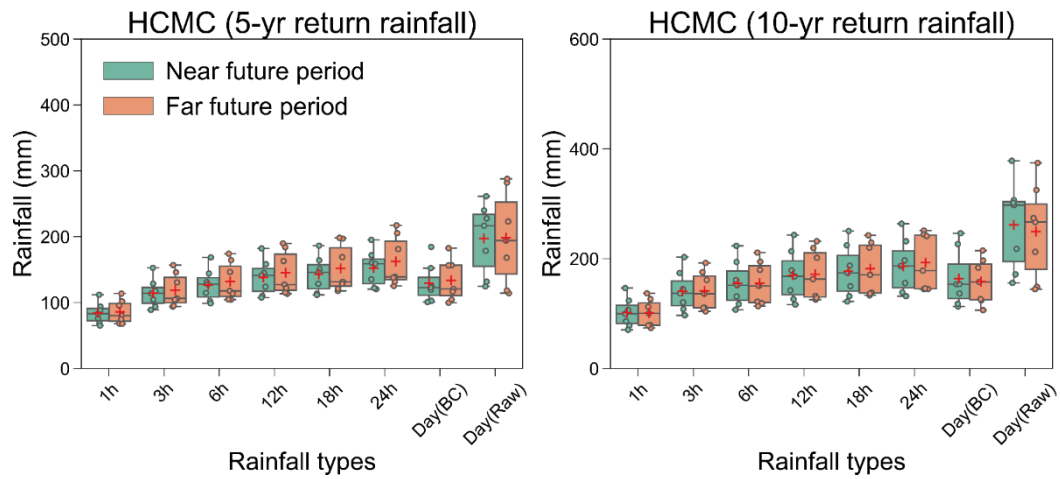


Fig. 2.11. Box plot showing the uncertainty in 2 and 25-yr return rainfalls with different rainfall durations in HCMC (top) and CTC (bottom). Day (BC) and Day (Raw) represent the daily return rainfall based on bias-corrected model datasets (downscaled by selected BC methods) and those divided by an areal reduction factor based on raw climate model datasets (that mean uncorrected models), respectively. The horizontal line within the box represents the median value, and each circle represents rainfall depth for a specific climate model. The red plus sign in the figure denotes the ensemble mean value. The extent of box signifies the first quartile (Q1) to the third quartile (Q3), which is also called interquartile range (IQR). The values outside the range of [Q1- 1.5IQR, Q3 + 1.5IQR] are considered as outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



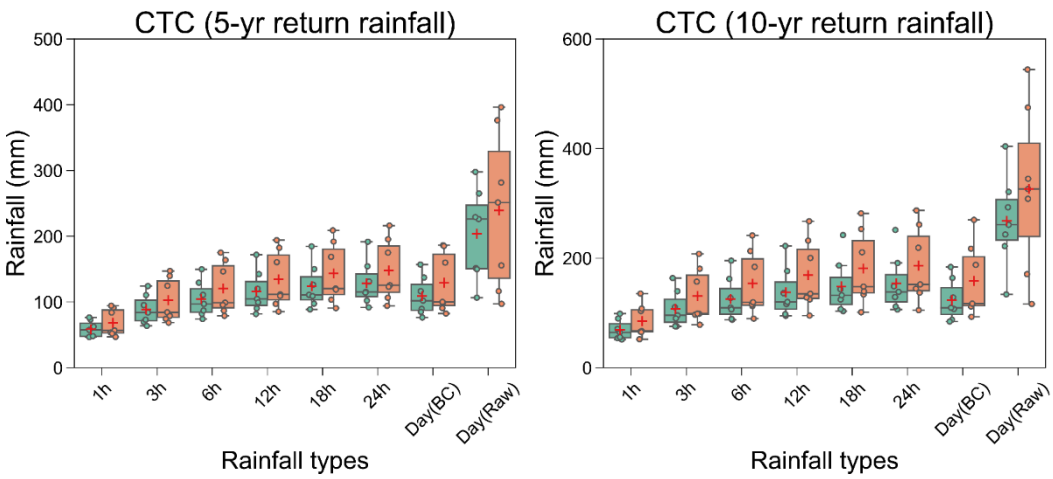


Fig. 2.12. Same as Fig. 2.11, but for 5 and 10-yr return rainfalls.

2.5 Summary

This chapter provides a framework for constructing IDF curves in the near future period (2026–2045) and far future period (2066–2085) for tropical cities in Southeast Asia based on the state of art climate models provided by CORDEX-SEA simulations, through the application in three cities in Southeast Asia. In this framework, we proposed a new BC method, NQM, and compared it with QDM and QM. Thereafter, we spatially interpolated the IDF values in a data-scarce city (i.e., PPC) by using the scaling property of heavy rainfall. Furthermore, we selected the best BC method for each dataset provided by CORDEX-SEA to ensure the removal of the systematic model biases by BC methods as well as the retainment of the raw modeled CCS of CORDEX-SEA simulations. An ANN-based disaggregation model was introduced and compared with a stochastic disaggregation model, Hyetosminute. The following conclusions can be drawn:

(1) In general, NQM and QDM performed better than QM. Given that for some datasets, NQM showed much better accuracy than QDM and QM, the climate change impact studies in other cities should consider using NQM and comparing it with other downscaling methods.

(2) CORDEX-SEA-simulated rainfall is always characterized by a significant bias and uncertainty, which precludes its direct use for constructing IDF curves. The selected BC methods eliminated the systematic model bias and reduced the associated model uncertainty for the construction of IDF curves for future periods effectively.

(3) The short-duration MMS generated by ANN was much more accurate than that generated by Hyetosminute. Future work should focus on reducing uncertainty, especially in short durations. In addition to utilizing the enhanced observed and modeled datasets, using the most accurate outputs in previous steps as the inputs in later steps is recommended.

(4) Projected IDF curves for future periods revealed that rainfall intensity for rarer rainfall events was expected to increase more, requiring that the existing urban drainage systems for higher return periods in HCMC, CTC and PPC be updated.

3. Establishment of IDF curves in data-scarce countries by using auxiliary datasets: changes, physical mechanisms, and associated risks

3.1 Background

Cambodia is regarded as one of the most climate-vulnerable countries worldwide, ranked as 14th out of 181 countries by the global climate risk index (CRI) for 2000-2019 (Germanwatch, 2021). The high CRI, representing high vulnerability and exposure to extreme weather events, is primarily due to the climate-sensitive fabric of the country (WBG and ADB, 2021), where about one-fourth of the country's population is likely to be affected in case of extreme flood events (Willner et al., 2018). Heavy reliance on rainfed (not irrigated) agriculture and inland fisheries (essentially surrounding Tonle Sap Lake) and their contribution to the country's GDP (25%) and labor force (49%) highlight the need for timely adaptation to mitigate the devastating effects of weather extreme on people and properties (WBG and ADB, 2021; Willner et al., 2018). The country has high exposure to flooding, mainly ascribed to the heavy rainfall, with an average of ~10% of the total population affected by flooding annually (Chen et al., 2020; Kundzewicz et al., 2014) and an average annual economic loss of $250 million (WBG and ADB, 2021). Therefore, quantifying the impact of climate change on flooding and putting forward subsequent adaption suggestions for decision-makers have become the substantial way forward to alleviate the flood risks effectively. Although heavy rainfall-induced flooding has been recognized as the dominant natural hazard in Cambodia (UNISDR and the World Bank, 2010), sustainable infrastructure design to withstand flooding remains largely unexplored due to the scarcity of rainfall dataset in terms of the record length, spatial coverage, and/or temporal resolution. Therefore, in the context of global warming, the question arises whether or not the current infrastructure design of Cambodia is resilient enough to withstand the future heavy rainfall and its resultant flooding. Since there is no previous work to answer this question, we hypothesize that the future flooding risk in Cambodia under climate change depends on the rainfall duration, geographical conditions and region-wide socio-economic situations, which needs to be explicitly quantified to prioritize the areas for adaptation strategies.

Extreme/heavy rainfalls with a given exceedance probability are commonly used for the infrastructure design (e.g., urban drainage system) to withstand flooding (Kourtis and Tsihrintzis, 2021). Although the conventional way of quantifying the extreme rainfall under different durations, which are further processed to obtain rainfall IDF curves, is based on the historical rain gauge-based rainfall datasets, non-availability of sufficient record length, impaired data quality, and inadequate spatiotemporal distribution impede the assessment solely based on the ground data, especially in developing countries (Ombadi et al., 2018). Therefore, researchers have recently been focusing on using auxiliary precipitation products, i.e., remotely sensed datasets, reanalysis datasets, or a combination thereof in conjunction with the gauge data (Courty et al., 2019; Hosseinzadehtalaei et al., 2020; Marra et al., 2017; Noor et al., 2021; Ombadi et al., 2018; Poschlod et al., 2021). These alternative datasets can provide rainfall information in the ungauged locations even in complex terrains, thus helping in quantifying the distributive heavy rainfall (which plays a vital role in the flood generation (Zhou et al., 2019)), associated risks and impacts, and subsequent implications for the infrastructure design.

Regarding future rainfall, GCMs and RCMs are commonly used for rainfall projections at various spatiotemporal scales. Compared with GCMs, RCMs included in the CORDEX, which is a part of Coupled Model Intercomparison Project 5 (CMIP5), have been extensively used in projecting changes in regional-scale rainfall (Hosseinzadehtalaei et al., 2020, 2021; Requena et al., 2021; Tangang et al., 2018; Villafuerte et al., 2020) owing to their added values in simulating the physical process of regional climate. CORDEX in Southeast Asia, launched in 2013 by international collaborations between several research institutions, represents the most comprehensive set of readily available regional climate simulations over the region so far (Tangang et al., 2020). However, sub-daily rainfall datasets from CORDEX in Southeast Asia are still limited due to the high demand for computational power (Westra et al., 2013). Unfortunately, sub-daily heavy rainfall, particularly 1-3h, can inflict severe damage on the urban area through rapidly developing flooding (Fowler et al., 2021). This rainfall information is essential for urban planning (e.g., drainage system design) in developing cities (e.g., Phnom Penh in Cambodia), where the drainage system's capacity is either not sufficiently well designed or not updated timely to match the increasing residential area and other urban infrastructure (Irvine et al., 2015; Yim et al., 2016). This problem has been aggravated by the transformation of the wetlands and existing

drainage systems into impervious surfaces (Mialhe et al., 2019).

Therefore, it is imperative to apply the temporal disaggregation models to obtain rainfall datasets with finer temporal resolutions. The widely used models can be mainly categorized into two types: stochastic models (e.g., Kossieris et al., 2018) and data-driven models (e.g., Nourani and Farboudfam, 2019), both of which are mainly used to generate the time-series of rainfall with relatively finer temporal scales. Other models mainly focus on the extreme rainfall, including disaggregating daily maximum to a sub-daily maximum based on equidistance quantile mapping (Srivastav et al., 2014) and establishing a relationship among rainfalls with different temporal resolutions based on the 'scale-invariance' (or 'scaling') property (Blanchet et al., 2016; Cannon and Innocenti, 2019). However, all of the abovementioned models typically suffer from the stationarity assumption that supposes the relationships between the covarying variables (i.e., rainfall with finer and coarser temporal resolutions) established based on historical datasets are also applicable for the future (Lu et al., 2016; Sane et al., 2018, among others). In other words, when obtaining the rainfall with finer temporal scales (e.g., sub-daily) only based on rainfall with coarser time scales (e.g., daily rainfall), these models may be problematic as the daily and sub-daily rainfall would not change in the same manner in the context of changing climate (Westra et al., 2013). Even under limited climate change and global warming influences, this problem can be recognized from the different CC scaling for the daily rainfall and the sub-daily rainfall (super-CC scaling) (Lenderink et al., 2017). Therefore, it is essential to have improved statistical methods (e.g., a temporal disaggregation model with less dependence on the stationarity assumption) utilizing the auxiliary datasets (e.g., reanalysis or modeled datasets).

This chapter has considered both aspects of accuracy in quantifying extreme rainfalls and developing IDF information, i.e., high-quality input data and more robust statistical inference methods. As a whole, this chapter investigates and bridges over the quantitative assessment of the heavy rainfall characteristics, followed by developing a province-level flood risk change map and proposing adaptation strategies. More specifically, the objectives are,

1) To propose a temporal disaggregation model which is less dependent on the stationarity assumption.
2) To project the global warming-induced change in heavy rainfall (referring to a rainfall intensity with a given return period and rainfall duration) over Cambodia and subsequently establish the past and future IDF relationships, followed by quantification and discussion of the governing physical

mechanisms.

3) To develop a province-level flood risk change map based on the projected heavy rainfall and various socioeconomic factors and provide strategies for sustainably mitigating the multidimensional adverse impacts of flood events in different provinces of Cambodia.

The remainder of this chapter is presented as follows. Section 3.2 describes the study area and datasets used in this chapter, including both the rainfall datasets and the atmospheric variables used for establishing the temporal disaggregation model. Section 3.3 introduces a framework applied to quantify the extreme rainfalls under different durations and return periods and develop the past and future IDF curves for Phnom Penh, based on the disaggregated (duration of 1 hour) and existing (durations longer than 1 hour) rainfall extremes. Thereafter, the way to develop the flood risk change map is introduced. Section 3.4 presents the results and discussion, including warming-induced changes in heavy rainfall, associated physical mechanisms, and the specific sustainable adaptation strategies for different provinces of Cambodia. At last, conclusions are summarized in Section 3.5.

3.2 Study area and datasets

3.2.1 Case study area

Cambodia is located in the southern area of Indochina Peninsular and is bordered by Thailand, Laos, Vietnam, and the Gulf of Thailand (Fig. 3.1). The topography consists of a central low-lying terrain surrounded by mountains and highlands in the coastal areas and at the eastern borders. The climate in Cambodia is governed by the tropical monsoon, which brings highly humid air from the sea during the rainy season (May to October). Approximately 80–90% of the rainfall occurs during the rainy season, with a distinctively large rainfall in the coastal area and the nearby highland regions (WBG and ADB, 2021). Cambodia's population is approximately 16.5 million, 76% of which live in rural areas (WBG and ADB, 2021). Since the beginning of the 21st century, Cambodia has been sustaining one of the fastest-growing economies in the world (average annual growth rate of 7.7%) (World Bank, 2021) and experiencing rapid urbanization that is likely to be accelerated in the future due to population growth and the associated socioeconomic developments (UNDESA, 2019). In addition to flash floods ascribed to the undeveloped drainage system (Section 3.1), river floods arising from the high moisture brought by monsoon winds and high water levels caused by heavy rainfall in

the upstream of the Mekong River are also causing severe damage to the country (JICA, 2015; WBG and ADB, 2021). For example, 19 out of 25 (76%) provinces in Cambodia were affected by the recent heavy rainfall-induced (river and flash) floods in October 2020, with over 175 thousand households affected and 300 thousand hectares of agricultural land damaged (HRF, 2020).

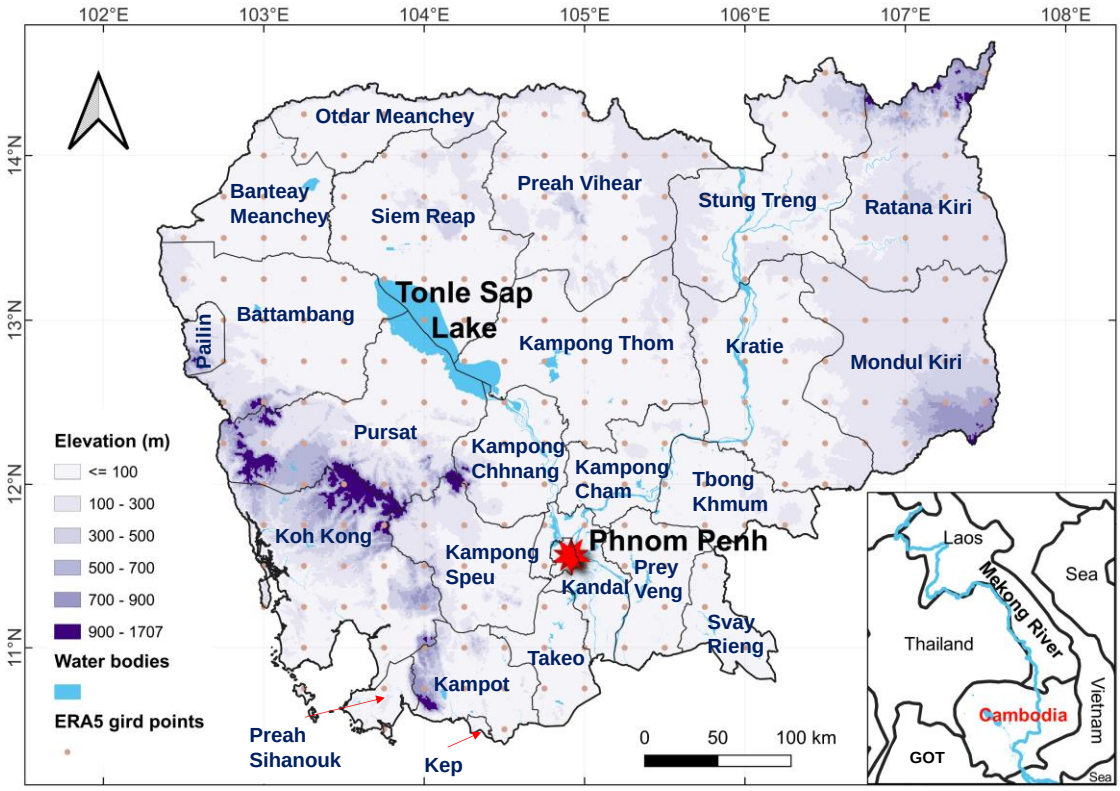


Fig. 3.1. The geographical location (inset), provinces, and elevation (above mean sea level) of Cambodia and ERA5 grid points used for obtaining IDF information. GOT: Gulf of Thailand.

3.2.2 Datasets

To get the input datasets for the temporal disaggregation model (Section 3.3.1), a range of atmospheric variables summarized in Table 3.1 from the CORDEX simulations (Table 3.2) was extracted from the ERA5 reanalysis (regarded as grid-scale observed datasets) and CORDEX simulations (regarded as grid-scale projected datasets) for each grid point in Fig. 3.1. The data extracted from ERA5 is from 1980 to 2005, at a spatial scale of 0.25° × 0.25° and a temporal scale of one hour. ERA5 is the new reanalysis that replaces the ERA-Interim reanalysis (spanning 1979

onwards) which was started in 2006. Benefitting from a decade of developments in data assimilation, model physics, and core dynamics, ERA5 possesses a significantly enhanced horizontal resolution of 31 km, compared to 80 km for ERA-Interim, hourly output throughout, and an uncertainty estimate from an ensemble (3-hourly at half the horizontal resolution) (Hersbach et al., 2020).

The one from CORDEX is from 1980 to 2005 (historical climate reproduction that ends in 2005), 2040-2065 (mid-century), and 2070-2095 (late-century) under the RCP 8.5 scenario (future climate projection that starts from 2006), at a spatial scale of 0.22° × 0.22° and temporal scale of three hours. RCP 8.5 scenario was used in the current study to represent a worst-case future situation (Hausfather and Peters, 2020). It should be noted that extraction of these variables was done by bilinear interpolation if the variables are not located exactly at the ERA5 grid points shown in Fig. 3.1. These variables represent the water vapor content (specific humidity) in the atmosphere, horizontal water vapor flux (e.g., wind speed, air pressure) from nearby regions to the heavy rainfall regions, and vertical water vapor transport (e.g., surface latent heat flux). Moreover, previous applications have successfully used these variables (e.g., Fowler et al., 2021, Westra et al., 2013) to study the atmospheric thermodynamics and dynamics that may trigger heavy rainfall events.

Table 3.1. Summary of the variables extracted from ERA5 reanalysis and CORDEX simulations.

Variable	Abbreviation	Unit
Precipitation (1h for ERA5, 3h for CORDEX)	pr	mm
2 m air temperature	tas	℃
2 m specific humidity	huss	g/kg
10 m eastward wind speed	uas	m/s
10 m northward wind speed	vas	m/s
Surface latent heat flux	hfls	W/m^2
Surface sensible heat flux	hfss	W/m^2
Sea level pressure	psl	Pa
Surface air pressure	ps	Pa

Table 3.2. List of CORDEX simulations for RCP 8.5 used in this chapter.

GCM	RCM	Developer	Resolution	Time
HadGEM2-ES	RegCM4-7	ICTP	0.22° (3h)	1980–2005 (Historical period) 2040–2065 (Mid-century) 2070–2095 (Late-century)
NorESM1-M				
MPI-ESM-LR				

Acronym: ICTP, International Centre for Theoretical Physics

3.3. Methods

The overall research methodology framework adopted in the current study consists of four main parts: temporal disaggregation, extreme value analysis, development of IDF relationships, and adaptation strategies discussion based on the developed flood risk change maps. A schematic of the framework is shown in Fig. 3.2.

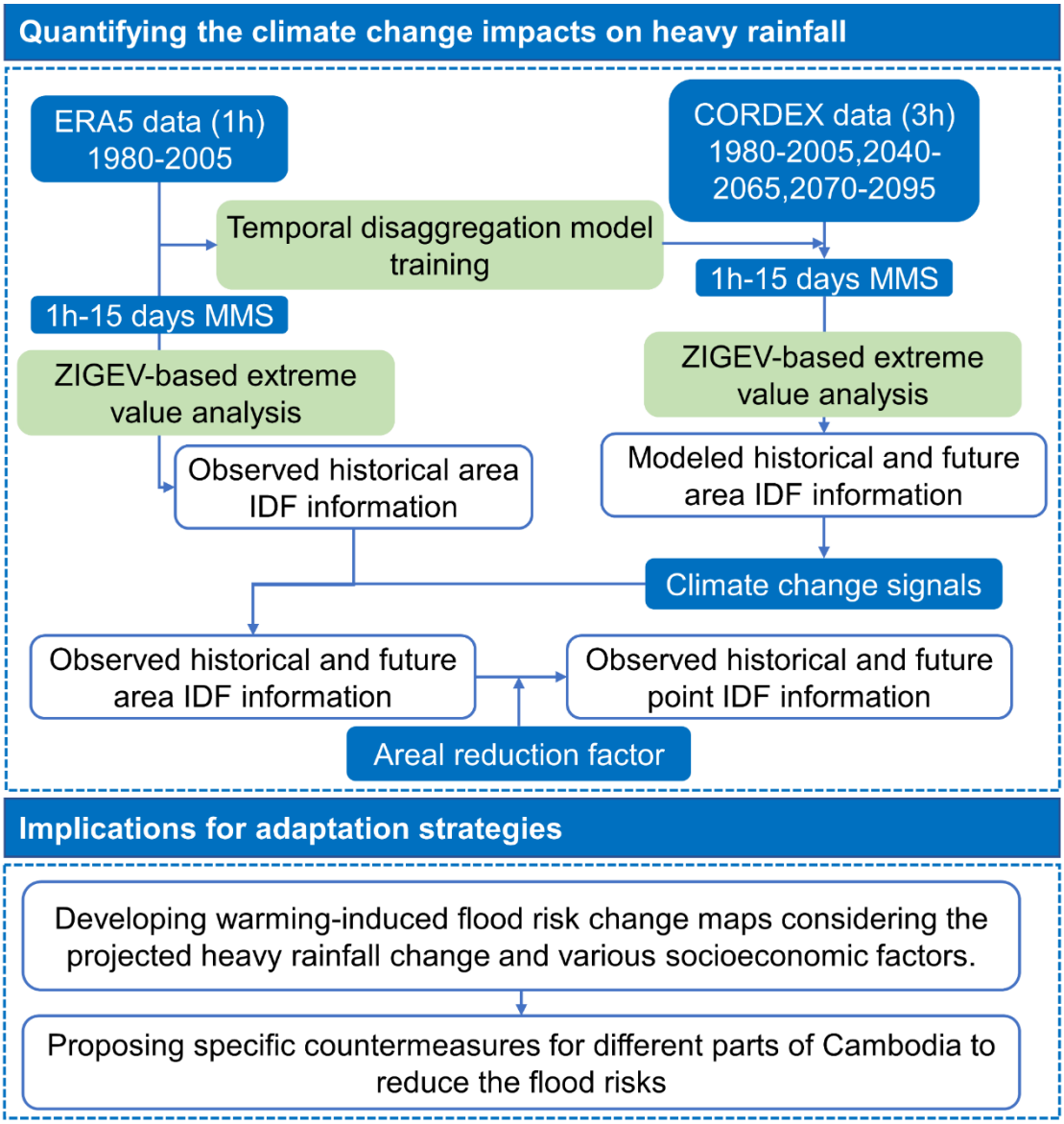


Fig. 3.2. Research methodology framework used in this chapter. (MMS: Monthly maximum series, ZIGEV: Zero-inflated generalized extreme value distribution).

3.3.1. Conditional temporal disaggregation

In this chapter, we built over and advanced an ANN-based temporal disaggregation model previously used in chapter 2 (Fig. 3.3). This improved ANN model uses multiple atmospheric variables and is designed to generate future finer temporal-scale rainfall extremes conditional on (in other words, based on) the historical atmospheric status, unlike the model in Section 2.3.4, which is based only on rainfall datasets. Therefore, this conditional ANN model (hereinafter, CANN), to some extent, is able

to reduce the stationarity uncertainty in past applications (e.g., Kossieris et al., 2018). The CANN model consists of three layers, i.e., input, hidden, and output layers, which are used to minimize the error function through the backpropagation. Owing to the good performance and simplicity of the ANN model with respect to traditional hypothesis-driven models, ANN models have been successfully utilized in various hydrological applications, such as streamflow forecasting (Cheng et al., 2020), flood frequency analysis (Kumar et al., 2015), temporal rainfall disaggregation (Mirhosseini et al., 2014), and data gap filling (Abhishek et al., 2021), particularly when the ample hydrological data of auxiliary variables are available. However, the establishment of the ANN model is usually time- and labor-consuming due to the flexibility of the ANN model's structure (i.e., hyperparameters such as learning rate). Therefore, we applied the grid search approach to tune the hyperparameters.

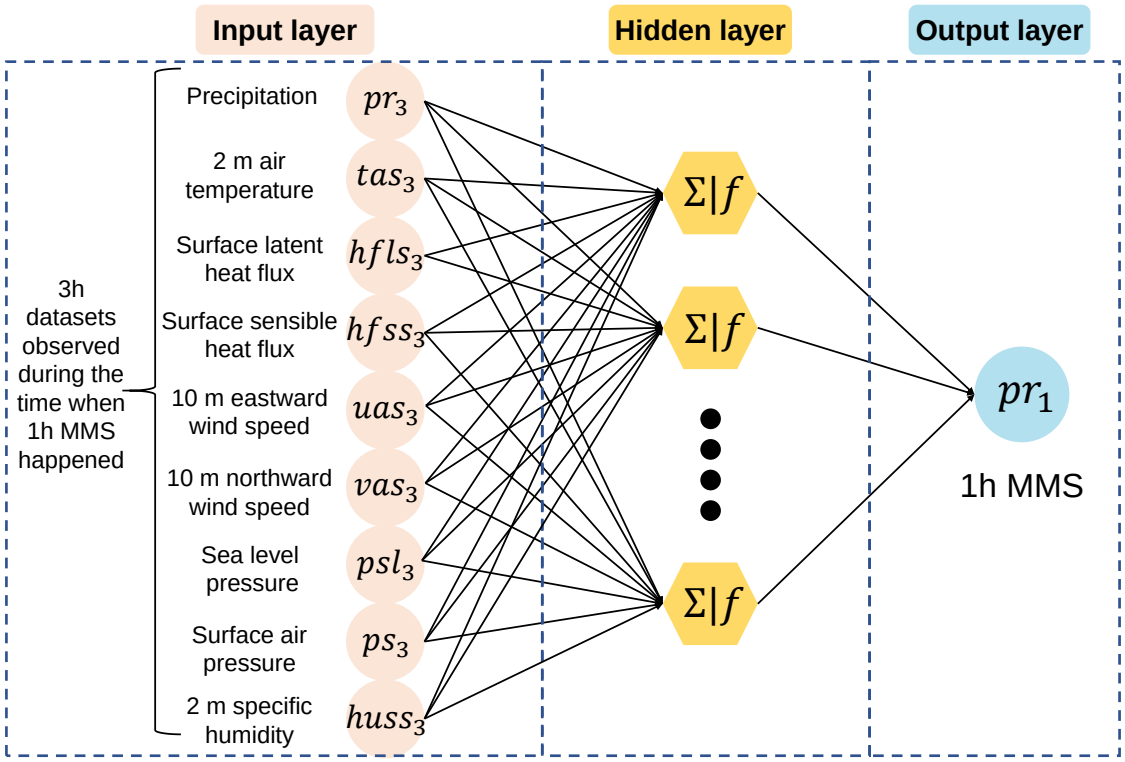


Fig. 3.3. The framework of the CANN temporal disaggregation model developed in the study. Various predictors (input layer), hidden layers, and the output variable are also shown.

Apart from the hyperparameter tuning, some studies suggested that selecting the suitable variables (i.e., predictors) to save the computing resources is critical for ANN models (Mirhosseini et al., 2014).

Nevertheless, the current CANN model does not require this input variable selection because we only input the concurrent atmospheric variables (i.e., information on the atmospheric status when heavy rainfall happens) for training instead of the whole time-series of atmospheric variables (Fig. 3.3). Our preliminary assessment using the complete time series did not improve the accuracy of the model, thus giving confidence in using the concurrent variables. More specifically, in the training processes, a MMS of hourly (1h) rainfall was firstly extracted from ERA5 reanalysis as the output variables. Secondly, 3h dataset of each atmospheric variable, aggregated from 1h ERA5 dataset of a corresponding atmospheric variable, were generated as the input variables if these 3h datasets were recorded during the 1h MMS events. For example, when one heavy rainfall event from 1h MMS happened at 10 a.m., the 3h datasets covering the period from 9 a.m – 11 a.m. would be extracted as the input variables (Fig. 3.3). When the CANN model was applied for disaggregating future rainfall in the CORDEX simulations, the 3h time series of rainfall along with other 3h time series of atmospheric variables was given as input in the model directly, and subsequently, 1h MMS was extracted. Afterward, this extracted (disaggregated) 1h MMS was compared with the observed 1h MMS (Section 3.4.1). Overall, for each grid cell over Cambodia (Fig. 3.1), 243 CANN models were established.

3.3.2. Extreme value analysis

In order to reduce the impacts of the bias in the disaggregated rainfall (Kossieris et al., 2018; Mirhosseini et al., 2014), we performed extreme value analysis (EVA) based on MMS directly, which is different from the traditional approach used in these past studies, where the AMS was extracted from the disaggregated MMS and fitted based on some commonly used probability distribution functions, such as GEV.

In this chapter, for 1h rainfall, EVA was applied directly to 1h MMS, which is the output of the CANN model (temporal disaggregation) explained in Section 3.3.1 to infer the rainfall intensity with a given return period (hereinafter, return rainfall). The same approach was also applied for a set of MMS (with the duration of 3, 6, 12, 18, 24, 36, 48 hours (2 days), 5, 10, and 15 days) provided by CORDEX simulations. Due to numerous zero-rainfall in MMS (especially during December, January, February, March), commonly used probability distribution functions, such as GEV and Pearson III distribution, cannot be used (Silva et al., 2014). Therefore, we applied a zero-inflated generalized

extreme value function (ZIGEV) proposed by Quadros Gramosa et al. (2019), which is a combination of zero-inflated Poisson Distribution (Lambert, 1992) and GEV. The parameters in ZIGEV were estimated using a Bayesian framework with the Markov Chain Monte Carlo algorithm.

The estimation of ZIGEV parameters was implemented by a joint application of Metropolis Block sampling and Gibbs sampling (Quadros Gramosa et al., 2019). Mathematically, the joint posterior probability distribution,π, is given by

$$\begin{aligned}\pi(\theta, w|D) \propto\ & w^{S+c-1}(1-w)^{n-S+d-1}\left(\frac{1}{\sigma}\right)^{m-S}\left[\left(1-\frac{\xi\mu}{\sigma}\right)^{-\frac{1}{\xi}-1}\right]^{m-S} \\ & \times \prod_{i=1}^{n-m}\left\{\frac{1}{\sigma}\left[1+\frac{\xi}{\sigma}(z_i-\mu)\right]^{-\frac{1}{\xi}-1} exp\left\{-\left[1+\frac{\xi}{\sigma}(z_i-\mu)\right]^{-\frac{1}{\xi}}\right\}\right\} \\ & \times \left[exp\left\{-\left(1-\frac{\xi\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\}\right]^{m-S} exp\left\{-\frac{\mu^2}{\sigma_{0\mu}^2}-\frac{\xi^2}{\sigma_{0\xi}^2}\right\}\sigma^{a_0-1}exp\{-\sigma b_0\}\end{aligned} \tag{3.1}$$

where $\theta = (\mu, \sigma, \xi)$, which specifies location (μ), scale (σ) and shape (ξ) parameters of GEV, w represents the proportion of zeroes responsible for data inflation, D denotes the variables for zero-inflated distribution, z_i is MMS, n is the number of MMS, m is the observed zeroes in MMS, S is the number of points inflating with zeroes, the hyper-parameters $\sigma_{0\mu}^2$ and $\sigma_{0\xi}^2$ are the prior variance for μ and ξ, a_0 and b_0 are hyper-parameters of Gamma prior distribution for σ, c and d are the hyper-parameters of Beta prior distribution for w. On the right-hand side of Eq. (3.1), the first two lines and the last line denote the likelihood and the prior distribution, respectively (Quadros Gramosa et al., 2019).

Based on the estimated parameters from Eq. (3.1), the annual return levels can be obtained by the following equation (Coles et al., 2001):

$$P_a = \left(\frac{P_m - w}{1-w}\right)^{1/12} \tag{3.2}$$

$$Z_{p_a} = \begin{cases}\mu + \frac{\sigma}{\xi}\left(\left(-log(p_a)\right)^{-\xi} - 1\right), \xi \neq 0 \\ \mu - \sigma log\left(-log(p_a)\right), \xi = 0\end{cases} \tag{3.3}$$

where p_m and p_a denote the monthly and annual non-exceedance probability, respectively, Z_{p_a} is

the rainfall with return period of $\frac{1}{1-P_a}$ years.

Apart from this approach for EVA (i.e., MMS combined with ZIGEV), we also employed the traditional approach (i.e., AMS combined with GEV) for comparison purposes (Section 3.4.1). Please note that ZIGEV was not applicable for some grid cells (<5%) because of some outliers in the datasets of MMS. Therefore, for these locations, we computed the return rainfall based on GEV through the L-moment method that usually performs well in the presence of outliers (Poschlod et al., 2021).

3.3.3. IDF curve development

IDF curves of the duration ranging from 1-hour to 48-hour were developed based on the estimated rainfall quantiles (with different durations) through ZIGEV. For example, to construct 2-year IDF curves, 50% of non-exceedance return rainfall of 1, …, 48h based on ZIGEV were estimated. For the historical period (i.e., 1980-2005), IDF curves were developed based on the 1, …, 48h MMS provided by ERA5. However, owing to the spatial resolution of ERA5, these IDF curves depict the extreme rainfall information on a large region (0.25° × 0.25°) (hereinafter, area IDF curves), which is usually not sufficient for the urban drainage design (Willems, 2013). Therefore, the areal reduction factor (ARF) was used to transform the area IDF curves to point IDF curves (or sometimes called gauge-based IDF curves) (Willems, 2013). Owing to the difficulty in estimating ARF, arising from a lack of gauge-based sub-daily rainfall datasets in Cambodia, this transformation was only performed in Phnom Penh based on the gauge-based datasets available at https://www.wunderground.com/dashboard/pws/IPHNOMPE8.

For the mid- and late-century (i.e., 2040-2065 and 2070-2095, respectively), climate change signals (future return rainfall/past return rainfall) based on CORDEX were computed following the approach explained in Section 3.3.2 and subsequently multiplied by the ERA5-based area IDF curves to get the future area IDF curves. Afterward, ARF, which is assumed to remain the same in the future, was applied to obtain future point IDF curves in Phnom Penh.

3.3.4 Warming-induced flood risk change index

Flood risk, i.e., the potential for the loss associated with a flood event, usually derives from the

interaction between hazard, exposure, and vulnerability (Kundzewicz et al., 2014; Shrestha et al., 2013; Zscheischler et al., 2018, among others).

In this chapter, instead of focusing on the flood risk itself, we developed a warming-induced flood risk change index as below,

$$Flood\ risk\ change_{warming-induced} = (Change\ in\ Hazard) \times Exposure \times Vulnerability \quad (3.4)$$

where, $Change\ in\ Hazard$ denotes the warming-induced change in the occurrence of flood events (mainly caused by heavy rainfall). In this chapter, the change in heavy rainfall obtained by the extreme value analysis was used as a proxy for the change in the occurrence of flood events, although multiple factors (e.g., antecedent soil moisture) also affect the flood generation. To account for the major components of $Exposure$, we considered the province-level population, economy, and agriculture. $Vulnerability$ is defined as the historical damage and loss in the affected provinces (which, to some extent, signify the inability to withstand the flood events), as suggested in Shrestha et al. (2013). $Exposure$ and $Vulnerability$ are assumed to be unchanged in the future due to the lack of detailed information on the population and development policy in Cambodia. Since agriculture is primarily affected by river flooding and massive inundation due to long-lasting (one day or longer) heavy rainfall, two different types of flood risk change index are quantified based on the $Change\ in\ Hazard$. Considering the heavy rainfalls of two different durations, i.e., the duration longer than one day, and less than or equal to one day, and the agriculture's exposure and vulnerability only for the long-lasting rainfalls, Eq. (3.4) takes the form,

$$Flood\ risk\ change_{warming-induced} = \begin{cases} (Change\ in\ heavy\ rainfall) \times (Population + Economy + Agriculture) \times (Damage), while\ \ rainfall\ duration > one\ day \\ (Change\ in\ heavy\ rainfall) \times (Population + Economy) \times (Damage)_{without\ agriculture}, while\ rainfall\ duration < one\ day \end{cases} \quad (3.5)$$

where $Population$ denotes the population of the respective province (NIS, 2019), $Economy$ is the relative number (ratio of each province to country) of economic units (NIS, 2012), which refers to all the economic activities, excluding the agriculture, forestry, fishing and public service, that are conducted in different provinces of Cambodia. $Agriculture$ stands for the total area used for the cropping of rice (the staple food in Cambodia) in the respective province (BDLINK, 2017). $Change\ in\ heavy\ rainfall$ in short and long durations is the change of return rainfall with the

duration of 1-24h and the duration of 2, 5, 10, and 15 days, respectively. $Damage$ refers to the total number of affected people, destroyed houses, affected paddy fields, and submerged roads due to flood events during 1996-2021 (available at http://camdi.ncdm.gov.kh/DesInventar/definestats.jsp). However, this information does not separate the damage resulting from different flood events, thus we use the same information on damage for both types of duration (short and long). Under the limited data availability, this approach would be acceptable considering that these two kinds of heavy rainfall sometimes occur simultaneously (e.g., floods in October 2020, Cambodia). To remove the influence of the areal size of different provinces (a province with the larger area tends to show more population, assets, and damages), each element included in $Exposure$ and $Vulnerability$ in Eq. (3.5) was divided by the total area of individual province. Moreover, these elements were transformed to the range of 0 (least) to 1 (most) by min-max rescaling, an approach also used in other flood risk assessments (e.g., Mondal et al., 2020; Shrestha et al., 2013). The detailed information of items considered in *Exposure* and $Vulnerability$ is summarized in Table 3.3.

In the calculation of the $Flood\ risk\ change_{warming-induced}$, for each type of rainfall durations, the occurrence probability of the rainfall events was also incorporated in $Change\ in\ heavy\ rainfall$ based on the discrete equation as below (Chow et al., 1988),

$$Change\ in\ heavy\ rainfall = \left\{\sum_{i=1}^{\infty}\left[\frac{Change_{T_i} + Change_{T_{i+1}}}{2}\right][P(T \geq T_i) - P(T \geq T_{i+1})]\right\} \div P(T \geq T_1) \quad (3.6)$$

where $Change_{T_i}$ denotes the change in return rainfall for the return period T_i, $P(T \geq T_i)$ is the exceedance probability of return rainfall (= $1/T_i$), and the subscript i (=1, 2, 3, 4,…) specifies the return period (5-year, 10-year, 25-year, 50-year,…, respectively). In Eq. (3.6), we assume that the change in return rainfall for T_i larger than 50-year is the same as that for T_i = 50-year. Furthermore, the change in return rainfall corresponding to T_i smaller than 5-year was omitted from Eq. (3.6) given the negligible damage caused by the return rainfall with a smaller than 5-year return period. For example, in the mid-century, the 5-year, 10-year, 25-year, and 50-year return rainfall change in Phnom Penh, averaged over 1h to 24h, are 40%, 50%, 60%, and 70%, respectively. Subsequently, the

$Change\ in\ heavy\ rainfall$ would be 45.5% ($\left(\left[\frac{40\%+50\%}{2}\right]\left(\frac{1}{5}-\frac{1}{10}\right)+\left[\frac{50\%+60\%}{2}\right]\left(\frac{1}{10}-\frac{1}{25}\right)+\left[\frac{60\%+70\%}{2}\right]\left(\frac{1}{25}-\frac{1}{50}\right)+70\%\times\frac{1}{50}\right)\div\frac{1}{5}=45.5\%$). At last, we quantified the relative contribution of each factor to the total flood risk change $\left(\frac{flood\ risk\ change\ \ from\ each\ factor}{total\ flood\ risk\ change}\right)$, where each factor was normalized by min-max scaling and $total\ flood\ risk\ change$ was computed by summarizing flood risk change from each factor. The detailed information of $Change\ in\ heavy\ rainfall$ is summarized in Table 3.4.

Table 3.3. Summary data of each element considered in $Exposure$ and $Vulnerability$ for each province of Cambodia (including the capital city). Tbong Khmum province, a newly formed province after 2013, was considered part of Kampong Cham province.

Region	Province/ Metropolis (number)	Area (km2)	Exposure			Vulnerability			
			Population (persons)	Economy (%)	Agriculture (hectares)	People-affected (persons)	Paddy field-affected (hectares)	House-destroyed	Road-submerged (m)
Southeast	Svay Rieng (1)	2964	524554	3.1	162989	41	84168	76	21192
	Prey Veng (2)	4873	1057428	6	309570	490	143689	11	28818
	Kampot (3)	4859	592845	3.6	240140	23	113383	75	42030
	Kandal (4)	3180	1195547	8	127240	134	29487	917	31104
	Takeo (5)	3569	899485	6.1	240224	72	9295	24	24280
	Kampong Cham Tbong Khmum (6)	9659	1671059	11	239609	809	40380	247	118951
	Phnom Penh (7)	679	2129371	19	12517	5	1857	0	2400
	Kep (8)	337	41798	0.3	3109	0	562	0	6780
Northwest	Banteay Meanchey (9)	6663	859545	4.2	307176	60	202011	0	228012

	Battambang (10)	11755	987400	6.4	513863	77	157970	98	157563
	Siem Reap (11)	10271	1006512	5.1	424458	337	39610	143	78094
	Pailin (12)	804	71600	0.7	10703	10	923	248	23545
Other region	Otdar Meanchey (13)	6220	261252	1.1	194889	5	14700	153	12877
	Kampong Chhnang (14)	5536	525932	4	155032	32	4260	40	18500
	Kampong Speu (15)	7034	872219	4.4	127380	7	32394	37	3800
	Kampong Thom (16)	13822	677260	4.4	305310	86	44651	53	63110
	Preah Sihanouk (17)	1942	302887	2.4	22278	11	73	0	1580
	Kratie (18)	10965	372825	2.4	59472	97	31974	46	110949
	Pursat (19)	12867	411759	2.7	147207	24	36835	79	12000
	Preah Vihear (20)	13964	251352	1.3	95833	14	29793	45	28215
	Stung Treng (21)	11398	159565	1.1	46923	15	50248	18	104655
	Ratanak Kiri (22)	10738	204027	1.2	34346	9	25660	29	17705
	Koh Kong (23)	10302	123618	1.1	12574	1	1519	47	170

	Mondul Kiri (24)	14775	88649	0.4	28150	1	2558	15	0

Table 3.4. Same as Table 3.3, but for $Change\ in\ heavy\ rainfall$ based on Eq. (3.6).

Region	Province/ Metropolis (number)	Mid-century (short-duration) (%)	Late-century (short-duration) (%)	Mid-century (long-lasting) (%)	Mid-century (long-lasting) (%)
Southeast	Svay Rieng (1)	74	143	38	62
	Prey Veng (2)	56	91	32	52
	Kampot (3)	79	176	11	58
	Kandal (4)	54	106	29	67
	Takeo (5)	61	120	22	66
	Kampong Cham Tbong Khmum (6)	57	102	29	58
	Phnom Penh (7)	52	110	25	72
	Kep (8)	64	167	6	46
Northwest	Banteay Meanchey (9)	54	84	29	45
	Battambang (10)	56	94	26	43
	Siem Reap (11)	51	57	29	30
	Pailin (12)	40	75	18	34
Other region	Otdar Meanchey (13)	78	78	34	40
	Kampong Chhnang (14)	63	127	28	69
	Kampong Speu (15)	61	128	21	71

	Kampong Thom (16)	47	98	27	49
	Preah Sihanouk (17)	56	156	7	37
	Kratie (18)	42	72	33	38
	Pursat (19)	38	82	16	42
	Preah Vihear (20)	42	64	31	37
	Stung Treng (21)	37	68	29	25
	Ratanak Kiri (22)	51	60	36	24
	Koh Kong (23)	58	115	9	41
	Mondul Kiri (24)	37	65	29	33

3.4. Results and discussion

3.4.1 Performance of CANN and ZIGEV

The CANN model was trained for each grid cell based on the disaggregated and observed (ERA5-based) 1h MMS for the period between January 1980 and December 1997 and validated for the period between January 1998 and December 2004 (Fig. 3.4). The results showed that the proposed CANN model performed well for training runs (RMSE = 0.97 mm, R = 0.95) and test runs (RMSE = 1.1 mm, R = 0.93), when the results for each grid cell were averaged over Cambodia (Fig. 3.4a). However, the extreme spatiotemporal variability of rainfall in the mountainous regions has compromised the CANN performance (RMSE = 1.4 mm, R = 0.86 for the test runs) in the region encompassing site A (Fig. 3.4a), which is in line with the observations by Nourani and Farboudfam, (2019). These observations are also visible in the time series of disaggregated and observed 1h MMS (left panel of Fig. 3.4b). Although the CANN model generally reproduced the MMS well, it tends to underestimate some large values in the MMS, which is evident in the extracted AMS (right panel of Fig. 3.4b). This underestimation, especially for the larger extreme rainfalls, is a common problem in temporal disaggregation models (Kossieris et al., 2018; Mirhosseini et al., 2014, among others), which motivated us to apply climate change signals (future return rainfall/past return rainfall) (Section 3.3.3).

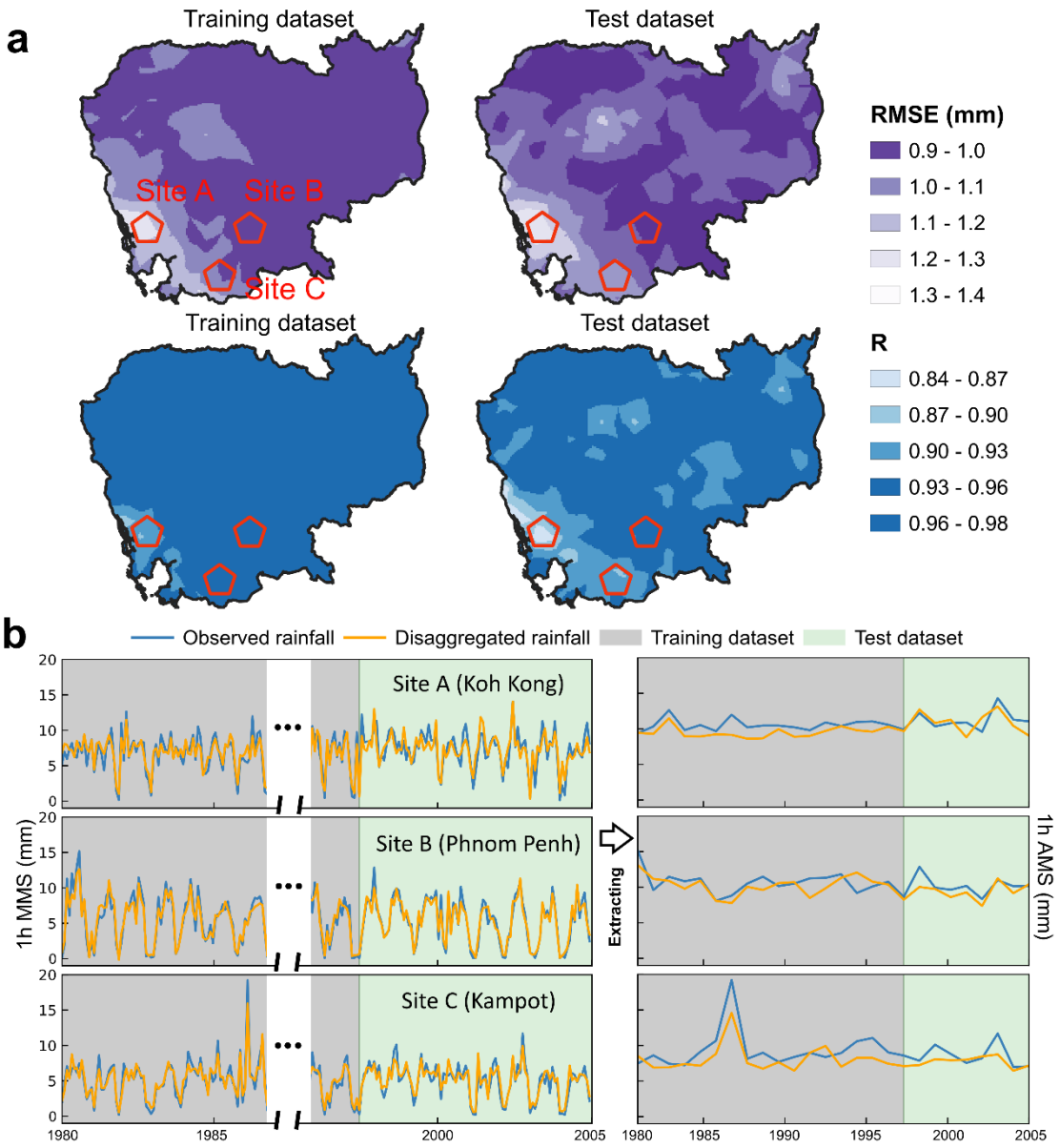


Fig. 3.4. Evaluation of CANN model based on training and test datasets provided by ERA5 for the entire country (a) and three selected locations (b), i.e., Site A: grid (lon=103.35, lat=11.57) in Koh Kong, Site B: grid (lon=104.92, lat=11.56) in Phnom Penh, and Site C: grid (lon=104.46, lat=10.84) in Kampot. MMS and AMS in (b) denote the monthly and annual maximum series, respectively. For better visualization, symbols denoting the three sites in (a) have been enlarged and the period from 1987 to 1996 is not shown in the left panel of Fig. 3.4(b).

The main reason for the underestimation is that the CANN model is specifically designed to

simulate the rainfall patterns (Section 3.3.1), and subsequently, to reproduce the rainfall. However, the rainfall patterns of extremely large rainfall (e.g., upper 80th percentile of the MMS) are sometimes different from those of normal rainfall (e.g., lower 20th percentile of the MMS) (see Fowler et al., 2021). Therefore, the CANN model trained by all those rainfalls may not reproduce the extreme rainfall well, which is a drawback common to other data-driven models. Therefore, to improve the accuracy, further efforts can focus on reproducing the AMS directly instead of MMS that comprises some zero-rainfall, when longer observed rainfall datasets and the concurrent atmospheric variables are available. As an alternate solution, instead of reproducing the Block Maxima (i.e., MMS and AMS), the current CANN model can be used to reproduce the 'Peak-Over-Threshold', where thresholds can be chosen to avoid simulating multiple rainfall patterns of these peaks (i.e., extremes) at the same time. Subsequently, these peaks can be fitted by the generalized Pareto distribution to estimate the return rainfall (Coles et al., 2001).

To apply ZIGEV for EVA at each grid cell, our approach to estimating return rainfall based on observed MMS through ZIGEV was compared with the return rainfall based on a standard approach, i.e., observed AMS through GEV in Phnom Penh (Table 3.5). The estimated return rainfall (mean value) by ZIGEV was slightly larger (6.2%-8.9%) than the return rainfall by GEV, which may be caused by different assumptions used in these two distributions (Quadros Gramosa et al., 2019). Further comparison of the return rainfall bias (difference between disaggregated and observed return rainfall) by ZIGEV and GEV is illustrated based on the outputs from CANN models (Fig. 3.5). ZIGEV was generally found to outperform the GEV, especially in 5- and 10-year return rainfall. Specifically, considering the absolute bias averaged over Cambodia, the application of ZIGEV decreased this bias of GEV by a minimum of 0.5% (5.6% for GEV and 5.1% for ZIGEV) for 25-year return rainfall and a maximum of 1.7% (5.1% for GEV and 3.4% for ZIGEV) for the 5-year one. This result demonstrates the overall better performance in EVA using CANN combined with ZIGEV. However, as the CANN models introduced in this chapter cannot provide a clear understanding of the causal relationships between input and output variables, a careful validation should be performed, especially for studies focusing on new areas

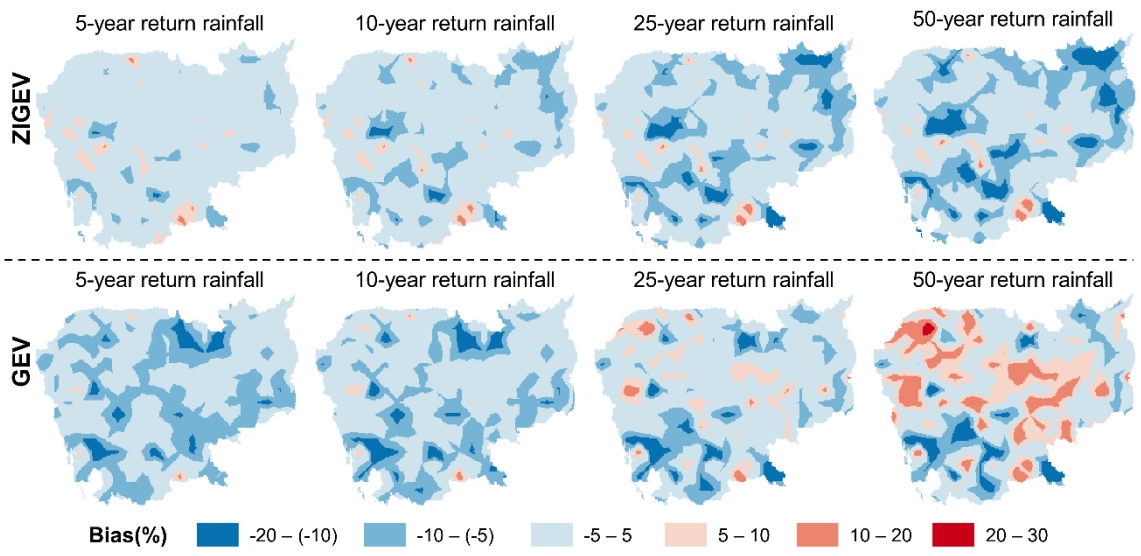


Fig. 3.5. Comparison between disaggregated (CANN) and observed (ERA5) hourly return rainfall based on GEV and ZIGEV, respectively, in the historical period (Bias = 100×(disaggregated - observed)/observed). Return rainfall computed by ZIGEV (or GEV) was based on MMS (or AMS extracted from this MMS). Please note that MMS and AMS were obtained by inputting 3h time-series datasets into CANN (Section 3.3.1)

Table 3.5. Comparison between return rainfall (mm) based on ZIGEV and that based on GEV in Phnom Penh during 1980-2005. GEV and ZIGEV parameters were computed by the maximum likelihood and Bayesian estimation, respectively. Lower and upper limits of 95% confidence intervals (CI) correspond to 2.5% and 97.5% of posterior distributions of the ZIGEV parameters, respectively.

Return period	**GEV (mean)**	**ZIGEV**		
		Lower bound (95% CI)	**Mean**	**Upper bound (95% CI)**
5-year	11.63	10.23	12.67	15.77
10-year	12.44	10.78	13.54	17.18

15-year	12.88	11.03	13.96	17.91
25-year	13.41	11.30	14.43	18.75
50-year	14.09	11.59	14.97	19.77

3.4.2 Change in return rainfall and underlying physical mechanism

Applying ZIGEV to the disaggregated 1h MMS and the existing MMS for the duration from 3h to 15 days based on three CORDEX simulation outputs (historical, mid-century, and late-century), we obtained the spatial distribution of relative change in extreme rainfalls for different durations and return periods (Figs. 3.6-3.9). An apparent increase in the return rainfall is observed over Cambodia in the future, where more increase is likely to occur in late-century (82.6%) than mid-century (48.5%), as compared to the historical period when averaged over Cambodia for four different durations (1, 3, 12, 24h) and all return periods (Figs. 3.6 and 3.7). This result is consistent with the larger increase in atmospheric water vapor content in the late-century than the mid-century (Fig. 3.11) due to the high-emission scenario (RCP 8.5) that projects a comparatively higher temperature rise in Cambodia (3.4 ℃ in the late-century and 1.8 ℃ in the mid-century). According to a simple idea that the extreme rainfall intensity is expected to increase at a rate proportional to the warming-induced moisture increase (i.e., CC scaling, amounting to approximately 7% per degree warming) (see Pendergrass, 2018), extreme rainfall intensity is expected to increase by 24% ≈ 3.4 ℃ × 7%/℃ in the late-century, which is much smaller than the values shown in Fig. 3.6. This difference can most likely be explained by two reasons. Firstly, 'extreme rainfall' in the study of CC scaling usually refers to the rainfall with a given percentile (e.g., 99%) of the annual rainfall series (including the normal rainfall) (see Lenderink et al., 2017), which is different from the 'extreme rainfall' (different percentiles of MMS or AMS) defined in Eqs. (3.2) and (3.3). Secondly, a physical argument implies that the more extreme rainfall (which is less frequent) is likely to increase more under global warming owing to the change in atmospheric circulation (Pendergrass, 2018). A similar result can also be seen in Fig. 3.12, where the stronger wind convergence (4.9×10^{-5}/s for the past, 6.1×10^{-5}/s for the mid-century, and 6.2×10^{-5}/s for the late century) brings more precipitable water to the heavy rainfall region (i.e., the southern part of Koh Kong province), causing a super-CC scaling for heavy rainfall events in the future. Some other factors (e.g.,

interplay between thermodynamic and dynamic factors, see Fowler et al., 2021), need a further investigation in the future work.

25-year return rainfall is found to increase by 69.7%, while 10-year one is around 61.5% (averaged over Cambodia for four different durations in both mid- and late-century) (Fig. 3.6), which indicates that rainfall with larger return periods is likely to intensify more in the future. This result is in good agreement with the previous studies focusing on the same or nearby regions (Ge et al., 2019; Truong Ha, 2018) and also following well the abovementioned physical mechanism (Pendergrass, 2018). Apart from this difference in rainfall with different return periods, the rainfall intensification is likely to abate with the increasing duration (Fig. 3.6). When averaged over Cambodia for different return periods (10-year and 25-year) during mid- and late-century, the relative changes are 76.2%, 72.35%, 64.5%, and 49.1% for 1h, 12h, 24h, and 48h, respectively, which indicates that rainfall with shorter duration is likely to intensify more in the future. The possible underlying reasoning may be the feedbacks in cloud-core updrafts that could cause more increased extreme rainfall in shorter durations (Lenderink et al., 2017; Loriaux et al., 2013). Moreover, the convective heavy rainfall is likely to be more frequent and intense in the future over the most area of Cambodia (Fig. 3.13) (albeit with a little decrease in the coastal areas), which also contributes to a higher increase in the shorter duration, especially sub-daily, extreme rainfall. However, there is a lack of broad consensus that extreme rainfall with a shorter duration will intensify more globally, which is due to different climate models and methods applied in different climate zones.

Spatially, it is evident that more increase is estimated to happen in Southern and Central Cambodia, mainly in the late century (Figs. 3.6-3.9). This spatial pattern can be explained by the synergy of prevailing westerly monsoon winds (Fig. 3.11) and the mountains in southwestern Cambodia (Fig. 3.1), bringing more water vapors from the sea to coastal areas and further to Southern and Central parts. In addition, some regions in Southern Cambodia are projected to experience a distinctively large intensification (over 200%) of extreme rainfall (Fig. 3.6), because of the combined effect of moisture increase in this region and a stronger (future > past) wind convergence (during the heavy rainfall events) which results in more water vapor transport from the sea to heavy rainfall areas (Fig. 3.12).

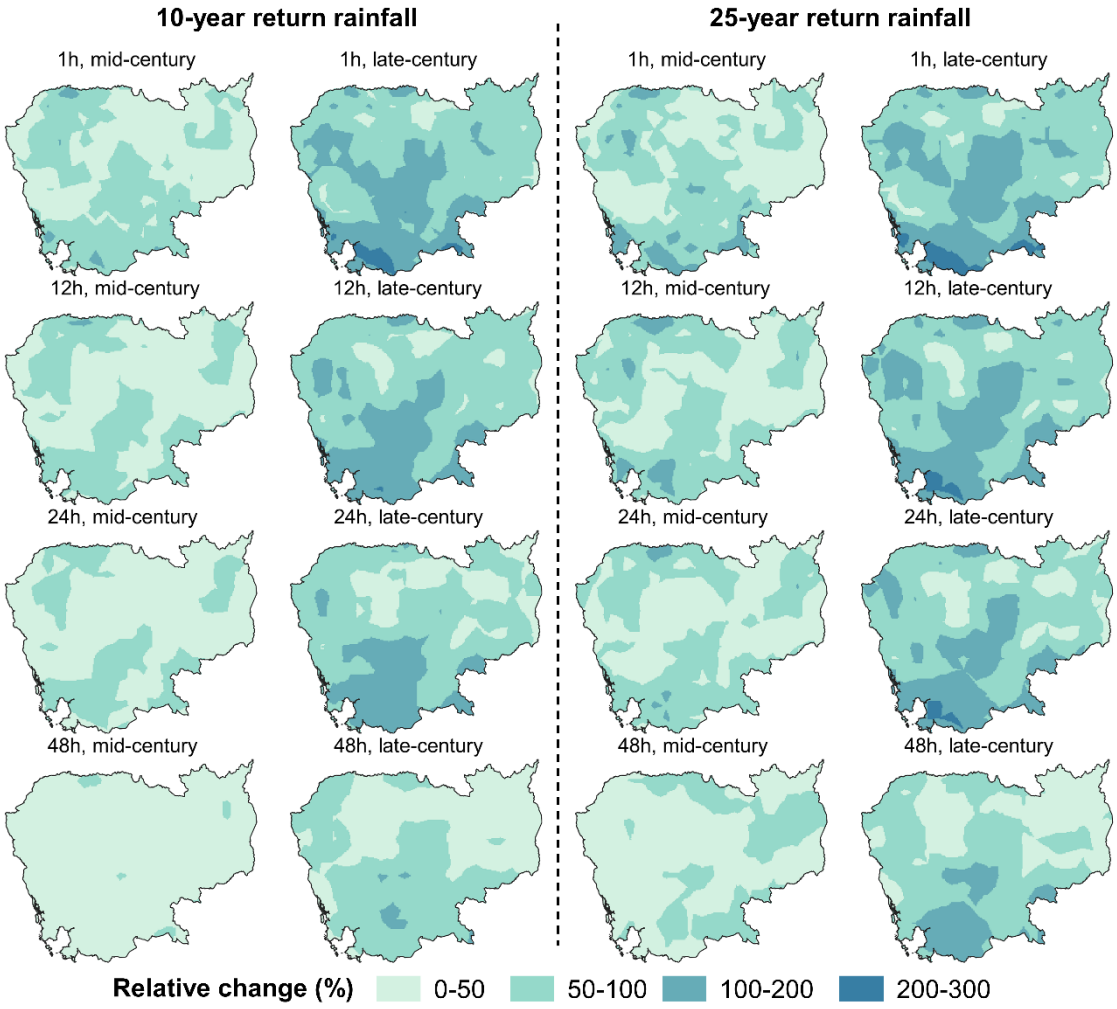


Fig. 3.6. Relative change of 1, 12, 24, and 48h return rainfall from the historical period to the mid- and late-century corresponding to the 10-year (left) and 25-year (right) return periods based on the ensemble mean of EVA results obtained from three CORDEX simulation outputs (Table 3.2). The historical return rainfall is shown in Fig. 3.10.

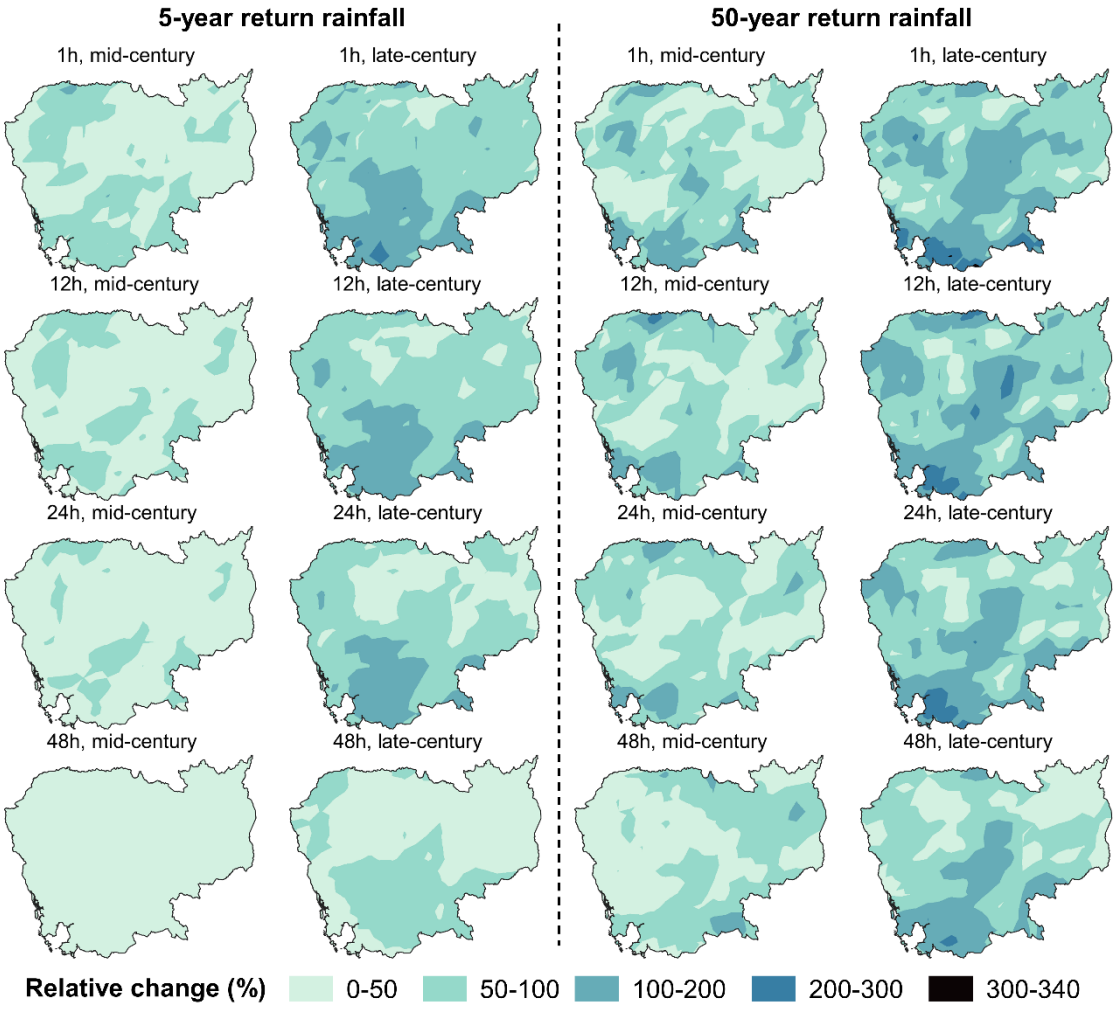


Fig. 3.7. Same as Fig. 3.6 but for 5-year (left) and 50-year (right) return periods.

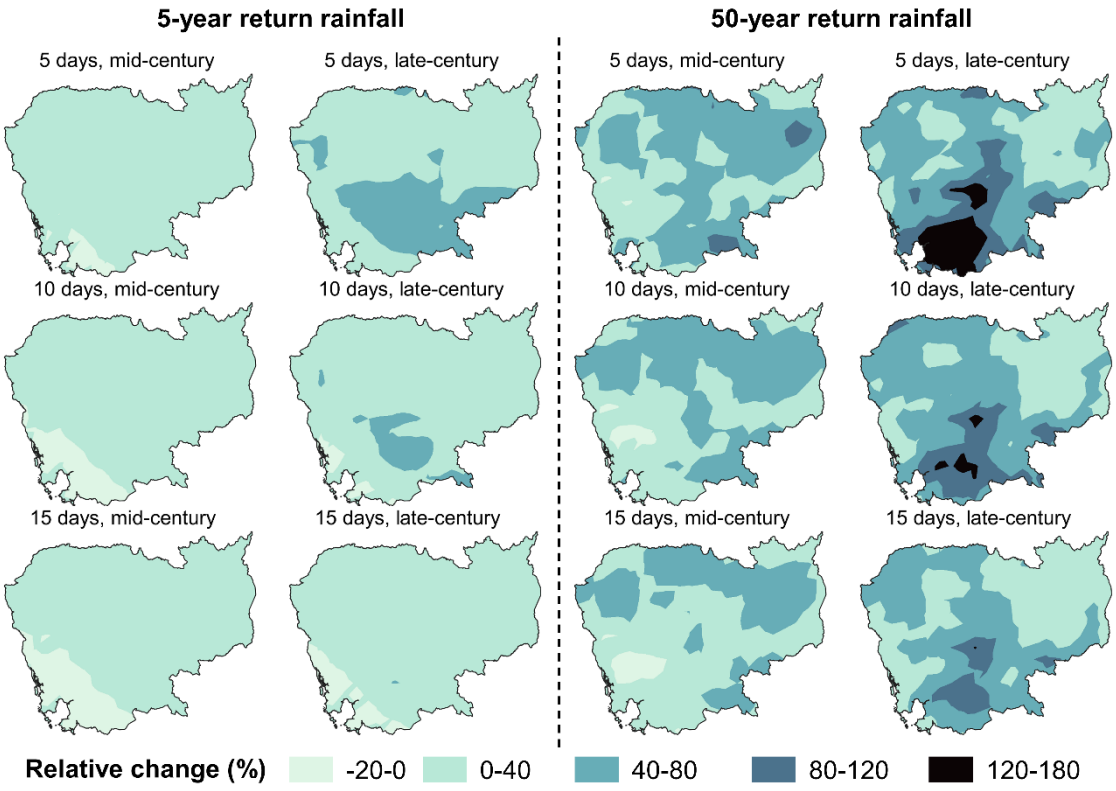


Fig. 3.8. Same as Fig.3.7 but for 5-, 10- and 15-day return rainfall. Please note that color scheme has been changed for better visualization.

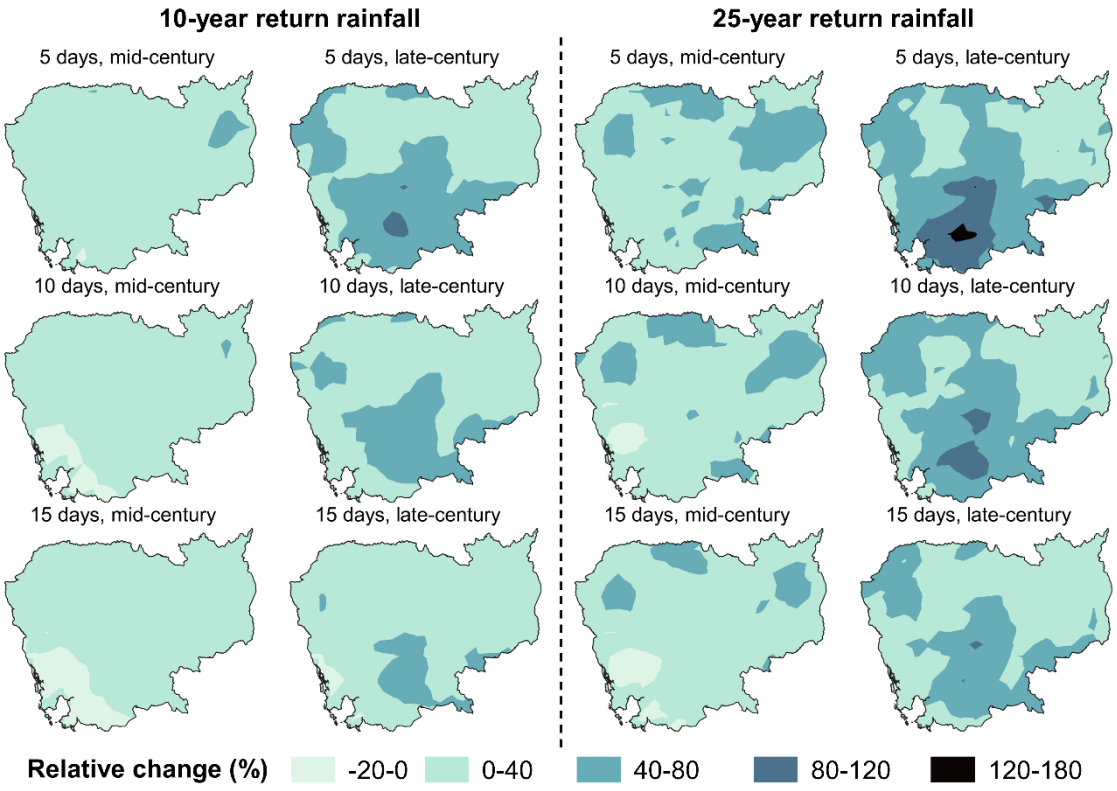


Fig. 3.9. Same as Fig. 3.8, but for 10- and 25-year return periods.

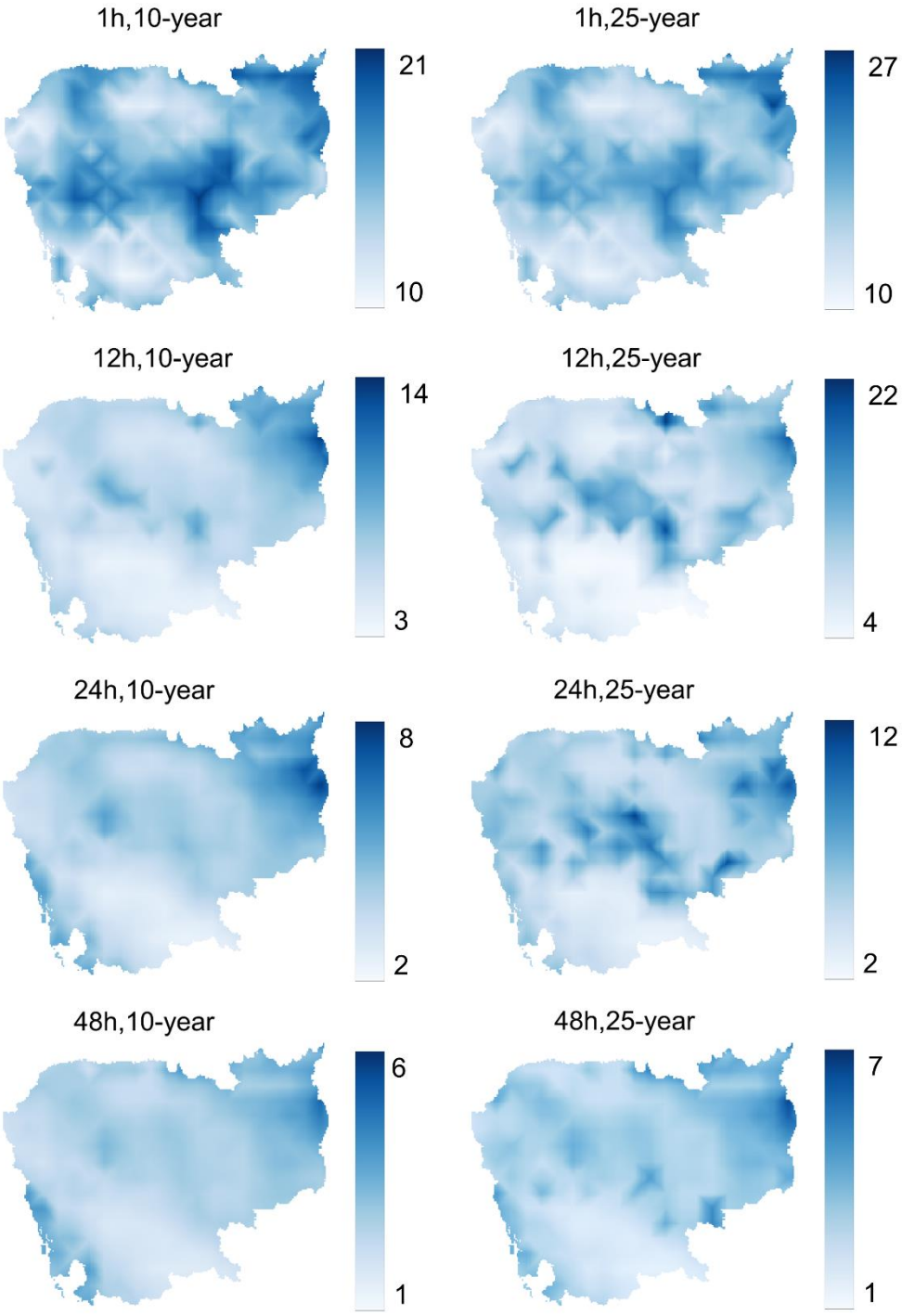


Fig. 3.10. 1, 12, 24, and 48h return rainfall intensity (mm/h) (corresponding to 10-and 25-year return period) in the historical period based on observed (ERA5) datasets by ZIGEV.

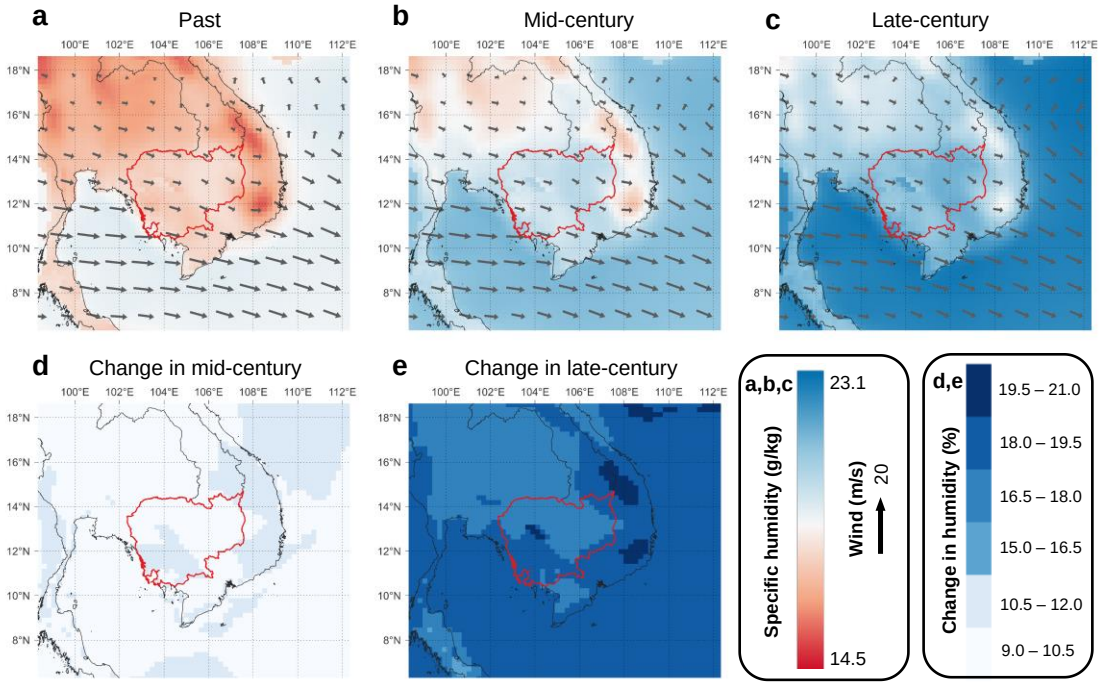


Fig. 3.11. Wind and specific humidity at 1000 hPa over Cambodia in the rainy season (May to Oct) during past (a), mid-century (b), and late-century (c), and change of specific humidity from the past to mid- (d) and late-century (e). For each period, monthly average values during the rainy season are used.

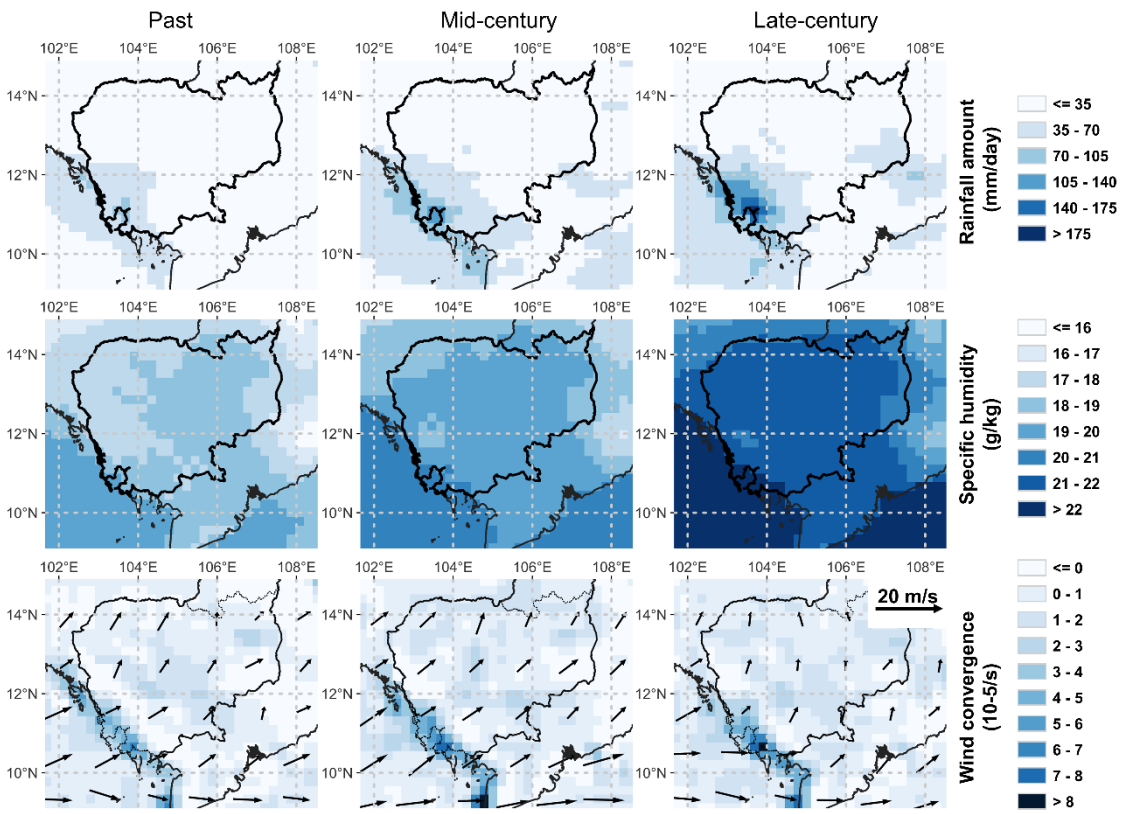


Fig. 3.12. Heavy rainfall and the concurrent weather conditions (i.e., specific humidity, wind, and its convergence) in the historical, mid- and late-century periods. In each period, the amount of rainfall and concurrent weather conditions were averaged for 26 events (one event per year) included in the annual maximum series. Please note that the convergence in the southern part of Koh Kong province (lat=11.00, lon=103.75; Fig. 3.1) was 4.9×10^{-5}/s for the past, 6.1×10^{-5}/s for the mid-century, and 6.2×10^{-5}/s for the late century.

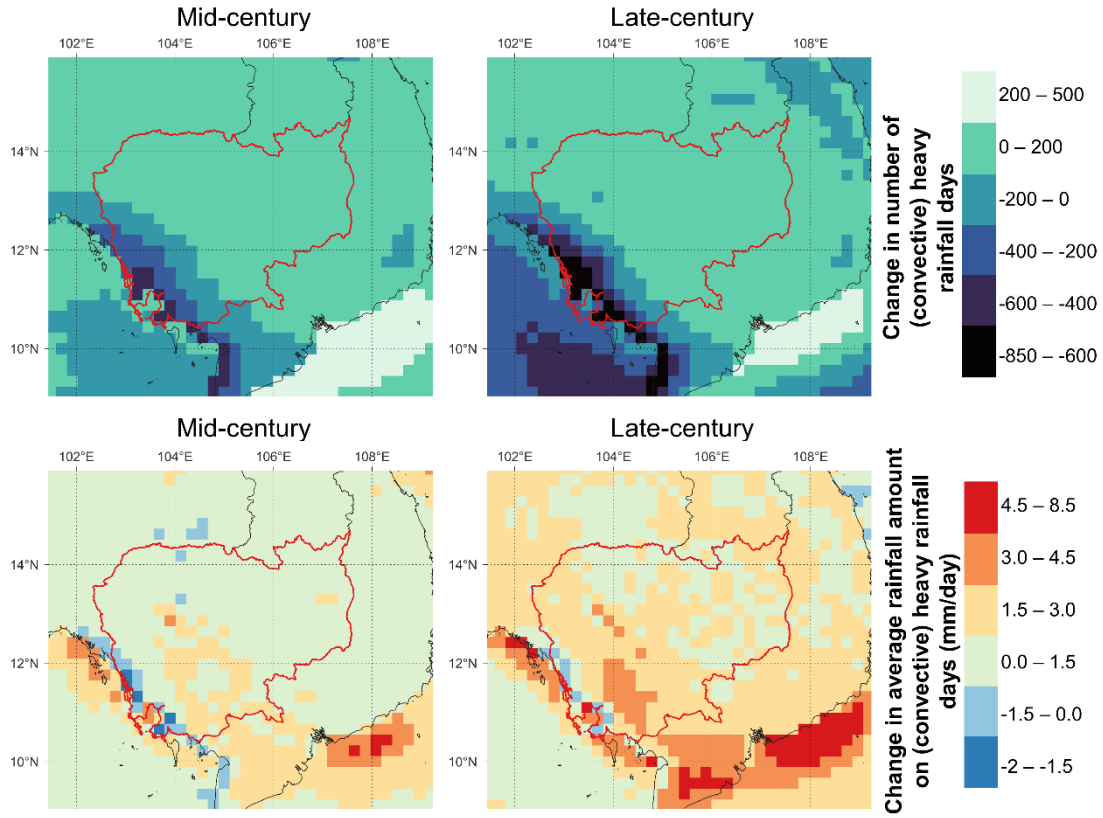


Fig.3.13. Change in the number of convective heavy rainfall days and the average rainfall amount on these days from the past to mid- and late- century. Heavy rainfall days denote the days when the daily convective rainfall amount is larger than 20 mm.

Based on the projected heavy rainfall change, the historical and future IDF curves (1-48h) in Phnom Penh city were developed (Fig. 3.14), revealing an apparent upshift of the IDF curves in the future periods compared with the curves in the historical period. Rainfall intensity with a return period of 5 years is estimated to increase by 26% (48h) - 50% (1h) in the mid-century and 75% (48h) - 128% (1h) in the late-century. Specifically, the 5-year 1h rainfall intensity is likely to increase from 110.5 mm/h in the past to 252.2 mm/h in the late-century, which would place high pressure on the drainage system, whose capability is already insufficient. Apart from the change in rainfall intensity, the occurrence probability of extreme rainfall is also estimated to become approximately 10-folds in the mid-century. For example, rainfall intensity from 50-year IDF curves in the historical period is similar to that from 5-year IDF curves in mid-century, meaning that a 50-year 48h extreme rainfall event in the past would likely be a 5-year 48h event in the mid-century.

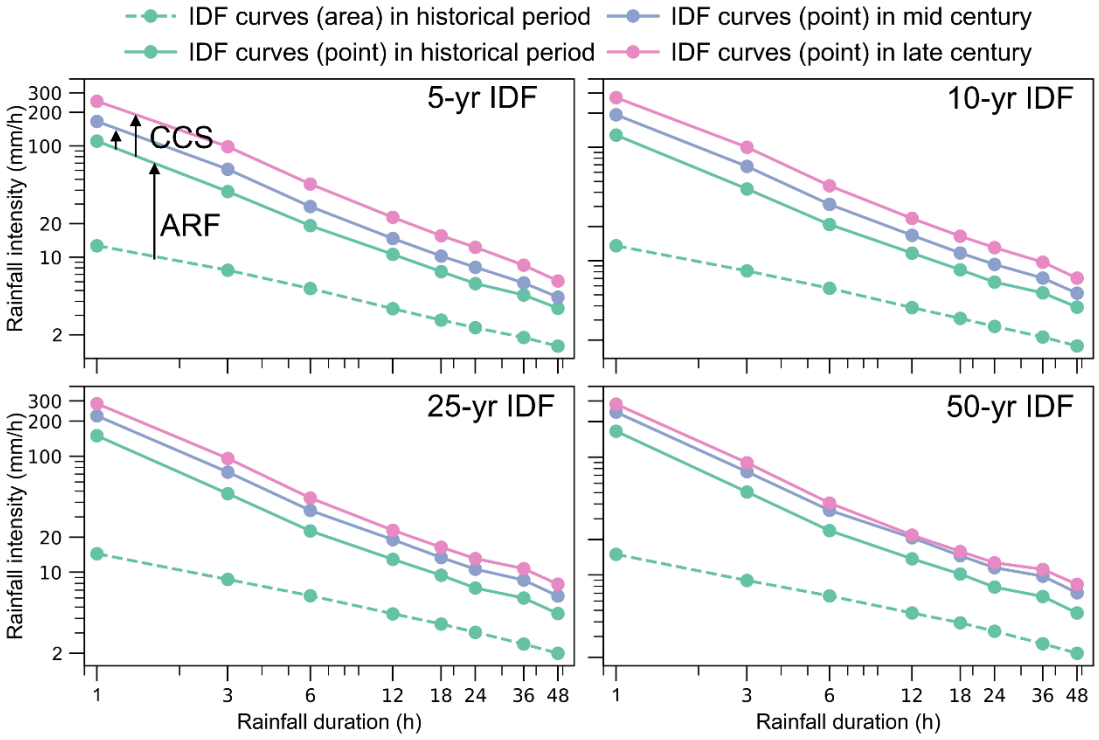


Fig. 3.14. IDF curves (1-48h) in Phnom Penh city in the historical, mid- and late-century period. CCS and ARF denote the climate change signals (i.e., rainfall intensity change shown in Figs. 3.6 and 3.7), and areal reduction factor. IDF curves (point) and (area) denote the IDF curves established based on point (gauge) datasets and area (ERA5) grid datasets, respectively.

3.4.3 Flood risk change map and suggestions for sustainable adaptation strategies

The projected increase in rainfall intensity is likely to exacerbate flooding over Cambodia at different spatiotemporal scales, while the influence of river flows from upstream regions of the Mekong River (i.e., China, Laos, and Thailand; see Fig. 3.1) and warming-induced sea level rise are the additional causes of flood hazard. For example, the more intensive sub-daily rainfall will trigger localized surface (pluvial) flooding, particularly in urban areas with rapidly increasing imperviousness. In the rainy season, the longer duration (2-day or longer) heavy rainfall in Cambodia may result in a prolonged inundation, especially in low-lying areas, and the rise in water level of rivers causing the

overflow in nearby land areas (i.e., fluvial flooding). Several suggestions made to reduce these flood risks are discussed as follows based on the flood risk change maps, embodying the information of the spatial distribution of risk changes in different provinces of Cambodia, developed in this chapter.

The maps of flood risk change caused by sub-daily heavy rainfall (Fig. 3.15) show that the Southeastern and Northwestern Cambodia are projected to encounter a larger risk change in the mid- and late-century than the other regions. Based on Figs. 3.15a and b, Phnom Penh city (labeled with number '7') and Kandal province ('4') (urban areas surrounding the capital city) are likely to face the most serious flash flooding damage, which is explained by manifold reasons; 1) the relatively larger increase of sub-daily heavy rainfall (Figs. 3.6 and 3.7), 2) the remarkably huge exposure of population and economic assets (especially in Phnom Penh city, mainly ascribed to the rapid urbanization) (Fig.3.16a), 3) the substantial flood damages in Kandal province (Fig. 3.16b), and 4) their contributions to the total risk change (over 60% attributed to 'Exposure' in Phnom Penh, and over 60% to 'Vulnerability' in Kandal province, see Fig. 3.15c). Considering that 'People' and 'Road' were mainly affected in Phnom Penh ('7' in Fig. 3.16d), two countermeasures, including improving early warning systems to avoid casualties or injuries and upgrading the drainage systems to reduce flood impacts on traffic system in the city, are suggested. To this end, the projected IDF curves in Phnom Penh (Fig. 3.14) can be applied by the governing agencies in Phnom Penh. For Kandal province, as 'House' was mostly affected ('4' in Fig. 3.16d), optimizing the urban development plan by determining the base elevation and guiding the location of newly-built buildings can be considered. For other regions in Southeastern and Northwestern Cambodia, those countermeasures can also be selected depending on the degree of risk change shown in Fig. 3.15 and the major contributing components in Fig. 3.16.

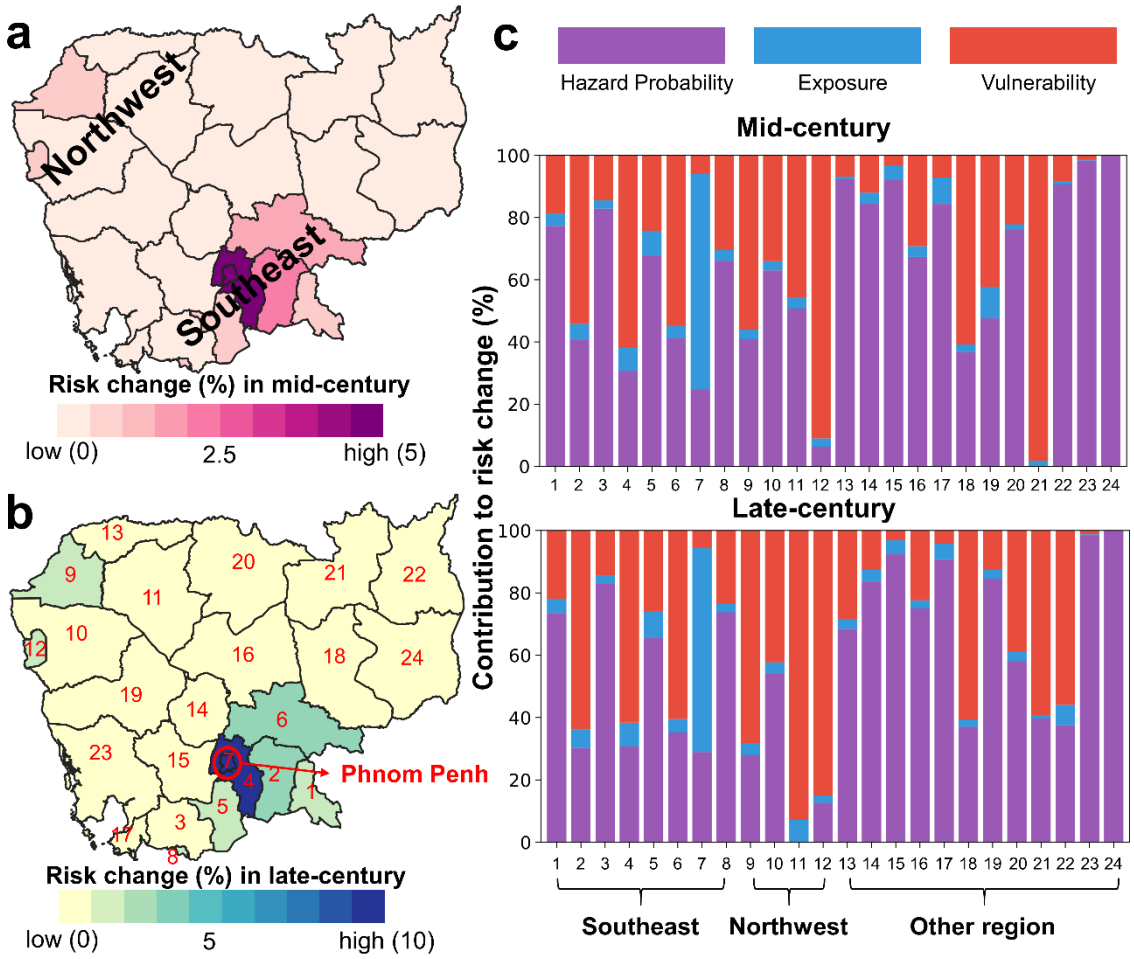


Fig. 3.15. Province-level (and the capital city, Phnom Penh) flood risk change caused by short-duration (1-24h) heavy rainfall in the mid-century (a), late-century (b), and contribution of various parameters to the total risk (c). The location and information of each province are shown in Fig. 3.1 and Tables 3.3 and 3.4, respectively.

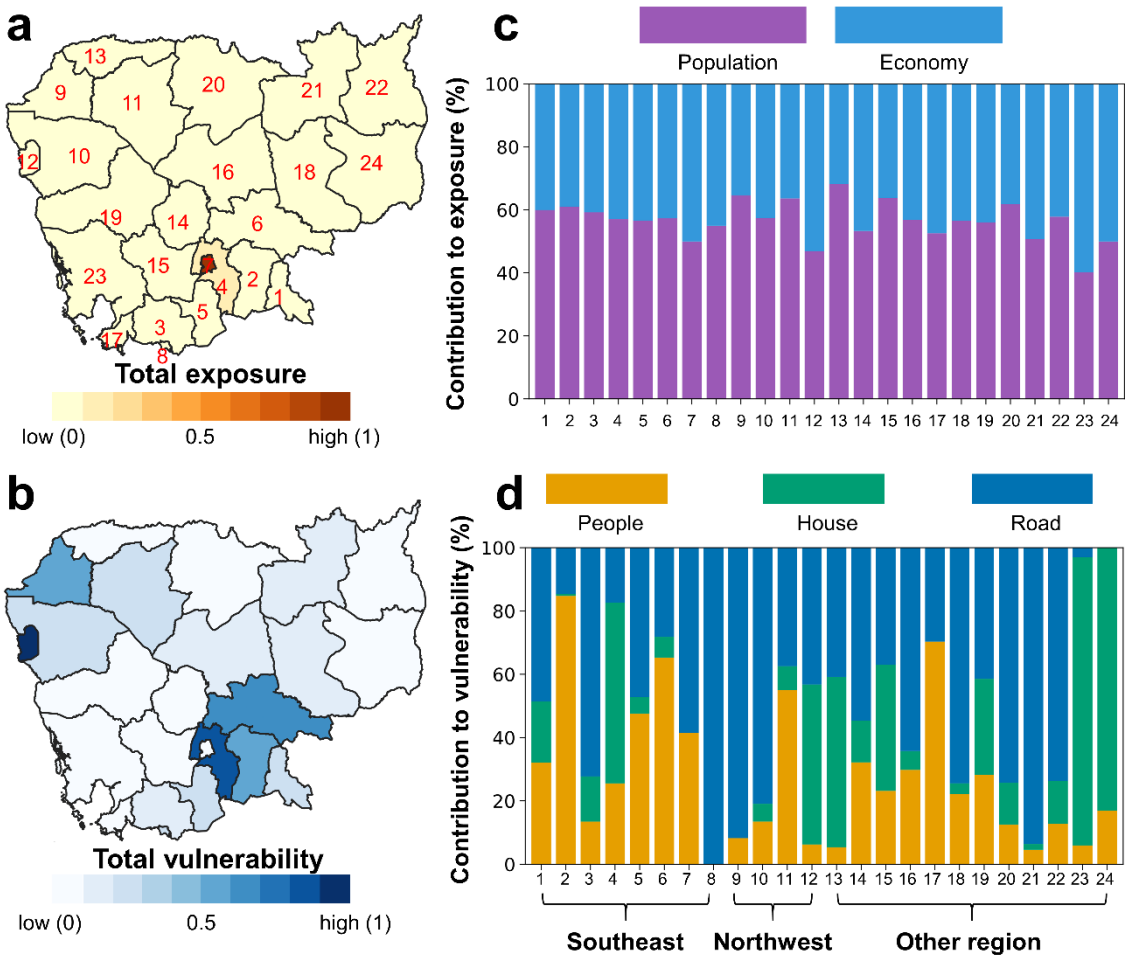


Fig. 3.16. The spatial distribution of the exposure (a) and vulnerability (b) for the short-duration rainfalls, as well as the contribution of various indicators to the total exposure (c) and vulnerability (d).

Apart from the pluvial (flash) flooding, provinces ('1-12') in Southeastern and Northwestern Cambodia are likely to experience an increased risk of massive inundation and river flooding caused by long-lasting heavy rainfall (Fig. 3.17). This projection mainly owes its existence to the heavy rainfall increase, especially in the late-century, albeit with a little decrease in the mid-century in the coastal areas (Figs. 3.8 and 3.9), and higher exposure of people and properties to flood as well as the associated damages than other provinces (Fig. 3.18a and b). In particular, agriculture, the major component of exposure (approximately 88% averaged over these provinces, except for '7'), was seriously affected by the flood events during the past 26 years ('Paddy field' accounts for 35% of total

flood damage on average, as indicated in Figs. 3.18c and d) and is likely to be aggravated in the future due to the increase in the amount of heavy rainfall, which could pose a significant threat to national food security. The long-lasting heavy rainfall can cause persistent inundation that submerges the rice fields, damages the soil structure, and further leads to nutrient losses which will take a long time to recover. Implementing or improving hydraulic structures or irrigation systems (e.g., ditch) in the paddy fields (mostly rainfed) is needed to avoid the potential inundation. Moreover, provinces closely located to the Mekong River or the Tonle Sap Lake (e.g., '2', '4', '6', '10') are likely to be affected by the river flooding due to the increase in long-lasting heavy rainfall amount and low-lying areas of these provinces (Figs. 3.8, 3.9 and 3.1). Therefore, considering the economic situation of Cambodia, making full use of the existing dams on the Mekong River (mainly in the Lancang River in China) by optimizing reservoir operation through international cooperation is pivotal to improving food security and possibly one of the cost-effective approaches (compared with other measures such as importing food from other countries). Besides agriculture, 'People', 'House' and 'Road' account for 24%, 7%, and 34%, respectively, of the flood damage (averaged over Southeastern and Northwestern Cambodia) (Fig. 3.18d). Adaptation strategies similar to flash flooding, such as determining the base elevation of newly-built buildings, can also be considered here to reduce the associated inundation and river flooding impacts.

Other regions in Cambodia, mainly in the highlands (except for provinces '14-16') with low population density, are found to face the lowest risk change (Figs. 3.15 and 3.17). These provinces have been reported to suffer from flash foods due to the massive deforestation in these provinces (JICA, 2015). Such deforestation combined with the heavier rainfall events will increase the risks of flash floods and may even lead to unprecedented severe landslides in the mountainous regions (e.g., Koh Kong province). Therefore, enacting legislation to control further deforestation, designing detailed flood and landslide risk maps, and relocating the inhabitant in vulnerable regions in advance can be considered as potential countermeasures. At last, due to human intervention (such as dam construction and operation) in the upstream countries, further research is required to quantify the socioeconomic loss better, design the flood risk map, and put forward the applicable countermeasures.

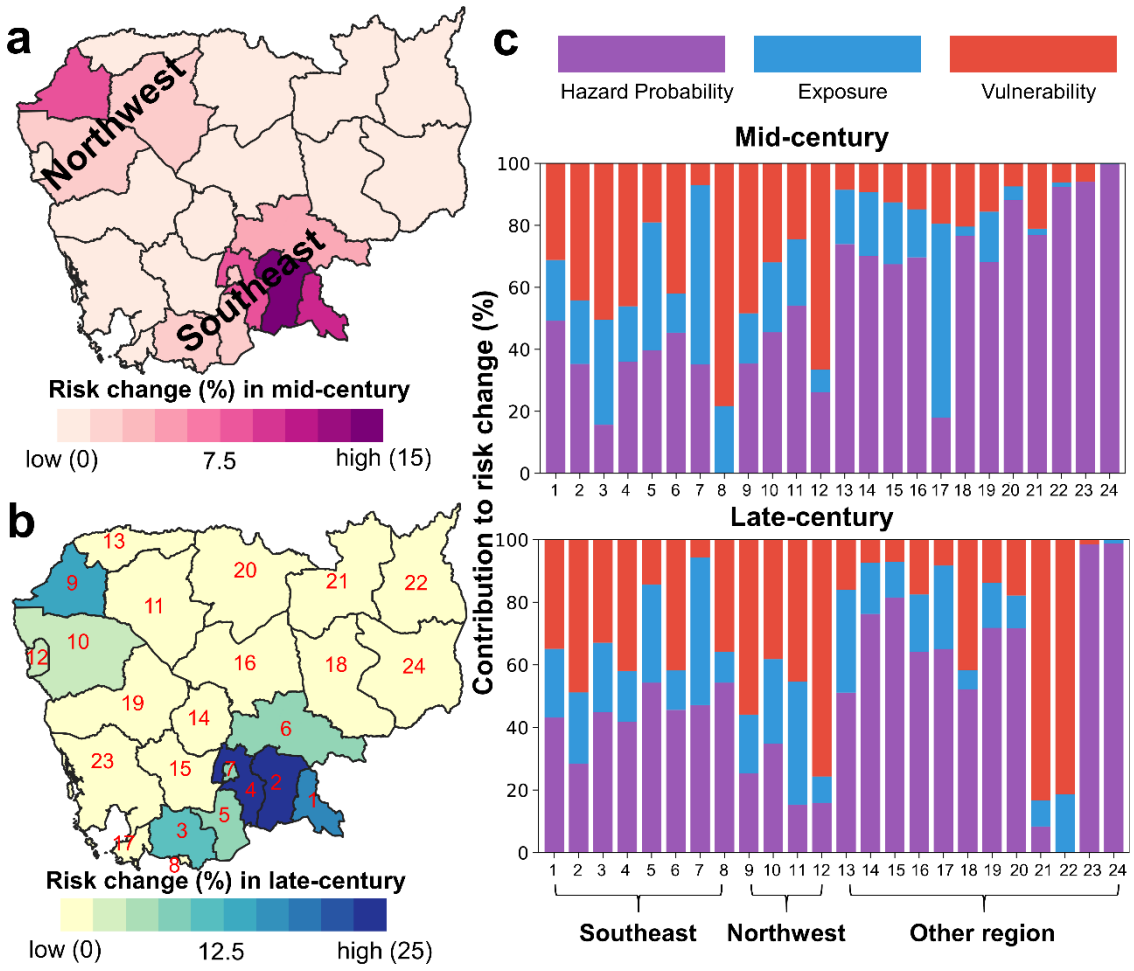


Fig. 3.17. Same as Fig. 3.15 but for flood risk change caused by long-lasting (>2 days) heavy rainfall.

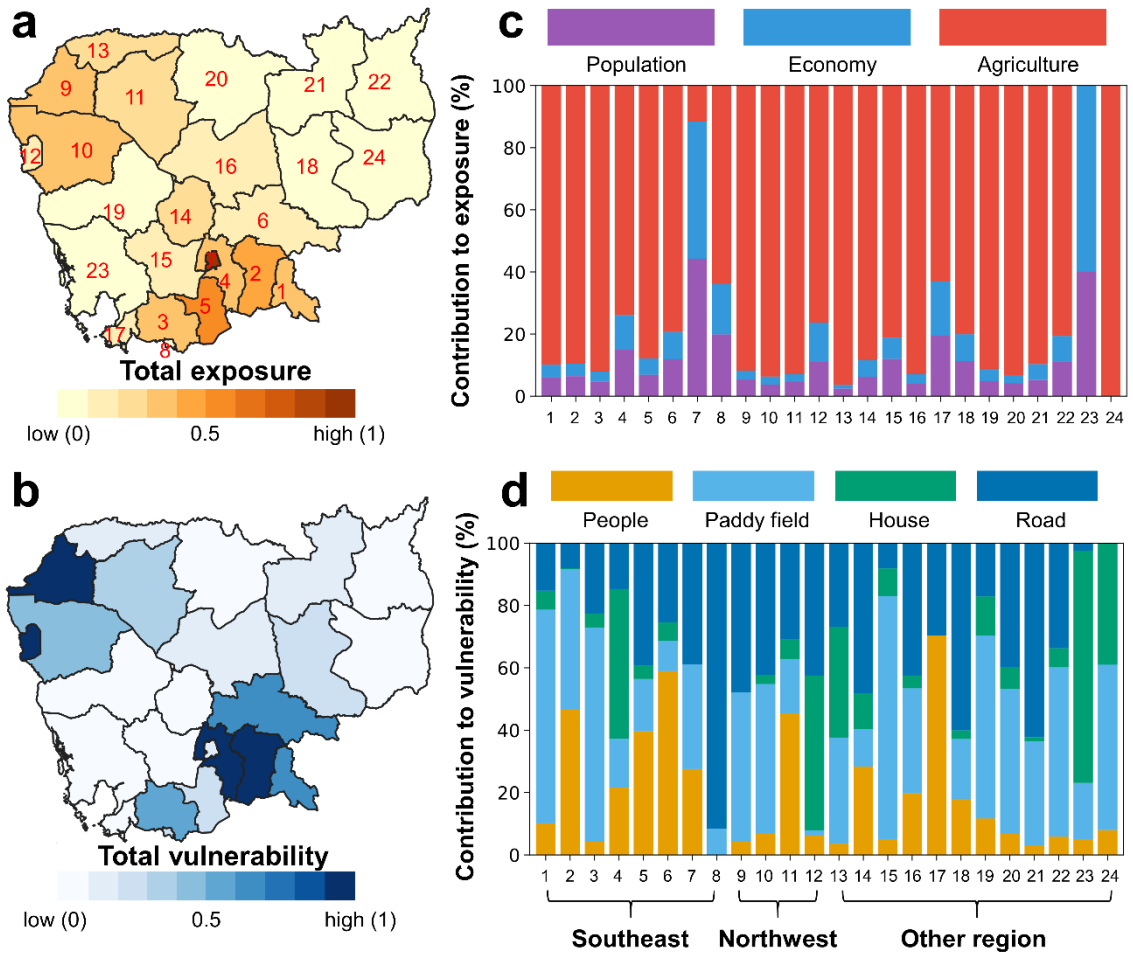


Fig. 3.18. Same as Fig. 3.16, but for the long-lasting rainfalls (taking agriculture into account).

3.4.4 Uncertainties in projecting future extreme rainfall

Rainfall data plays a crucial role in understanding the future extreme rainfall events, assessing the associated risks, and subsequently designing the appropriate infrastructure to mitigate the adverse effects. Using an ensemble of a large number of climate models is important to reduce the impacts of model uncertainty. However, we used only three RCMs provided by CORDEX simulations due to the limited availability of high-quality datasets (multiple variables with high spatial and temporal resolution are needed, as shown in Fig. 3.3) that are not accessible from other RCMs. Future extreme rainfall projection based on three CORDEX models may not be sufficient enough because the previous assessment of CORDEX models has indicated a large variance among different model outputs (Tangang et al., 2020), especially for the extreme rainfall (e.g., RX1day), some model outputs can be

up to five times of the others for the same location (Supari et al., 2020). The application of ERA5 can also introduce additional uncertainty owing to the sparse and uneven gauge-based observed rainfall datasets (that are needed in developing ERA5 datasets) in Cambodia and some simplified physical processes (e.g., parameterized convection scheme) (Hersbach et al., 2020)

In terms of the accuracy of CORDEX, R50mm (annual number of days when rainfall ≥ 50mm) and RX1day (annual maximum daily rainfall) of CORDEX were much greater than those of ERA5 (Fig. 3.19), which is mainly due to the wet bias produced by the parameterization schemes adopted in RegCM4 (Tangang et al., 2020). Such overestimation of extreme rainfall in CORDEX outputs was also reported in Supari et al. 2020. Although some studies applied bias correction methods (e.g., quantile mapping) for removing the model bias (e.g., Choubin et al., 2019; Hosseinzadehtalaei et al., 2020), these methods are only applied in rainfall and temperature datasets so far and usually, they neglect the interdependence that exists among the multiple atmospheric variables (Su et al., 2020). While bias problems can be better resolved in climate models at less than 4 km spatial resolution (e.g., convection-permitting climate models) (Tabari et al., 2016), the spatiotemporal data records of these models are still limited for use by the research community. Nevertheless, the application of climate change signals (i.e., the relative change of return rainfall amount) from CORDEX instead of absolute value (i.e., return rainfall amount) (Section 3.3.3), to some extent, avoids the problem of wet bias in our study although this approach could introduce uncertainty by implicitly assuming the bias of CORDEX models is stationary.

Regarding the temporal disaggregation, the CANN model was trained by the near-surface (~1000 hPa) atmospheric variables provided by ERA5 that might not well represent the atmospheric circulation that triggers extreme rainfall (Fowler et al., 2021). For example, the mountainous terrain may affect the movement of near-surface moisture. Although the higher altitude atmospheric variables (e.g., moisture and wind at 850 hPa) can better depict the moisture movement, they are not concurrently available in ERA5 and CORDEX. The aggregation of multiple datasets, such as CORDEX and RCMs under CMIP6 may help solve this problem. Still, this task imposes a significant challenge due to the difference in standards and accuracy (both spatially and temporally) of these datasets, particularly when the CANN model is highly dependent on the concurrency among these atmospheric variables (as shown in Fig. 3.3). Moreover, someone could argue whether the CANN model trained and verified

based on ERA5 reanalysis can be used for CORDEX. Here, the CANN model is specifically designed to simulate the near 'observed' precipitation mechanisms embedded in ERA5; an approach similar to the Massachusetts Institute of Technology-Emanuel scheme, albeit dealing with different mechanisms. Therefore, the CANN model built based on ERA5 gives confidence in its use for CORDEX, regardless of the inherent biases in the latter, which could inevitably result in uncertainties. In other words, the discrepancy between ERA5 and CORDEX outputs, mainly ascribed to the bias in CORDEX, should not be the reason for us not to apply the mathematical relationship obtained either the data-driven model (i.e., CANN) or other mathematical derivation through physical modeling. Nevertheless, future studies can reduce this uncertainty by using climate models with higher accuracy. At last, the use of ARF to transform area IDF curves to point IDF curves can also cause some uncertainty (ascribed to the homogeneity assumption (Peleg et al., 2018) and stationarity assumption of ARF, which can be reduced in the future by using climate models with finer spatial resolution.

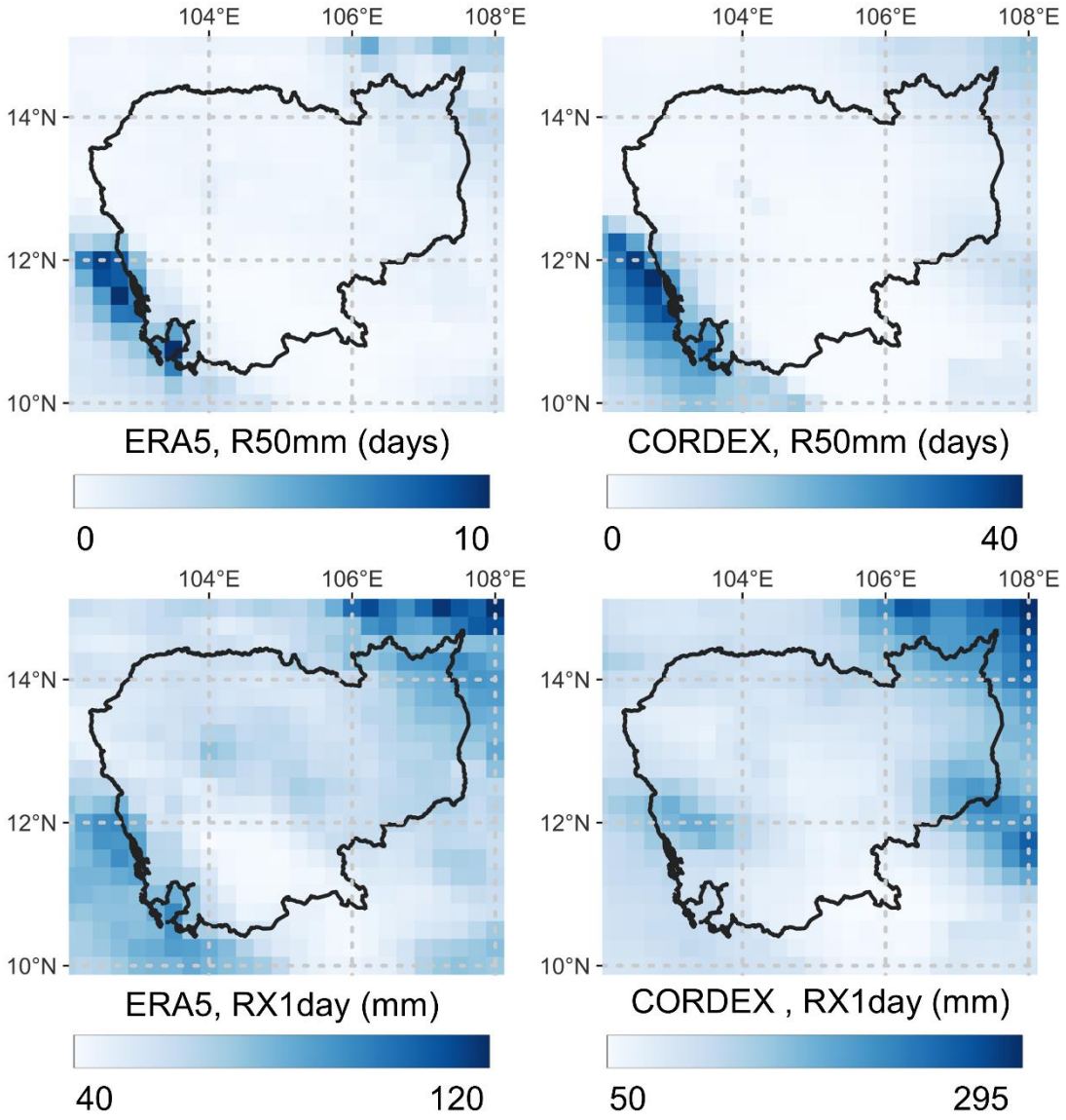


Fig. 3.19. Spatial distribution of R50mm (annual number of days when rainfall≥50mm) and RX1day (annual maximum daily rainfall) for the historical period based on ERA5 (left column) and the ensemble mean of CORDEX (right column)

3.5. Summary

In this chapter, we proposed a framework for quantifying the climate-informed changes in heavy rainfall and assessed the associated adaptation strategies related to flood risk change in our case study of the flood-prone country, Cambodia. The framework was developed based on improved temporal disaggregation models (i.e., CANN), and the outputs were statistically analyzed based on ZIGEV. A novel flood risk change index was built on the information of various socioeconomic factors, including

the population, agriculture, economy, flood impacts on population, etc., as well as the information on projected heavy rainfall change. Subsequently, we developed a province-level flood risk change map in Cambodia and discussed countermeasures to reduce the multidimensional flood-induced risks. The following conclusions can be drawn.

(1) The improved CANN model can reproduce the hourly extreme rainfall well (training, RMSE = 0.97 mm, R = 0.95; test, RMSE = 1.1 mm, R = 0.93) in Cambodia, and when used in combination with ZIGEV for EVA, it can reduce the uncertainty arising from temporal disaggregation bias. Future research can further reduce this uncertainty by selecting the suitable output variables for CANN (e.g., reproducing AMS instead of MMS).

(2) The rainfall intensity change obtained in Chapter 3 is a little different from the change in Chapter 2, which is owing to the use of different CORDEX models (that can cause large difference in projecting future heavy rainfall). However, similar to Chapter 2, the rainfall intensity is projected to increase more in higher return periods (i.e., rarer rainfall events). Moreover, rainfall intensity is projected to increase more in the shorter durations, higher return periods, late-century (than mid-century), especially over Southern and Central Cambodia, which is consistent with the underlying governing physical arguments (e.g., the convective heavy rainfall is estimated to intensify in the future, explaining the higher increase of rainfall intensity in the shorter durations). However, when averaging over the whole of Cambodia, only 10~30% of the rainfall intensity increase can be attributed to warming-induced moisture increase (according to CC scaling, Section 3.4.2). The remaining increase can be explained by the change in wind convergence (Fig. 3.12) and some other atmospheric activities (Fowler et al., 2021) that need further investigation in future work.

(3) For flood risk change caused by sub-daily heavy rainfall, Southeastern and Northwestern Cambodia, especially in Phnom Penh city and Kandal province, are detected as the priority regions for strategy adaptation (e.g., optimizing the urban development planning to avoid property loss) and taking mitigation actions to reduce the potential urban (flash) flood risks in the future.

(4) For flood risk change caused by long-lasting heavy rainfall, Southeastern and Northwestern Cambodia, especially for the provinces which play a major role in food (rice) production, should update field drainage systems and enhance the international cooperation, especially with countries

in the upper Mekong River basin to avoid the adverse impacts of inundation and river flooding on the national food security.

4. Uncertainty quantification in the development of IDF curves

4.1 Background

Updating the existing IDF curves to reduce the global warming risks are typically based on RCMs. However, climate change predictions are subject to the model uncertainty arising from the unresolved physical processes, parameterization schemes, and coarse spatial resolution (usually 25-50 km) (Flato et al., 2013; Volosciuk et al., 2017). The uncertainty is found to be larger for projecting the future rainfall extremes (Ning et al., 2015; Ning and Bradley, 2016). Therefore, reducing the model uncertainty is essential in the development of IDF curves to help identify and act upon suitable adaptation measures in the face of global warming. Climate model uncertainty can be reduced in two ways. The first is to develop high-resolution convection-permitting models; however, convection-permitting models are still limited in their spatiotemporal data records and access to the research community due to their high demand in computational power (Berg et al., 2019; Tabari et al., 2016). The second is to use statistical downscaling methods to remove the model bias by tailoring the datasets for specific locations, scales (Pour et al., 2018). Among these statistical downscaling methods, bias correction (BC) methods are most commonly used owing to the ease of their implementation and good performance (Chen et al., 2013a; Hosseinzadehtalaei et al., 2020; Maraun, 2013; Requena et al., 2021; Srivastav et al., 2014). Specifically, Li et al. (2017) found that the application of BC significantly improved the accuracy of extreme rainfall and was able to estimate the IDF curves accurately. However, BC methods have been criticized for introducing additional uncertainty by altering the raw modeled climate change signals (Cannon et al., 2015). Some studies also claimed that as climate models can preserve the climate change signals, the application of BC is unnecessary if we only focus on the change of the target variables (e.g., IDF curves) (Mamoon et al., 2016; Supari et al., 2020; Tangang et al., 2018, 2020). Currently, there is no consensus on whether BC should be used or not. BC methods are generally used in daily or longer time scale datasets of climate models, with a few studies focusing on the sub-daily datasets (Hosseinzadehtalaei et al., 2020; Requena et al., 2021; Srivastav et al., 2014). Several studies have assessed the impacts of BC on daily and longer time scale rainfall extremes (Cannon et al., 2015; Chen et al., 2013a; Ji et al., 2020; Maraun 2013), yet the effects are rarely

assessed on the sub-daily datasets due to the limited availability of sub-daily datasets from climate models and observations. Considering that BC methods are gaining an increasing interest in climate change impact studies, there is a need to explore and expand our understanding of the impacts of BC on sub-daily rainfall extremes as well as the associated uncertainties.

In addition to the model uncertainty, the future emission scenarios, which are represented by RCP scenarios, are highly uncertain owing to a lack of collective political will to efficiently control the emissions, which can be seen in the deep contradictions in UN Climate Change Conference of the Parties (COP26) (Masood and Tollefson, 2021). The natural variability of the climate system ('noise', i.e., the inherent chaotic nature of atmospheric system) limits the predictability of climate change signals when a limited number of climate models are used (Aalbers et al., 2018). Estimation of distribution parameters also introduces uncertainty for a small record length (Poschlod et al., 2021). Although these uncertainties are difficult to reduce without access to new information (e.g., more observed datasets), quantifying them is also pertinent to better understand the possible climate change impacts. Therefore, this chapter 's objectives are two-fold.

1. Quantifying different uncertainty sources, including the BC method, RCP scenarios, climate models, and parameter estimation methods, in the development of IDF curves, with a special focus on whether the BC method should be used or not.

2. Qualitative discussion on the adaptive measures, which can be updated with additional information, under deep uncertainty quantified in this chapter.

The remainder of this chapter is organized as follows. Section 4.2 describes the study area and various datasets and introduces the methods used in this chapter. The results and discussion are described in Section 4.3 followed by the future scope of the adaptive measures that can be updated in face of deep uncertainty in Section 4.4. Section 4.5 presents the conclusion of this chapter.

4.2. Material and methods

4.2.1 Study area and datasets

To investigate the uncertainty in the development of IDF curves, we used three flood-prone cities introduced in Chapter 2, namely, CTC, HCMC, and PPC (Fig. 2.1). One-hour (1h) observed rainfall datasets in HCMC and CTC were collected at Tan Son Hoa gauge station and Can Tho gauge station,

respectively. The future rainfall datasets were provided by CORDEX simulations in Southeast Asia (Tangang et al., 2020). Although the CORDEX simulations have shown the added values compared with the GCM, multiple uncertainty sources still exist (Tangang et al., 2020). To reduce these uncertainties, a multi-model ensemble is commonly used to better represent the study parameters (e.g., extreme rainfall) than the individual model (Supari et al., 2020). Therefore, our study used all the available CORDEX simulations (as of May 26, 2021) (https://esg-dn1.nsc.liu.se/projects/cordex/) that provide 3h rainfall datasets in Southeast Asia under two Representative Concentration Pathway scenarios (i.e., RCP 4.5 and 8.5) as listed in Table 2.2.

4.2.2. Framework for developing IDF curves

The framework of constructing IDF curves (focusing on 1-24h) for historical and future periods is shown in Fig. 4.1. In HCMC and CTC (Fig. 4.1a), IDF curves in the historical period were constructed based on observed 1h rainfall datasets. In the future period, the 3h rainfall datasets provided by RCMs were firstly spatially downscaled based on the BC method and secondly temporally disaggregated to 1h extreme rainfall datasets. Subsequently, these datasets (from 1h to 24h) were used to estimate the return rainfall (rainfall event with a given return period, i.e., rainfall event that is expected to occur once in this given period), which was later applied to construct IDF curves in future periods. A change factor approach was applied to remove the bias generated in temporal disaggregation. To quantify the uncertainty introduced by BC method, rainfall intensity change with BC was obtained in these two cities and compared with the change with BC, as described in Fig. 4.1. In HCMC and CTC, the influence of other uncertainty sources was also quantified. To avoid redundancy in this chapter, the methods and results for developing IDF curves in PPC were discussed in detail in Chapter 2.

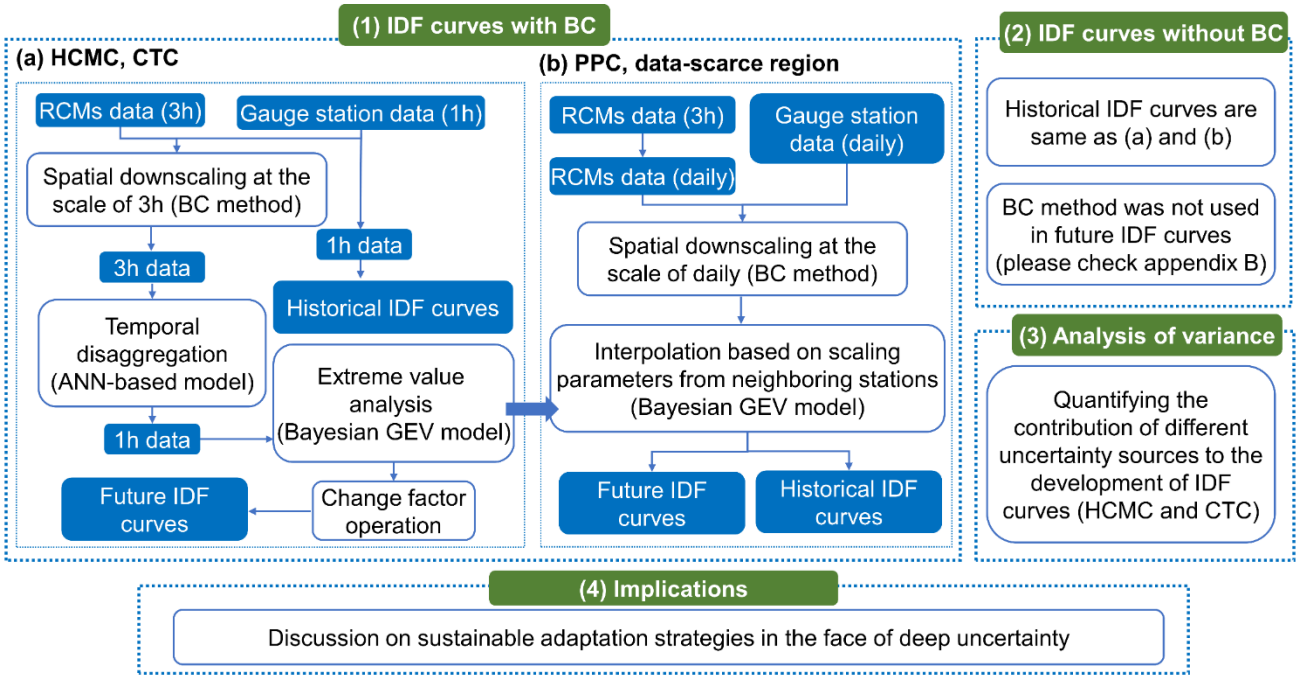


Fig. 4.1. Research methodology framework for obtaining IDF curves with bias correction (1) for (a) HCMC, CTC and (b) PPC, as well as without bias correction (2), quantification of different uncertainty sources (3) and discussion on adaption strategies under deep uncertainty (4). (Acronyms: RCMs, Regional climate models, Bayesian GEV, Bayesian generalized extreme value, BC method, Bias correction method, ANN, Artificial neural network)

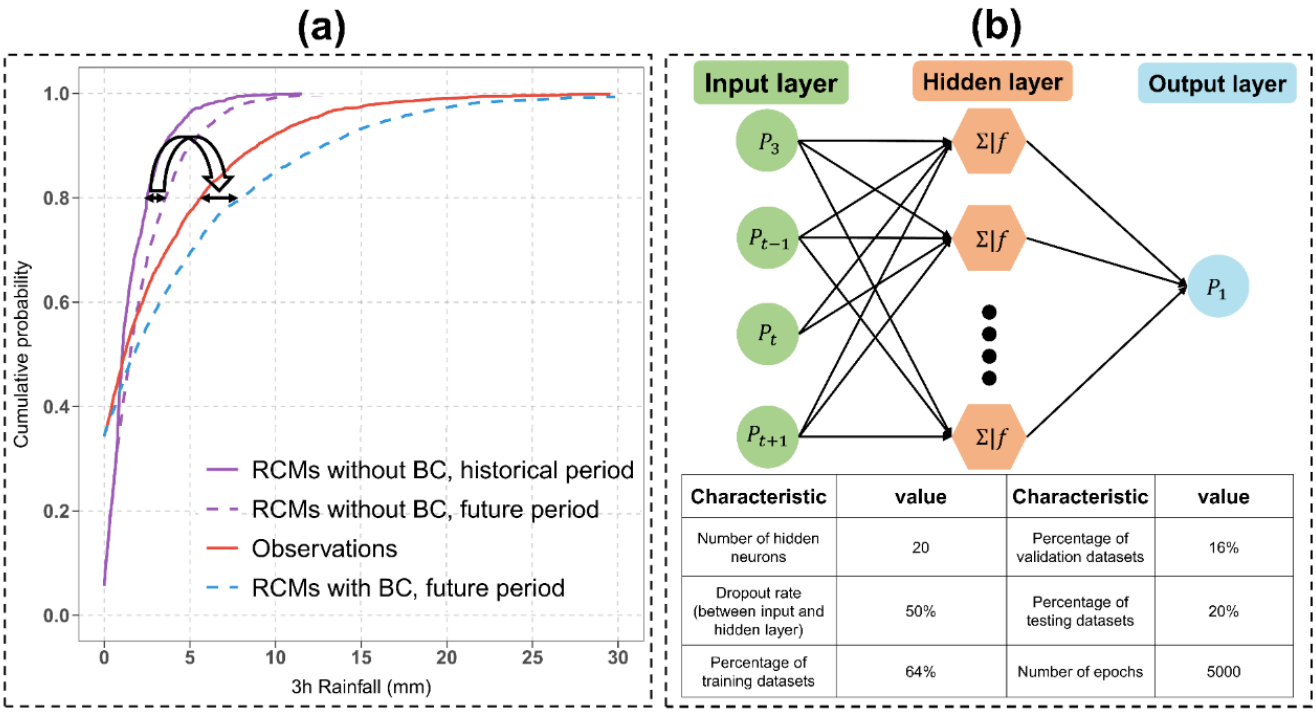
4.2.3. Bias correction method

Quantile delta mapping, which is an adaptation of the well-known Quantile mapping method and has been proven to be more reliable in removing the systematic model bias (Cannon et al., 2015), was used to spatially downscale the 3h rainfall datasets from CORDEX simulations for HCMC and CTC. Although some other BC methods, such as Normalized Quantile Mapping (see in Section 2.3.2) and Scaled Distribution Mapping (Switanek et al., 2017), also show superior performance over Quantile mapping, we did not consider those in this chapter to avoid the redundancy in achieving the objectives (to compare the results with BC and without BC instead of comparing results with different BC methods) of the study. Mathematically, Quantile delta mapping can be represented as below,

$$p_{cor,fut} = p_{obs} \times CV \quad (4.1)$$

$$CV = \frac{F_{mod,fut}^{-1}(\tau)}{F_{mod,his}^{-1}(\tau)} \tag{4.2}$$

where p_{obs} and $p_{cor,fut}$ denote the observed rainfall for the historical period and modeled rainfall after BC for the future period, respectively, and CV represents the correction value. $F_{mod,fut}^{-1}$ and $F_{mod,his}^{-1}$ denote the inverse CDF of modeled rainfall for the historical period and the future period before BC, respectively, and τ is the non-exceedance probability. Fig. 4.2a gives an example of removing the model bias at $\tau = 0.8$. In this chapter, Quantile delta mapping was performed in a seasonal moving window and used to downscale the modeled datasets from the nearest neighboring grid that contains the targeted rain gauge. To remove the drizzle-effects of modeled datasets, we set a threshold of 0.1 mm considering the minimum detectable hourly rainfall of most rain gauges (Pang et al., 2021), below which all rainfall values were set to zero.



Characteristic	value	Characteristic	value
Number of hidden neurons	20	Percentage of validation datasets	16%
Dropout rate (between input and hidden layer)	50%	Percentage of testing datasets	20%
Percentage of training datasets	64%	Number of epochs	5000

Fig. 4.2. (a) Schematic of BC based on quantile delta mapping and (b) schematic diagram of the ANN-based temporal disaggregation model. Here, P_1 denotes the 1h monthly maximum series and P_3 represents the 3h rainfall depth happening in the same time of P_1. P_{t-1}, P_t , P_{t+1} denote the daily rainfall depth centered by the day that records P_1. Characteristic features of various statistics

associated with the ANN model are also shown in the table.

4.2.4 Temporal disaggregation method

Since the CORDEX simulations provide only 3h rainfall datasets (for both RCP 4.5 and 8.5), which is not (temporally) sufficient enough to update the IDF curves (focusing on 1-24h rainfall), the temporal disaggregation methods are necessary to estimate 1h rainfall (Mirhosseini et al., 2014). Here, we used an ANN-based model similar to Section 2.3.4 but using extreme 3h rainfall datasets instead of extreme daily rainfall datasets, which can reduce the uncertainty caused by the stationarity assumption to some extent (the model trained by the historical observations can be applied to the future period). This is because compared with extreme daily rainfall, extreme 3h rainfall is more likely to occur in the same rainfall event (e.g., convective process) as that of extreme 1h rainfall (Fowler et al., 2021). Thus, the characteristics of extreme 1h rainfall can be better represented by extreme 3h rainfall than the extreme daily rainfall. The architecture of this ANN-based model is shown in Fig. 4.2b. To estimate the observed 1h rainfall extremes (i.e. P_1), the 3h observed rainfall (P_3), which accumulated in the same rainfall events that recorded P_1, and the three consecutive days of rainfall amounts (i.e., P_{t-1}, P_t, P_{t+1}) centered by the day that recorded P_1 were used as the input variables. Thus, the model was trained based on the extreme rainfall information (observed during the rainfall events that recorded P_1) instead of the whole time series of rainfall, which improves the accuracy and efficiency. However, direct application of this ANN-based model for obtaining the future 1h rainfall extremes is problematic as we don't know when 1h rainfall extremes will happen in the future, given only a time series of 3h rainfall is available. Thus, the input variables for the ANN-based model cannot be determined by the time series of 3h rainfall provided by CORDEX simulations. To resolve this problem, firstly, we regarded time series of 3h rainfall as P_3 and associated three consecutive daily rainfall as P_{t-1}, P_t, P_{t+1}. Thus, we obtained 7305*8=58440 (8 sets of 3h per day for a total of 7305 days during 2026–2045 or 2066-2085) sets of (P_3, P_{t-1}, P_t, P_{t+1}). Secondly, (P_3, P_{t-1}, P_t, P_{t+1}) was input into the ANN-based model to generate 58440 output variables. Next, we extracted the monthly maximum series from these output variables. We tested this procedure in the historical period (Section 4.3.1).

4.2.5. Extreme value analysis

The Bayesian GEV through MCMC was described in detail in Section 2.3.5. Apart from the application of MCMC, this chapter also employed two other parameter estimation methods (PEMs), i.e., maximum likelihood estimation (MLE) and L-moment methods to examine the impacts of different parameter estimation methods on IDF curves. As Bayesian GEV can only be estimated by MCMC, MLE and L-moment were mainly used to estimate the GEV parameters.

4.2.6. Constructing IDF curves based on bias-corrected RCMs

In HCMC and CTC, the IDF curves for the historical period were constructed based on the return rainfall of different durations computed by Bayesian GEV parameters. Regarding the IDF curves in the future period in HCMC and CTC, we found that the estimation of Bayesian GEV parameters (through MCMC)was slow and hard to achieve for modeled rainfall datasets across two gauge stations. This may be because the CORDEX simulations still have a limited ability to simulate the rainfall (especially the spatial information of the rainfall) very well (Tangang et al., 2020). Therefore, for the future case, we used the Bayesian GEV for each station separately (which means N=1 in Eq. 2.9). Subsequently, IDF curves focusing on the duration of 3h to 24h can be obtained directly from return rainfall based on Eq. 2.9. For the IDF curves of 1h rainfall, due to the deviation in the disaggregated 1h extreme rainfalls, a change factor operation (denoting the ratio of future return rainfall to the historical one) was used, represented as below,

$$Change\ factor = \frac{RL_{mod,fut}}{RL_{mod,his}} \tag{4.3}$$

$$RL_{cor,fut} = RL_{obs,his} \times Change\ factor \tag{4.4}$$

where, $RL_{mod,fut}$ and $RL_{mod,his}$ denote the ANN model-based (or raw RCMs-based) return rainfall for the future and historical period, respectively. $RL_{obs,his}$ is the observation-based return rainfall and $RL_{cor,fut}$ denotes the corrected/improved future return rainfall. This change factor operation has been applied in several climate change impact studies (e.g., Liew et al., 2014; Tabari et al., 2016; among others) and is superior to the approach used in Lima et al. (2018), which suffers more from a stationarity assumption of the scaling process (from daily rainfall to hourly rainfall).

4.2.7. Impacts of multiple uncertainty sources on IDF curves

To evaluate the impacts of the BC on constructing IDF curves, two approaches, including the one described in Fig. 4.1, with and without BC, were used/explored as summarized in Table 4.1. The methods used for obtaining IDF curves in the future periods without BC were also based on the change factor operation. In specific, the change factor, in this case, denotes the ratio of future return rainfall to the historical one, based on RCMs without BC. Subsequently, this change factor was multiplied by the historical IDF curves at the targeted rain gauge station to obtain the future IDF curves. This change factor operation was mainly used to eliminate the discrepancy between modeled datasets and observations, which is similar to that used in Hosseinzadehtalaei et al. (2021). In the data-sufficient region (i.e., HCMC and CTC), apart from the impacts of the BC, the impacts of different parameter estimation methods, different RCP scenarios, and RCMs on the uncertainty in IDF curves were also assessed. Specifically, to quantify different uncertainty sources linked with the choices of different RCMs, RCP scenarios, parameter estimation methods, and with and without BC, the 'analysis of variance' method was used (Wang et al., 2016). The total variance was computed as the sum of the standard deviation of each uncertainty source. For each uncertainty source, the standard deviation was calculated after taking the average value of rainfall intensity change derived from the remaining uncertainty sources. For example, the standard deviation across RCMs was calculated after taking the average from all parameter estimation methods, both RCP scenarios and datasets with and without BC. Finally, the fraction of total variance was calculated as the ratio of the variance of each uncertainty source to the total variance.

Table 4.1. Summary of the method used to estimate sub-daily return rainfall for quantifying the influence of BC

City	Time	Datasets	Bias correction	Temporal disaggregation	Extreme value analysis	Major assumptions
HCMC & CTC	Historical period	Observed 1h rainfall	-	-	Bayesian GEV	-
	Future periods	3h rainfall from RCMs	Yes	Obtaining 1h rainfall extremes from 3h rainfall based on ANN	Bayesian GEV	ANN holds true from 3h to 1h rainfall in the future periods* (Section 4.2.4)
	Future period	3h rainfall from RCMs	No	Obtaining 1h rainfall extremes from 3h rainfall based on ANN	Bayesian GEV	1. ANN holds true from 3h to 1h rainfall in the future periods 2. Return rainfall change based on RCMs without BC represents the return rainfall change in the gauge stations** (Section 4.2.6)

PPC	Historical period	Observed daily rainfall	-	Obtaining sub-daily return rainfall from daily return rainfall based on scaling property	Bayesian GEV	Bayesian GEV parameters from neighboring stations (i.e., HCMC & CTC) can be used in PPC (Section 4.2.5)
	Future period	Daily rainfall from RCMs	Yes			
	Future period	3h rainfall from RCMs	No	Obtaining 1h return rainfall from 3h return rainfall based on scaling property	Bayesian GEV	1. Scaling property holds true from 3h to 1h rainfall in the future periods*** (Section 4.2.5) 2. Return rainfall change based on RCMs without BC represents the return rainfall change in the gauge stations

* This ANN model trained based on historical observations can be applied to the future period.

** From a statistical perspective, the percentage change of the areal (return) rainfall provided by RCMs without BC equals that change of the point (return) rainfall in gauge stations.

*** To estimate the GEV parameters of 1h rainfall extremes, which can be not obtained based on 3h rainfall provided by RCMs, we used the scaling property.

4.3. Results and discussion

4.3.1 Bias correction and ANN-based temporal disaggregation

Based on the framework described in Fig. 4.1a, Quantile delta mapping was applied to remove the bias of 3h rainfall datasets provided by RCMs (Fig. 4.2a), after which ANN-based model was applied to obtain 1h rainfall extreme datasets for the future periods (Fig. 4.2b). Before applying these two methods, we validated them using the observed data in the historical period as shown in Fig. 4.3a and Fig. 4.3b, respectively. Compared with observations, the RCMs without BC apparently underestimated the 3h rainfall extremes, with the mean value of -59% relative to the observations, averaged over two cities (Fig. 4.3a), which is similar to the previous studies using CORDEX simulations (e.g., Berg et al., 2019). After BC, RCMs agreed favorably well with the observations, with a weak underestimation of the mean value (-9%) averaged over two cities, but the ensemble spread was found to increase slightly. The strong underestimation of rainfall extremes in original outputs of RCMs (without BC) is mainly attributed to convective parameterization (Berg et al., 2019; Kendon et al., 2017; Li et al., 2012), coarse spatial resolution (i.e., 25 km^2 areal average rainfall) used in these RCMs, and some simplified atmospheric forcing mechanisms (e.g., aerosol forcing) employed in the driving GCMs (Lin et al., 2018). The slight increase in ensemble spread is also due to the abovementioned model simplification of physical processes that cannot be resolved by the BC (Maraun et al., 2019).

In Fig. 4.3b, the disaggregated 1h rainfall extremes (i.e., P_1), specifically 1h monthly maximum series, showed a good agreement with the observed P_1, but deviated from the large values in the tail, which is a common problem for many temporal disaggregation models (Burian et al., 2000; Burian and Durrans, 2002; Kossieris et al., 2018; Mirhosseini et al., 2014). This deviation is mainly due to some large 1h rainfall values generated from rainfall events that lasted within one hour, which makes 3h and daily rainfall values (input layer in Fig. 4.2b) difficult to reproduce the characteristics of P_1. Therefore, we applied the change factor operation to correct the deviation, although the uncertainty still persists. Apart from this deviation problem, the stationarity assumption of the ANN-based model, which means that the model trained in the historical period can be used for the future period, may add additional uncertainty. Such stationarity assumption of the temporal disaggregation model is commonly used in climate change impact studies (e.g., Hosseinzadehtalaei et al., 2021; Lombardo et al., 2017) and is difficult to validate due to the complexity of the rainfall generating system (Moustakis

et al., 2021), and the lack of observed rainfall datasets for future periods. The possible way to reduce the uncertainty from the deviation problem and stationarity assumption is to use more climate variables., e.g., temperature, humidity, wind velocity, to reproduce the rainfall generating system (Westra et al., 2013). An alternate solution is to use the high-accuracy convection-permitting models; however, these convection-permitting models are very limited due to their high demand for computational power and the limited understanding of (or a lack of focus on) this system in a small spatial scale (e.g., cloud-aerosol interactions and urban heat island impacts) (Fowler et al., 2021; Zhang et al., 2018). In general, the application of the BC and ANN-based temporal disaggregation can be an efficient way to provide the 1h extreme rainfall datasets based on 3h datasets.

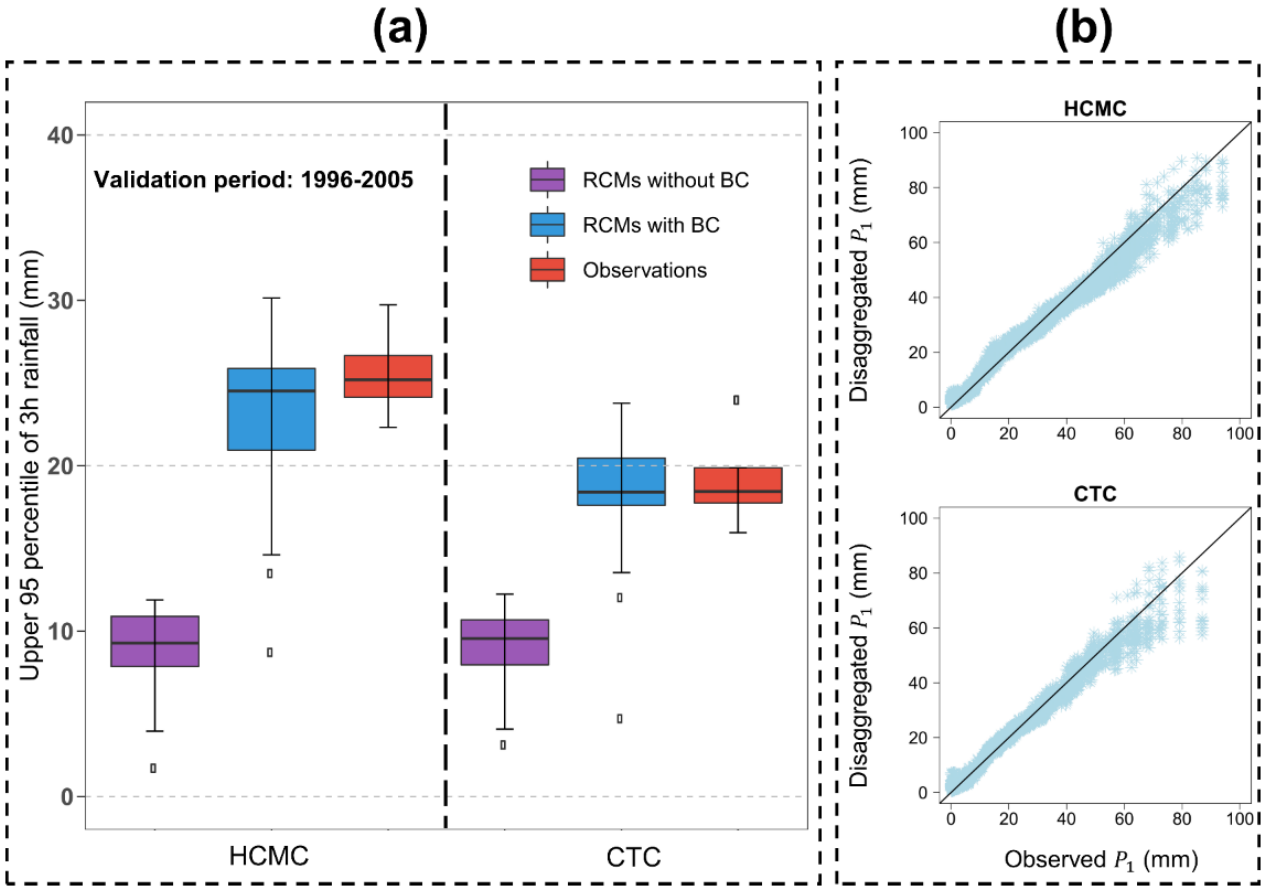


Fig. 4.3. (a) Comparison between RCMs with and without BC, and observations. For comparison purposes, the BC method calibrated in 1986-1995 was used to remove the bias of RCMs in 1996-2005, and subsequently, the RCMs with and without BC were compared with the observations in 1996-2005. (b) Comparison of the quantile-quantile plots of the observed and disaggregated (wet-

day) 1h monthly maximum time series based on testing datasets. The training datasets were randomly selected 50 times from the observed rainfall series and the unselected 50 sets of remaining rainfall series were regarded as the testing datasets.

4.3.2 Impacts of bias correction on rainfall intensity change

The relative changes in rainfall intensity from historical to future periods for different return periods and durations are shown in Fig. 4.4. Apparent inflation of relative change by BC was observed in three cities. For example, in HCMC during the far future period (Figs. 4.4c and 4.4d), a relative change in excess of 25% with respect to the observations was observed by applying BC, while the maximum relative change without BC hardly exceeded 0%. Nevertheless, this large increase (indicated by the RCMs with BC in Fig. 4.4d) during the far future was also found in some previous studies for the same or nearby research regions based on the same or similar scenarios (RCP 8.5 and A2). (Liew et al., 2014; Noor et al.,2018; Truong Ha, 2018), albeit using different methods to remove the model bias. Moreover, in CTC and PPC, RCMs with BC projected a smaller increase in the far future than near future under RCP 4.5 (Fig. 4.4b) and a larger increase in the far future than near future under RCP 8.5 (Fig. 4.4d). However, in HCMC (Figs. 4.4b and 4.4d), the projected change in rainfall intensity showed a (weak) opposite trend to those in CTC and PPC, which could be explained by a multitude of factors, including the limited number of available CORDEX simulations used in the current research, the inherent model biases of these simulations (Section 4.3.1), and some unconsidered factors (e.g., spatial and seasonal water availability; Tabari, 2020). Owing to these factors, this kind of disagreement between neighboring regions was also found in previous studies (e.g., Trinh-Tuan et al., 2019; Tangang et al., 2018; Supari et al., 2020).

Based on the RCMs with BC in these three cities (Figs. 4.4b and 4.4d), generally more increase was found in rainfall with higher return periods, suggesting that a rarer event (denoting a relatively higher rainfall depth distributed in 24 hours) is likely to increase more in the future, which compares well to the previous research on the same or nearby research regions (Ge et al., 2019; Truong Ha, 2018). This observation follows well the laws of physics where, atmospheric circulation tends to drive a larger increase in rarer rainfall events than less rare ones in the context of global warming (Pendergrass, 2018). However, such a trend was not observed in most cases based on RCMs without

BC (Figs. 4.4a and 4.4c). In general, with respect to the change in rainfall intensity based on the raw outputs of RCMs, the application of BC corresponds well to the previous research and physical reality. Further investigation on the impacts of the BC by using more high-accuracy RCMs could be a subject of future research efforts.

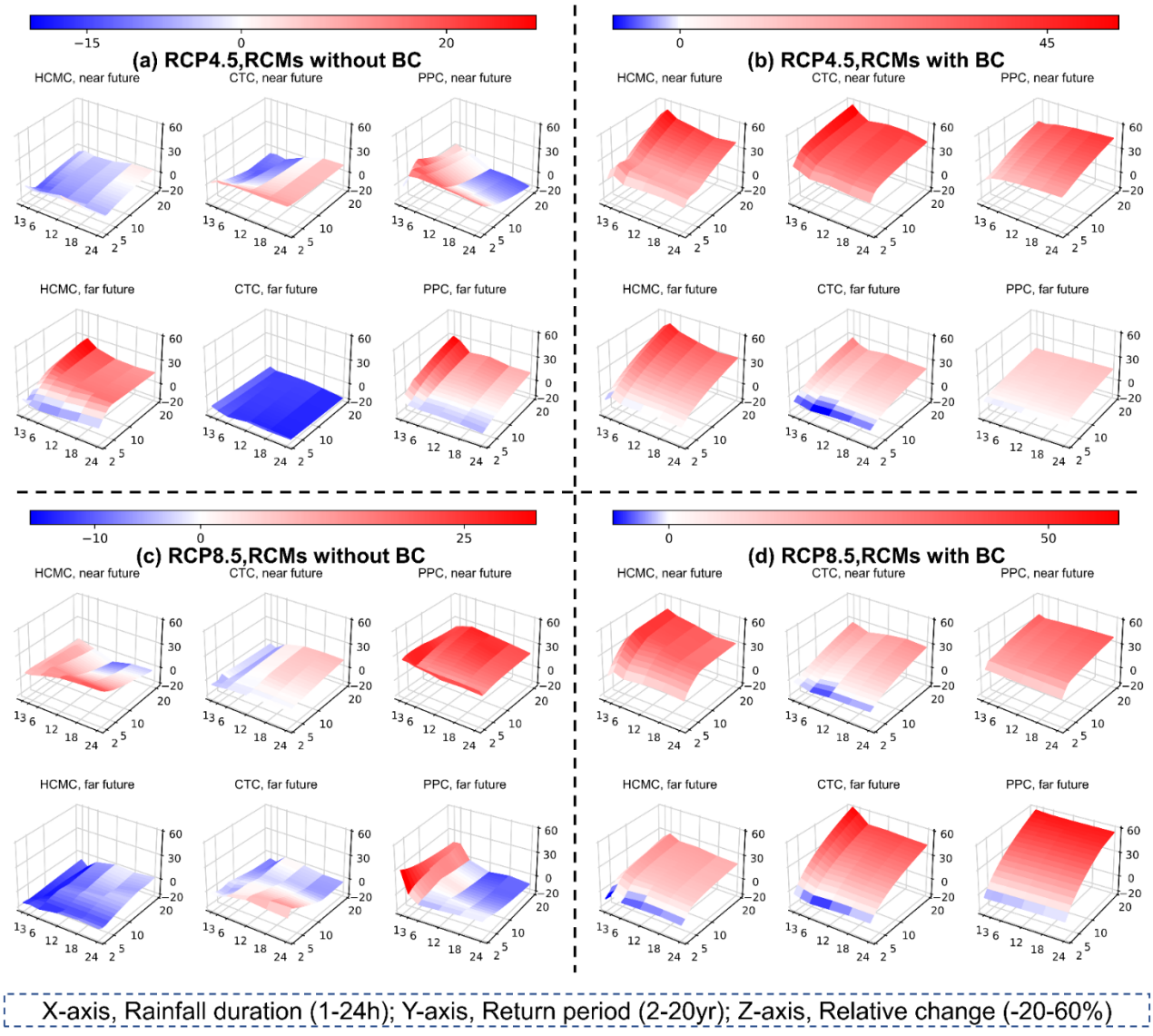


Fig. 4.4. Change in rainfall intensity from historical period to future periods based on the ensemble median of RCMs without and with BC. The rainfall intensity was calculated by Bayesian GEV through Markov Chain Monte Carlo (MCMC). Please note that the legend bars correspond to different ranges for each plot for better visualization.

4.3.3 Impacts of bias correction and parameter estimation methods on the uncertainty

In addition to the abovementioned impacts on the projected change in the ensemble median of rainfall intensity, the impacts of BC on the standard deviation in the change of rainfall intensity among RCMs were analyzed to evaluate the uncertainty related to the BC (Fig. 4.5). A comparison between the size (i.e., diameter) of circles (with BC) and that size (i.e., side) of squares (without BC) in Fig. 4.5 shows that the results with BC sometimes (e.g., rainfall duration of 1h, return period of 20-yr and RCP 8.5 in Fig. 4.5a) had a larger standard deviation (regarded as the uncertainty) than the results without BC. This behavior could be caused by the simplified physical processes that cannot be resolved by BC (Section 4.3.1). Apart from the size, in HCMC (Fig. 4.5a), the circles (with BC) and squares (without BC) displayed different colors under RCP 4.5 (with hatching line) (except for a few cases with a 2-yr return period) and similar colors under RCP 8.5 (without hatching line). A similar observation was also found in CTC (Fig. 4.5b). These results indicate that BC may have a larger impact on the uncertainty change under RCP 4.5 than RCP 8.5.

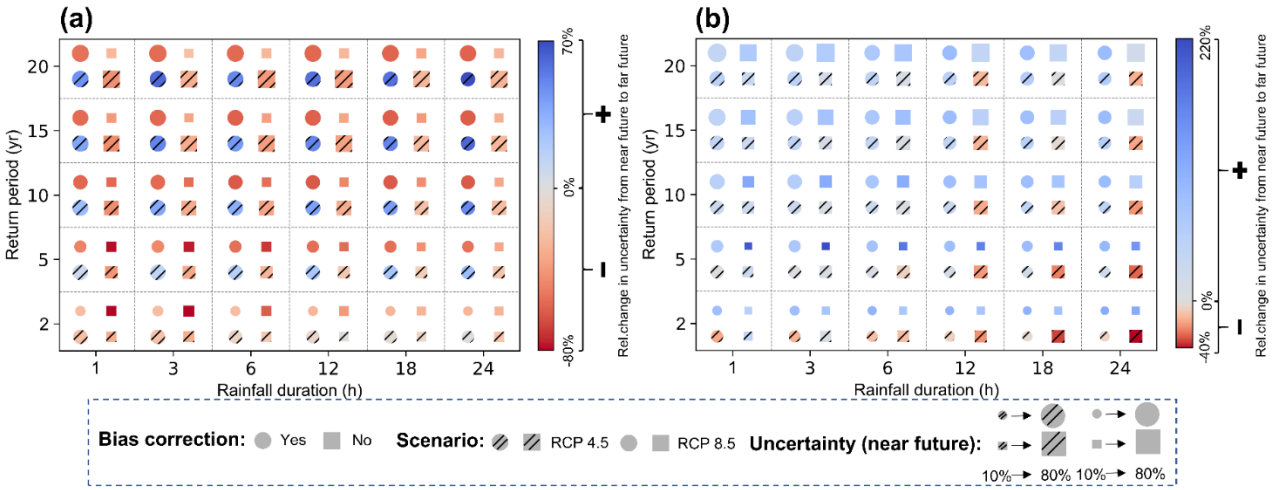


Fig. 4.5. The impacts of BC on the uncertainty of the change in rainfall intensity in the future periods in HCMC (a) and CTC (b) based on Bayesian GEV through MCMC. The uncertainty, which is represented by the standard deviation in the change of rainfall intensity among RCMs for the near future, is signified by the size (i.e., diameter) of the circles (with BC) and the size (i.e., side) of

squares (without BC). The color indicates the relative change of the standard deviation from the near future to the far future.

Apart from the impacts of BC, different parameter estimation methods affected the uncertainty, as shown in Fig. 4.6. A comparison of the size (signified by the diameter in circles and side in squares and diamonds) among circles (MCMC), squares (L-moment), and diamonds (MLE) shows that the results based on MLE had larger uncertainty than MCMC and L-moment, especially in the larger return periods. This behavior is mainly because MLE cannot work well for some unreliable values at the edges of the ensemble in a small sample size (Poschlod et al., 2021). Nevertheless, L-moment is less influenced by these unreliable values owing to its stable algorithm, and MCMC makes use of more information (i.e., prior information of GEV parameters and scaling property of rainfall extremes) to reduce the impact of these values. Focusing on the change of uncertainty from the near future to the far future, we found that generally, all three methods had similar colors in each city despite some exceptions (e.g., rainfall duration of 6h and return period of 20-yr in Fig. 4.6a). Given that MCMC delivered reliable results and Bayesian GEV through MCMC was only the method applicable to PPC, we presented the IDF curves based only on MCMC in these these cities.

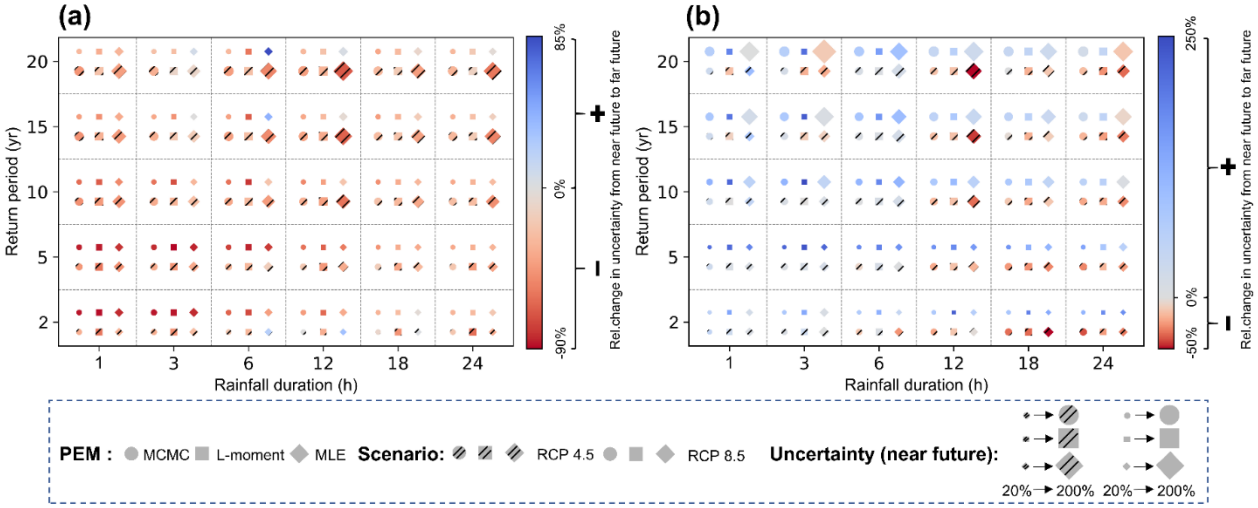


Fig. 4.6. The impacts of parameter estimation method (PEM) on the uncertainty of the change in rainfall intensity in the future periods in HCMC (a) and CTC (b) based on MCMC, maximum likelihood method (MLE), and L-moment method, respectively. The uncertainty, which is

represented by the standard deviation in the change of rainfall intensity among RCMs without BC in the near future, is signified by the size (i.e., diameter) of the circles (MCMC), size (i.e., side) of squares (L-moment), and size (i.e., side) of diamonds (MLE). The color indicates the relative change of the standard deviation from the near future to the far future.

4.3.4 Contribution of different uncertainty sources to rainfall intensity change

The variance of each uncertainty source is by and large proportional to the increasing return period, albeit with a varying contribution to the total variance (Figs. 4.7). The contribution of RCMs to total variance decreases with the return period, while that of the parameter estimation method generally increases with the return period (Fig. 4.7). RCMs are found to be the dominant uncertainty source, of which the contribution to the total variance (averaged for all the cases of the return period, duration, and future period) accounts for 41.0% at HCMC and 47.9% at CTC. BC accounts for the second-largest uncertainty, with an average value of 39.4% at HCMC and 29.7% at CTC. The contribution of RCP scenarios and parameter estimation methods are 16.5% (13.8%) and 3.2% (8.6%) at HCMC (CTC), respectively, the latter providing the lowest uncertainty. Moreover, the contribution of RCP scenarios is found to be approximating to zero in the near future in some cases (e.g., 3h rainfall in Fig. 4.7a), while the contribution is generally more stable and larger in the far future than the near future. Specifically, the contribution of RCP scenarios is, on average, 25.8% (14.5%) for the far future relative to 7.2% (13.2%) for the near future in HCMC (CTC). This behavior indicates that the rainfall intensity differs between RCP scenarios more in the far future than near future, which corresponds well to a large difference in accumulated CO_2 emissions between RCP 4.5 and 8.5 in the late century (Hausfather and Peters, 2020). The larger uncertainty in the projected change of rainfall intensity is likely to be linked with more investment in updating the infrastructure. This is because, for example, under a high-end case of larger uncertainty, to mitigate the possible severe damage from urban flooding, more cost is needed to build the drainage system with high capability.

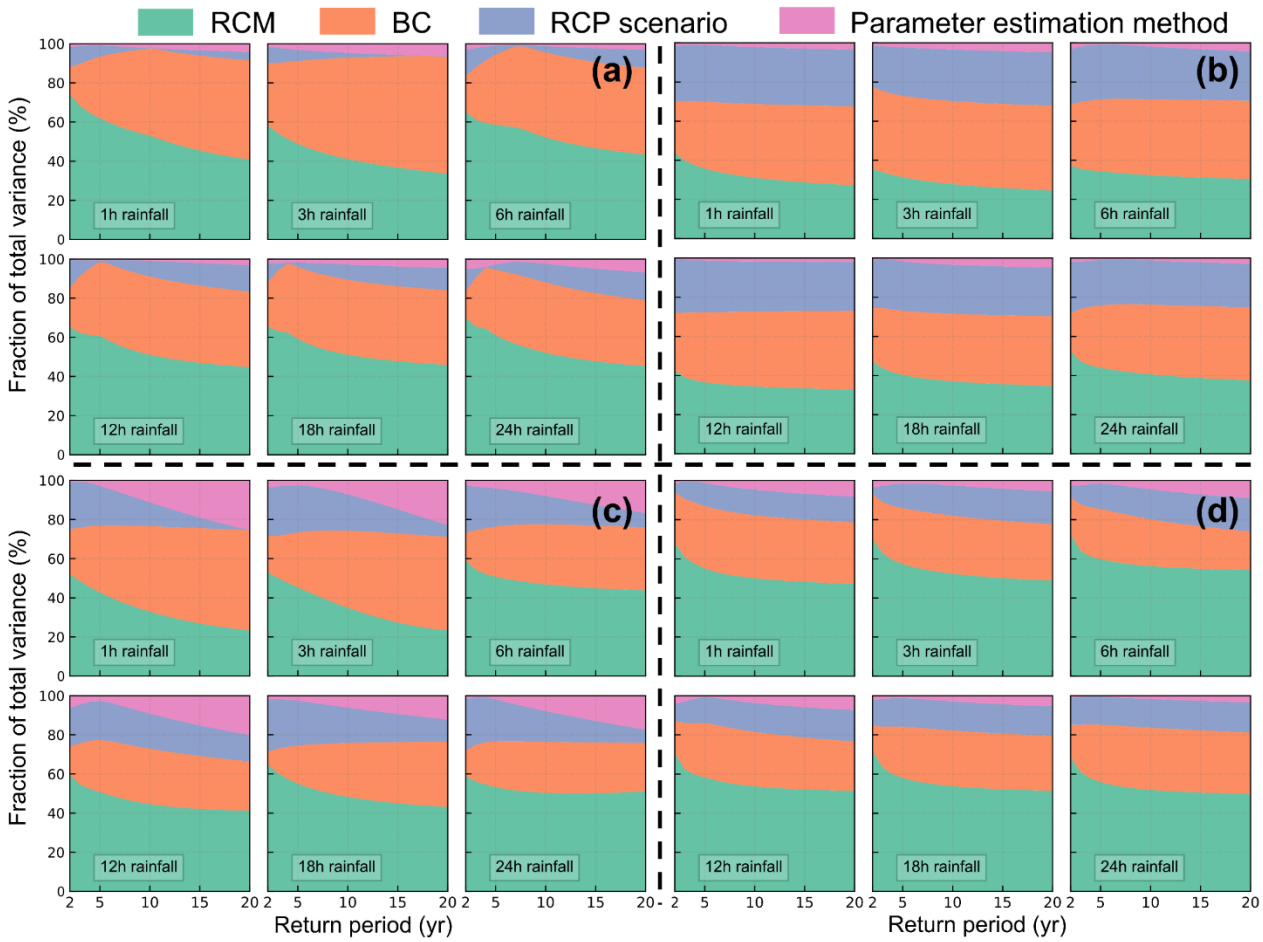


Fig. 4.7. Fraction of total variance in the projected rainfall intensity change in (a) near future, (b) far-future period in HCMC, and (c) near future, and (d) far-future period in CTC. The variance was calculated as a function of return periods explained by the standard deviation of the RCMs, with and without BC, different parameter estimation methods, and RCP scenarios.

4.3.5 Change in the IDF curves and convection rainfall

IDF curves for both the historical period (based on observations) and the future periods (based on the RCMs with BC) are shown in Fig. 4.8. Based on the ensemble median, the rainfall intensity of 5-yr (20-yr) IDF curves in the near future in HCMC is projected to increase by 13.4–23.5% (37.8–52.3%) under RCP 4.5, and 22.6 – 31.0% (30.7– 47.1%) under RCP 8.5. This result reveals an apparent upshift of the IDF curves in the future periods, which can also be seen in CTC and PPC.

Other characteristics of extreme rainfall, besides the rainfall intensity, are also important for us to better understand the extreme rainfall but have rarely been assessed so far. For example, the density

estimation of scaling parameters (H and θ) in the historical period and future periods are shown in Figs. 4.9 and 4.10 (for CTC). Generally, large H values and small θ values are regarded as a proxy for the strong convective activities due to the large variations (usually caused by the convective rainfall) in rainfall distributions with different durations (Eqs. 3 and 4) (Innocenti et al., 2017). It is worth mentioning that the results based on RCMs without BC represent the rainfall pattern in a spatial scale of 25 km instead of the local scale (e.g., 4 km). Compared with the scaling parameters in the historical period based on observations, the parameters based on RCMs without BC are more dispersed and give smaller H and larger θ values (Fig. 4.9). Specifically, in the historical period, the results based on observations display a mode value of H being 0.93 (θ being 0.2), with the standard deviation of 0.04 (0.21) while these values are 0.77 (0.31) and 0.10 (0.80), respectively, based on RCMs without BC. This observation is also because of the abovementioned deficiency in these RCMs (Section 4.3.1). Nevertheless, RCMs with BC have apparently larger H values than RCMs without BC in the future periods, which indicates that, to a certain extent, the BC may improve the simulation of the convection process. A comparison between the historical and future periods shows that in both RCMs with BC and without BC, there were no obvious changes in these scaling parameters (a small change in θ does not impact the results (Lima et al., 2018)), which implies that from a statistical perspective, there is no apparent trend in convection rainfall in the future periods. Further investigation on the trend in convection rainfall based on the physical mechanisms will be the subject of future study.

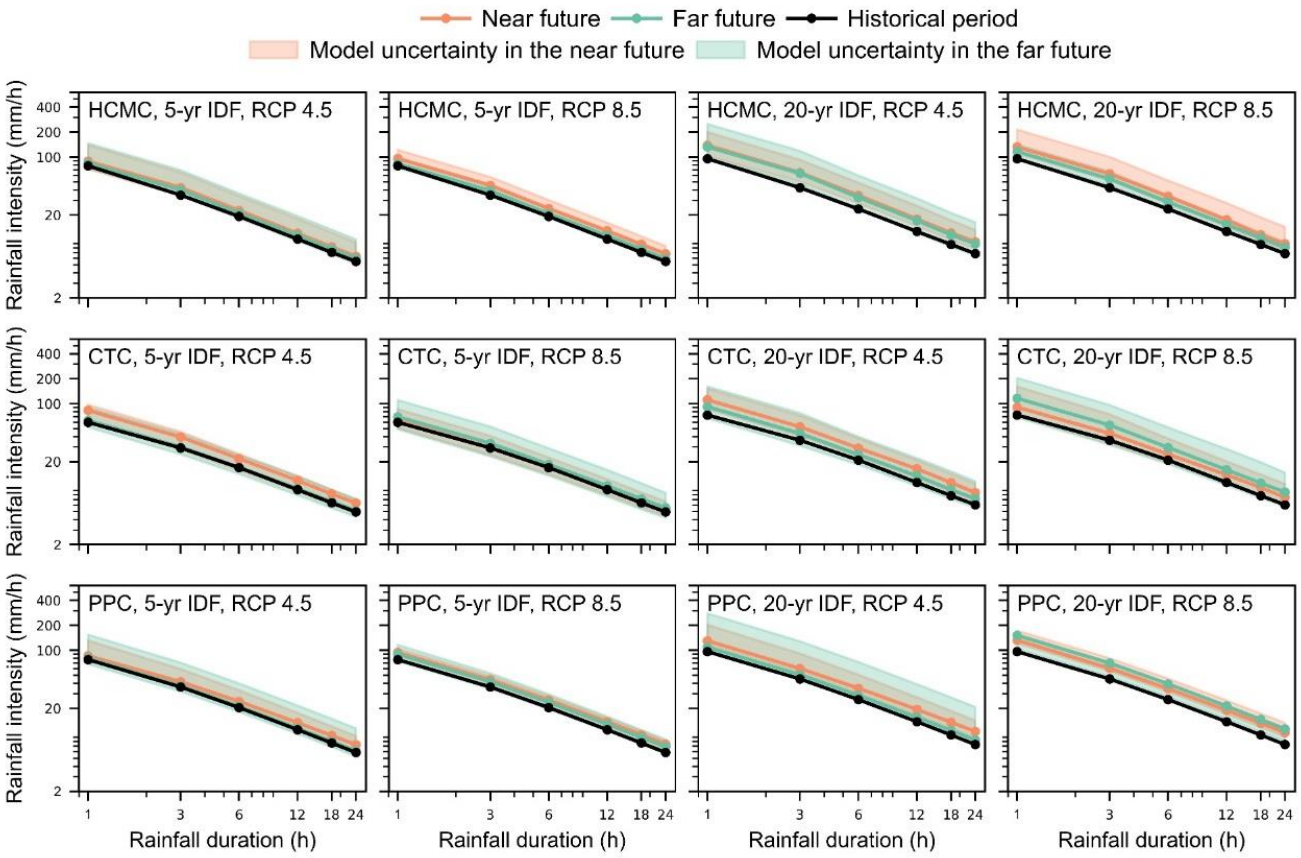


Fig. 4.8. 5- and 20-yr IDF curves in the historical period and future periods in all three cities. The rainfall intensity in the historical period was based on observations, and that in the future periods was calculated based on the ensemble median of the RCMs (with BC) by using Bayesian GEV through MCMC. The model uncertainty was based on the 90% of the ensemble spread.

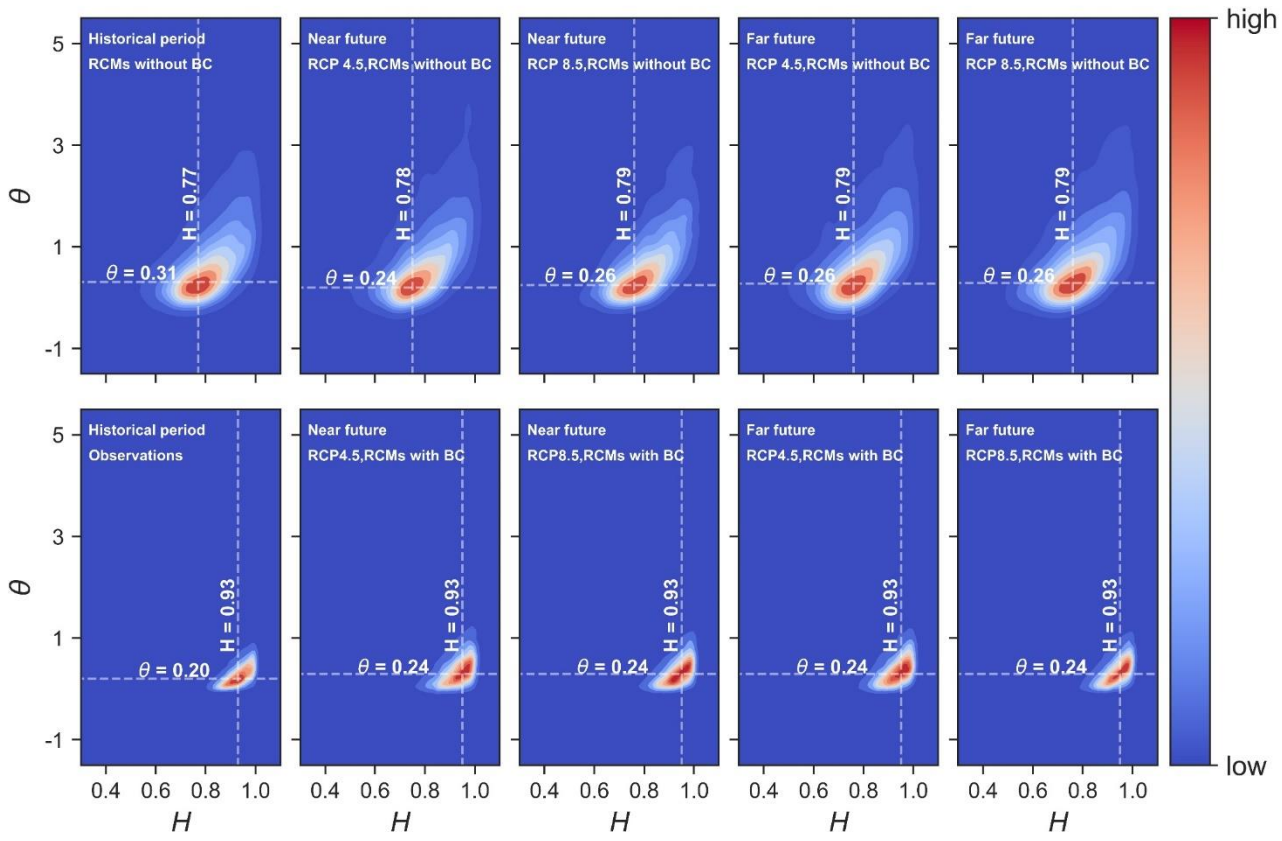


Fig. 4.9. Kernel density estimation of scaling parameters (H and θ) in the historical period and future periods in HCMC.

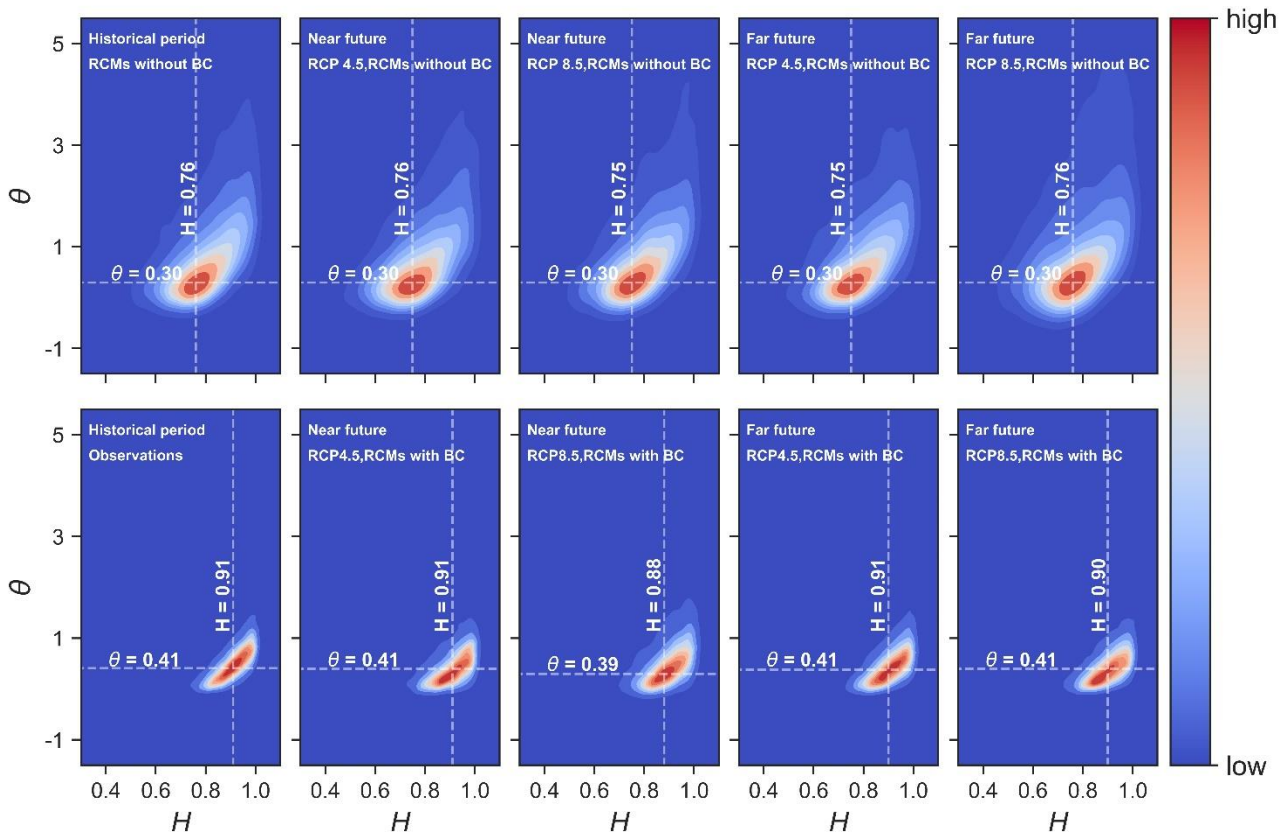


Fig. 4.10 same as Fig. 4.9 but for CTC.

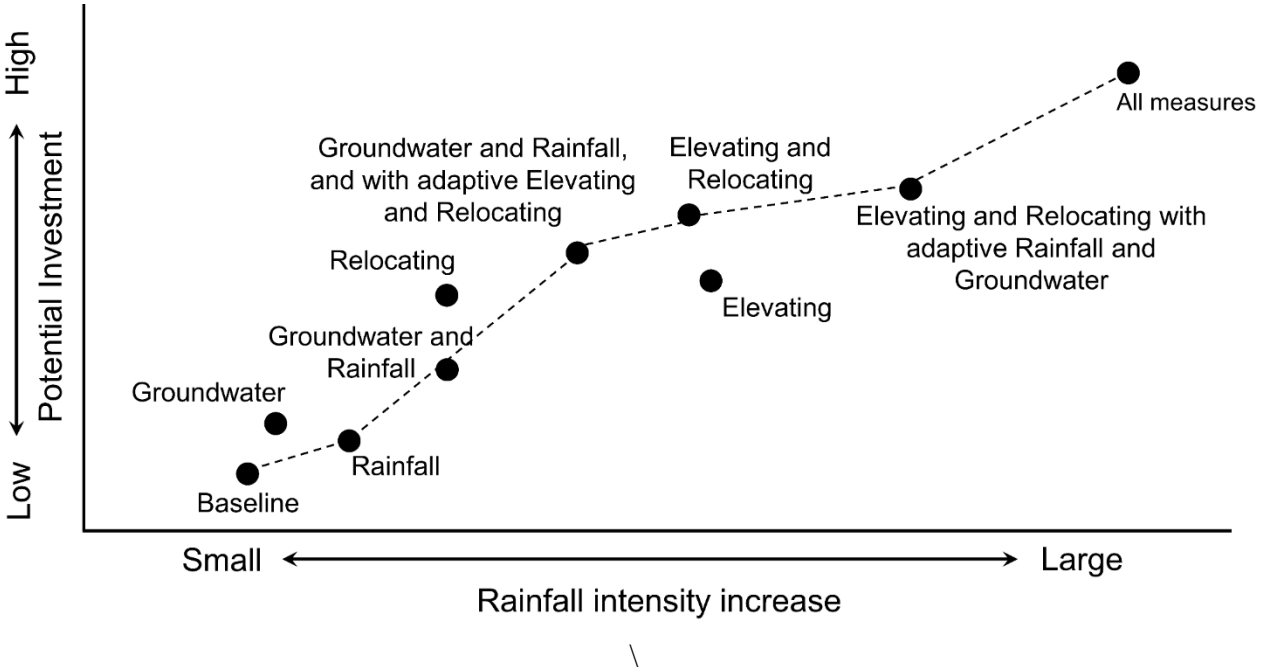
4.4. Way forward to institutional measures for climate change adaptation under deep uncertainty

The changes in extreme rainfall cannot be directly related to the changes in pluvial floods due to some other factors (e.g., imperviousness ratio and channelization) that affect the extreme rainfall responsiveness (Westra et al., 2014). However, given the rapid urbanization, the abovementioned increase in IDF curves generally can account for increasing flood risk in these urban regions. To reduce this risk, updating the existing drainage system and other flood-prevention infrastructures with the help from the World Bank, among others, is about to end or in process in these cities (JICA, 2016; GFDRR, 2016; Tran Ngoc et al.,2016). However, recent evidence shows that yearly floods (including the pluvial floods and river floods) caused by extreme rainfall continue to cause serious damage in these cities (https://floodlist.com/tag/cambodia). For example, at least 1251 households in PPC were affected by

the flood events in October 2020 (HRF, 2020). The failure in preventing the flood risk is, to a certain extent, owing to the aforementioned deep uncertainty in the projected extreme rainfall and its associated floods. To efficiently and effectively reduce this flood risk, it is necessary to reduce the uncertainty as quantified in this chapter, especially the uncertainty from climate models. Although using high-accuracy convection-permitting models could help address this problem, it is computationally expensive, requires substantial expertise from researchers with different fields, and needs further validation before their application for future projections (Kendon et al., 2017; Berg et al., 2019).

If uncertainty is difficult to reduce by the current technology, then one alternative adaptation strategy to accordingly incorporate the uncertainty can be envisaged. For example, adopting a flexible and adaptive flooding prevention strategy, based on an approach for decision making under deep uncertainty, is being increasingly applied (Marchau et al., 2019). This strategy characterizes the vulnerability (failing to meet the policymakers' needs, e.g., minimizing the worst outcome of substantial importance) of candidate adaptation measures under multiple future scenarios (usually hundreds to thousands of future conditions, owing to the deep uncertainty), and identifying new options to reduce the vulnerability (Marchau et al., 2019). An example of this strategy for reducing the flood risk in HCMC, proposed by Rand Corporation and other partners (Lempert et al., 2013), involving various stakeholders is conceptually shown in Fig. 4.11. However, this strategy is qualitative and just for the illustration purpose, and therefore needs a quantitative assessment before a pragmatic application. The main idea is that some measures like groundwater replenishment can be implemented immediately to reduce the flood risk, while some effective but costly measures like elevating buildings and relocating residents in the flood-prone areas should be implemented only if needed. Specifically, when higher flood risk is quantified based on some observed evidence, these costly measures can be implemented to reduce the flood risk and minimize the investment simultaneously. Thus, this kind of dynamic strategy should be iteratively updated based on additional new information, e.g., more advanced climate models under the Shared Socioeconomic Pathways and feedbacks from the success rate of this strategy. To this end, our newly found results in HCMC (CTC and PPC) could provide a new perspective for updating (proposing) this strategy, which can be a subject of future research. However, the main challenge of this strategy remains that it requires great computing power and much

information on the demographic and socio-economic conditions of the target area (Lempert et al., 2013).



\

Fig. 4.11. A qualitative tradeoff plot of increase in the rainfall intensity, possible preventive measures, and various associated potential investments (adapted from Lempert et al., 2013). (Baseline: Master Plan adopted by JICA in 1999; Groundwater: groundwater replenishment & Baseline; Rainfall: rainfall capture & Baseline; Relocating: relocating residents in flood-prone areas & Baseline; Elevating: elevating some buildings in flood-prone areas & Baseline; All measures: Implementing all measures immediately & Baseline; Adaptive measures (e.g., elevating) implementing such measure when needed).

4.5. Summary

In this chapter, by developing IDF curves, we quantified the impacts of different uncertainty sources associated with the choices of different RCMs, RCP scenarios (i.e., RCP 4.5 and 8.5), parameter estimation methods (i.e., MCMC, L-moment, and MLE), and with and without BC on the rainfall intensity change, followed by focusing on potential countermeasures to mitigate the multidimensional adverse impacts of global warming under the deep uncertainties. The main conclusions are listed as follows:

(1) A comparison between the RCMs with and without BC revealed that BC can efficiently remove the model bias in the historical and future periods. The changes of rainfall intensity with BC were consistent with the literature and the underlying physical mechanisms (i.e., more increase in higher return periods), supporting the application of BC before developing IDF curves.

(2) RCMs accounted for the largest uncertainty in the rainfall intensity change, followed by the BC. RCP scenarios and parameter estimation methods provided the comparable and lowest uncertainty. Moreover, these uncertainties (signified by the variance) were by and large proportional to the increasing return period.

(3) Due to the deep uncertainty, one may hold the opinion that decision-making should be delayed until sufficient knowledge (e.g., convection-permitting models) becomes available. However, deep uncertainty should not be the reason under the condition of increasing urban flood risk, and it is more reasonable to have a flexible and adaptive strategy that can evolve over time with regard to new information, e.g., feedback from the current strategy. For the future study, more high-accuracy climate models such as CPMs will be explored, and the physical reasons for these observed changes in the IDF curves should be focused on.

5. Conclusion and future study

5.1 Conclusion

This study investigated the global warming-induced extreme rainfall changes in Southeast Asia through the establishment of IDF curves, proposed specific adaptation strategies to reduce the potential risks associated with the changes, and further quantified the uncertainty in the development of IDF curves.

In chapter 2, we proposed a framework to develop the IDF curves in data-sufficient and -scarce cities by using a novel ANN model for temporal disaggregation, selected BC methods for spatial downscaling, and the scaling properties of extreme rainfall for interpolation. Rainfall intensity for rarer rainfall events is found to increase more in these cities. However, the large uncertainty arising from the climate models requires us to update the results by using more high-accuracy models. And homogeneity assumption needs further investigation in the next step.

In chapter 3, we developed the IDF curves in a data-scarce country-Cambodia mainly by using

reanalysis datasets and climate model datasets, and further quantified the warming–induced flood risks based on the IDF information as well as various socioeconomic factors. The detailed summaries are as follows.

1. A temporal disaggregation model (CANN), designed to simulate the heavy rainfall pattern based on a range of atmospheric variables, was proposed to reduce the uncertainty arising from stationarity assumption.

2. MSS from CANN combined with ZIGEV outperforms AMS coupled with GEV in reducing the uncertainty from biased disaggregated rainfall.

3. Flood risk change maps were developed based on a novel flood risk change index.

4. Southeastern and Northwestern Cambodia are prioritized for adaptation strategies.

In chapter 4, impacts of different uncertainty sources on updating the IDF curves were evaluated. Applying BC methods to remove the climate model bias is suggested although it can introduce some uncertainties. RCMs account for the largest uncertainty in the rainfall intensity change, followed by the BC. RCP scenarios and parameter estimation methods. Dynamic adaptation that can be iteratively updated based on new information is needed under deep uncertainty.

These Southeast Asian cities and countries are projected to face heavier rainfall events in the future, as revealed by the rainfall intently increase, requiring the governing agencies in our study area to update the infrastructure design. Our proposed methods (ANN, CANN, NQM) and approaches (ZIGEV&MMS, selected BC methods) can be easily validated and then applied in other southeast Asia countries (especially for the data-limit region like Cambodia) owing to the readily accessible global datasets (i.e., ERA5, CORDEX) and simplicity of these methods and approaches.

5.2 Future study

Quantification of the climate change impacts on extreme rainfall and the associated risk needs reducing the uncertainty from the datasets and methods, listed as below.

Uncertainty from datasets

More enhanced datasets are needed, including the climate model datasets from CMIP6, combination of different reanalysis datasets and satellite datasets.

Stationarity assumption of data-driven models

Physics needs considering in the data-driven architecture to reduce the stationarity assumption of the temporal disaggregation models.

Multi-scale spatial downscaling

The relationship among different variables (precipitation & temperature) and different locations (study area & nearby regions) needs considering in the spatial downscaling.

References

Abhishek, Kinouchi, T., Sayama, T., 2021. A comprehensive assessment of water storage dynamics and hydroclimatic extremes in the Chao Phraya River Basin during 2002–2020. J. Hydrol. 603, 126868, doi:10.1016/J.JHYDROL.2021.126868.

Aalbers, E.E., Lenderink, G., van Meijgaard, E., van den Hurk, B.J.J.M., 2018. Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability? Clim. Dyn. 50, 4745–4766. https://doi.org/10.1007/s00382-017-3901-9

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen, H., Nguyen, V.-T.-V., 2013. Impacts of climate change on rainfall extremes and urban drainage systems: a review. Water Sci. Technol. 68, 16–28. https://doi.org/10.2166/wst.2013.251

BDLINK, 2017. Agriculture and agro-processing sector in Cambodia. http://www.ukabc.org.uk/wp-content/uploads/2017/04/AgriProject_Reporting_FINAL-VERSION-copy.pdf

Berg, P., Christensen, O.B., Klehmet, K., Lenderink, G., Olsson, J., Teichmann, C., Yang, W., 2019. Summertime precipitation extremes in a EURO-CORDEX 0.11° ensemble at an hourly resolution. Nat. Hazards Earth Syst. Sci. 19, 957–971. https://doi.org/10.5194/nhess-19-957-2019

Blanchet, J., Ceresetti, D., Molinié, G., Creutin, J.D., 2016. A regional GEV scale-invariant framework for Intensity–Duration–Frequency analysis. J. Hydrol. 540, 82–95. https://doi.org/10.1016/j.jhydrol.2016.06.007

Burian, S.J., Durrans, S.R., 2002. Evaluation of an artificial neural network rainfall disaggregation model. Water Sci. Technol. 45, 99–104. https://doi.org/10.2166/wst.2002.0033

Burian, S.J., Durrans, S.R., Tomić, S., Pimmel, R.L., Wai, C.N., 2000. Rainfall Disaggregation Using Artificial Neural Networks. J. Hydrol. Eng. 5, 299–307. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:3(299)

Cannon, A.J., Innocenti, S., 2019. Projected intensification of sub-daily and daily rainfall extremes in convection-permitting climate model simulations over North America: implications for future intensity–duration–frequency curves. Nat. Hazards Earth Syst. Sci. 19, 421–440. https://doi.org/10.5194/nhess-19-421-2019

Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile

mapping: How well do methods preserve changes in quantiles and extremes? J. Clim. 28, 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Cannon, A.J., 2018. Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Clim. Dyn. 50, 31–49. https://doi.org/10.1007/s00382-017-3580-6

Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013a. Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour. Res. 49, 4187–4205. https://doi.org/10.1002/wrcr.20331

Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013b. Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. J. Hydrol. 479, 200–214. https://doi.org/10.1016/j.jhydrol.2012.11.062

Chen, J., Chen, H., Guo, S., 2018. Multi-site precipitation downscaling using a stochastic weather generator. Clim. Dyn. 50, 1975–1992. https://doi.org/10.1007/s00382-017-3731-9

Chen, A., Giese, M., Chen, D., 2020. Flood impact on Mainland Southeast Asia between 1985 and 2018—The role of tropical cyclones. J. Flood Risk Manag. 13, 83–89. https://doi.org/10.1111/jfr3.12598

Chen, J., Zhang, X.J., Brissette, F.P., 2014. Assessing scale effects for statistically downscaling precipitation with GPCC model. Int. J. Climatol. 34, 708–727. https://doi.org/10.1002/joc.3717

Cheng, L., AghaKouchak, A., Gilleland, E., Katz, R.W., 2014. Non-stationary extreme value analysis in a changing climate. Clim. Change 127, 353–369. https://doi.org/10.1007/s10584-014-1254-5

Cheng, M., Fang, F., Kinouchi, T., Navon, I.M., Pain, C.C., 2020. Long lead-time daily and monthly streamflow forecasting using machine learning methods. J. Hydrol. 590, 125376. https://doi.org/10.1016/j.jhydrol.2020.125376

Choubin, B., Khalighi-Sigaroodi, S., Mishra, A., Goodarzi, M., Shamshirband, S., Ghaljaee, E., Zhang, F., 2019. A novel bias correction framework of TMPA 3B42 daily precipitation data using similarity matrix/homogeneous conditions. Sci. Total Environ. 694, 133680. https://doi.org/10.1016/j.scitotenv.2019.133680

Coles, S., Bawa, J., Trenner, L., Dorazio, P., 2001. An Introduction to Statistical Modeling of Extreme Values, vol. 208. Springer

Courty, L.G., Wilby, R.L., Hillier, J.K., Slater, L.J., 2019. Intensity-duration-frequency curves at the global scale. Environ. Res. Lett. 14, 084045. https://doi.org/10.1088/1748-9326/ab370a.

Doyle, S.E., 2012. City of Water: Architecture, Urbanism and the Floods of Phnom Penh. Nakhara : Journal of Environmental Design and Planning, 8, 135-154. https://ph01.tci-thaijo.org/index.php/nakhara/article/view/104950

Endreny, T.A., Imbeah, N., 2009. Generating robust rainfall intensity-duration-frequency estimates with short-record satellite data. J. Hydrol. 371, 182–191. https://doi.org/10.1016/j.jhydrol.2009.03.027

Fischer, E.M., Knutti, R., 2015. Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. Nat. Clim. Chang. 5, 560–564. https://doi.org/10.1038/nclimate2617

Fischer, E.M., Sedláček, J., Hawkins, E., Knutti, R., 2014. Models agree on forced response pattern of precipitation and temperature extremes. Geophys. Res. Lett. 41, 8554–8562. https://doi.org/10.1002/2014GL062018

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 9, edited by: Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., Cambridge University Press, Cambridge, UK and New York, NY, USA, 741–866, 2013

Flower, B., Fortnam, M., 2015. Urbanising disaster risks: vulnerability of the urban poor in Cambodia to flooding and other hazards. https://www.preventionweb.net/publications/view/47109

Fowler, H.J., Lenderink, G., Prein, A.F., Westra, S., Allan, R.P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Do, H.X., Guerreiro, S., Haerter, J.O., Kendon, E.J., Lewis, E., Schaer, C., Sharma, A., Villarini, G., Wasko, C., Zhang, X., 2021. Anthropogenic intensification of short-duration rainfall extremes. Nat. Rev. Earth Environ. https://doi.org/10.1038/s43017-020-00128-6

Ge, F., Zhu, S., Peng, T., Zhao, Y., Sielmann, F., Fraedrich, K., Zhi, X., Liu, X., Tang, W., Ji, L., 2019. Risks of precipitation extremes over Southeast Asia: Does 1.5 °c or 2 °c global warming make a difference? Environ. Res. Lett. 14. https://doi.org/10.1088/1748-9326/aaff7e

Germanwatch, 2021. Global climate risk index 2021.https://www.developmentaid.org/api/frontend/cms/file/2021/03/Global-Climate-Risk-Index-2021_1.pdf

Giorgi, F., Jones, C., Asrar, G.R., others, 2009. Addressing climate information needs at the regional level: the CORDEX framework, World Meteorological Organization (WMO) Bulletin. Available from < https://public.wmo.int/en/bulletin/addressing-climate-information-needs-regional-level-cordex-framework>

Global Facility for Disaster Reduction and Recovery (GFDRR), 2016. Can Tho Urban Development & Resilience.https://www.gfdrr.org/sites/default/files/Can%20Tho%20Urban%20Development%20and%20Resilience%20-%20Land%20Sale%20and%20Development%20Rights.pdf

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen-Skaugen, T., 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. Hydrol. Earth Syst. Sci. 16, 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Gutowski, W.J., Giorgi, F., 2020. Coordination of Regional Downscaling. Oxford Res. Encycl. Clim. Sci. https://doi.org/10.1093/acrefore/9780190228620.013.658

Hausfather, Z., Peters, G.P., 2020. Emissions – the 'business as usual' story is misleading. Nature 577, 618–620. https://doi.org/10.1038/d41586-020-00177-3

Hawkins, E., Sutton, R., 2009. The Potential to Narrow Uncertainty in Regional Climate Predictions. Bull. Am. Meteorol. Soc. 90, 1095–1108. https://doi.org/10.1175/2009BAMS2607.1

He, X., Chaney, N.W., Schleiss, M., Sheffield, J., 2016. Spatial downscaling of precipitation using adaptable random forests. Water Resour. Res. 52, 8217–8237. https://doi.org/10.1002/2016WR019034

Hewitson, B., Crane, R., 1996. Climate downscaling: techniques and application. Clim. Res. 7, 85–95. https://doi.org/10.3354/cr007085

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803

Hijioka, Y., Lin, E., Pereira, J.J., Corlett, R.T., Cui, X., Insarov, G.E., Lasco, R.D., Lindgren, E., Surjan, A., 2014. Asia. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York. Available from <https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-Chap24_FINAL.pdf>.

Hnilica, J., Hanel, M., Puš, V., 2017. Multisite bias correction of precipitation data from regional climate models. Int. J. Climatol. 37, 2934–2946. https://doi.org/10.1002/joc.4890

Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Others, 2018. Impacts of 1.5ºC Global Warming on Natural and Human Systems. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, Special Report, Intergovernmental Panel on Climate Change. Available from <https://www.ipcc.ch/site/assets/uploads/sites/2/2019/02/SR15_Chapter3_Low_Res.pdf>.

Hosseinzadehtalaei, P., Tabari, H., Willems, P., 2020. Climate change impact on short-duration extreme precipitation and intensity–duration–frequency curves over Europe. J. Hydrol. 590, 125249. https://doi.org/10.1016/j.jhydrol.2020.125249

Hosseinzadehtalaei, P., Tabari, H., Willems, P., 2018. Precipitation intensity–duration–frequency curves for central Belgium with an ensemble of EUROCORDEX simulations, and associated uncertainties. Atmos. Res. 200, 1–12. https://doi.org/10.1016/j.atmosres.2017.09.015

Hosseinzadehtalaei, P., Ishadi, N.K., Tabari, H., Willems, P., 2021. Climate change impact assessment on pluvial flooding using a distribution-based bias correction of regional climate model simulations. J. Hydrol. 598, 126239. https://doi.org/10.1016/j.jhydrol.2021.126239

Huong, H.T.L., Pathirana, A., 2013. Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. Hydrol. Earth Syst. Sci. 17, 379–394. https://doi.org/10.5194/hess-17-379-2013

Humanitarian Response Forum (HRF) ,2020. Situation Report No. 2 – Floods in Cambodia. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/hrf_sitrep_no2_14-oct-20.pdf

Innocenti, S., Mailhot, A., Frigon, A., 2017. Simple scaling of extreme precipitation in North America. Hydrol. Earth Syst. Sci. 21, 5823–5846. https://doi.org/10.5194/hess-21-5823-2017

Irvine, K., Sovann, C., Suthipong, S., Kok, S., Chea, E., 2015. Application of PCSWMM to Assess Wastewater Treatment and Urban Flooding Scenarios in Phnom Penh, Cambodia: A Tool to Support Eco-City Planning. J. Water Manag. Model. 87, 104061. https://doi.org/10.14796/JWMM.C389

Ivanov, M.A., Kotlarski, S., 2017. Assessing distribution-based climate model bias correction methods over an alpine domain: added value and limitations. Int. J. Climatol. 37, 2633–2653. https://doi.org/10.1002/joc.4870

Japan International Cooperation Agency (JICA), 2015. Country Report: Cambodia. https://openjicareport.jica.go.jp/pdf/1000023400.pdf

Japan International Cooperation Agency (JICA), 2016, The study on drainage and sewerage improvement project in Phnom Penh metropolitan area. https://openjicareport.jica.go.jp/pdf/12270294_01.pdf

Ji, X., Li, Y., Luo, X., He, D., Guo, R., Wang, J., Bai, Y., Yue, C., Liu, C., 2020. Evaluation of bias correction methods for APHRODITE data to improve hydrologic simulation in a large Himalayan basin. Atmos. Res. 242, 104964. https://doi.org/10.1016/j.atmosres.2020.104964

Johnson, F., Sharma, A., 2012. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. Water Resour. Res. 48, 1–16. https://doi.org/10.1029/2011WR010464

Kendon, E.J., Ban, N., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C., Evans, J.P., Fosser, G., Wilkinson, J.M., 2017. Do Convection-Permitting Regional Climate Models Improve Projections of Future Precipitation Change? Bull. Am. Meteorol. Soc. 98, 79–93. https://doi.org/10.1175/BAMS-D-15-0004.1

Kharin, V. V., Flato, G.M., Zhang, X., Gillett, N.P., Zwiers, F., Anderson, K.J., 2018. Risks from Climate Extremes Change Differently from 1.5°C to 2.0°C Depending on Rarity. Earth's Futur. 6, 704–715. https://doi.org/10.1002/2018EF000813

Kossieris, P., Makropoulos, C., Onof, C., Koutsoyiannis, D., 2018. A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures. J. Hydrol. 556, 980–992. https://doi.org/10.1016/j.jhydrol.2016.07.015

Kourtis, I.M., Tsihrintzis, V.A., 2021. Adaptation of urban drainage networks to climate change: A review. Sci. Total Environ. 771, 145431. https://doi.org/10.1016/j.scitotenv.2021.145431

Koutsoyiannis, D., 2003. Rainfall disaggregation methods: Theory and applications, in: Proceedings,Workshop on Statistical and Mathematical Methods for Hydrological Analysis. Università Degli Studi Di Roma ‘‘La Sapienza, Rome, pp. 1–23. https://doi.org/https://doi.org/10.13140/RG.2.1.2840.8564

Koutsoyiannis, D., Kozonis, D., Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. J. Hydrol. 206, 118–135. https://doi.org/10.1016/S0022-1694(98)00097-3

Koutsoyiannis, D., Onof, C., 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. J. Hydrol. 246, 109–122. https://doi.org/10.1016/S0022-1694(01)00363-8

Kumar, R., Goel, N.K., Chatterjee, C., Nayak, P.C., 2015. Regional Flood Frequency Analysis using Soft Computing Techniques. Water Resour. Manag. 29, 1965–1978. https://doi.org/10.1007/s11269-015-0922-1

Kundzewicz, Z.W., Kanae, S., Seneviratne, S.I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L.M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G.R., Kron, W., Benito, G., Honda, Y., Takahashi, K., Sherstyukov, B., 2014. Flood risk and climate change: global and regional perspectives. Hydrol. Sci. J. 59, 1–28. https://doi.org/10.1080/02626667.2013.857411

Lambert, D., 1992. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing.

Technometrics 34, 1. https://doi.org/10.2307/1269547

Lempert, R.J., Kalra, N., Peyraud, S., Mao, Z., Tan, S.B., Cira, D., Lotsch, A., 2013. Ensuring Robust Flood Risk Management in Ho Chi Minh City, World Bank. https://openknowledge.worldbank.org/handle/10986/15603

Lenderink, G., Barbero, R., Loriaux, J.M., Fowler, H.J., 2017. Super-Clausius–Clapeyron Scaling of Extreme Hourly Convective Precipitation and Its Relation to Large-Scale Atmospheric Conditions. J. Clim. 30, 6037–6052. https://doi.org/10.1175/JCLI-D-16-0808.1

Li, F., Rosa, D., Collins, W.D., Wehner, M.F., 2012. "Super-parameterization": A better way to simulate regional extreme precipitation? J. Adv. Model. Earth Syst. 4, n/a-n/a. https://doi.org/10.1029/2011MS000106

Li, J., Johnson, F., Evans, J., Sharma, A., 2017. A comparison of methods to estimate future sub-daily design rainfall. Adv. Water Resour. 110, 215–227. https://doi.org/10.1016/j.advwatres.2017.10.020

Li, X., Babovic, V., 2019. Multi-site multivariate downscaling of global climate model outputs: an integrated framework combining quantile mapping, stochastic weather generator and Empirical Copula approaches. Clim. Dyn. 52, 5775–5799. https://doi.org/10.1007/s00382-018-4480-0

Li, Z., Shi, X., Li, J., 2017. Multisite and multivariate GCM downscaling using a distribution-free shuffle procedure for correlation reconstruction. Clim. Res. 72, 141–151. https://doi.org/10.3354/cr01460

Li, Y., Jones, D.B.A., Zarzycki, C.M., 2019. Evaluating the Performance of VR-CESM for Modeling Precipitation in Southeast Asia, In: GESM Working Group Meetings. Available from <https://www.cesm.ucar.edu/events/wg-meetings/2019/presentations/AMWGChemWAWG/li.pdf>

Liew, S.C., Raghavan, S. V., Liong, S.Y., 2014. How to construct future IDF curves, under changing climate, for sites with scarce rainfall records? Hydrol. Process. 28, 3276–3287. https://doi.org/10.1002/hyp.9839

Lima, C.H.R., Kwon, H.H., Kim, Y.T., 2018. A local-regional scaling-invariant Bayesian GEV model for estimating rainfall IDF curves in a future climate. J. Hydrol. 566. https://doi.org/10.1016/j.jhydrol.2018.08.075

Lin, L., Xu, Y., Wang, Z., Diao, C., Dong, W., Xie, S.-P., 2018. Changes in Extreme Rainfall Over India and China Attributed to Regional Aerosol-Cloud Interaction During the Late 20th Century Rapid Industrialization. Geophys. Res. Lett. 45, 7857–7865. https://doi.org/10.1029/2018GL078308

Llasat, M.-C., 2001. An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of spain. Int. J. Climatol. 21, 1385–1400. https://doi.org/10.1002/joc.692

Lombardo, F., Volpi, E., Koutsoyiannis, D., Serinaldi, F., 2017. A theoretically consistent stochastic cascade for temporal disaggregation of intermittent rainfall. Water Resour. Res. 53, 4586–4605. https://doi.org/10.1002/2017WR020529

Loriaux, J.M., Lenderink, G., De Roode, S.R., Siebesma, A.P., 2013. Understanding Convective Extreme Precipitation Scaling Using Observations and an Entraining Plume Model. J. Atmos. Sci. 70, 3641–3655. https://doi.org/10.1175/JAS-D-12-0317.1

Lu, Y., Qin, X.S., 2014. Multisite rainfall downscaling and disaggregation in a tropical urban area. J. Hydrol. 509, 55–65. https://doi.org/10.1016/j.jhydrol.2013.11.027

Lu, Y., Qin, X.S., Xie, Y.J., 2016. An integrated statistical and data-driven framework for supporting flood risk analysis under climate change. J. Hydrol. 533, 28–39. https://doi.org/10.1016/j.jhydrol.2015.11.041

Lyna, K., 2019. An economic assessment of urban flooding in Cambodia: A case study of Phnom Penh. Cambodia J. Basic Appl. Res. 1, 125–149. http://www.rupp.edu.kh/CJBAR/files/Vol-1-Issue-1/6-KHAN-2019.pdf

Mamoon, A. Al, Joergensen, N.E., Rahman, A., Qasem, H., 2016. Design rainfall in Qatar: sensitivity to climate change scenarios. Nat. Hazards 81, 1797–1810. https://doi.org/10.1007/s11069-016-2156-9

Macukow, B., 2016. Neural networks--state of art, brief history, basic models and architecture, in: IFIP International Conference on Computer Information Systems and Industrial Management. pp. 3–14. Available from https://hal.inria.fr/hal-01637477/document

Maraun, D., 2016. Bias Correcting Climate Change Simulations - a Critical Review. Curr. Clim. Chang. Reports 2, 211–220. https://doi.org/10.1007/s40641-016-0050-x

Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. J. Clim. 26, 2137–2143. https://doi.org/10.1175/JCLI-D-12-00821.1

Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M., Thiele-Eich, I., 2010. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Rev. Geophys. 48, RG3003. https://doi.org/10.1029/2009RG000314

Maraun, D., Widmann, M., Gutiérrez, J.M., 2019. Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. Int. J. Climatol. 39, 3692–3703. https://doi.org/10.1002/joc.5877

Marchau, V., Walker, W., Bloemen, P., Popper, S., 2019. Decision Making Under Deep Uncertainty. From Theory to Practice, Springer.

Marra, F., Morin, E., 2015. Use of radar QPE for the derivation of Intensity-Duration-Frequency curves in a range of climatic regimes. J. Hydrol. 531, 427–440. https://doi.org/10.1016/j.jhydrol.2015.08.064

Marra, F., Morin, E., Peleg, N., Mei, Y., Anagnostou, E.N., 2017. Intensity-duration-frequency curves from remote sensing rainfall estimates: Comparing satellite and weather radar over the eastern Mediterranean. Hydrol. Earth Syst. Sci. 21, 2389–2404. https://doi.org/10.5194/hess-21-2389-2017

Maurer, E.P., Pierce, D.W., 2014. Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. Hydrol. Earth Syst. Sci. 18, 915–925. https://doi.org/10.5194/hess-18-915-2014

Masood, E., Tollefson, J., 2021. ‘COP26 hasn’t solved the problem’: scientists react to UN climate deal. Nature 599, 355–356. https://doi.org/10.1038/d41586-021-03431-4

Mehrotra, R., Sharma, A., 2016. A multivariate quantile-matching bias correction approach with auto- and cross-dependence across multiple time scales: implications for downscaling. J. Clim. 29, 3519–3539. https://doi.org/10.1175/JCLI-D-15-0356.1

Mehrotra, R., Sharma, A., 2015. Correcting for systematic biases in multiple raw GCM variables across a range of timescales. J. Hydrol. 520, 214–223.

https://doi.org/10.1016/j.jhydrol.2014.11.037

Mélèse, V., Blanchet, J., Molinié, G., 2018. Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis. J. Hydrol. 558, 579–591. https://doi.org/10.1016/j.jhydrol.2017.07.054

Mendenhall, J., Meiler, J., 2016. Improving quantitative structure–activity relationship models using Artificial Neural Networks trained with dropout. J. Comput. Aided. Mol. Des. 30, 177–189. https://doi.org/10.1007/s10822-016-9895-2

Mialhe, F., Gunnell, Y., Navratil, O., Choi, D., Sovann, C., Lejot, J., Gaudou, B., Se, B., Landon, N., 2019. Spatial growth of Phnom Penh, Cambodia (1973–2015): Patterns, rates, and socio-ecological consequences. Land use policy 87, 104061. https://doi.org/10.1016/j.landusepol.2019.104061

Mirhosseini, G., Srivastava, P., Fang, X., 2014. Developing Rainfall Intensity-Duration-Frequency Curves for Alabama under Future Climate Scenarios Using Artificial Neural Networks. J. Hydrol. Eng. 19, 04014022. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000962

Mirhosseini, G., Srivastava, P., Stefanova, L., 2013. The impact of climate change on rainfall Intensity-Duration-Frequency (IDF) curves in Alabama. Reg. Environ. Chang. 13, 25–33. https://doi.org/10.1007/s10113-012-0375-5

Moustakis, Y., Papalexiou, S.M., Onof, C.J., Paschalis, A., 2021. Seasonality, Intensity, and Duration of Rainfall Extremes Change in a Warmer Climate. Earth's Futur. 9, 1–15. https://doi.org/10.1029/2020EF001824

Mullan, D., Chen, J., Zhang, X.J., 2016. Validation of non-stationary precipitation series for site-specific impact assessment: comparison of two statistical downscaling techniques. Clim. Dyn. 46, 967–986. https://doi.org/10.1007/s00382-015-2626-x

Muller, A., Bacro, J.N., Lang, M., 2008. Bayesian comparison of different rainfall depth-duration-frequency relationships. Stoch. Environ. Res. Risk Assess. 22, 33–46. https://doi.org/10.1007/s00477-006-0095-9

Müller, H., Haberlandt, U., 2018. Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. J. Hydrol. 556, 847–864. https://doi.org/10.1016/j.jhydrol.2016.01.031

Myhre, G., Alterskjær, K., Stjern, C.W., Hodnebrog, Ø., Marelle, L., Samset, B.H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., Stohl, A., 2019. Frequency of extreme precipitation increases extensively with event rareness under global warming. Sci. Rep. 9, 16063. https://doi.org/10.1038/s41598-019-52277-4

National Institute of Statistics (NIS), Ministry of Planning Phnom Penh, Cambodia, 2012. Economic Census of Cambodia

2011. https://www.nis.gov.kh/nis/EC2011/EC2011_Final_Results_Revised.pdf

National Institute of Statistics (NIS), Ministry of Planning Phnom Penh, Cambodia, 2019. The General Population Census of Cambodia 2019. https://www.nis.gov.kh/nis/Census2019/Provisional%20Population%20Census%202019_English_FINAL.pdf

Ngai, S.T., Juneng, L., Tangang, F., Chung, J.X., Salimun, E., Tan, M.L., Amalia, S., 2020. Future projections of Malaysia daily precipitation characteristics using bias correction technique. Atmos. Res. 240, 104926. https://doi.org/10.1016/j.atmosres.2020.104926

Nguyen-Thi, T., Ngo-Duc, T., Tangang, F.T., Cruz, F., Juneng, L., Santisirisomboon, J., Aldrian, E., Phan-Van, T., Narisma, G., 2021. Climate analogue and future appearance of novel climate in Southeast Asia. Int. J. Climatol. 41, E392–E409. https://doi.org/10.1002/joc.6693

Ning, L., Bradley, R.S., 2016. NAO and PNA influences on winter temperature and precipitation over the eastern United States in CMIP5 GCMs. Clim. Dyn. 46, 1257–1276. https://doi.org/10.1007/s00382-015-2643-9

Ning, L., Riddle, E.E., Bradley, R.S., 2015. Projected Changes in Climate Extremes over the Northeastern United States. J. Clim. 28, 3289–3310. https://doi.org/10.1175/JCLI-D-14-00150.1

Nissen, K.M., Ulbrich, U., 2017. Increasing frequencies and changing characteristics of heavy precipitation events threatening infrastructure in Europe under climate change. Nat. Hazards Earth Syst. Sci. 17 (7), 1177–1190. https://doi.org/10.5194/nhess-17-1177-2017

Noor, M., Ismail, T., Chung, E.-S., Shahid, S., Sung, J., 2018. Uncertainty in Rainfall Intensity Duration Frequency Curves of Peninsular Malaysia under Changing Climate Scenarios. Water 10, 1750. https://doi.org/10.3390/w10121750

Noor, M., Ismail, T., Shahid, S., Asaduzzaman, M., Dewan, A., 2021. Evaluating intensity-duration-

frequency (IDF) curves of satellite-based precipitation datasets in Peninsular Malaysia. Atmos. Res. 248, 105203. https://doi.org/10.1016/j.atmosres.2020.105203

Nourani, V., Farboudfam, N., 2019. Rainfall time series disaggregation in mountainous regions using hybrid wavelet-artificial intelligence methods. Environ. Res. 168, 306–318. https://doi.org/10.1016/j.envres.2018.10.012

O'Grady, 2021. The new climate pact is more ambitious. But hopes dim for limiting warming to 1.5°C. https://doi.org/10.1126/science.acx9667

Ombadi, M., Nguyen, P., Sorooshian, S., Hsu, K. lin, 2018. Developing Intensity-Duration-Frequency (IDF) Curves From Satellite-Based Precipitation: Methodology and Evaluation. Water Resour. Res. 54, 7752–7766. https://doi.org/10.1029/2018WR022929

Quadros Gramosa, A.H., Ferraz do Nascimento, F., Castro Morales, F.E., 2019. A Bayesian approach to zero-inflated data in extremes. Commun. Stat. - Theory Methods 0, 1–12. https://doi.org/10.1080/03610926.2019.1594305

Ouali, D., Cannon, A.J., 2018. Estimation of rainfall intensity–duration–frequency curves at ungauged locations using quantile regression methods. Stoch. Environ. Res. Risk Assess. 32, 2821–2836. https://doi.org/10.1007/s00477-018-1564-7

Panchal, G., Ganatra, A., Kosta, Y.P., Panchal, D., 2010. Searching Most Efficient Neural Network Architecture Using Akaike's Information Criterion (AIC). Int. J. Comput. Appl. 1, 54–57. https://doi.org/10.5120/126-242

Pang, Z., Shi, C., Gu, J., Pan, Y., Xu, B., 2021. Assessment of a gauge-radar-satellite merged hourly precipitation product for accurately monitoring the characteristics of the super-strong meiyu precipitation over the yangtze river basin in 2020. Remote Sens. 13. https://doi.org/10.3390/rs13193850

Papalexiou, S.M., Koutsoyiannis, D., Makropoulos, C., 2013. How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrol. Earth Syst. Sci. 17, 851–862. https://doi.org/10.5194/hess-17-851-2013

Peleg, N., Marra, F., Fatichi, S., Paschalis, A., Molnar, P., Burlando, P., 2018. Spatial variability of extreme rainfall at radar subpixel scale. J. Hydrol. 556, 922–933. https://doi.org/10.1016/j.jhydrol.2016.05.033

Pendergrass, A.G., 2018. What precipitation is extreme? Science (80-.). 360, 1072–1073. https://doi.org/10.1126/science.aat1871

Poschlod, B., Hodnebrog, Ø., Wood, R.R., Alterskjær, K., Ludwig, R., Myhre, G., Sillmann, J., 2018. Comparison and Evaluation of Statistical Rainfall Disaggregation and High-Resolution Dynamical Downscaling over Complex Terrain. J. Hydrometeorol. 19, 1973–1982. https://doi.org/10.1175/JHM-D-18-0132.1

Poschlod, B., Ludwig, R., Sillmann, J., 2021. Ten-year return levels of sub-daily extreme precipitation over Europe. Earth Syst. Sci. Data 13, 983–1003. https://doi.org/10.5194/essd-13-983-2021

Pour, S.H., Shahid, S., Chung, E.-S., Wang, X.-J., 2018. Model output statistics downscaling using support vector machine for the projection of spatial and temporal changes in rainfall of Bangladesh. Atmos. Res. 213, 149–162. https://doi.org/10.1016/j.atmosres.2018.06.006

Rajagopalan, B., Salas, J.D., Lall, U., 2010. STOCHASTIC METHODS FOR MODELING PRECIPITATION AND STREAMFLOW, in: Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting. WORLD SCIENTIFIC, pp. 17–52. https://doi.org/10.1142/9789814307987_0002

Räty, O., Räisänen, J., Ylhäisi, J.S., 2014. Evaluation of delta change and bias correction methods for future daily precipitation: intermodel cross-validation using ENSEMBLES simulations. Clim. Dyn. 42, 2287–2303. https://doi.org/10.1007/s00382-014-2130-8

Requena, A.I., Nguyen, T.H., Burn, D.H., Coulibaly, P., Nguyen, V.T. Van, 2021. A temporal downscaling approach for sub-daily gridded extreme rainfall intensity estimation under climate change. J. Hydrol. Reg. Stud. 35, 100811. https://doi.org/10.1016/j.ejrh.2021.100811

Roelfsema, M., van Soest, H.L., Harmsen, M., van Vuuren, D.P., Bertram, C., den Elzen, M., Höhne, N., Iacobuta, G., Krey, V., Kriegler, E., Luderer, G., Riahi, K., Ueckerdt, F., Després, J., Drouet, L., Emmerling, J., Frank, S., Fricko, O., Gidden, M., Humpenöder, F., Huppmann, D., Fujimori, S., Fragkiadakis, K., Gi, K., Keramidas, K., Köberle, A.C., Aleluia Reis, L., Rochedo, P., Schaeffer, R., Oshiro, K., Vrontisi, Z., Chen, W., Iyer, G.C., Edmonds, J., Kannavou, M., Jiang, K., Mathur, R., Safonov, G., Vishwanathan, S.S., 2020. Taking stock of national climate policies to evaluate implementation of the Paris Agreement. Nat. Commun. 11, 2096. https://doi.org/10.1038/s41467-020-15414-6

Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H., Quantin, G., Wilcox, C., Ndiaye, O., Diongue-Niang, A., Diop Kane, M., 2018. Intensity–duration–frequency (IDF) rainfall curves in Senegal. Nat. Hazards Earth Syst. Sci. 18, 1849–1866. https://doi.org/10.5194/nhess-18-1849-2018

Shah, S.A., Jehanzaib, M., Yoo, J., Hong, S., Kim, T.-W., 2022. Investigation of the Effects of Climate Variability, Anthropogenic Activities, and Climate Change on Streamflow Using Multi-Model Ensembles. Water 14, 512. https://doi.org/10.3390/w14040512

Sharma, A., Mehrotra, R., 2010. Rainfall generation, in: Geophysical Monograph Series. pp. 215–246. https://doi.org/10.1029/2010GM000973. Available from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2010GM000973>

Sharma, A., Srikanthan, S., 2006. Continuous rainfall simulation: a nonparametric alternative, in: 30th Hydrology \& Water Resources Symposium: Past, Present \& Future. p. 86. Available from https://search.informit.org/doi/10.3316/INFORMIT.499059223202660

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F.-J., Ganguly, S., Hsu, K.-L., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., Tsai, W.-P., 2018. HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community. Hydrol. Earth Syst. Sci. 22, 5639–5656. https://doi.org/10.5194/hess-22-5639-2018

Shrestha, B.B., Okazumi, T., Tanaka, S., Sugiura, A., Kwak, Y., Hibino, S., 2013. Development of Flood Vulnerability Indices for Lower Mekong Basin in Cambodain Floodplain. J. Japan Soc. Civ. Eng. Ser. B1 (Hydraulic Eng. 69, I_1-I_6. https://doi.org/10.2208/jscejhe.69.I_1

Silva, A.T., Portela, M.M., Naghettini, M., 2014. On peaks-over-threshold modeling of floods with zero-inflated Poisson arrivals under stationarity and nonstationarity. Stoch. Environ. Res. Risk Assess. 28, 1587–1599. https://doi.org/10.1007/s00477-013-0813-z

Srivastav, R.K., Schardong, A., Simonovic, S.P., 2014. Equidistance Quantile Matching Method for Updating IDFCurves under Climate Change. Water Resour. Manag. 28, 2539–2562. https://doi.org/10.1007/s11269-014-0626-y

Su, T., Chen, J., Cannon, A.J., Xie, P., Guo, Q., 2020. Multi-site bias correction of climate model outputs for hydro-meteorological impact studies: An application over a watershed in China. Hydrol. Process. 34, 2575–2598. https://doi.org/10.1002/hyp.13750

Supari, Tangang, F., Juneng, L., Cruz, F., Chung, J.X., Ngai, S.T., Salimun, E., Mohd, M.S.F., Santisirisomboon, J., Singhruck, P., PhanVan, T., Ngo-Duc, T., Narisma, G., Aldrian, E., Gunawan, D., Sopaheluwakan, A., 2020. Multi-model projections of precipitation extremes in Southeast Asia based on CORDEX-Southeast Asia simulations. Environ. Res. 184, 109350. https://doi.org/10.1016/j.envres.2020.109350

Switanek, M.B., Troch, P.A., Castro, C.L., Leuprecht, A., Chang, H.-I., Mukherjee, R., Demaria, E.M.C., 2017. Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes. Hydrol. Earth Syst. Sci. 21, 2649–2666. https://doi.org/10.5194/hess-21-2649-2017.

Takagi, H., Ty, T.V., Thao, N.D., Esteban, M., 2015. Ocean tides and the influence of sea-level rise on floods in urban areas of the Mekong Delta. J. Flood Risk Manag. 8, 292–300. https://doi.org/10.1111/jfr3.12094

Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., Brisson, E., Van Lipzig, N., Willems, P., 2016. Local impact analysis of climate change on precipitation extremes: are high-resolution climate models needed for realistic simulations? Hydrol. Earth Syst. Sci. 20, 3843–3857. https://doi.org/10.5194/hess-20-3843-2016

Tabari, H., 2020. Climate change impact on flood and extreme precipitation increases with water availability. Sci. Rep. 10, 1–10. https://doi.org/10.1038/s41598-020-70816-2

Tabari, H., Hosseinzadehtalaei, P., AghaKouchak, A., Willems, P., 2019. Latitudinal heterogeneity and hotspots of uncertainty in projected extreme precipitation. Environ. Res. Lett. 14, 124032. https://doi.org/10.1088/1748-9326/ab55fd

Tangang, F., Chung, J.X., Juneng, L., Supari, Salimun, E., Ngai, S.T., Jamaluddin, A.F., Mohd, M.S.F., Cruz, F., Narisma, G., Santisirisomboon, J., Ngo-Duc, T., Van Tan, P., Singhruck, P., Gunawan, D., Aldrian, E., Sopaheluwakan, A., Grigory, N., Remedio, A.R.C., Sein, D. V., Hein-Griggs, D., McGregor, J.L., Yang, H., Sasaki, H., Kumar, P., 2020. Projected future changes in rainfall in Southeast Asia based on CORDEX–SEA multi-model simulations. Clim. Dyn. 55, 1247–1267. https://doi.org/10.1007/s00382-020-05322-2

Tangang, F., Santisirisomboon, Jerasorn, Juneng, L., Salimun, E., Chung, J., Supari, S., Cruz, F., Ngai, S.T., Ngo-Duc, T., Singhruck, P., Narisma, G., Santisirisomboon, Jaruthat, Wongsaree,

W., Promjirapawat, K., Sukamongkol, Y., Srisawadwong, R., Setsirichok, D., Phan-Van, T., Aldrian, E., Gunawan, D., Nikulin, G., Yang, H., 2019. Projected future changes in mean precipitation over Thailand based on multi-model regional climate simulations of CORDEX Southeast Asia. Int. J. Climatol. 39, 5413–5436. https://doi.org/10.1002/joc.6163

Tangang, F., Supari, S., Chung, J.X., Cruz, F., Salimun, E., Ngai, S.T., Juneng, L., Santisirisomboon, Jerasorn, Santisirisomboon, Jaruthat, Ngo-Duc, T., Phan-Van, T., Narisma, G., Singhruck, P., Gunawan, D., Aldrian, E., Sopaheluwakan, A., Nikulin, G., Yang, H., Remedio, A.R.C., Sein, D., Hein-Griggs, D., 2018. Future changes in annual precipitation extremes over Southeast Asia under global warming of 2°C. APN Sci. Bull. 8. https://doi.org/10.30852/sb.2018.436

Themeßl, M.J., Gobiet, A., Heinrich, G., 2012. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. Clim. Change 112, 449–468. https://doi.org/10.1007/s10584-011-0224-4

Trinh-Tuan, L., Matsumoto, J., Tangang, F.T., Juneng, L., Cruz, F., Narisma, G., Santisirisomboon, J., Phan-Van, T., Gunawan, D., Aldrian, E., Ngo-Duc, T., 2019. Application of Quantile Mapping bias correction for mid-future precipitation projections over Vietnam. Sci. Online Lett. Atmos. 15, 1–6. https://doi.org/10.2151/SOLA.2019-001

Tran Ngoc, T.D., Perset, M., Strady, E., Phan, T.S.H., Vachaud, G., Quertamp, F., Gratiot, N., 2016. Ho Chi Minh City growing with water-related challenges. http://eaumega.org/wp-content/uploads/2016/05/HCMC-MonographyEN.pdf

Trenberth, K.E., Dai, A., Rasmussen, R.M., Parsons, D.B., 2003. The Changing Character of Precipitation. Bull. Am. Meteorol. Soc. 84, 1205–1218. https://doi.org/10.1175/BAMS-84-9-1205

Trinh-Tuan, L., Matsumoto, J., Tangang, F.T., Juneng, L., Cruz, F., Narisma, G., Santisirisomboon, J., Phan-Van, T., Gunawan, D., Aldrian, E., Ngo-Duc, T., 2019. Application of Quantile Mapping bias correction for mid-future precipitation projections over Vietnam. Sci. Online Lett. Atmos. 15, 1–6. https://doi.org/10.2151/SOLA.2019-001

Truong Ha, M., 2018. Climate Change Impact on Intensity-Duration-Frequency Curves in Ho Chi Minh city, In: Workshop on Disaster Resilient Cities. http://ancst.org/wp-content/uploads/2018/07/4.-Mr.-Minh-Truong-Ha-Vietnam-Institute-of-Meteorology-

Hydrology-and-Climate-Change.pdf

Trzaska, S., Schnarr, E., 2014. A review of downscaling methods for climate change projections, United States Agency for International Development by Tetra Tech ARD. Available from http://www.ciesin.org/documents/Downscaling_CLEARED_000.pdf

Ung, P., Peng, C., Yuk, S., Tan, R., Ann, V., Miyanaga, K., Tanji, Y., 2019. Dynamics of bacterial community in Tonle Sap Lake, a large tropical flood-pulse system in Southeast Asia. Sci. Total Environ. 664, 414–423. https://doi.org/10.1016/j.scitotenv.2019.01.351

UNISDR and the World Bank, 2010. Synthesis Report on Ten ASEAN Countries Disaster Risks Assessment. http://www.unisdr.org/files/18872_asean.pdf

United Nations Department of Economic and Social Affairs (UNDESA), 2019. World Population Prospects 2019. https://population.un.org/wpp/Download/Standard/Population/

Vachaud, G., Quertamp, F., Phan, T.S.H., Tran Ngoc, T.D., Nguyen, T., Luu, X.L., Nguyen, A.T., Gratiot, N., 2019. Flood-related risks in Ho Chi Minh City and ways of mitigation. J. Hydrol. 573, 1021–1027. https://doi.org/10.1016/j.jhydrol.2018.02.044

Villafuerte, M.Q., Macadam, I., Daron, J., Katzfey, J., Cinco, T.A., Ares, E.D., Jones, R.G., 2020. Projected changes in rainfall and temperature over the Philippines from multiple dynamical downscaling models. Int. J. Climatol. 40, 1784–1804. https://doi.org/10.1002/joc.6301

Volosciuk, C., Maraun, D., Vrac, M., Widmann, M., 2017. A combined statistical bias correction and stochastic downscaling method for precipitation. Hydrol. Earth Syst. Sci. 21, 1693–1719. https://doi.org/10.5194/hess-21-1693-2017

Wang, L., Ranasinghe, R., Maskey, S., van Gelder, P.H.A.J.M., Vrijling, J.K., 2016. Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: a case study of the Huai River Basin, China. Int. J. Climatol. 36, 145–164. https://doi.org/10.1002/joc.4334

Westra, S., Evans, J.P., Mehrotra, R., Sharma, A., 2013. A conditional disaggregation algorithm for generating fine time-scale rainfall data in a warmer climate. J. Hydrol. 479, 86–99. https://doi.org/10.1016/j.jhydrol.2012.11.033

Westra, S., Fowler, H.J., Evans, J.P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E.J., Lenderink, G., Roberts, N.M., 2014. Future changes to the intensity and frequency of short-duration extreme

rainfall. Rev. Geophys. 52, 522–555. https://doi.org/10.1002/2014RG000464

Wilby, R.L., Troni, J., Biot, Y., Tedd, L., Hewitson, B.C., Smith, D.M., Sutton, R.T., 2009. A review of climate risk information for adaptation and development planning. Int. J. Climatol. 29, 1193–1215. https://doi.org/10.1002/joc.1839

Willems, P., 2013. Revision of urban drainage design rules after assessment of climate change impacts on precipitation extremes at Uccle, Belgium. J. Hydrol. 496, 166–177. https://doi.org/10.1016/j.jhydrol.2013.05.037

Willner, S., Levermann, A., Zhao, F., Frieler, K. (2018). Adaptation required to preserve future high-end river flood risk at present levels. Science Advances: 4:1. https://doi.org/10.1126/sciadv.aao1914

World Bank, 2021. The World Bank In Cambodia. https://www.worldbank.org/en/country/cambodia/overview#1. Accessed on 24th November 2021.

World Bank Group (WBG) and Asian Development Bank (ADB), 2021. Climate Risk Country Profile: Cambodia. https://www.adb.org/publications/climate-risk-country-profile-cambodia.

World Meteorological Organization (WMO), 2021. State of the Global Climate 2021. https://library.wmo.int/doc_num.php?explnum_id=10859

Xu, Z., Han, Y., Yang, Z., 2019. Dynamical downscaling of regional climate: A review of methods and limitations. Sci. China Earth Sci. 62, 365–375. https://doi.org/10.1007/s11430-018-9261-5

Yim, S., Aing, C., Men, S., Sovann, C., 2016. Applying PCSWMM for Stormwater Management in the Wat Phnom Sub Catchment, Phnom Penh, Cambodia. J. Geogr. Environ. Earth Sci. Int. 5, 1–11. https://doi.org/10.9734/JGEESI/2016/23525

Young, C.B., McEnroe, B.M., 2003. Sampling Adjustment Factors for Rainfall Recorded at Fixed Time Intervals. J. Hydrol. Eng. 8, 294–296. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:5(294)

Zhang, W., Villarini, G., Vecchi, G.A., Smith, J.A., 2018. Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. Nature 563, 384–388. https://doi.org/10.1038/s41586-018-0676-z

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 Resour. Res. 55, 1871–1889. https://doi.org/10.1029/2018WR023567

Zscheischler, J., Westra, S., van den Hurk, B.J.J.M., Seneviratne, S.I., Ward, P.J., Pitman, A., AghaKouchak, A., Bresch, D.N., Leonard, M., Wahl, T., Zhang, X., 2018. Future climate risk from compound events. Nat. Clim. Chang. 8, 469–477. https://doi.org/10.1038/s41558-018-0156-3

Copyright

Acknowledgement

This research is partly funded by the Japan Society for the Promotion of Science Bilateral Joint Research Program. I also thank the financial support from the China Scholarship Council (grant number = 201906260299). Without this financial help provided by my motherland, I would not have been able to come to Japan. I would like to express my sincere gratitude to my supervisor Prof. Tsuyoshi Kinouchi for his continuous help in my doctoral career. Thank Prof. Kinouchi for providing me a chance to study in this lab, for sharing his immense knowledge in paper writing and revision with me, for allowing and helping me to conduct the doctoral research freely and smoothly, for his patience and motivation in our discussions. A similar gratitude belongs to my lab-mates, who have helped me and shared me with their knowledge and happiness. Last but not least, I would also like to thank my parents, relatives and friends who support me in all aspects.