THE IMPACT OF CLIMATE CHANGE ON HYDROLOGICAL REGIMES IN THE SAIGON-DONGNAI RIVER BASIN

Submitted in partial fulfillment of the requirement for the award of the degree of

Doctor of Philosophy

by

Dang Dong Nguyen

Roll No. 715014

Supervisor

Prof. K.V. Jayakumar

DEPARTMENT OF CIVIL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
WARANGAL-506004, INDIA

APRIL 2019

NATIONAL INSTITUTE OF TECHNOLOGY

WARANGAL

CERTIFICATE

This is to certify that the thesis entitled "**The Impact of Climate Change on Hydrological Regimes in The Saigon-Dongnai River Basin**" being submitted by **Mr. Dang Dong Nguyen** for award of the degree of Doctor of Philosophy to the Faculty of Engineering and Technology of the **National Institute of Technology Warangal** is a record of bonafide research work carried out by him under my supervision and it has not been submitted elsewhere for award for any degree.

K.V. Jayakumar

Thesis supervisor

Professor

Department of Civil Engineering

National Institute of Technology Warangal (T.S.) INDIA

Declaration

This is to certify that the work presented in the thesis entitled "**The Impact of Climate Change on Hydrological Regimes in The Saigon-Dongnai River Basin**" is a bonafide work done by me under the supervision of **Prof. K.V. Jayakumar** and was not submitted elsewhere for the award of any degree.

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Dang Dong Nguyen

Roll No. 715014

Date:

Approval Sheet

This thesis entitled "**The Impact of Climate Change on Hydrological Regimes in The Saigon-Dongnai River Basin**" by **Dang Dong Nguyen** is approved for the degree of Doctor of Philosophy

Examiners

Supervisor

Chairman

Date:

Abstract

Extreme events are becoming more intense, more frequent and more destructive. Changes in extreme events such as flood and droughts are the primary ways that most people experience climate change. Flooding may intensify in many regions in the world particularly in regions of South Asia. Study for prediction of future flood risks in catchment scale using hydrological models along with climate change projection has played a considerable role in recent years.

Vietnam is one of the countries severely impacted by climate change. Trian watershed is located in the upper Saigon-Dongnai River basin and it is one of the biggest sub-basins of this river. Besides, this region is also the economic center in the south of Vietnam. However, not many studies have been conducted or reported in the literature to assess the impact of climate change on this region. It is, hence, necessary to evaluate the potential impact of climate change in future on this watershed, particularly on flood frequency, because flood events cause negative impacts on economic and social aspects.

Peak over Threshold (POT) approach uses the available flood data more efficiently and this approach can estimate return level more accurate. The Generalized Pareto Distribution (GPD) is fitted with POT magnitude as a default in extreme value analysis. However, there could be more than one distribution that can be fitted to the data sample. Therefore, it is important that POT approach for testing numerous distributions should be considered in assessing the changes of flood frequency.

The downscaled atmospheric data are used as input for a physically-based hydrological model to simulate future streamflow data. The changes in the frequency of flood peak extracted by the POT approach is compared between historical and future periods. The results indicate that there is a significant increase in flood magnitude under climate change for Trian catchment. To be more specific, the 100-year return level of Trian catchment is increasing up to 32.34 % in one of future scenario. Moreover, the results of this study also indicate that directly using the asymptotic distribution to model the POT dataset sometimes provides wrong insights.

The flood characteristics namely, peak, duration and volume for a given frequency provide important knowledge for the design of hydraulic structures, water resources planning,

reservoir management and flood hazard mapping. Flood is a complex phenomenon defined by strongly correlated characteristics. If univariate frequency analysis is used to assess the behaviour of each flood characteristics, it will lead to over or under estimation of associated flood risk. In these cases, multivariate probability approach, which provides a comprehensive understanding of flood characteristics and their relationship, may provide better estimate of the flood magnitude when compared to the univariate approach.

Traditional multivariate parametric distributions have widely been applied for hydrological applications. However, this approach has some drawbacks such as the dependence structure between the variables, which depends on the marginal distributions or the flood variables have the same type of marginal distributions. Copulas are applied to overcome the restriction of classical multivariate flood frequency analysis by choosing the marginal distribution from different types of the probability distribution function for flood characteristics. The most important step in the modelling process using copula is the selection of copula function which is the best fit to data sample. The choice of copula may significantly impact on the bivariate quantiles.

From the results of the study, it is observed that the result from tail dependence test is useful in selecting the appropriate copula for modelling the joint dependence structure of flood variables. The extreme value copulas with upper tail dependence have proved that they are appropriate models for the dependence structure of the flood characteristics. Frank, Clayton and Gaussian copulas have been identified as the appropriate copula models in case of variables, which are diagnosed to have asymptotic independence.

Flood hazard mapping is one of the important aspects of flood risk assessment which has a significant implication on the planning of social and economic development activities. It also provides useful information to operate the flood warning system as well as to prepare the emergency evacuation plans. However, the development of the hydrodynamic models for the large river system is a very challenging task. The hydrodynamic models have to be large enough to cover the entire river basin and it must be sufficiently detailed to represent smaller features. Therefore, the latest advances in flood modelling techniques, such as flexible meshes generation and the advantages of coupled hydrodynamic model with the high-resolution of topography data will be applied in this study.

MIKE FLOOD, which is a coupled hydrodynamic model, is used to simulate the flood regime. The coupled hydrodynamic model has been developed for cross-sections based on channel modelling with 1D model and linking these floodplain modelling with 2D model. The design flood hydrograph is estimated using bivariate flood frequency analysis, high-quality topography data (i.e., DEM and LiDAR) and flexible meshes generation are used as the input data for hydrodynamic model to simulate the flood regime for the study area. Two parameters namely flood depth and flow velocity, which are obtained from a coupled hydrodynamic model, have been used for developing the high-resolution flood hazard maps.

This study presents the results of the assessment of the changes in the flood hazard and the duration of inundation under climate change context for Saigon-Dongnai River basin, Vietnam. From the results, it is observed that 54.47 km^2 of the study area is subjected to H6 hazard index under 100-year return period in the present flood. 140.62 km^2 and 50.90 km^2 are under H6 hazard index for RCPs 4.5 and 8.5 scenarios respectively. This study indicated that the duration of inundation is not only controlled by flood magnitude but also by the volume of flood. Further, this study showed that most of the agricultural areas located downstream of Trian catchment will be severely inundated under climate change context.

Content

List of Figures

List of Tables

Nomenclature

The following list gives the notations used in chapters of the thesis.

$A_n(t)$	Non-parametric of Pickands dependence function
$A_{\theta n}(t)$	Parametric estimator of Pickands dependence function
F_1^{-1}, F_2^{-1}	Quantile function
$T_{X,Y}^{OR}, T_{X,Y}^{AND}$	OR and AND joint return periods
μ_T	Mean inter-arrival time
£(z)	Slowly varying function and
C	Copula function
c	Threshold
CF	Change factor
F(x,y)	Bivariate distribution function
$F_1(x)$, $F_2(y)$	Marginal distribution function
Log L(Θ|X)	Log-likelihood function
N	Number of observations
n_{ml}	Rank of the combination of (x_j, y_j)
P	Probability
Q	Discharge
r	Time interval
T_{lag}	Lag time
η	Coefficient of tail dependence
λ_U	Non-parametric tail dependence estimator
ϕ	Cumulative density function of standard normal
χ, $\bar{\chi}$	Extreme dependence measurement
Φ	Standard normal density distribution function
$F(t)$	Standard uniform distribution

Abbreviations

1D	One-dimensional
2D	Two-dimensional
AD	Anderson-Darling Criterion
ADB	Asian Development Bank
AIC	Akaike Information Criterion
AM	Annual Maxima
ANN	Artificial Neural Network
CDF	Cumulative Distribution Function
CN	Curve Number
CORDEX	Coordinated Regional Downscaling Experiment
CVM	Cramer-von Mises
D	Flood Duration
DEM	Digital Elevation Models
DHI	Danish Hydraulic Institute
DRH	Direct Runoff Hydrograph
ELM	Extreme Learning Machine
FFA	Flood Frequency Analysis
GAM	Generalised Addictive Model
GCM	Global Climate Models
GEV	Generalised Extreme Value
GIS	Geographic Information System
GL	Generalised Logistic
GoF	Goodness-of-fit
GPD	Generalised Pareto Distribution
HCA	Hierarchical Cluster Analysis
HCMC	Ho Chi Minh City
HEC-GeoHMS	Geospatial Hydrological Modelling Extension
HEC-HMS	Hydrologic Modelling System
IID	Independent and Identically Distributed
IPCC	Intergovernmental Panel on Climate Change

KS	Kolmogorov–Smirnov
LARS-WG	Long Ashton Research Station Stochastic Weather Generator
LiDAR	Light Detecting and Ranging
LLH	Log-Likelihood
LLHR	Log-Likelihood Ratio
LN	Log Normal
LP3	Log Pearson Type 3
LULC	Land Use Land Cover
MAE	Mean Absolute Error
MK	Mann-Kendall
ML	Maximum Likelihood
MONRE	Ministry of Natural Resources and Environment
MOS	Model Output Statistics
MPL	Maximum Pseudo Likelihood
MRL	Mean Residual Life
NHMS	National Hydro-Meteorological Service
NSE	Nash and Sutcliffe
P	Flood Peak
P3	Pearson Type 3
PBIAS	Percent Bias
PDF	Probability Density Function
POT	Peak over Threshold
PP	Perfect Prognosis
P-P	Probability-Probability
Q-Q	Quantile-Quantile
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RMSE	Root Mean Square Error
RSR	Root Mean Square Error-observation Standard Deviation Ratio
SCS	Soil Conservation Service
SMA	Soil Moisture Accounting
SRTM	Shuttle Radar Topography Mission

SWM	Stanford Watershed Model
V	Flood Volume
WGs	Weather Generators

Chapter 1

Introduction

1.1. Flood Frequency Analysis

Among the worst natural disasters, floods cause huge damages annually including loss of property and human lives. The damage to properties and loss of life caused by floods could be higher in the future due to changing of climate. Assessment of changes of flood characteristics under climate change context plays a considerable role in managing the risk of flood. Quantifying the vulnerable areas associated with the changes of climate allows local authorities to provide a good future development planning. Therefore, the quantifying the impact of climate change on flood risk is necessary to be carried out on a river basin scale.

The design and assessment of flood risk of hydraulic structures, water resources planning, reservoir management and flood hazard maps involve the identifying the given flood events with a low probability of exceedance. Flood Frequency Analysis (FFA) seeks to connect the magnitude of extreme events with their frequency of occurrence via probability distribution.

The objective of FFA is to estimate the return period associated with a flood of a given magnitude. Use of the return period as a standard criterion is common in the design of hydraulic structures and flood control. It is, hence, necessary in most cases, to obtain the frequency curve fitting the Probability Distribution Function (PDF) to the observed data to estimate flood quantiles associated with given return periods.

Annual Maxima (AM) and Peak Over Threshold (POT) are the commonly used approaches to extract the flood events in the investigation of the changes in flood magnitude. The maximum peak flow of each year defines the AM sample. However, AM cannot be used in the case of data of short length, because such data will not provide adequate information (Bezak et al., 2014, Lang. et al., 1999). Unlike the AM, which only extracts one event per year, POT considers a broader range of events and provides more information than AM, primarily for estimation of flood magnitude (Kay et al., 2009, Roth et al., 2012). Many researchers have investigated the choice between AM and POT. Some have emphasized that POT approach is more suitable for extreme value analysis and provide better estimates of flood quantiles than corresponding AM approach (Bezak et al., 2014, Saf, 2009b).

The next step is to select an appropriate PDF that has the best fit for the data sample. Many PDFs have been considered in different situations for the probability modelling of flood events. Malamud and Turcotte (2006) and El Adlouni et al. (2008) divided the widespread distribution in FFA into four groups such as the Generalized Extreme Value (GEV) family, the Normal family, the Pearson Type III family and the Generalized Pareto Distribution (GPD) family.

The GPD is fitted with POT magnitudes as a default in extreme values analysis (Coles et al., 2001, Davison and Smith, 1990, Katz et al., 2002). However, there could be more than one distribution that can be fitted to the sample data. Hence, identifying the best fitting distribution to the sample need to be tested with several distributions (Lang. et al., 1999). Furthermore, many studies are reported which used AM to assess the impact of climate change on flood frequency at global as well as regional scale (Hirabayashi et al., 2013, Jung et al., 2011), while there are only a few studies which used POT approach. Even the studies which used POT approach mostly considered only a single distribution (i.e., GPD). Thus, it becomes imperative that many distributions which are normally used in extreme values analysis should

be tested and the best fitting distribution needs to be identified for each POT dataset.

FFA can be classified into univariate and multivariate analyses. Although univariate frequency analysis has been widely used to quantify the behaviour of each flood characteristics, only a limited assessment of flood events is obtained (Yue, 1999). In fact, univariate frequency analysis can be useful, if the infrastructure design is based on a single flood characteristic (i.e., peak). Otherwise, univariate frequency analysis may not provide the complete behaviour of flood characteristics (Chebana and Ouarda, 2011). Additionally, the flood is a multivariate natural calamity characterized by its peak, volume and the duration. Hence, it is essential to study the multivariate probabilistic behaviour of flood characteristics simultaneously. The joint probability to be incorporated into flood risk analysis involving two or more flood characteristics have received significant attention in recent years. However, many aspects need to be solved related to this subject.

Traditional multivariate FFA does not allow using different marginal distributions and full coverage of dependence structure between the variables. To overcome these shortcomings, copulas, one of the promising mathematical tools for investigating multivariate problems, have been widely applied in hydrological studies in recent times (Li et al., 2012). For example, Reddy and Ganguli (2012) used a copula approach for flood frequency analysis of Godavari River, India. Their study indicated that Frank copula was the best-fit copula for bivariate models (i.e, flood peak and volume, volume and duration pairs). Additionally, climate change impact on the flood characteristics for the northeast Canadian basin was evaluated using copula-based bivariate flood frequency analysis in a study of Jeong et al. (2014). The results of projected changes indicated that an increase in the joint return period of flood characteristics. Similarly, in their work, Duan et al. (2016) used copula-based bivariate frequency analysis to investigate the changes in flood characteristics in the Huai River, China under climate change context. Their study showed that Archimedean copulas were more appropriate to model the dependence structure of flood characteristics in the study area. Therefore, bivariate flood frequency analysis based on copula approach is considered to evaluate the inherent flood characteristics in this study.

1.2. Flood Hazard Mapping

A warmer climate is already causing extreme weather events that affect the lives of millions of people around the world (Schiermeier, 2011). Specifically, as per IPCC (2014), extreme climate events are likely to occur more frequently in different parts and during different seasons in Asia in the future, particularly in East Asia regions. Brunner et al. (2017) indicated that climate change has an impact not only on the peak but also on the volume and the shape of the flood hydrograph. Hence, assessment of climate change impacts on floods should consider all the flood characteristics (i.e., peak, volume and duration) rather than only the flood peak. Therefore, to make appropriate adaptation strategies, decisions and policies under climate change context, it is essential to understand the change of the flood characteristics and potential flood risks on a river basin scale.

Flood hazard mapping is one of the critical aspects of flood risk assessment which has a significant implication on the planning of social and economic development. The information of flood hazard is also essential to provide various strategies for mitigating the flood risk, which in turn, can reduce the losses of human life and damages in urban and rural sectors (Pappenberger et al., 2012, Sampson et al., 2015). Assessing the flood risk at the river basin is not a simple task, because of the complex nature of flood generation caused by a combination of precipitation, river basin characteristics and human activities. However, the development of numerical flood modelling in recent years, namely the availability of advanced flood modelling and modern survey techniques for collection of high-quality input data for those models allow to simulate flood behaviour and to study the characteristics of future floods (Alkema, 2007).

To prepare reliable flood hazard maps, a methodology that combines the advantageous features of 1D and 2D hydraulic models and also the high-resolution of topographic data, are typically applied. Flood hazard maps show the intensity of floods and their associated exceedance probability (Di Baldassarre et al., 2010). One of the common approaches of flood inundation modelling is use of deterministic approach based on single simulation (Ali, 2018). In deterministic approach, three main issues in developing the flood hazard maps using hydrodynamic models such as the topography data resolution, the hydrodynamic model simulation and the design flood hydrograph estimation are considered in the study.

There are numerous studies related the assessing the impact climate change on floods have been reported in the literature. Most of the studies on flood frequency analysis focus only on the flood peak (Camici et al., 2014, Dobler et al., 2012, Qin and Lu, 2014). However, flood is a complex phenomenon defined by the strong correlation between its characteristics such as peak, duration and volume. If univariate frequency analysis is used to assess the behaviour of each flood characteristic, it will lead to over or underestimation of associated flood risk. To develop flood hazard maps, only the flood peak cannot give a reliable evaluation of hazard. Consequently, it is also essential to consider simultaneously the flood peak along with other flood characteristics in developing flood hazard mappings. Furthermore, faster and accurate flood modelling at high spatial-temporal resolutions remains a significant challenge in hydrologic and hydraulic studies. Therefore, it is necessary to establish an advanced deterministic approach, including bivariate frequency analysis, efficient and flexible hydrodynamic models and high-resolution data to develop the flood hazard maps under climate change context.

1.3. Motivation for the Study

Floods are one of the most commonly occurring natural disasters in the word. In the past decade, floods have caused devastating damage to property and loss of life across the world. For example, tens of billions of US dollars were spent and thousands of people were killed in every year (Hirabayashi et al., 2013). Climate change is widely recognized to affect flood regimes in many parts of the world. Thus, the losses are expected to be massive in the future due to climate change. Studies related to future flood risk provide information on the frequencies and magnitudes of possible floods in the future. Numerous studies for assessing the changes of flood hazard at many scales are available in the literature. However, the impact of climate change on flood has not been sufficiently understood at a catchment scale in Asia, and in particular for Vietnam.

Flood event is multivariate in nature and hence it is necessary to consider simultaneously the various component processes in some situations. For example, determination of the occurrence of the flood peak and volume is necessary to design the hydraulic infrastructure along a river. To develop the flood hazard maps, information on flood peak alone will not give a reliable assessment of hazard. It is, therefore, also essential to consider flood peak along with

other flood characteristics simultaneously.

Flood risk plays an essential role in the planning of water infrastructure projects, reservoir management and flood hazard mapping in river basin scale too. Therefore, the evaluation of changing of flood characteristics, flood inundation areas and preparation of possible flood hazard mapping due to climate change are essential to help the policymakers and stakeholders for social and economic development planning in the river basin scale.

1.4. Objectives of the Study

With this background and appreciating the significance of the studies on flood risk, flood zone mapping and flood frequency analysis, the objectives of the study have been formulated.

The objectives of this study are listed as follows:

- Evaluating the correctness of directly using the asymptotic distribution to model the future POT dataset
- Assessing the potential impact of climate change on flood magnitude
- Investigating the potential of performing the tail dependence tests for pairs of flood characteristics
- Evaluating the choice of copula based on the tail dependence test
- Estimating of the flood design hydrographs using copula theory of flood variables (i.e., volume and peak) and shapes through historical observed flood hydrograph combined with cluster analysis
- Developing the computationally efficient flood model using advanced deterministic approach based on coupled 1D-2D hydraulic model and high quality of topography data
- Developing of flood hazard maps, which is quantified by considering the flood depth and velocity in combination
- Assessing the changes of flood risk under climate change context.

1.5. Contribution from the Study

Hydrologists have widely used flood frequency analysis to evaluate the potential flood risk. This information can provide valuable knowledge for designing infrastructure, reservoir system operational plans, and flood hazard assessment. The GPD is fitted with POT magnitudes as a default in extreme values analysis. However, there could be more than one distribution that can be fitted to the sample data. Furthermore, POT approach is better suited for extreme value analysis and to arrive at better and reliable accurate estimates of flood quantiles than corresponding AM approach. However, many studies used AM to assess the impact of climate change on flood frequency at global as well as regional scale while there are only a few studies which used POT approach. Even the studies which used POT approach mostly considered only a single distribution. Thus, it is vital that several distributions generally used in extreme values analysis should be tested, and the appropriate distribution need to be identified for each POT dataset.

Single variable flood frequency analysis does not give a comprehensive understanding and assessment of the actual behaviour of flood phenomena. This approach can lead to high uncertainty or failure of guidelines in water resources planning, operation and design of hydraulic structure and floodplain zoning. Therefore, it is essential to study the multivariate probability behaviour of flood characteristics. Copulas are widely used for multivariate analysis in various fields. The main advantage of copulas is that the dependence structure is independently modelled with the marginal distribution that allows for multivariate distribution with different margins and full coverage of dependence structure. The essential step in the modelling processing copula is the selection of copula function, which is the best fit for the data sample. This study suggests that the copula function should be selected based on the dependence structure of the variable. Furthermore, the performance of extreme value copulas for asymptotic dependence variable and Clayton, Frank and Gaussian copula for an asymptotic independent variable are also assessed.

To simulate accurately the spatial and temporal dynamics of the flood process, the three main issues in creating the flood hazard maps using hydrodynamic models (i.e., design flood hydrograph estimating, the resolution of topography data and hydraulic model selection) are carefully assessed. The same joint probability of occurrence of flood peak and volume are used

to estimate the design flood hydrograph.

The surveyed cross-sections and high-resolution LiDAR along with design flood hydrograph are used as the input data for the coupled hydrodynamic model to simulate the flood regime. Furthermore, the multi-scale mesh modelling approach, where fine resolution is applied for channel and raised embankment areas and coarser resolution is developed for uniform topographic height are used to develop the hydrodynamic model in this study. The coupled hydrodynamic model in which channel flow is linked to floodplain flow using lateral connection is used to improve accuracy the flood inundation results without the significant increasing computational requirement of the hydraulic model. The flood depth and velocity obtained from the hydrodynamic model are used to develop high-resolution flood hazard maps for the study area under climate change context.

1.6. Outline of the Study

After introducing the problem taken up for the study and discussing about the significance of the problem, the objectives of the study is introduced in this chapter. Chapter 2, a detailed review of the literature related to various methods of transferring climate change, hydrological modelling, flood frequency analysis and flood hazard mapping is also presented in this Chapter.

Chapter 3 presents the methodology involved in addressing the research objectives. Further more. the summaries of all case studies as well as the availbel sofwares are also presented in this Chapter.

The impact of climate change on flood magnitude is examined and presented in Chapter 4. Besides, the correctness of directly using the asymptotic distribution to model the future POT dataset is also assessed.

Chapter 5 presents the potential of performing the tail dependence tests for the pairs of flood characteristics. The choice of copula based on the tail dependence test is also explored. Besides, assessing the impact of the different copula functions on bivariate quantiles is also considered in this Chapter.

The design flood hydrographs obtained using copula approach combined with cluster

analysis is reported in Chapter 6. The high-resolution flood hazard maps, which is quantified by considering the flood depth and velocity in combination, are also established.

Chapter 7 presents the summary of the study, the conclusions arrived and some recommendation for further research activities based on the conclusions from study on the impact of climate change on flood risk.

Chapter 2

Literature Review

2.1 Introduction

This chapter consists of six sections. The first part, dealing with downscaling methods (i.e., dynamical and statistical) of climate data for use in hydrological models is presented in Section 2.2, while Section 2.3 discusses the different categories of hydrological models, which are used to obtain the future discharge time series. In the third part, the focus is on the flood frequency analysis. The flood hazards mapping estimation based on the deterministic approach is presented in Section 2.5. The conclusion highlighting the findings from the literature review is presented in the last section.

2.2 Downscaling Methods Climate Projections

Global Circulation Models (GCMs) are widely used to predict the changes in atmospheric variables under climate change scenarios (Anandhi et al., 2011). GCMs are physically-based meteorological models, which represent atmospheric and oceanic dynamics

(Angeles et al., 2007). However, the output from GCMs has typically a low spatial resolution of approximately 100-250 km and it is inadequate for regional impact studies especially for analyzing the changes in extreme precipitation and floods because it lacks detailed regional information which is needed to resolve various features at the catchment scale (Fowler et al., 2007). This obstacle can be solved using the downscaling method, which can be used to derive local to regional scale information from large-scale spatial and temporal scales. The downscaling methods have been classified into two types, namely dynamic and statistical. The dynamical downscaling generates finer resolution output based on atmospheric physics over a region using GCM as a boundary. Statistical downscaling method establishes an empirical relationship between the GCMs output with observed climate data. Table 2.1 summaries some advantages and disadvantages of both dynamical and statistical downscaling methods. The following section presents more detail information both downscaling methods.

2.2.1 Dynamical downscaling

Dynamic downscaling method refers to the use of high-resolution regional simulations to dynamically extrapolate the effect of large-scale climate processes to regional or local scales. Dynamical downscaling uses a limited area, a high-resolution model such as Regional Climate Models (RCMs) driven large scale and lateral boundary conditions from a GCM to produce higher resolution outputs (Fowler et al., 2007). RCMs are frequently used to analyze the impact of climate change on hydrology in the watershed because of its higher resolution. The resolution of RCMs is around 12-50 km and it accounts for the sub-GCM grid scale forcing (e.g., complex topographical features and land cover heterogeneities in a physically-based way).

As a consequence of the higher spatial resolution output, RCMs provide a better description of topographic phenomena. Further, the finer dynamical processes in RCMs produce more realistic mesoscale circulation pattern. There are numerous RCMs, which are widely used in climate change downscaling studies. They are the East and South Asia Regional Climate Model (RegCM4), Canadian Regional Climate Model (CRCM), Dutch Regional Atmospheric Climate Model (RACMO), UK Met Office Hadley Center's Regional Climate Model version 3 (HadRM3), German Regional Climate Model (REMO), the U. S Regional Climate Model version 3 (RegCM3) and the Hadley Center Global Environment Model version 3 Regional Climate Model (HadGEM3-RA). Some of these studies which used dynamical

downscaling in hydrological researches are discussed as follows.

Van Roosmalen et al. (2010) used the climate outputs of HIRHAM4 to generate future time series of precipitation, temperature and evapotranspiration for hydrological impact assessment in Denmark. The result showed that the HIRHAM4 model simulated the current output climate which are not too far from observed value and it does not show a significant difference from ensemble RCMs. Gu et al. (2012) predicted the future climate change by using regional climate model (RegCM4) for East and South Asia. The result indicated that the Yangtze river basin will witness the changes of extreme events (i.e., precipitation and drought) and it indicates potentially increased risks of both floods and droughts at the same time. Bárdossy and Pegram (2011) used the output of three different regional climate models, namely HadRM3, RACMO2 and REMO modelled the future climate scenarios for Rhine River catchment. The results indicated that the climate in the Rhine River basin is likely to be wetter than in the past for the future climate scenarios.

However, RCMs inherit the biases due to systematic model errors caused by imperfect conceptualization, discretization and spatial averaging within grid cells. Andréasson et al. (2004) showed that these biases are not only for precipitation but also for temperature. These biases can affect the result of the hydrological simulation. The statistical downscaling method, which is computationally inexpensive in comparison to RCMs, is a viable and sometimes advantageous alternative for an institution that does not have the computational capacity and technical expertise require in dynamical downscaling. Therefore, the following section will discuss the statistical downscaling methods.

2.2.2 Statistical downscaling

The statistical downscaling establishes an empirical relationship between GCM resolution climate variable and local climate. There are several statistical downscaling approaches which establish statistical links between large-scale climate and the observed local-scale climate data. Maraun et al. (2010) classified statistical downscaling approaches into Perfect Prognosis (PP), Model Output Statistics (MOS), and Weather Generators (WGs). PP method establishes statistical relationships between variables at large scales and local scales. MOS establishes statistical relationships between variables simulated by the RCM and local scale observation to correct RCM errors. WGs approach generates local scale climate time

series resembling the statistical properties of observed climate.

The statistical downscaling methods are also classified based on the techniques such as regression methods, weather classification scheme (i.e., weather pattern) and stochastic weather generator. Table 2.2 summaries some advantages and disadvantages of several statistical downscaling methods.

Regression methods represent the linear and nonlinear relationships between predictand and predictors. A simple linear regression is widely used to establish the relationship between one large-scale predictor and one local predictand. This relationship is obtained using observed local scale data and GCM/RCMs climate output data. Multiple regression methods establish the relationship between single predictand with two or more of predictor variables. Canonical correlation and singular value decomposition are widely used for the study of the interrelationship among spatially distributed coarse simulations and observed local scale variables by determining the sets of patterns that have strong correlation over time.

Table 2.1: Comparison of dynamical and statistical downscaling methods (Trzaska and Schnarr, 2014)

	Dynamical downscaling	Statistical downscaling
Requires	• High computational resources and expertise • High volume of data input • Reliable GCM simulations	• Medium/low computational resources • Medium/low volume of data inputs • Sufficient amount of good quality observation data • Reliable GCM simulations
Advantages	• Based on consistent, physical mechanism • Resolves atmospheric and surface processes occurring at sub-GCM grid scale • Not constrained by historical record so that novel scenarios can be simulated • Experiments involving an ensemble of RCMs are becoming available for uncertainty analysis	• Computationally inexpensive and efficient, which allows for many different emissions scenarios and GCM pairing • Methods range from simple to elaborate and are flexible enough to tailor for specific purposes • The same method can be applied across regions or the entire globe, which facilitates comparisons across different case studies • Relies on the observed climate as a basis for driving future projections • Can provide point-scale climatic variables for GCM scale output • Tools are freely available and easy to implement and interpret; some methods can capture extreme events
Disadvantages	• Computationally intensive • Due to computational demands, RCMs are typically driven by only one or two GMC/ emission scenarios simulations • A limited number of RCMs available and no model results for many parts of the globe • May require further downscaling and bias correction of RCM outputs • Results depend on RCM assumptions; different RCMs will give different results • Affected by the bias of driving GCM	• High quality observed data might be unavailable for many areas or variables • Assumes that relationships between large and local-scale processes will remain to be the same in the future (stationarity assumptions) • The simplest methods may only provide projections at a monthly resolution

Table 2.2: Advantages and disadvantages of different of statistical downscaling methods (Trzaska and Schnarr, 2014)

Statistical downscaling methods	Advantages	Disadvantages	Categories
Regression methods	• Straightforward to apply • Employ full range of available predictor variables.	• Not suitable for extreme events • Inefficient for non-normal distributed data • Poor representation of observed variance.	• Simple and multiple regression • Canonical correlation analysis and singular value decomposition • Artificial neural networks.
Weather pattern approach	• Yields physically interpretable linkages to surface climate • Apply for both normal and non-normal distributed data • Provides better understanding of the climate sensitivity and variability.	• Requires additional step of weather type classification • Unable to predict the new value that are outsides of the range of the historical data • Requires large amount of data and intensively computational capacities • Circulation-based schemes may be insensitive to future climate forcing.	• Cluster analysis • Analog method • Monte Carlo method • Principle components.
Weather generator	• Provides sub-daily information • Simulates length of wet and dry spells • Obtains weather time series in region of scares data by using interpolating technique • Produces large number of series which is valuable for uncertainty analysis.	• Requires large amount of observed data • Takes little into account of spatial correlation of climate • Sensitive to missing or erroneous data in the calibration set.	• Markov chain approach • The spell length approach • Mixture models • Stochastic methods.

Weather pattern approaches involve grouping of local meteorological variables about different weather classification schemes. This method can be applied to variables that have normal and non-normal distributions. The analog is the simplest of the weather classification methods. In this method, the large-scale atmospheric circulation simulated by GCM/RCMs is compared to historical observed and the most similar is chosen as its analog. The simultaneously observed local weather is then associated with the projected large-scale pattern. To estimate future values of local predictand, the GCM/RCMs output is compared with the large-scale observed data over the historical period. Once a large-scale simulation is aggregated to a cluster, a random observation from the batch of data associated with this cluster is chosen as the local scale prediction. Artificial Neural Network (ANN) is an established technique with a flexible mathematical structure that is capable of identifying complex non-linear relationships between input and output data (Vu et al., 2016). In order to statistically downscale climate variables, ANN establishes a non-linear relationship between atmospheric and local scale climate variables.

Weather generator is a statistical model used to generate sequences of daily variables using GCM/RCM output. It produces multiple daily weather series, which is natural and logically consistent because any number of small-scale weather sequences may be associated with a given set of larger scale values. These generators frequently simulate meteorological variables at the daily or annual time scales on the basis of empirical statistical models. There are a number of well-known weather generators, which are widely used in agricultural, water resources and flood risk analysis, such as the Long Ashton Research Stochastic Weather Generator (LARS-WG), ClimGen, CLIGEN, WXGEN and Met&Roll (Fatichi et al., 2011).

2.3 Hydrological Modelling

Hydrological modelling is considered as an important tool for water resources planning and management. A hydrological model is a valuable tool for studying the impact of climate change on water resources from future scenarios of future change. Many river basins have experienced a change in the frequency and magnitude of hydrological extreme events (Arnell and Gosling, 2016, Milly et al., 2002). Numerous hydrological models have been developed and applied to assess the impact of climate change on water resources, particularly for floods. This section is devoted to a comprehensive review on hydrological models.

Also referred to as rainfall-runoff models, hydrological models can be classified into a number of categories, based on model input parameters as well as the physical principles used within the models. They can also be classified based on the model parameters as a function of time and space (i.e., lumped and distributed). Besides, a model is considered as deterministic if a set of input values will always produce the same output values. A model is stochastic if the input values need not to produce the same output values. The event and continuous simulation models are distinguished based on the specific and continuous period of output respectively (Sharma et al., 2008). The popular classifications, which are widely used in the literature, are empirical, parameters and physically-based models. Table 2.3 introduces the brief characteristics of three types of hydrological models.

Table 2.3: Characteristics of hydrological models (Devia et al., 2015)

Empirical model	Conceptual model	Physically-based model
• Metric or black box model • Described by mathematical structure using time series information • Does not need prior knowledge about hydrology process • High predictive power but low explanatory capacity • Cannot be generated to another catchment	• Parametric or grey box model • Based on modelling of reservoir and include semi empirical equations with a physical basis • Model parameters are calibrated • Simple and easily calculated using computer code • Require large hydro-meteorological data	• Mechanistic or white box model • Based on spatial distribution, evaluation of parameters describing physical characteristics • Model parameters can be measured • Complex model and requires high computational demand • Suffer from extreme data demand, scale related problems and overparameterization

2.3.1 Empirical models

Empirical models are derived from experiments or observed input-output relationships. Although they do not consider the physical law, the empirical models contain parameters that may have little direct physical significance and can be estimated only using the concurrent measurement of inputs and outputs (Xu, 2002). The rating curves, unit hydrograph, statistical models (i.e., linear and non-linear regressions) and machine learning methods are popularly used in the empirical models. Machine learning methods for empirical models are widely used to predict streamflows in recent years. There are many machine learning techniques that are

used for streamflow prediction (like ANN, support vector machines, random forest and k nearest neighbour, etc.). They have been proved to be appropriate tools for hydrological modelling and exploratory data analysis, particularly in systems that exhibit complex and non-linear behavior (Abrahart and See, 2007, Solomatine and Ostfeld, 2008).

Shortridge et al. (2016) used ANN, random forest, Generalized Additive Models (GAM), multivariate adaptive regression splines and M5 cubist models to simulate monthly streamflows in the highland of Ethiopia. The results indicated that random forest and GAM were useful in providing insights into physical watershed function. Yaseen et al. (2016) applied the Extreme Learning Machine (ELM) method to forecast the monthly streamflow discharge rate in the Tigris River, Iraq. The results showed a good improvement using ELM model than support vector regression and generalized regression neural network in forecasting the streamflow. A study by Badrzadeh et al. (2015) confirmed the robustness of hybrid wavelet-based models for real-time runoff forecasting at Casino station on Richmond River, Australia compared to ANN, adaptive neuro-fuzzy inference systems, wavelet neural networks. However, these models do not add any scientific knowledge or improved understanding in the field of hydrology. Therefore, parameter and physically-based models are adopted as useful alternative approaches to able to consider the physical law in these models.

2.3.2 Parameter models

Unlike the empirical models, the structure of parameter models is defined by the modeller's understanding of the hydrological system. Parameter models are formulated with some conceptual elements that are a simple representation of a reference system. Conceptual models also considered physical law but in a profoundly simplified form. One of the advantages of conceptual models is its non-linearity, which reflects the threshold presence in the hydrological system. Conceptual models describe all the component of hydrological processes. It consists of several interconnected reservoirs representing the physical elements in a catchment. Conceptual models are useful for various purposes and they can be used to infill the lost data or reconstruction of flow sequences.

One of the well-known conceptual models is the Stanford Watershed Model (SWM) elaborated by Crawford and Linsley (1966). The Sacramento model (Bergstrom, 1976) and the GR4J model (Perrin et al., 2003) are other well-known conceptual rainfall-runoff models with

different complexities. The GR4J model uses two unit hydrographs and two storages for the production and routing of water. The storage of rainfall, evapotranspiration and percolation in the surface soil are controlled by the production storage and the routing of effective rainfall is controlled by the routing storage. The Sacramento model has five runoff components (i.e., direct runoff, surface runoff, interflow, supplementary base flow and primary base flow). These models have been widely applied in many studies (Petheram et al., 2012, Shin et al., 2015, Shin and Kim, 2017).

Conceptual models simulate a behaviour of a system based on some perception while physically-based models represent the relevant process by physically considering the meaning of the full procedure in a hydrological system. These models, which have been widely used to simulate streamflows, are discussed in the following section.

2.3.3 Physically-based models

Physically-based models are based on physical laws and theoretical principles. These models are characterized by parameters derived by field measurements and have a direct physical significance. The models use a spatial discretization based on grid, hillslopes or some hydrologic response units. Therefore, these models can be highly appropriate when a high level of spatial discretization is needed in modelling. The physically-based models can have many advantages compared to other models because of the use of parameters having a physical interpretation. The limitation of these models is that large data needed, scale-related problems and overparameterization.

One of the best-known physically-based models is the MIKE SHE, which was developed by a consortium of European institutes such as Danish Hydraulic Institute (DHI), British Institute of hydrology and French consulting agency SOGREAH. MIKE SHE is a fully distributed, physically-based, distributed model capable of both single event and continuous simulations. The model can simulate hydrology in plot field and watershed scales (Frana, 2012). The physically-based nature of the model lends inclusion of topography and watershed characteristics (i.e., soil, vegetation and weather parameter sets).

Larsen et al. (2014) presented the results from coupling of the HIRAM RCM at 11 km resolution and MIKE SHE-SWET hydrology and land surface models over the 2,500 km^2

Skjern River catchment, Denmark. Golmohammadi et al. (2014) used MIKE SHE, Soil and Water Assessment Tool (SWAT) and Agricultural Policy Environment extender models to simulate the streamflows of the Canagagigue watershed in the Grand River basin, Canada. The results indicated that the mean daily and monthly flow simulated by MIKE SHE was much better than other models. Three hydrological model (i.e., NAM, SWAT and MIKE SHE) were used to model the combined impact of climate change and land use change on hydrology for a catchment in Denmark in a study of Karlsson et al. (2016). The results indicated that substantial changes in discharge extreme due to the changing of land use.

Physically-based models, including SWAT and the Hydrologic Engineering Center-Hydrologic Modelling System (HEC-HMS) are also used universally to estimate runoff in both gauged and ungauged watersheds. SWAT is a complex physically-based, continuous model and was designed to forecast the impact of watershed management practices on hydrology, sediment, water quality and agriculture production on the gauge and ungauged basins. The model simulates a watershed by dividing it into sub-basins which are further subdivided into Hydrologic Response Units (HRU). For each HRU in every sub-basin, SWAT simulates the soil water balances, groundwater flow, lateral flow, channel routing, evapotranspiration, crop growth and nutrient uptake, pond and wetland balances, soil pesticide degradation and in-stream transformation nutrients and pesticides (Vazquez-Amábile and Engel, 2005).

Abbaspour et al. (2015) used SWAT model to simulate the hydrologic regime for sub-basin scale of Europe. This study contributed essential understanding into continental water resources quantity and water quality at a sub-basin scale with a monthly time interval. An improved version of SWAT model was used to predict the impacts on watershed hydrology and water quality for two watersheds in the Midwest USA (Raj et al., 2016). The study of Lin et al. (2015) showed a varying change in runoff among three time scale (i.e., daily, monthly and annual) and three catchments in the Jinjiang River basin under land use change scenarios using calibrated SWAT model.

HEC-HMS model, a physically-based distributed model, designed to simulate the rainfall-runoff process of dendritic watershed systems, has been widely used to simulate and forecast streamflows in humid, tropical, subtropical and arid watersheds (Ibrahim-Bathis and Ahmed, 2016). HEC-HMS models including Soil Moisture Accounting (SMA) and snow

algorithms for assessing the climate and land use changes of three watersheds in the Great Lakes was used by Gyawali and Watkins David (2013). Their study showed that HEC-HMS model provides a reasonably good estimate of runoff and shows some modest improved results compared with large basin runoff model. HEC-HMS and Watershed Bounded Network Model (WBNM) were used to predict runoff hydrograph from the small urban catchments in Azzaba city. The results indicated that HEC-HMS provided acceptable simulations of the flood events and simulated flood hydrograph was fitted with the realistic situation (Laouacheria and Mansouri, 2015). Performance of two hydrological models, namely HEC-HMS and PRMS, were evaluated by simulating the storm event in Taunton River Basin. Results from a study by Teng et al. (2018) showed that both models could provide flood predictions of rainfall runoff during the storm event. Kabiri et al. (2015) combined SDSM and HEC-HMS models to project the discharge of Klang River, Malaysia under climate change context. An increasing trend in the discharge during the months of June, September and October in three future periods under A2 scenario was indicated in their study.

2.4 Flood Frequency Analysis

Flood Frequency Analysis (FFA) can be used for understanding of the probabilistic behaviour of flood events. Further, FFA is also used for establishing a relationship between flood magnitude and frequency of occurrence (return period) and providing the flood quantile estimate at a given location of interest (Castellarin et al., 2012). FFA can be classified into univariate and multivariate analyses, depending on whether one flood variable or several flood characteristics are considered.

Univariate flood frequency analysis, in term of flood peaks, are considered as a common design criterion in flood control engineering. However, in hydrologic planning and design for flood management, it is used not only to know information about flood peak but also to determine the characteristics of flood volume and duration. Furthermore, flood events with associated flood characteristics can be considered as multivariate events. Therefore, instead of focusing on one flood characteristic, which has been done traditionally, the flood event can be modelled using the joint distribution of several flood characteristics (Karmakar and Simonovic, 2008, Shiau et al., 2006). A review of univariate and bivariate flood frequency analysis is presented in the following sub-sections.

2.4.1 Univariate frequency analysis

The frequency of occurrence of the extreme events is analyzed using statistical probability distribution fitted to the flood series. The first step is to extract the flood series from flood data. AM and POT are commonly used to investigate the changes in flood frequency. AM sample is defined by the maximum peak discharge of each year. However, AM cannot be used in the case of short data series because it provides less information (some peak values, which are still relatively high are not considered in AM series while some low values can be part of the AM series sample) (Bezak et al., 2014, Lang. et al., 1999). Unlike the AM, which extracts only one event per year. POT considers a wider range of events and provides more information than AM, especially for flood magnitude estimation (Kay et al., 2009, Roth et al., 2012). Although POT approach is also widely applied where flood variables have small sample sizes, AM approach is still used in FFA by several researchers. Bezak et al. (2014) compared the results of FFA for data from Litija on the Sava river in Slovenia by using both POT and AM. Their study indicated POT gave better results than AM approach. Similarly, Karim et al. (2017) indicated that frequency estimates based on POT approach are better than AM for small and medium floods while both approaches gave the same results for large floods.

Another important step in FFA is the choice of a Probability Distribution Function (PDF) for the fitting of extreme flood series. There have been several studies which compared the various probability distributions for FFA. These studies indicated that selection of appropriate PDF has an important role in FFA, as a wrong choice could lead to significant error and bias in flood quantile estimation, particularly at higher return periods (Rahman et al., 2013). However, the choice of an appropriate probability distribution is still one of the major issues in FFA. There are many probability distributions which are widely used in FFA for extreme events studies such as Gumbel, Log-Normal (LN), Pearson Type 3 (P3), Log-Pearson Type 3 (LP3), GPD and Generalized Logistic (GL).

Rahman et al. (2013) used five PDFs to analyze the frequency of flood series in Australia. The results showed that LP3, GEV and GPD are the three of the best-fit distributions. Seven different PDFs were considered for Tasmania and the results indicated that LN distribution provided the best-fit to the observed flood data in a study carried by Haddad and Rahman (2011). Gumbel distribution provided a better fit to the flood data than LN, P3, LP3,

and 3-parameters LN distributions for FFA in the Balu-Tongikhal River system in Dhaka city which was indicated by a study of Gain and Hoque (2013).

2.4.2 Bivariate frequency analysis

In contrast to the univariate frequency analysis in which only one flood variable is used in the analysis, multivariate frequency analysis simultaneously considered flood characteristics (i.e., flood volume, duration and peak). Multivariate parametric distributions (e.g., bivariate normal, bivariate gamma, bivariate extreme value distributions, etc.), which have been extended from univariate distribution, have been used to model multivariate flood characteristics for different purposes.

Sheng (2001) used bivariate gamma distribution to model the joint probability behaviour of bivariate flood characteristics in the Madawask River basin in Quebec, Canada. The results indicated that bivariate gamma distribution could be useful for multivariate extreme events. A study using bivariate extreme value distribution in FFA was conducted by Shiau (2003) and another study was done by Escalante-Sandoval (2007). These two studies concluded that bivariate extreme value distribution showed a good agreement between theoretical model output and observed data. They also provided more useful information than univariate frequency analysis. Similarly, under the assumption that flood peak and volume have the same type of marginal distributions, Yue and Wang (2004) used bivariate extreme value distribution for FFA of these variables. However, bivariate distribution functions cannot model the joint probability behaviour of the flood variables that are inter-correlated. Furthermore, for the flood events, all flood characteristics do not have the same type of marginal probability distribution. Therefore, the copula approach may provide a flexible solution (Genest et al., 2007).

The copula is a function that links univariate distribution functions to form bivariate distribution functions. The main advantage of this method is that the dependence structure is independently modelled with the marginal distributions allowing for different type of distribution (Dupuis, 2007, Zhang and Singh, 2007). Several researchers have used copulas approach to investigate the bivariate frequency analysis.

For example, Duan et al. (2016) used copula approach to investigate the flood characteristics in the Huai River basin under climate change condition. Their study indicated

that copula can be a viable and flexible tool for FFA and may provide useful information for risk-based flood control. Filipova et al. (2018) applied copulas to present a method for selecting different copulas for modelling the joint probability of flood peak and volume for 27 catchments in Norway. They suggested that two parameter copulas BB1 and BB7 should be selected in catchments with high steepness, high mean annual runoff and rainfall flood regime. Additionally, Sraj et al. (2015) applied different bivariate copulas from three families (i.e., Archimedean, extreme value and elliptical) to carry out the bivariate flood frequency analysis of flood characteristics of Sara River in Slovenia. Their study found that Gumbel-Hougaard copula as the most appropriate for the pair of flood characteristics. Similarly, Karmakar and Simonovic (2009) indicated that Gumbel-Hougaard copula is the best approach for modelling the dependence structure between flood characteristics of Red River at Grand Forks, Dakota by using Archimedean family copula.

2.5 Flood Hazard Maps

All parts of any river basin are vulnerable to floods in different degrees under different cases and situations. Flood hazard maps are used most commonly for flood risk communication and management (Luke et al., 2018). A flood hazard map is a useful tool for decision and policymakers and local authorities to design protection measures in the river basin. Hazard maps provide information on the probability of the flooding for different return periods as well as the depth and extent of the spread of floodwater in the affected areas. Therefore, it is important to identify potential inundation areas with high appropriate level of accuracy. This information also provides emergency organizations to calibrate and adjust their warning system and prepare priority evacuation plans. It may be used to find the best strategies for risk reduction.

There are two main approaches to develop the flood hazard maps, namely deterministic and probability approaches. The most common representation of simulation results is a deterministic flood inundation map based on a single simulation. Probabilistic flood mapping designed to incorporate uncertainty from input data and model parameters, represent spatial and temporal risk and present flood maps in terms of probabilities and percentages.

2.5.1 Probabilistic approach

In the probabilistic approach, the process of floodplain mapping requires certain steps. These steps include: (i) the setting up of flood inundation models; (ii) sensitivity analysis of the model using historical flood data and (iii) ensemble simulation using an uncertainty design event. The probabilistic approach, which is based on ensemble simulation, does not necessarily require the use of physical behaviour of the river and floodplain models. Di Baldassarre et al. (2010) compared two different methods (i.e., deterministic and probabilistic) for flood hazard mapping using 2D hydrodynamic hydraulic model. Their study indicated that flood hazard mapping using probability approach seems to be more reliable. Kalyanapu et al. (2012) used Monte Carlo based 2D flood inundation framework for estimating flood hazard mapping. Their study showed that the probability-weighted flood risk approach provides improved accuracy of flood risk estimation.

However, the main disadvantage of using physically-based 2D hydraulic models in probability frameworks, has been the simulation time required for each simulation. Simulating hundreds of flood events with these computational speeds would take large computer time making 2D model application counterproductive (Timbadiya et al., 2015). A probability analysis with 2D hydraulic models has been limited to a smaller number of scenarios and smaller spatial domains. Additionally, Aronica et al. (2012b) suggested that flood inundation probability alone may be insensitive to discharge in relatively steep urban catchments and maybe a limited measure of flood hazard. Furthermore, Thompson and Frazier (2014) supposed that a few probabilistic flood hazard maps were limited with respect to the hazard behaviour they modelled. These models could also be computationally expensive and parameterization was difficult to compute for forces that were not fully predictable.

The deterministic approach, which used physically-based 2D flood modelling to simulate synthetic design flood events, will overcome these problems to develop flood hazard maps. A review of a deterministic approach will discuss in the following part.

2.5.2 Deterministic approach

In a deterministic approach, floodplain maps consist of construction of a physically-based fully 2D hydraulic model, calibration and validation of the model using historical flood

event, using the best-fit statistical model to generate the design flood hydrograph and elaboration of the model results to generate flood hazard maps. Deterministic modelling tools have widely been applied because they are capable of translating changes in input parameters into a change in flood characteristics.

Flood inundation depth and inundation extent can be computed using computational models based on solutions of the full or approximate form of the shallow water equations. 2D hydrodynamic models are identified as the appropriate tools for simulating the flow of water over flat terrain and complex topography. Such model results provide further opportunity to develop more meaningful hazard maps by incorporating additional hazard parameters.

The high-resolution flood hazard maps, which was developed using the advanced deterministic and probability approaches, can provide complete information about the physical hazard and reduce uncertainty found in traditional approaches. Masood and Takeuchi (2012) developed flood hazard maps using 1D hydrodynamic model for the city of Dhaka in Bangladesh. These studies used a simple form of deterministic approach in establishing flood hazard maps. Mazzoleni et al. (2014) suggested a semi-probabilistic approach to develop the hazard map due to embankment-overtopping for the Po River basin. They used 1D and 2D hydrodynamic models to simulate the hydrodynamic regime (i.e., water depth and flow velocity) and the flood hazard maps were obtained using the hazard curves, which combined different flood parameters (i.e., flood extent, water depth and flow velocity). Mosquera-Machado and Ahmad (2007) used flood frequency analysis and 1D hydraulic model (HEC-RAS) to create three flood hazard maps with different return periods for Atrato River, Columbia. Their study provided useful information in evacuation planning as well as damages estimating.

2.6 Conclusions

In this chapter, an overview of the literature on downscaling methods for climate change projections, hydrological modelling, flood frequency analysis and flood hazard maps estimation are presented. It is seen that extreme hydrological events (i.e., floods and drought) are likely to increase in frequency, duration and magnitude in the sensitive climate regions. Besides, climate change is expected to increase the magnitude and frequency of extreme events and likely to

cause more intense and frequent floods. Therefore, it is crucial to examine whether the magnitude of flood characteristics will remain the same or will change under climate change scenarios at river basin scale. This study assesses the potential impact of climate change on flood frequency using the POT approach.

One of the measures to mitigate the flood damage is providing useful information through floodplain areas, the spatial distribution of flood hazard. Therefore, it is of great importance for understanding flood hazard at river scale. This research aims to determine the flood hazard maps for Saigon-Dongnai River system, Vietnam under climate change context in formulating climate adaptation and risk mitigation strategies. Combining the downscaling methods, which are used to downscale climate variables from climate model outputs, hydrological models, flood frequency analysis and hydrodynamic models, the flood hazard maps have been developed.

Chapter 3

Methodology

3.1 Introduction

The methodology is presented in this chapter aims to address all the research objectives. The description and linking of each objective corresponding to the methodology are also introduced in this chapter.

3.2 Methodology

In order to achieve all the research objectives listed in section 1.4, a framework of the methodology adopted is presented in Fig. 3.1. The aim of this study is to assess the changes of flood risk under impact of climate change context. In order to achieve the general objective, the specific objective along with the corresponding actual methodology are presented as follow. The first objective is obtained by using the univariate flood frequency analysis. For each POT dataset, the best-fitting distribution is selected by testing several distributions

(i.e., GPD, P3, LP3, Gumbel, LN and GL) and they are presented in Chapter 4. Downscaling, hydrological modelling and flood frequency analysis methods are used to achieve the second objective in this study (Chapter 4). The third and fouth objectives can be acquired by using the LLHR, tail dependence test and copula approach respectively and they are presented in Chapter 5.

Bivariate flood frequency analysis using copula and hierarchical cluster analysis are performed to achieve a fifth objective (Chapter 5). The sixth objective and final objective can be obtained by simulate the flood depth and velocity variables using coupled hydrodynamic and they are presented in Chapter 6.

The methodology proposed for each part of the present research to achieve a specific objective is applied for the study area and presented as case studies. Three case studies are considered. Each case study is summarised below and detailed in its corresponding section. The details of the available software which are used in this study is also presented in the following sections.

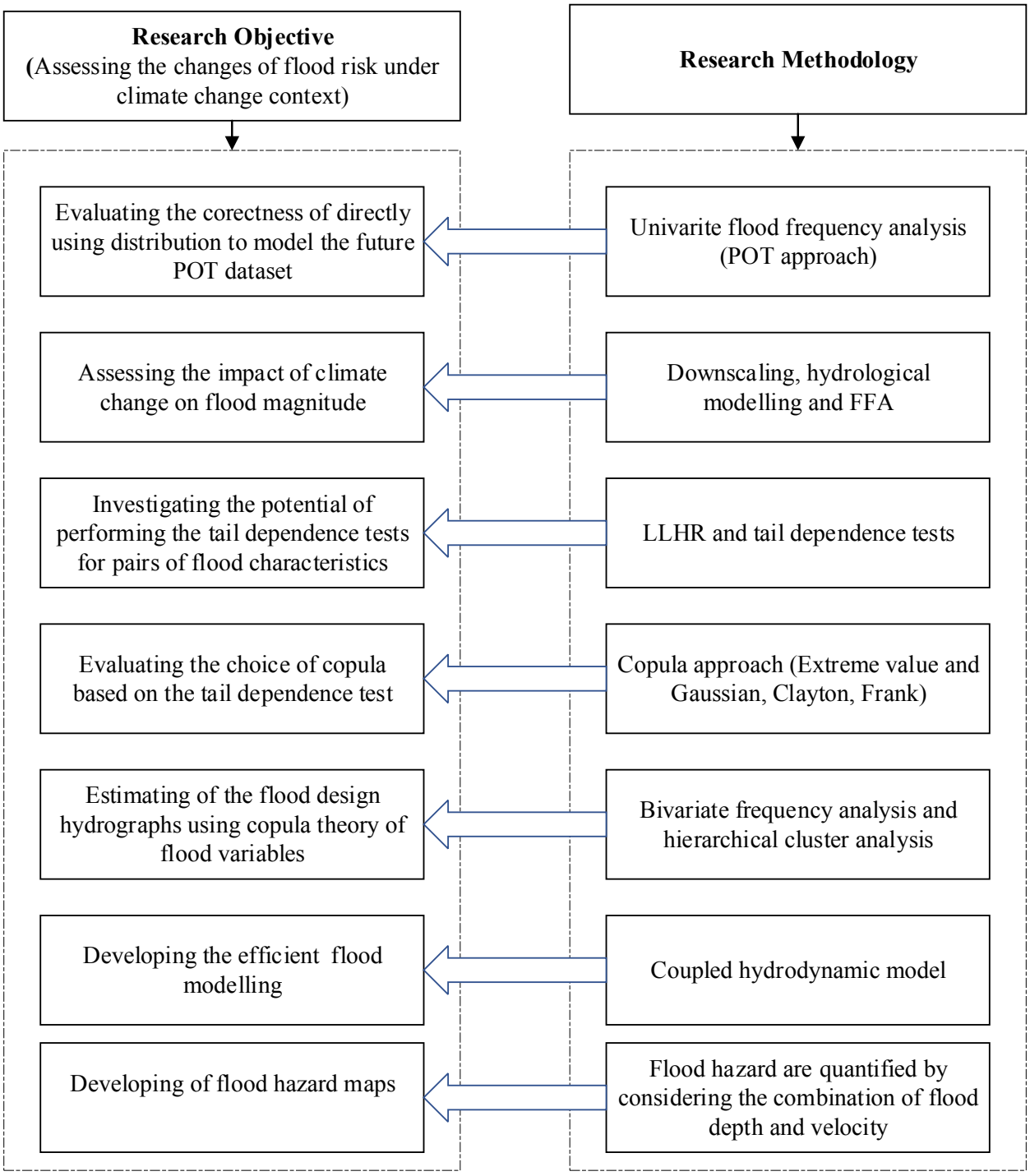


Fig. 3.1: A methodology schematic

3.2.1 Case studies

To investigate the changes of flood frequency, flood series extracted using POT approach are applied to the data of the Trian gauging station in the Saigon-Dongnai River in the South of Vietnam. In this case study, the monthly change factors concerning the baseline period of 1980-2005 are calculated for the future period (2020-2045) by using RCMS outputs. These change factors are used as the input data for LARS-WG to generate the future climate time series. The continuous hydrological model (i.e., HEC-HMS) is used to generate future discharge time series based on the future climate series. The best fitting distribution is selected by testing several distributions that are normally used in extreme value analysis, namely GPD, P3, LP3, Gumbel, LN and GL distributions. This case study also evaluates the correctness of directly using the asymptotic distribution to model the future POT dataset.

The bivariate flood frequency analysis based on copulas approach is performed on a Trian streamflow gauge located in the South of Vietnam. To find the best fit marginal distributions both parametric and nonparametric family of distribution are used in this study. Gumbel, GPD, LN, P3, LP3 and GEV belonging to the parametric distributions along with nonparametric kernel distributions are evaluated. The selection of copulas in this study is decided based on the tail dependence test. Three extreme value family of copulas (i.e., Gumbel-Hougaard, Galambos and Husler-Reiss) are evaluated to model asymptotically dependence pair of flood characteristics. Clayton, Frank and Gaussian copulas are used to assess the potential their application in case of variables are diagnosed as asymptotic independence.

The flood hazard maps are quantified by considering the water depth and velocity in combination for the downstream of Trian catchment. In this case, the coupled 1D and 2D hydrodynamic model (MIKE FLOOD) is used to simulate the flow regime. The design flood hydrograph obtained from bivariate flood frequency analysis. is used as the input for the hydrodynamic model. Besides, high-quality topographic data (i.e., DEM and LiDAR) of input for hydraulic models, collected using modern survey techniques, are also used to improve the accuracy of flood hazard maps. The latest advancement in flood modelling technique is the development of flexible meshes. These meshes allow constructing a finer resolution of smaller features to reflect the changes in the topography and coarser resolution at the broader floodplain without resorting model grid nesting to save computational effort. Therefore, 2D flexible mesh

is selected for the hydrodynamic modelling in this study.

3.2.2 Available software

The statistical community is providing a great number of functions for helping in performing flood frequency analysis using the R software. R is a free and open sources software for statistical computing and graphics. Several well-known statisticians and computational scientists have reviewed it. A summary of the R-packages used in performing a specific computation in this research is included below: (i) the copula (Kojadinovic and Yan, 2010b) and Vine copula (Schepsmeier et al., 2012) packages for copula modelling, (ii) ismev (Heffernan et al., 2012), extRemes (Gilleland and Katz, 2016) and nsRFA (Viglione et al., 2018) packages for marginal distribution modelling; (iii) Kendall (McLeod and AI, 2011) package for trend analysis; (iv) POT (Ribatet, 2007) package for testing tail independence test and plotting of parameters estimated at various thresholds.

The free software LARS-WG is used to downscale the climate data that are obtained from the output of GCMs and RCMs. LARS-WG is a stochastic weather generator which can be used for the simulation of weather data at the single site. LARS-WG has been considered as a computationally inexpensive tool to produce daily site-specific climate scenarios for impact assessment of climate change (Semenov, 2008). LARS-WG can be downloaded from http://resources.rothamsted.ac.uk/mas-models/larswg.html.

The free software HEC-HMS and HEC-GeoHMS are used to simulate the hydrological modelling. HEC-HMS is used for modelling rainfall-runoff processes for a dendritic watershed while HEC-GeoHMS is used to analyze digital terrain data and transforms the drainage paths and watershed boundaries into a hydrological data structure. The HEC-HMS and HEC-GeoHMS software can be downloaded from the USACE website at http://www.hec.usace.army.mil/software/hec-hms.html.

The commercial flood modelling package MIKE FLOOD is used to simulate the flood propagation, flood depth, flood velocity and flood inundations extent in this study. MIKE FLOOD dynamically links two independently software, namely MIKE 11 and MIKE 21 and they have been developed by the DHI, Denmark. MIKE 11 and MIKE 21 are coupled by using links in transferring water between the channels and the overland of the model domain. There

are several types of links that can be used in various situations. Standard link, which was commonly applied in flood modelling, is used in this study. The licensed of this model is available with the Department of Civil Engineering of National Institute of Technology, Warangal, India.

Chapter 4

Impact of Climate Change on Flood Frequency of the Trian Reservoir in Vietnam Using RCMS

4.1 Introduction

An increasing trend in frequency as well as intensity of extreme rainfall events is being observed in Southeast Asia and it is projected that more frequent and intense extreme events will occur in this region (IPCC, 2014). These changes may have a significant impact on the hydrological cycle of this region (Setegn et al., 2011). The changes in precipitation combined with evaporation have potential impacts on the runoff, especially, increasing the frequency and intensity of flood. Flood events may cause negative impacts on economic and societies aspects. For example, the flood can lead to lower crop production, human and material losses as well as cause a negative impact on the natural ecosystem. Therefore, it is necessary to analyse the changes in flood magnitude under climate change context to provide suitable measures in mitigating climate change impacts.

Vietnam is one of the countries that is severely impacted by climate change (Rutten et al., 2014, Trinh et al., 2013). The Saigon-Dongnai River basin is one of the largest river basins as well as the economic center in the South of Vietnam. Trian watershed is located in the upper Saigon-Dongnai River basin and it is the biggest sub-basin of this river. In order to meet water and energy demands as well as flood control for cities located in the downstream (i.e., Hochiminh, Bienhoa, Vungtau, etc.), Trian reservoir is being operated for multiple purposes such as hydropower production, water supply and flood control. However, till date, no study has been conducted to assess the impact of climate change on this basin. Hence, it is necessary to evaluate the potential future impact of climate change on this watershed, particularly on flood frequency.

Flood estimation cannot be done purely by statistical analysis because the characteristics of the flood events can change in the future due to climate change. Therefore, the physically-based meteorological and hydrological modelling should be used (Booij, 2005). The framework for evaluating flood frequency under climate change contexts related to physically-based approach are widely applied by many researchers (Arnell and Gosling, 2013, Hirabayashi et al., 2013, Kay and Jones, 2012). This framework includes three steps: (i) extraction of climate data series from the future climate change scenarios; (ii) simulation of future discharge series using any hydrological model; (iii) analyzing the flood frequency based on the projected discharge series using the statistical approach.

GCMs are widely used to predict the changes in atmospheric variables under climate change scenarios (Anandhi et al., 2011). GCMs are the physics-based meteorological models, which represent atmospheric and oceanic dynamics (Angeles et al., 2007). However, the output from GCMs normally has a low spatial resolution of approximately 100-250 km and it is inadequate for regional impact studies especially for analyzing the changes in extreme precipitation and floods because it is lacking detailed regional information which is needed to resolve various features at the catchment scale (Fowler et al., 2007).

In order to overcome these obstacles, statistical and dynamical downscaling approaches are used as a common approach (Prudhomme et al., 2002). On the other hand, RCMs are frequently used to analyse the impact of climate change on hydrology in the watershed because of its higher resolution. The resolution of RCMs is around 12-50 km and it accounts for the

sub-GCM grid scale forcing (e.g., complex topographical features and land cover heterogeneities in a physics-based way). However, RCMs inherit the biases due to systematic model errors caused by imperfect conceptualization, discretization and spatial averaging within grid cells. Therefore, statistical downscaling is also performed for RCM projection (Chen et al., 2011, Sunyer et al., 2012).

There are several statistical downscaling methods which establish statistical links between large-scale weather and observed local-scale weather. Maraun et al. (2010) classified statistical downscaling approaches into PP, MOS, and WGs. PP method establishes statistical relationships between variables at large scales and local scales. MOS establishes statistical relationships between variables simulated by the RCM and local scale observation to correct RCM errors. WGs approach generates local scale weather time series resembling the statistical properties of observed weather. WGs are extensively used to generate the daily time series statistically similar to the observed climate data, which are used as an input data for the hydrological model to assess the impact of climate change on hydrological risk. Furthermore, it is a useful approach for assessing the change in extreme events (Semenov, 2008, Wilks and Wilby, 1999).

Among several WGs, LARS-WG, based on semi-empirical distributions, is probably the best semi-empirical weather generators (Semenov and Stratonovitch, 2010). More importantly, the semi-empirical distribution in LARS-WG is very flexible because it can be fitted to several probability distributions (Mikhail et al., 1998). Further, this weather generator correctly reproduced most of the characteristics of the observed data in the Asian region (Semenov and Stratonovitch, 2010). All semi-empirical distribution parameters of weather variables are determined using observed daily climate data. These parameters combined with the relative change factor of length of wet and dry spell and mean of precipitation amount as well as an absolute change in temperature derived from GCMs or RCMs outputs are used to generate time series for future.

To investigate the changes in flood frequency, AM and POT are commonly used to extracts the flood events. AM sample is defined by the maximum peak flow of each year. However, AM cannot be used in the case of short data series because it provides less information (some peak values, which are still relatively high are not considered in AM series

while some low values can be part of the AM series sample) (Bezak et al., 2014, Lang. et al., 1999). Unlike the AM, which only extract one event per year, POT considers a wider range of events and provides more information than AM, especially for estimation of flood magnitude (Kay et al., 2009, Roth et al., 2012).

Normally, the GPD is fitted with POT magnitudes as a default in extreme values analysis (Coles et al., 2001, Davison and Smith, 1990, Katz et al., 2002). However, there could be more than one distribution that can be fitted to the sample data. Hence, identifying the best fitting distribution to the sample need to be tested with several distributions (Lang. et al., 1999). The LN, GPD, P3, LP3, Gumbel, and GL have been widely used for modelling extreme values (Bezak et al., 2014, Lang. et al., 1999, Saf, 2009a, Salas et al., 2012). Furthermore, many studies used AM to assess the impact of climate change on flood frequency at global as well as regional scale (Hirabayashi et al., 2013, Jung et al., 2011) while there are only a few studies which used POT approach. Even the studies which used POT approach mostly considered only a single distribution (i.e., GPD). However, in addition to the GPD, researchers have shown that the peak values can be fitted with many distributions such as Gumbel, LP3, LN, exponential distributions and etc. (Sarhadi et al., 2012, Seckin et al., 2011, Zaman et al., 2012). Thus, it is important that numerous distributions which are normally used in extreme values analysis should be tested and the best fitting distribution needs to be identified for each POT dataset.

Consequently, this study will evaluate the potential impact of climate change on flood frequency of the Trian reservoir belonging to the Saigon-Dongnai River basin, Vietnam. In addition, this study also evaluates the correctness of directly using the asymptotic distribution to model the future POT dataset.

4.2 Study Area and Data

The Trian catchment is one of the largest catchments in the south of Vietnam, which is a part of the Saigon-Dongnai River basin. The total area of the Trian catchment is 14,200 km^2. The basin lies between latitudes of 10°53'46" N and 12°22'08" N and longitudes of 107°01'52" E and 108°46'55" E (Fig. 4.1). The Trian basin lies in the monsoon tropical zone, which is affected by the North-East and South-West monsoon. The climate regime is divided into two distinct seasons. The rainy season is from April to November and the dry season is from

December to March of the following year. The annual average rainfall in the Trian basin is about 2,200 mm, while the annual average temperature is 20.6ºC. The maximum temperature is 36.6ºC in the month of April while the minimum temperature of 4.5ºC was recorded in the month of January. The mean monthly evaporation is 80.3 mm. The seasonal variation shows that evaporation is high in the dry season, especially in the month of March with the value of 118.6 mm per month while the evaporation is low during the rainy season which occurs in September (48.5 mm per month).

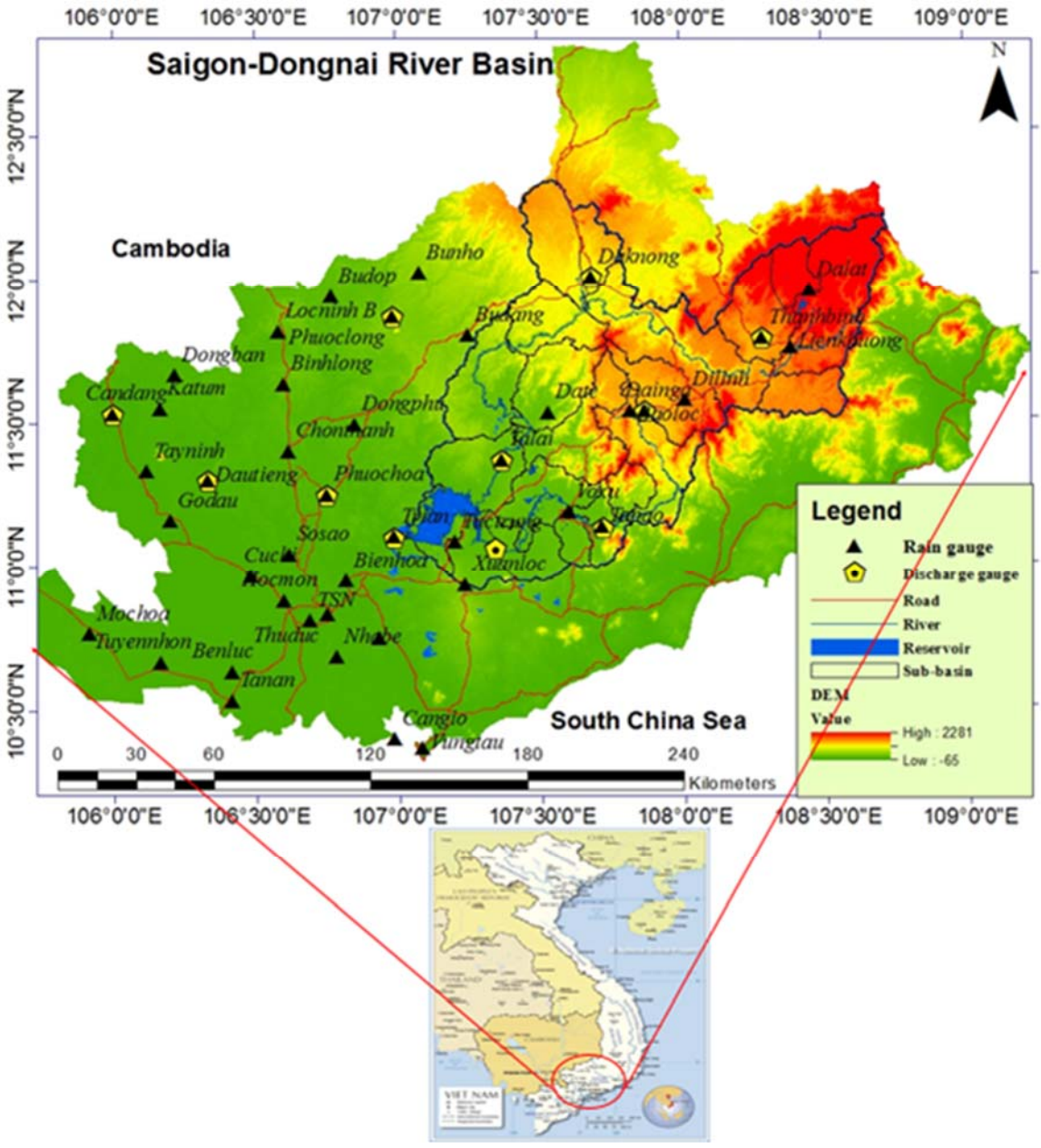


Fig. 4.1: The hydro-meteorological network in the study area

There are 13 rain gauges, six stations located in the Langa River basin and the others lying in the Dongnai River (Fig. 4.1). Baoloc, Daknong, Dalat, and Lienkhuong are the four weather stations which record all meteorological variables (i.e., rainfall, temperature, relative humidity, sunshine hour data, speed and direction of wind). The remaining stations only measure the precipitation data except for Xuanloc station which records temperature, evaporation, and humidity. There are seven streamflow gauges which cover the entire basin. Daknong, Talai, Trian, and Thanhbinh are located in the Dongnai River in which Trian is the main outlet. Tapao, Phudien, and Dainga are located in the Langa River. However, three main streamflow gauges (i.e., Trian, Tapao and Talai) which are located in the main river are used for calibration and validation procedure. Rainfall and runoff data are collected from the National Hydro–Meteorological Service (NHMS) of Vietnam. Additionally, a Digital Elevation Model (DEM) with 30 m spatial resolution is downloaded from Shuttle Radar Topography Mission (SRTM) while the Land Use Land Cover (LULC) and soil type digital maps are collected from the Ministry of Natural Resources and Environment (MONRE), Vietnam.

4.3 Methodology

The methodology used in this study is shown as a flowchart in Fig. 4.2. LARS-WG is used to downscale the climate data which are obtained from the RCMs and GCM for 10 stations in this study. The LARS-WG parameters in each station are determined based on the observed climate data (1980-2005). The monthly change factors with respect to the baseline period of 1980-2005 are then calculated for the future period (2020-2045) using RCMs and GCM outputs. These change factors are used as input data in LARS-WG to generate the future climate time series. Next, the future discharge series are simulated using a continuous hydrological model (HEC-HMS) and the future climate data are simulated using LARS-WG. The impact of climate change on flood frequency is analysed using the extreme discharge series extracted using the POT approach. For each POT series, the best fitting distribution is selected by testing several distributions which are normally used in extreme values analysis such as, GPD, P3, LP3, Gumbel, LN and GL distributions. In this study, the parameters of the distribution are estimated using the method of maximum likelihood and the return levels of flood magnitude are determined for different return periods ranging from 2 to 100 years.

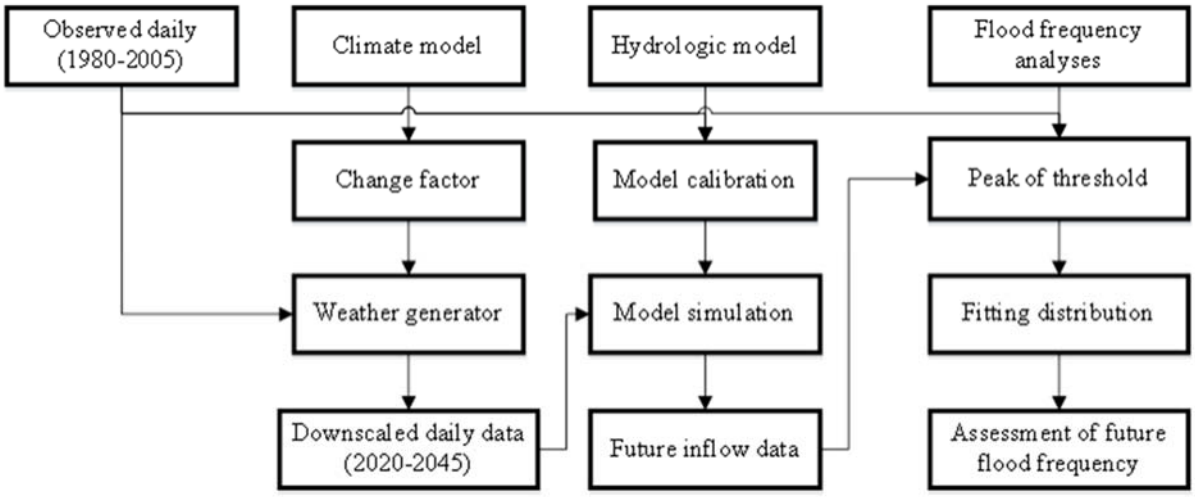


Fig. 4.2: The framework for assessing the changes in flood frequency

4.3.1 Statistical downscaling

CORDEX project provides a quality controlled dataset of Regional Climate Downscaling (RCD)-based information for historical and 21st-century projections (Park et al., 2016). Two scenarios (i.e., RCP 4.5 and 8.5) from five RCMs, namely, HadGEM3-RA (Had), SNU-MM5 (MM5), SNU-WRF (WRF), RegCM4 (Reg) and YSU-RSM (YSU) which are from the same GCM (HadGEM2-AO) belonging to the CORDEX-EA projects are used in this study. Table 4.1 lists the details of CORDEX-EA RCMs used in this study.

Table 4.1: The description of GCM and RCMs

Expansion	GCM	RCMs	Scenarios
National Institute of Meteorological Research (MOHC)	HadGEM2-AO	HadGEM3-RA	RCP 4.5&8.5
Kongju National University (ICTP)		RegCM4	RCP 4.5&8.5
Seoul National University Meso-scale Model version 5 (MM5)		SNU-MM5	RCP 4.5
Seoul National University (WRF)		SNU-WRF5	RCP 4.5&8.5
Yonsei University (Regional Spectral Model)		YSU-RSM	RCP 4.5&8.5

Fig. A. 1 (Appendix A) shows the boxplots of observed, different RCMs and GCM simulated annual precipitation during the base period (1980-2005) for four rainfall stations. There is a significant difference between the observed and raw RCMs/GCM annual rainfall for

rain gauges of the study area.

The Q-Q plots compare the quantiles of observed and simulated data. The Q-Q plots (Fig. A. 2) show a large bias between observed and RCMs/GCM simulated daily rainfall of four stations in the study area. Note that similar results are observed at the remaining six stations of the study area but, in the interest of brevity, boxplots and Q-Q plots are not shown for those stations. The simulated daily rainfall was underestimated in both GCM and Had while other RCMs overestimated the daily rainfall. These plots (Fig. A. 1 and Fig. A. 2) show that direct output of RCMs and GCM cannot be used as input data for the hydrological model. Therefore, statistical downscaling should be performed to adjust the GCM and RCMs simulated precipitation and temperature.

To generate future weather details (such as precipitation and temperature), first the change factors are calculated using the outputs from the RCMs/GCM. Then, along with observed data, these change factors are used as input in weather generator to simulate future weather details (Fowler et al., 2007, Kilsby et al., 2007, Sunyer et al., 2012). In this study, the same methodology is used for simulating future weather details for the study area.

LARS-WG is a stochastic weather generator which is generally used to downscale the climate data. There are several studies that compared LARS-WG to other statistical downscaling models (Duan and Mei, 2014, Hassan et al., 2014, Mehan et al., 2017, Semenov et al., 1998, Sunyer et al., 2012). Most of these studies indicated that the LARS-WG can be used as an effective tool for assessing the impact of climate change at local scale. In additional, LARS-WG is widely used in studying the impacts of climate change on hydrological variables. For example, Agarwal et al. (2014) used LARS-WG to downscale precipitation in assessing the future precipitation in Nepal. Qin and Lu (2014) assessed the impact of climate change on flood frequency in China region by combining the LARS-WG and the hydrological model. More importantly, LARS-WG is good in modelling the extreme rainfall events (Lu et al., 2015), which is the main cause of floods.

LARS-WG is used to downscale the climate data which are obtained from the RCMs and GCM for 10 stations in this study. The LARS-WG parameters (i.e., wet/dry spell length, daily precipitation, minimum and maximum temperatures) are analyzed for each station using the baseline period (1980-2005) observed climate data. KS, t and F tests are used to ensure the

results of this model are reliable. The change factors for two future scenarios (i.e., RCP 4.5 and 8.5) are calculated using the output of five RCMs and one GCM. In particular, relative changes in precipitation amount, wet/dry durations and temperature variability (standard deviation) are calculated using Eq. (4.1). For minimum and maximum temperature amount, the absolute changes are calculated using Eq. (4.2).

$$CF_i = \frac{Future_i}{Baseline_i} \tag{4.1}$$

$$CF_i = Future_i - Baseline_i \tag{4.2}$$

where, CF_i indicate the change factor for the i^{th} month. For example, the change factor for precipitation amount in the month of January is calculated by dividing average January month rainfall amount in future time period (2020-2045) with the average January month rainfall amount in baseline time period (1980-2005). These change factors are used as input for LARS-WG to generate the daily precipitation and daily temperature for both future scenarios (i.e., RCP 4.5 and 8.5).

4.3.2 Hydrological model

The hydrological model used in this study is HEC-HMS version 3.5 developed by the US Army Corps of Engineers. HEC-HMS is a well-known model for use in the long-term continuous hydrological modelling. It is a semi-distributed model with horizontal structure realized via sub-basins. This model consists of four main components (i.e., basin component, metrological, control specification and summary and display model outputs) which are combined for a simulation run. Besides, the Geospatial Hydrological Modelling Extension (HEC-GeoHMS) is also used to create river and basin properties and cross-sections used in the Muskingum-Cunge routing with the help of 30m SRTM DEM.

HEC-HMS provides the Soil Moisture Accounting (SMA) module for continuous modelling. SMA, which is designed to compute basin surface runoff, groundwater flow, losses attributable to potential evapotranspiration, and deep percolation over the entire basin, is used in this study. Fig. 4.3 shows the conceptual diagram of HEC-HMS SMA continuous algorithm.

SMA uses five tank storages to simulate the different parts of the rainfall-runoff process (Dariane et al., 2016). The Soil Conservation Service (SCS) dimensionless unit hydrograph is used to transform the excess precipitation into a flow hydrograph at the outlet of each basin. The potential evapotranspiration is calculated by the Priestley-Taylor method.

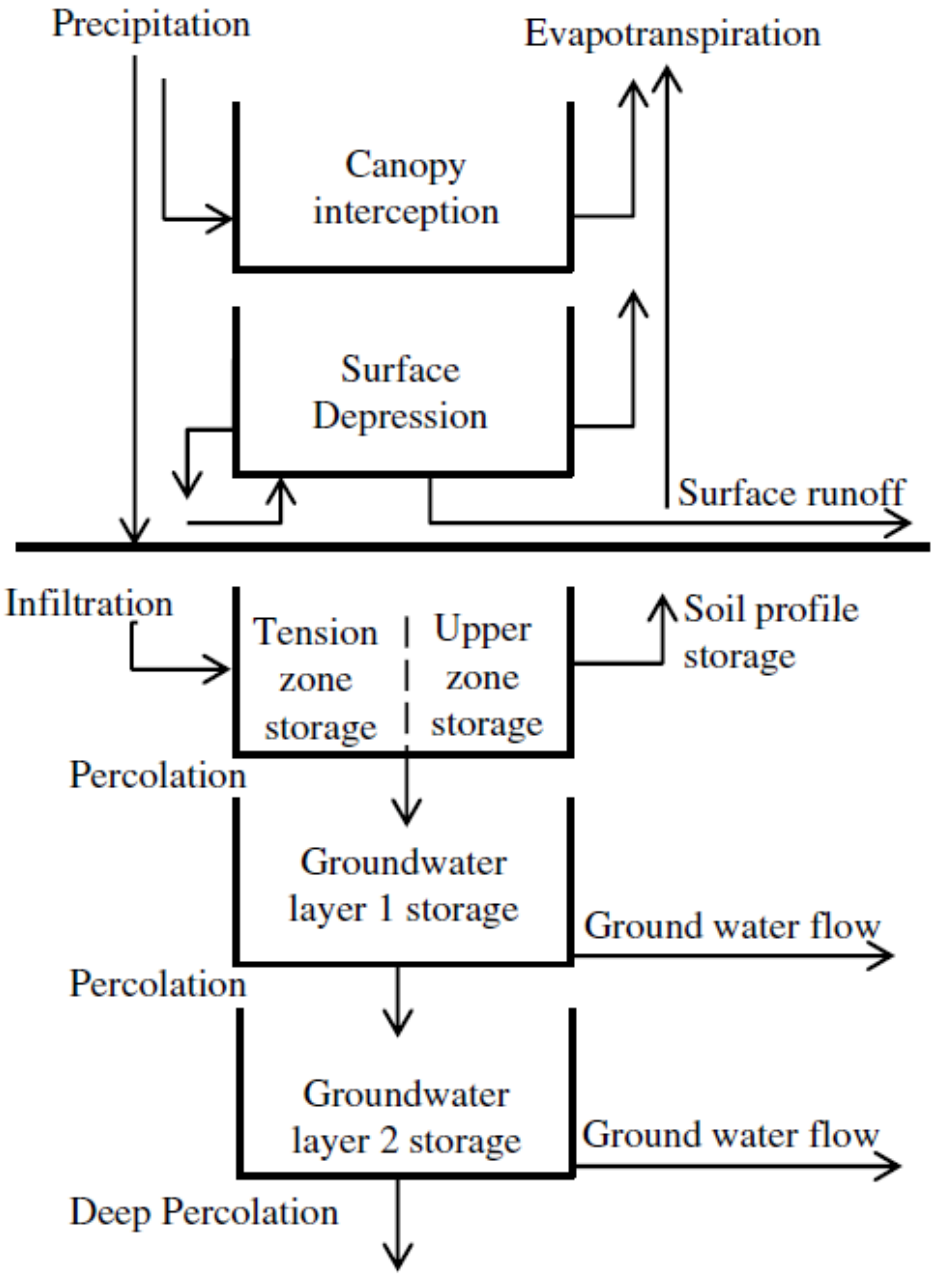


Fig. 4.3: The conceptual diagram of SMA continuous algorithm in HEC-HMS (Gyawali and Watkins David, 2013)

Precipitation, evapotranspiration, infiltration excess overland flow, saturation excess overland flow, subsurface storm flow, subsurface flow and river flow are important physical processes which play a key role in assessing the impact of climate change on flood (Booij, 2005). A conceptual model is a suitable choice to meet the above requirements because all important physical processes are considered as well as because of its simplicity (fewer data requirements). Continuous simulation estimates the losses and generates the streamflow by simulating the wetting and drying of a catchment at daily, hourly time steps (Boughton and Droop, 2003, Pathiraja et al., 2012). Further, the continuous hydrological models are widely used to assess the effect of climate change as well (Booij, 2005, Cameron, 2006, Raff et al., 2009). The main advantage of the continuous hydrological model is that the soil moisture condition is continuously simulated. It is to be noted that the soil moisture condition is an important component in flood modelling, especially for catchment with large storages as well as large difference in runoff between rainy and dry seasons (Pathiraja et al., 2012). SMA embedded in HEC-HMS is designed to compute basin surface runoff, groundwater flow, losses due to potential evapotranspiration and deep percolation over the entire basin (Feldman and Center, 2000).

The HEC-GeoHMS is used to create river network, basin properties and cross-sections which are used in the Muskingum-Cunge routing with the help of 30m SRTM DEM. Meteorological data recording sites in each sub-basin (22 sub-basins) are presented in Fig. 4.1. Theissen polygon approach is used to compute the areal average precipitation based on the available rain gauges. The SCS dimensionless unit hydrograph is used to transform the excess precipitation into a flow hydrograph at the outlet of each basin. The lag time (T_{lag}) is the main parameter for this method. Lag time is the time difference between the centroid of rainfall excess and the centroid of the Direct Runoff Hydrograph (DRH). Lag time can be estimated from the watershed characteristics using Curve Number (CN) by the SCS formula (Ponce, 1994) and it is given by Eq. (4.3).

$$T_{lag} = \left(\frac{L^{0.8} \times (2540 - 22.86 \times CN)^{0.7}}{14{,}104 \times CN^{0.7} \times Y^{0.5}} \right) \tag{4.3}$$

where, T_{lag} is the catchment lag time in hours, L is the hydraulic length measured along the main river in meters and Y is the average catchment slope in meter per meter. The hydraulic

length and the average catchment slope are derived from river and basin properties and the CN is determined based on the several factors such as hydrologic soil groups, LULC and antecedent moisture condition. Twelve parameters and five initial conditions of SMA algorithm are required to characterize the canopy, surface, soil and groundwater storage units.

The estimation of these parameters is based on the processing of the LULC, soil map using GIS and streamflows analysis. Penman-Monteith and Priestly-Taylor formulations are normally used to estimate the potential evapotranspiration (Alfieri et al., 2015, Kay et al., 2006, Taye et al., 2011). Arnell (1999) indicated that there was no difference in estimating runoff for several regions in Europe using two potential evapotranspiration formulations. Booij (2005) suggested that the Priestey-Taylor formulation should be preferable to calculate the potential evapotranspiration if meteorological data is not available. Hence, the Priestey-Taylor formulation is used in this study for the future climate change scenarios.

In this study, both manual and automatic methods are used to calibrate the hydrologic model. Manual calibration is used to determine a practical range of parameter values, while automated calibration is used to refine parameter values. The objective function (percent error peak) is used in automating calibration steps. Based on the river systems and available hydrological data in this region, a period of 1985-2005 is selected for the calibration of parameters in the model which are adjusted using three streamflow gauges in Talai, Tapao and Trian stations. The remaining streamflow data from 2006 to 2013 is used for the validation step.

Normally, the performance of the model is evaluated by comparing observed and model discharge in the catchment outlet (Krause et al., 2005). Moriasi et al. (2007) analysed numerous of model evaluation technique for watershed model evaluation guidelines for streamflow and suggested the use of quantitative statistics such as Root Mean Square Error-observation Standard Deviation Ratio (RSR), Nash and Sutcliffe (NSE) and Percent Bias (PBIAS). The formulation of RSR and PBIAS are presented in Eq. (4.4) and Eq. (4.5). The range of NSE, PBIAS and RSR values are used to assess the performance ratings. For example, the model performance is evaluated as satisfactory if NSE>0.5, RSR≤0.7 and PBIAS≤25%.

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(Q_i^{obs} - Q^{mean}\right)^2}} \quad (4.4)$$

$$PBIAS = \frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)*100}{\sum_{i=1}^{n} Q_i^{obs}} \quad (4.5)$$

where Q_i^{obs} and Q_i^{sim} are the *i*th observed and simulated discharge, Q^{mean} is the mean of observed discharge.

4.3.3 Flood frequency analysis

As mentioned earlier, in this study, POT method is used to assess the impact of climate change on flood frequency. The threshold estimation is the most difficult part in the POT method (Lang. et al., 1999, Scarrott and MacDonald, 2012). Threshold choice involves balancing between bias and variance. Too low a threshold may violate the asymptotic basis of the model, leading to bias while too high a threshold will reduce sample size, leading to high variance of the parameter estimates (Coles et al., 2001). Three different approaches, namely, the Mean Residual Life (MRL) plot, threshold stability plots and fitting distribution diagnostics (P-P, Q-Q, return level and density plots) are used in this study to decide the threshold value.

Since the POT series may have dependence, to deal with the temporal dependence within each peak flood series, declustering is used to filter the dependent peak flood series. The time interval between the peaks is chosen based on the catchment area. Lang. et al. (1999) suggested that five days plus the natural logarithm of square miles of basin area (r< 5 days+ log(A)) as a time interval. Svensson et al. (2005) used 5 days for catchments <45,000 km^2, 10 days for catchments between 45,000 km^2 and 100,000 km^2 20 days for catchments >100,000 km^2. Based on the actual basin area, the time spans for Trian, Talai and Tapao stations are 10, 10 and 5 respectively.

Once the independence condition of POT series is satisfied, the next steps involve the choice of appropriate distribution and method of parameter estimation. GPD, which is an

asymptotic model for POT series, is widely used in earlier studies (Coles et al., 2001, Davison and Smith, 1990, Scarrott and MacDonald, 2012). However, there is no guarantee that the GPD is the best distribution for all POT datasets, especially the POT datasets of future time (because threshold values are selected using the observed data). For example, from the P-P and Q-Q plots (Figs. 4.4a and b), it is observed that the P3 distribution captured the quantiles of flood magnitude better than GPD distribution for the POT series which is based on Had_85 of the Tapao station. Fig. 4.3c shows the return levels of flood magnitude, which are determined, based on GPD, P3 distributions and observed data in Tapao station. The flood magnitude appears to be overestimated by GPD while P3 is found as a good fit to the data. Therefore, the best fitting distribution to the sample needs to be tested with several distributions, especially for future POT samples.

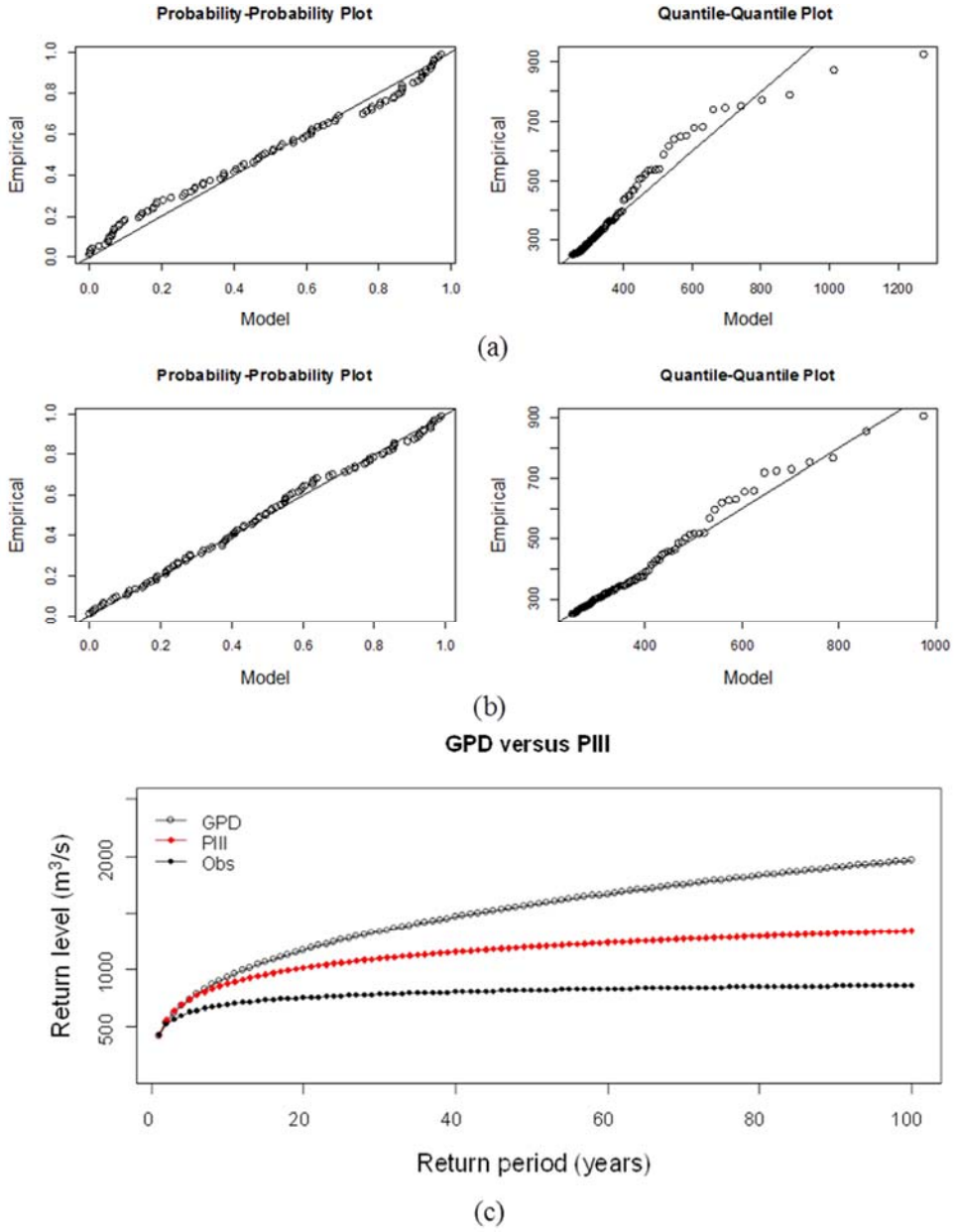


Fig. 4.4: P-P and Q-Q plots: (a) GPD, (b) P3 and (c) return level of Had_85 calculated using GPD, P3 and observed calculated using GPD in the Tapao station

The most important aspects for FFA is to find an appropriate distribution. Several two and three parameters probability distribution functions are commonly used in hydrological studies such as Gumbel, GEV, GPD, GL, P3, LN and LP3. These distributions are used to identify the appropriate distribution that best-fit to data sample. The Cumulative Distribution Functions (CDF) of these distributions and Log-Likelihood (LLH) used in this study are presented in Table 4.2.

Table 4.2: CDF and LLH of distributions

Distribution		Cumulative distribution function (CDF) and log-likelihood function (LLH)
LN	CDF	$\mathrm{F(x)} = \int_0^x \frac{1}{x\sigma_Y(2\pi)^{\frac{1}{2}}} e^{-\frac{(lnx-\mu_Y)^2}{2\sigma_Y^2}} dx$
	LLH	$L(\mu_Y, \sigma_Y) = -nln(2\pi) - nln\sigma_Y - \sum_{i=1}^{n} lnx_i - \frac{1}{2\sigma_Y^2}\sum_{i=1}^{n}(lnx_i - \mu_Y)^2$
P3	CDF	$\mathrm{F(x)} = \int_c^x \frac{1}{\beta\Gamma(\alpha)}\left(\frac{x-c}{\beta}\right)^{\alpha-1} e^{-\frac{(x-c)}{\beta}} dx$
	LLH	$L(\alpha,\beta,c) = -nln(\beta) - nln\big(\Gamma(\alpha)\big) - n(\alpha-1)\ln(\beta) + (\alpha-1)\sum_{i=1}^{n} ln(x_i - c) - \frac{1}{\beta}\sum_{i=1}^{n}(x_i - c)$
LP3	CDF	$\mathrm{F(y)} = \int_0^y \frac{1}{\Gamma(\alpha)}\left(\frac{y-c}{\beta}\right)^{\alpha-1} e^{-\frac{(y-c)}{\beta}} dy\,; y = logx$
	LLH	$L(\alpha,\beta,c) = -n\alpha ln(\beta) - nln\big(\Gamma(\alpha)\big) - \sum_{i=1}^{n} y_i + (\alpha-1)\sum_{i=1}^{n} ln(y_i - c) - \frac{1}{\beta}\sum_{i=1}^{n}(y_i - c)$
Gumbel	CDF	$\mathrm{F(x)} = e^{-e^{-\frac{(x-u)}{\alpha}}}$
	LLH	$L(u,\alpha) = -nln(\alpha) - \frac{1}{\alpha}\sum_{i=1}^{n}(x_i - u) - \sum_{i=1}^{n} e^{\frac{-(x_i-u)}{\alpha}}$
GL	CDF	$\mathrm{F(x)} = \left(1 + \left\{1 - \frac{k}{\alpha}[x-\xi]^{\frac{1}{k}}\right\}\right)$
	LLH	$L(k,\xi,\alpha) = -nln(\alpha) + \left(\frac{1}{k} - 1\right)\sum_{i=1}^{n} ln\left(1 - k\frac{x_i - \xi}{\alpha}\right) - 2\sum_{i=1}^{n} ln\left(1 + \left(1 - k\frac{x_i - \xi}{\alpha}\right)^{\frac{1}{k}}\right)$
GPD	CDF	$\mathrm{F(x)} = \begin{cases} 1 - [1 + \xi\left(\frac{\mathrm{x}-\mu}{\sigma}\right)^{-\frac{1}{\xi}}], \xi \neq 0 \\ 1 - \exp[-\frac{\mathrm{x}-\mu}{\sigma}], \xi = 0 \end{cases}$
	LLH	$L(\sigma,\xi) = -nln(\sigma) - \left(1 + \frac{1}{\xi}\right)\sum_{i=1}^{n} ln\left(1 + \frac{\xi x_i}{\sigma}\right), \xi \neq 0$
		$L(\sigma) = -nln(\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n} x_i\,, \xi = 0$

Different methods have been introduced to estimate the parameters of univariate distribution functions. The well-known methods are Maximum Likelihood (ML), the method of moment and method of L-moments. The method is one of the most widely used for fitting probability distributions to data (Strupczewski et al., 2001). ML method can be developed for a large of estimator situations. Another advantage of this method is that the estimates are consistent and asymptotically normally distributed. The probability of the observed data as a function of estimated parameters is called the likelihood function. Values of estimated parameters that have high likelihood correspond to models which give high probability to the observed data (Coles et al., 2001).

There are several approaches for helping in the selection of the appropriate PDF that are fitted to data sample. Graphical method, the goodness-of-fit (GoF) and model selection criteria have been widely used in hydrological studies. A graphical method is used to view the difference between theoretical and empirical of distributions. In this method, Probability-Probability (P-P) and Quantile-Quantile (Q-Q) plots are usually used. P-P plot compares the empirical cumulative distribution function and theoretical cumulative distribution of sample data. Q-Q plot where the quantiles obtained through the observed data and the distribution fitted are drawn together.

There are many GoF tests that used in hydrological studies to identify the appropriate distribution for fitting data, namely, Anderson-Darling (AD), Cramer-von Mises (CVM) and Kolmogorov-Smirnov (KS) tests. Besides, model selection criteria can be applied to find the appropriate distribution. Akaike Information Criteria (AIC) is the standard model selection technique commonly used in hydrology. Additionally, GoF tests can be used to select the distributions that fitted to the data sample while model selection criteria can be applied for obtaining the best distribution among those passing the GoF. Therefore, various approaches are used in this study in selecting the appropriate distributions. Table 4.3 summarises the description of all the tests that are used in this study.

Table 4.3: The list of test statistic used in this study

Test	Formulation
KS	$D = \max\left\|F_x(Q_{max,i}) - F_i(x)\right\|$
AD	$A^2 = -n - \sum_{i=1}^{n} \frac{(2i-1)\ln F_x(Q_{max,i}) + \ln\{1 - F_x(Q_{max,n-i+1})\}}{n}$
AIC	$AIC = -2\log L(\beta\|X) + 2k$
CVM	$W^2 = \sum_{i=1}^{n} \left[F_x(Q_{max,i}) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}$

Note: F_i and F_x are empirical and tested distribution functions; $Q_{max,i}$ is discharge; k is number of distribution parameters and n is sample size.

4.4 Results and Discussion

4.4.1 Hydrological model

The performance indicators calculated between HEC-HMS simulated and observed discharge data are presented in Table 4.4. For the Trian station, the NSE values are 0.72 and 0.63 for calibration and validation respectively. It indicates that the model performance is very good and satisfactory in calibration and validation respectively. Other statistical values (i.e, RSR and PBIAS) also show that the model performance is good (Moriasi et al., 2007). During the period 2002-2013, the upstream reservoir operation in Tapao station has changed the natural flow while this hydrological model only simulated flow in the natural condition. Therefore, the hydrological model performance in the validation step is unsatisfactory for Tapao station. However, the main purpose of this reservoir is to supply power and not for flood control (Babel et al., 2012). In the wet season, the peak inflow of this reservoir is released through the spillway. Hence, flood value in the downstream has not been affected by reservoir operation. Besides, there is a good agreement between observed flow and simulated flow by the model for Tapao station (Fig. 4.5). Furthermore, the model performance is very good and good in the calibration step. Therefore, the hydrological model can be used to simulate the future streamflow in this

watershed.

Table 4.4: Model statistics of HEC-HMS for three locations

Stations	NSE		RSR		PBIAS (%)	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
Talai	0.65 (Good)	0.6 (Satisfactory)	0.59 (Good)	0.63 (Satisfactory)	-2.5 (Very good)	-9.8 (Very good)
Tapao	0.67 (Good)	0.43 (Unsatisfactory)	0.57 (Good)	0.75 (Unsatisfactory)	-3.1 (Very good)	-13.3 (Good)
Trian	0.72 (Good)	0.63 (Satisfactory)	0.52 (Good)	0.6 (Good)	-8.9 (Very good)	-11.4 (Good)

Overall, the hydrologic model performance ranges from very good to satisfactory in both calibration and validation steps for the three locations in this basin, except for Tapao location in the validation step. It can be inferred from this that the continuous hydrological model is suitable to simulate and evaluate the changes of flood frequency due to climate change for this region as well as for the ungauged catchments belonging to the Saigon-Dongnai River basin.

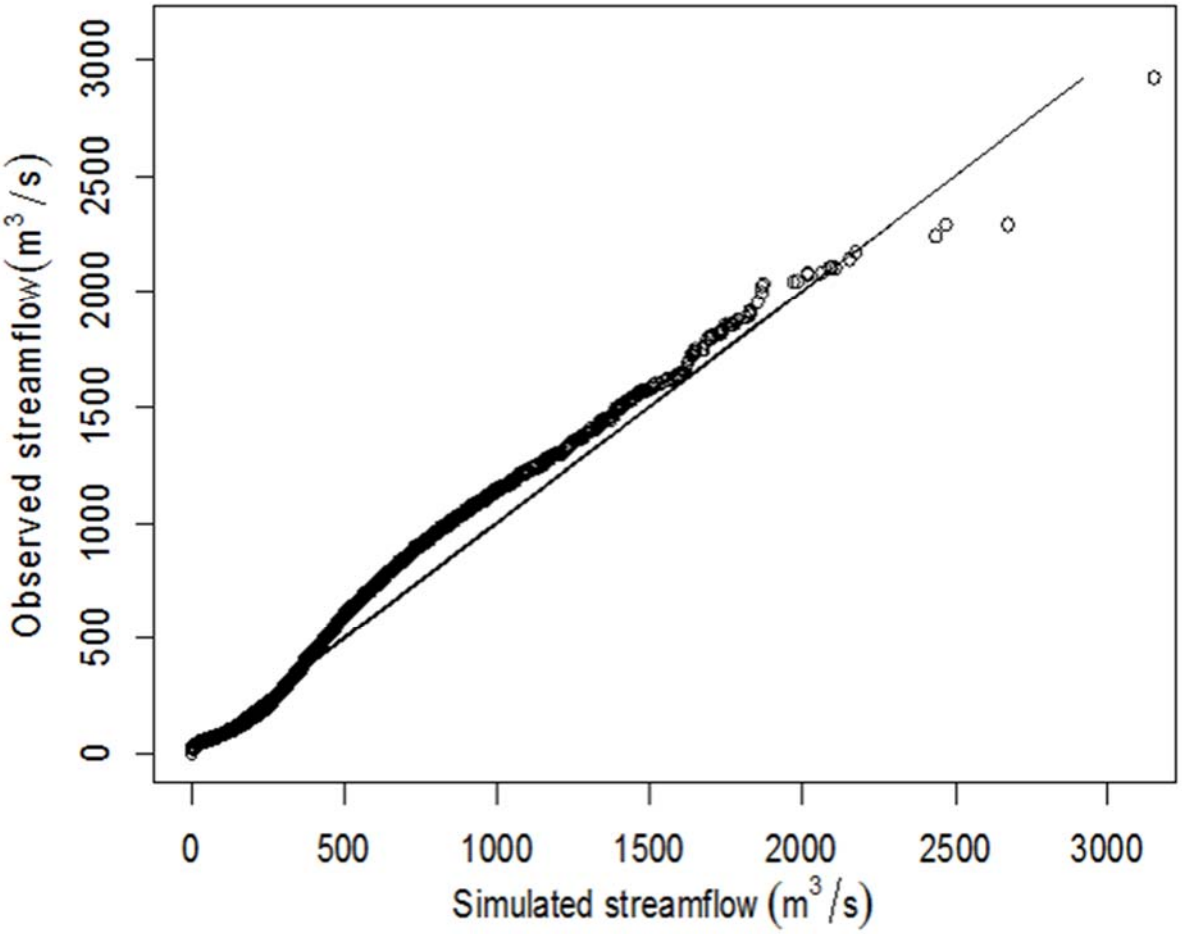


Fig. 4.5: Simulated and observed discharge for Tapao station during 1985-2013

4.4.2 Flood frequency analysis

Three criteria such as MRL plot, threshold stability plot and distribution fit diagnostics are used to search the appropriate threshold for each station. In order to reduce the variance of parameters estimation, the threshold is chosen in such a way that the length of each POT data sample after declustering is greater than 2 times the number of years. For example, Fig. A. 3 shows the MRL plot for Tapao location. The threshold value of u = 250 m^3/s gives over 574 exceedances before declustering. Fig. A. 4 shows the decluster run of Tapao station with a threshold u=250 m^3/s and r =5 (time interval for declustering) days. After declustering, the number of exceedance is 92 which meets the required condition. Besides, the diagnostic plots (i.e., P-P, Q-Q, return level and density) for the fitted GPD distribution with a threshold of 250 m^3/s are shown in Fig. A. 5 and they show a good agreement between model and empirical

values. Therefore, the threshold value of u = 250 m^3/s with GPD distribution is a suitable choice for Tapao location. Similarly, the selected threshold values for Trian and Talai are 950 m^3/s and 650 m^3/s respectively. After choosing the appropriate threshold, the set of distributions such as LN, GPD, P3, LP3, Gumbel, and GL distributions are tested to select the appropriate distribution.

The results of GoF tests for POT datasets of future (2020-2045) scenarios are given in Tables A.1, A.2, A.3 and A.4. Based on the AD test, most of the scenarios follow the GPD except for MM5_45 scenario which fits with P3 distribution in Talai location. 55 % of scenarios data follow the P3 distribution while 45 % of the remaining data are best fitted with GPD in Trian location. The AD test shows that, for Tapao station, only the Had_45 scenario follows the Gumbel distribution while 72.7 % and 18.2 % of samples fit the GPD and P3 distribution respectively. According to the KS, and CVM tests, the results are nearly similar.

Fig. 4.6 presents the boxplots of flood magnitude for the future time period (2020-2045) estimated from RCMs and GCM simulations. The results indicate an increasing trend in flood magnitude for Trian, Talai and Tapao locations under climate change context. For example, for 50-year return period, the percentage changes for the median value of flood magnitude are +29.68%, +48.76% and +24.22% for Talai, Tapao and Trian respectively. The changes in flood magnitude for all locations in this catchment are significant. Therefore, flood risk management strategies and hydraulic structure guidelines for this river basin under climate change context should be considered with importance.

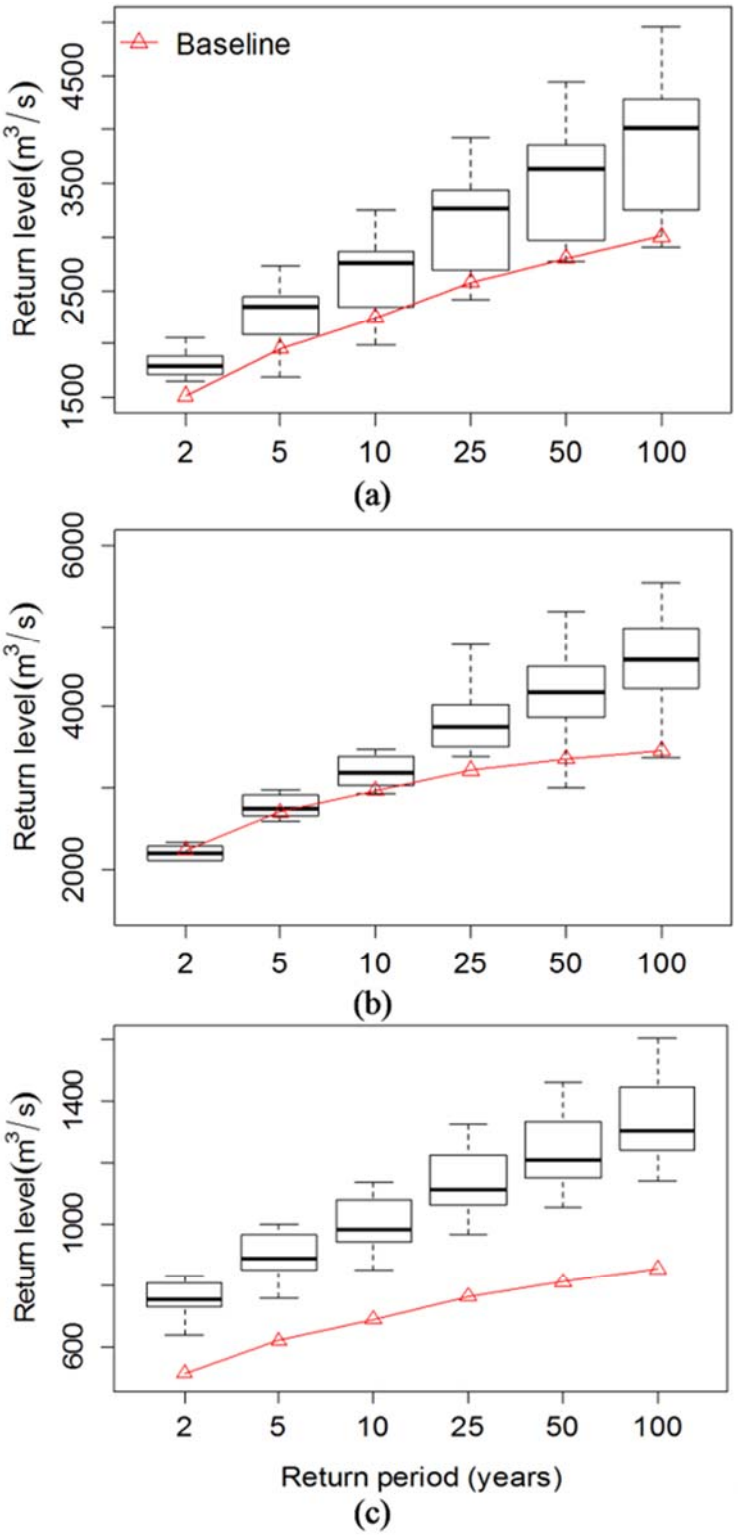


Fig. 4.6: Boxplots of flood magnitude with RCMs and GCM future simulations for (a) Talai, (b) Trian and (c) Tapao

4.5 Summary and Conclusions

This study used five RCMs and one GCM to assess the climate change impacts on the flood frequency of three sub-catchments in the Trian watershed for the future period (2020-2045). A combination of LARS-WG and HEC-HMS approach is used for studying the impact of climate change on flood frequency. LARS-WG and HEC-HMS are calibrated and validated based on the observed data. The performance of these models is found to be satisfactory and, therefore, used for generating and simulating daily future climate and streamflow data.

The POT approach is used to extract flood series and these extreme flood series are fitted with six different distributions namely, GPD, Gumbel, LN, LP3, GL and P3. In this study, the parameters of the distributions are estimated using the method of maximum likelihood and the best distribution for each POT dataset is selected using the different GoF tests such as AD, AIC, CVM and KS. Moreover, results of this study also reveal that directly using the asymptotic distribution to model the POT dataset sometimes provides wrong insights.

The results of five RCMs and one GCM suggest that flood magnitudes increase significantly in the future period (2020-2045) for three stations in Trian catchment. To be specific, the 100-year return level of Trian reservoir is increasing up to 32.34 % in one of future scenarios. Hence, planning or investment for flood management in Trian basin is highly necessary. Note that this study evaluates only the potential impact of climate change on flood frequency of the Trian reservoir belonging to the Saigon-Dongnai River basin, Vietnam. However, the changes in flood frequency owing to the river basin characteristic change (e.g., land use change, land cover change, etc.) can also be studied and it would be potential future work.

Chapter 5

Assessing the Selection of Copula for Bivariate Frequency Analysis Based on the Tail Dependence Test

5.1 Introduction

Single variable flood frequency analysis provides limited understanding and assessment of the true behaviour of flood phenomena which are often characterized by a set of correlated random variables like, peak, volume and duration (Favre et al., 2004, Yue et al., 2001). Univariate frequency analysis methods cannot describe the random correlated variables (Sarhadi et al., 2016). This approach can lead to high uncertainty or failure of guidelines in water resources planning, operation and design of hydraulic structures or creating the flood risk mapping (Chebana and Ouarda, 2011). Additionally, flood is a multivariate natural calamity characterized by peak, volume and duration. Hence, it is important to study the simultaneous, multivariate, probabilistic behaviour of flood characteristics. Multivariate parametric distributions (e.g., bivariate normal, bivariate gamma, bivariate extreme value distributions, etc.), which have been extended from univariate distribution, have been used to model

multivariate flood characteristics for different purposes (Adamson et al., 1999, Yue, 1999, Yue et al., 2001). However, this approach has some drawbacks such as the dependence structure between the variables, which depend on the marginal distributions or the flood variables have the same type of marginal distributions (Poulin et al., 2007, Zhang and Singh, 2007).

In order to overcome the limitation of multivariate distributions, a copula is a very versatile approach for simulating joint distribution in a more realistic way (Favre et al., 2004). The main advantage of this method is that the dependence structure is independently modelled with the marginal distribution that allows for multivariate distribution with different margins and dependence structures to be built (Dupuis, 2007, Zhang and Singh, 2007). Several researchers have used copulas to perform the bivariate frequency analysis (Dung et al., 2015, Reddy and Ganguli, 2012, Sraj et al., 2015). The most important step in the modelling process using copula is the selection of copula function which is the best fit to the data sample (Favre et al., 2004). The chosen copulas should include several classes of copulas and several degrees of tail dependence (Dupuis, 2007, Poulin et al., 2007).

Tail dependence characteristics constitute important features that differentiate extreme value copulas from other copula structures (Chowdhary et al., 2011). Therefore, the extreme value copulas with upper tail dependence are considered as suitable dependence structure models for the flood characteristics (Genest and Favre, 2007, Gudendorf and Segers, 2011, Poulin et al., 2007, Vittal et al., 2015). On the other hand, in the multivariate frequency analysis, the variables can be dependent or independent of each other. The relationship between flood characteristics (i.e., peak, volume and duration) are analyzed by several researchers. However, most of the results of the dependence between different pairs of flood variables were not consistent (Karmakar and Simonovic, 2009, Reddy and Ganguli, 2012, Sraj et al., 2015). Indeed, identification of the degree of dependence between flood variables is a difficult step, because the dependence of pairs of flood characteristics is controlled by different climate features and catchment properties (Gaál et al., 2015, Viglione and Blöschl, 2009).

Most studies used Pearson's linear correlation coefficient (r), Kendall's (τ) and Spearman's rank correlation (ρ) for measuring dependence among different flood variables. However, these measures are based on the association of the entire distributions but do not reveal the dependence in the specific part of the distribution (Aghakouchak et al., 2010). When

dealing with extreme events like floods, extreme values will appear in the tail of the distributions. Hence, the tail dependence, which describes the dependence in the tail of a multivariate distribution, can be a suitable measure (Aghakouchak et al., 2010, Coles et al., 1999, Hao and Singh, 2016, Serinaldi et al., 2015).

To describe dependence in multivariate extreme values, there are two possible situations, namely, asymptotic dependence or asymptotic independence (Coles et al., 1999). Diagnostic analysis to determine whether the variables have asymptotic dependence or asymptotic independence is very important in multivariate extreme analysis. In fact, in a situation where diagnostic checks suggest data to have asymptotically independence, modelling with the classical families of bivariate extreme value distribution is likely to lead to misleading results (Coles et al., 2001, Ledford and Tawn, 1996). Different measures of extremal dependence have been developed. Coles et al. (1999) proposed two measures of extreme dependence (χ and $\bar{\chi}$) for bivariate random variables. Nevertheless, recent studies show that there are still difficulties to detect between asymptotic dependence and independence in many cases (Bacro et al., 2010, Coles et al., 1999, Serinaldi et al., 2015, Weller et al., 2012).

Apart from these, several parametric and non-parametric approaches are suggested to determine the tail dependence. Non-parametric tail dependence estimator (λ_U), namely, λ_U^{LOG}(Coles et al., 1999, Frahm et al., 2005), λ_U^{SEC}(Joe et al., 1992), λ_U^{CFG}(Capéraà et al., 1997) and λ_U^{SS} (Schmidt and StadtmÜLler, 2006) have been preferred by most researchers in hydrological analysis (Li et al., 2009, Requena et al., 2016). However, Villarini et al. (2008) indicated that these tail dependence estimator have some drawbacks (bias, uncertainty). Furthermore, all tail dependence estimators exhibit very poor performance when the underlying upper tail dependence coefficient is null. It is, therefore, important to test for tail dependence before applying the estimator (Frahm et al., 2005, Poulin et al., 2007).

Consequently, upper tail (in)dependence testing is a useful alternative approach. Serinaldi et al. (2015) suggested that test for tail (in)dependence is mandatory because: (i) samples exist which seem to fail dependency but they are realizations of a tail dependent distribution; (ii) the use of misspecified parametric marginals instead of empirical marginals may lead to wrong interpretations of the dependence structure; and (iii) the tail dependence estimators can be insensitive to upper tail dependence, thus indicating upper tail dependence

even if none exist. Similarly, if data are to be independent in the upper tail, then modelling with dependence will lead to overestimation of probability of extreme joint events. Hence, Falk and Michel (2006) emphasized that testing for tail (in)dependence is essential in data analysis of extreme values.

Several recent studies indicated that Gumbel-Hougaard copula belonging to extreme value copulas work well when variables are asymptotically dependent (Dung et al., 2015, Karmakar and Simonovic, 2009, Poulin et al., 2007, Zhang and Singh, 2006). However, there are few studies which suggest that what is the best copula for modelling the dependence structure where variables have a strength of dependence but weaken at high levels or asymptotically independence. Therefore, it is important to find the appropriate copula to derive joint distribution of flood variables where the pair of flood characteristics have asymptotically independent or weak dependent at a high threshold.

The difference between extreme value copulas and Gaussian copula is that the Gaussian copula becomes independent at the high threshold. Furthermore, Gaussian copula, which is characterized by a correlation matrix, generates a wider range of dependence behaviour (Bortot et al., 2000). Studies by Renard and Lang (2007) also have proved the usefulness of Gaussian copula in hydrological extreme events analysis. In fact, they suggested that Gaussian copula can be reasonably well used for field significance determination, regional risk estimation, discharge-duration-frequency curves and regional frequency analysis. Frank and Clayton copulas, belonging to the Archimedean family, have been widely used in hydrology analysis because it can be modelled both negative and positive associated variables. Furthermore, Frank and Clayton copulas, which have zero dependency in both tails, are suitable in case of tail dependence does not exist (Dung et al., 2015, Poulin et al., 2007, Sraj et al., 2015). Therefore, Clayton, Frank and Gaussian copulas are used to assess the potential their application in case of variables are diagnosed as asymptotic independence for frequency analysis of flood in Trian watershed, Vietnam. The difference between this study and other studies on bivariate flood frequency is that this study proposed a copula selection procedure focus on tail dependence test. If there exists asymptotic dependence in the tail, extreme value copulas including Gumbel-Hougaard, Husler-Reiss and Galambos copulas should be chosen, Gaussian, Frank and Clayton copulas might be a good choice when variables are diagnosed as asymptotic independence.

This study aims to address the following issues: (i) investigating the potential of performing the tail dependence tests for the pairs of flood characteristics; (ii) evaluating the performance of extreme value copula for asymptotic dependence variables and Clayton, Frank and Gaussian copulas for asymptotic independent variables; and (iii) estimating joint return period of flood characteristics.

5.2 Study Area and Data

The detail of the study area is already presented in section 4.2. The description of the catchment development is presented in this section. There are two main tributaries of the Dongnai River (i.e., Dongnai and Langa). There are nine reservoirs, which are operating to supply water for drinking, irrigation, flood control and hydropower production, in the upstream of Trian gauge. Most of them began to operate in recent years except for Hamthuan-Dami and Daininh reservoirs which were operation from 2001 and 2008 respectively. In the Dongnai tributary, Daininh and Dakrtik reservoirs provide energy with capacity 300 MW and 144 MW respectively. Dongnai 2, Dongnai 3, Dongnai 4 and Dongnai 5 supply water to hydropower plants which have installed capacity of 70, 180, 340 and 150 MW respectively. Hamthuan and Dami reservoirs, located in the Langa tributary are a cascade of two hydropower plants with installed capacity of 300 MW and 175 MW. Tapao weir, located the downstream of Hamthuan and Dami reservoirs, is constructed to supply water for drinking and for irrigation of around 20,340 ha (Government, 2016). However, all reservoirs are located far away from the Trian gauge (Fig. 5.1). The flood from Trian station has significant impacts on the downstream areas (e.g., Bienhoa, Vungtau, Hochiminh cities, etc.). Therefore, this study mainly focused on the flood in the Trian gauge.

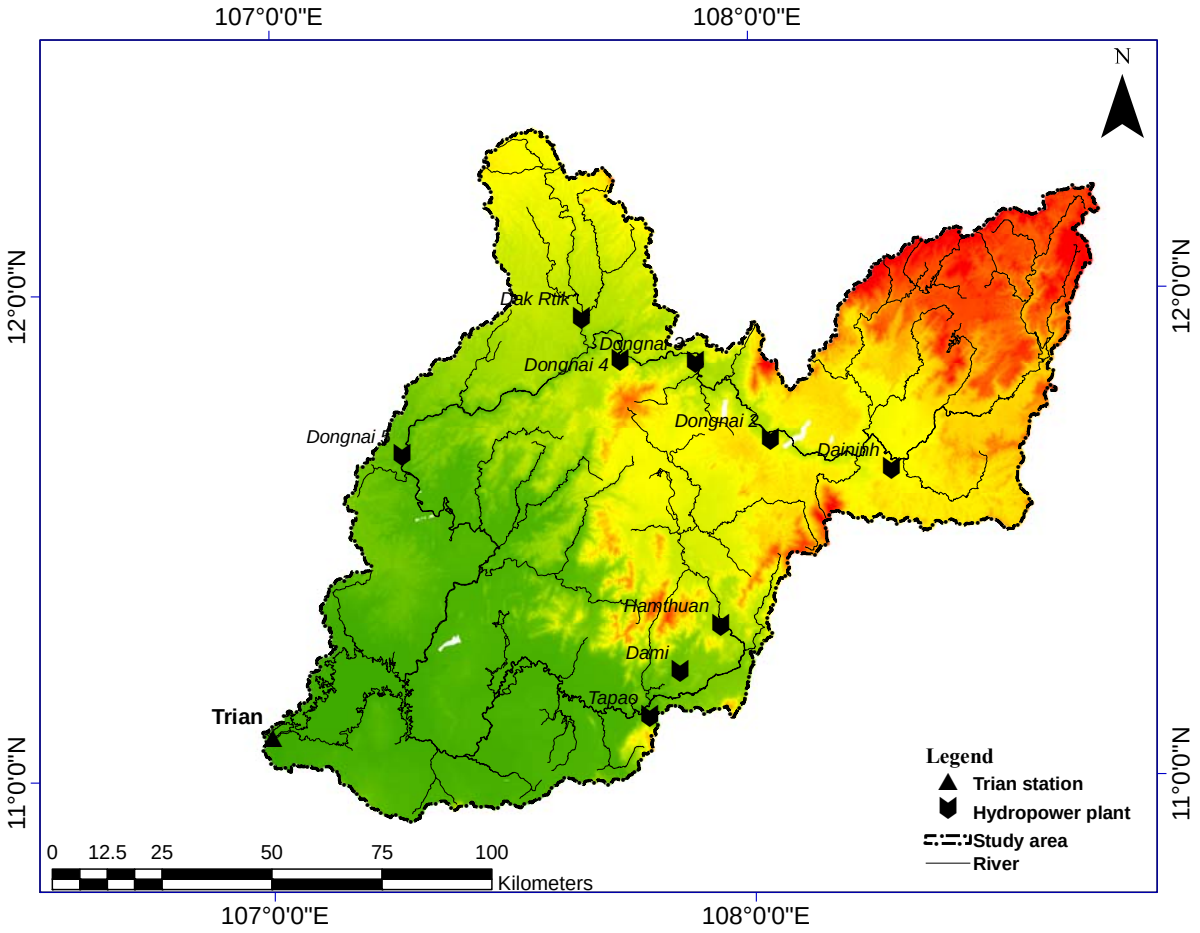


Fig. 5.1: Hydropower plants in the study area

Daily discharge data for the period 1978-2013 are available for the study from the Trian station on the Dongnai river, which is a part of the Saigon-Dongnai River basin and this data is used for flood frequency analysis. Trian station is located at 106º59'08" E and 11º06'16" N and it is at the confluence of two Dongnai and Langa rivers. Numerous researchers suggested that the length of data record should be at least 30 years for extreme value modelling (Bonnin et al., 2006, Jeong et al., 2014, Kioutsioukis et al., 2010, Yilmaz et al., 2017). Further, there are several multivariate frequency analysis studies using observed data of less than 35 years of data (Aissia et al., 2012, Zhang and Singh, 2006). Moreover, several studies suggested that the main advantage of the POT approach, which is for smaller sample sizes, is also used to increase the sample sizes (Beguería, 2005, Bezak et al., 2014, Lang. et al., 1999).

Based on the 35 years of observed data, the sample size of the flood variables in this study is 68. This meets the minimum requirement of the sample size (n=30) for the extreme

value modelling. Therefore, the length of the observed data is significant for the analysis of the tail dependence. The mean of daily discharge of Trian stream gauge from 1978 to 2013 is 527.4 m^3/s and the observed maximum daily discharge is 3,910 m^3/s. The daily time series of river discharge data are collected from the NHMS of Vietnam.

5.3 Methodology

The methodology used in this study is shown in the form of a flowchart (Fig. 5.2). Firstly, the identification of flood characteristics (i.e., peak, volume and duration) from the observed daily discharge time series is carried out. Next is the check whether flood variables time series are stationary or nonstationary. Tail dependence tests are then performed to diagnose whether the flood variables have asymptotic dependence or asymptotic independence. Finally, if the flood variables are having an asymptotic dependence, the extreme value copula is used for estimation of joint return periods. Otherwise, Gaussian, Frank and Clayton copulas are used.

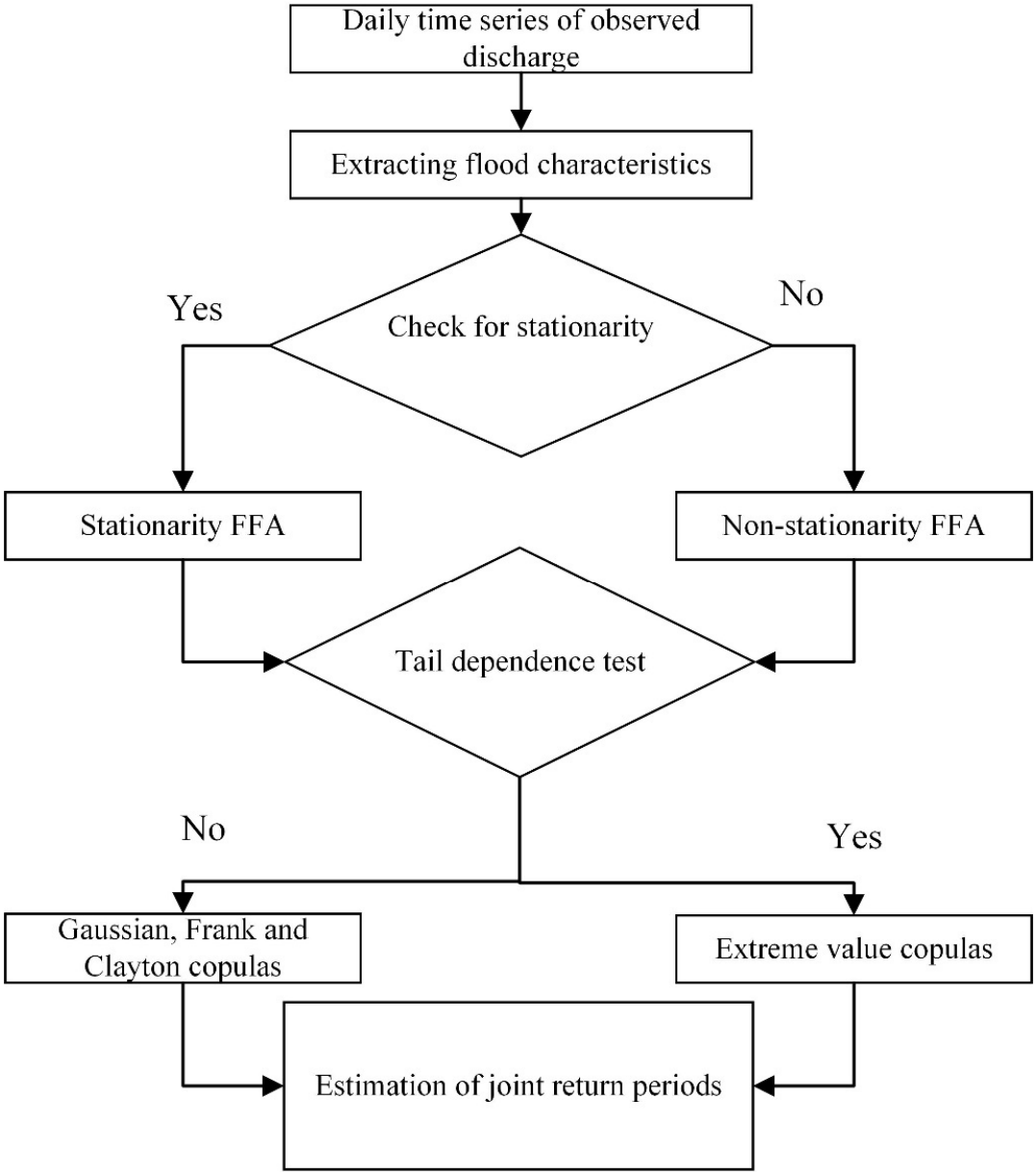


Fig. 5.2: Flowchart of methodology

5.3.1 Extracting flood characteristics

AM and POT approaches are widely used to extract flood characteristics. However, AM cannot consider multiple occurrences of flood events. (Bezak et al., 2014, Lang. et al., 1999). Unlike the AM, which only extracts one event per year, POT considers a wider range of events and provides more information than AM. The threshold estimation is the most difficult part in the POT approach (Lang. et al., 1999, Scarrott and MacDonald, 2012). Threshold choice involves balancing between bias and variance. Too low a threshold may violate the asymptotic

basis of the model, leading to bias, while too high a threshold will reduce sample size, leading to high variance of the parameter estimates (Coles et al., 2001).

There are two common approaches of choosing a threshold, namely, fixed quantile corresponding to a high non-exceedance probability (95%, 99% or 99.5%) and graphical method (Vittal et al., 2015). Three different techniques belonging to the graphical method, namely, the MRL, threshold stability plots and fitting distribution diagnostics (Solari and Losada, 2012, Thompson et al., 2009) are used in this study to decide the threshold value. In addition, the lag-autocorrelation plot is used to verify the assumption of the Independence and Identically Distributed (IID) flood variables (i.e., peak, volume and duration).

5.3.2 Diagnostic test to exanimate nonstationary

The extreme events, particularly for flood events, are intensifying due to global climate change, urbanization and anthropogenic activities. Therefore, the flood time series can have a nonstationary component. The flood frequency analysis which considers time series as stationary may lead to misleading results in estimation flood quantile. Checking nonstationarity of flood series in flood frequency analysis should be considered as an important initial step (Vittal et al., 2015). Trend analysis is normally used to detect nonstationarity in the flood variables. The Mann-Kendall (M-K) test is a non-parametric statistical test which is used to examine trends in time series and has been widely applied in the hydrological analysis (Lima et al., 2015, Sun et al., 2015, Villarini et al., 2009).

5.3.3 Tail dependent test

Coles et al. (1999) proposed two measures of extreme dependence (χ and $\bar{\chi}$) for bivariate random variables, as shown in Eq. (5.1) and Eq. (5.2).

$$\chi = 2 - \frac{logP(F_1(x)<u,F_2(y)<u)}{logu} \tag{5.1}$$

$$\bar{\chi} = \frac{2log(1-u)}{logP(F_1(x)>u,F_2(y)>u)} - \tag{5.2}$$

With a pair of complementary measure (χ,$\bar{\chi}$), a summary of multivariate extremal dependence can be determined.

- ✓ If $\bar{\chi}$=1 and 0<χ<1, the variables are asymptotically dependent and χ is a measure of strength of dependence within the class of asymptotic dependence distribution

- ✓ If -1<$\bar{\chi}$<1 and χ=0, the variables are asymptotically independent and $\bar{\chi}$ is a measure of strength of dependence within the class of asymptotically independence distribution.

There are still difficult to detect differentiate between asymptotic dependence and dependence in many cases using these extremal dependences. Furthermore, non-parametric tail dependence estimator mentioned earlier exhibit very poor performance when the underlying upper tail dependence coefficient is null. Therefore, the tail dependence test is used in this study. A description of tail dependence test is presented in the following sections.

The coefficient of tail dependence (η) introduced by Ledford and Tawn (1996) is used to detect asymptotically dependent and independent variables. They assumed that the joint survivor function of the pair (X, Y) with unit Frechet distribution is a regularly varying function, as shown in Eq. (5.3).

$$P(X>z, Y>z)=£(z)z^{-1/\eta} \quad (5.3)$$

where £(z) is a slowly varying function and η is the coefficient of tail dependence.

- ✓ If η=1 and $\lim_{z\to\infty} £(z) = c$ for some 0<c≤1, the variables are asymptotically dependent with a degree c

- ✓ If η<1, the variables are asymptotically independent.

The coefficient of tail dependence can be estimated by univariate theory because the joint survivor function can be reduced to univariate survivor function T=min (X, Y). The coefficient of tail dependence will be equal to shape parameter if T is fitted with GPD. The Log-likelihood Ratio (LLHR) test can be used to test for asymptotic dependence against the asymptotic independence. The null hypothesis of asymptotic dependence is tested comparing the log-likelihood of asymptotic dependence and asymptotic independence. Under the null hypothesis η=1 versus the alternative η<1, the LLHR test statistic, based on twice the difference between the log-likelihood of asymptotic dependence and asymptotic independence, has approximate Chi-square distribution with certain degrees of freedom. The significant of

asymptotic independence can be measured from the p-value of Chi-square distribution.

One of the most well-known approaches was proposed by Falk and Michel (2006) for testing tail dependence. Their test is based on the following theorem Eq. (5.4). With c→0, we have uniformly for t ϵ [0,1]:

$$P(X+Y>ct \mid X+Y>c)=\begin{cases}F(t)=t^2; \text{there is no } tail\ independence\\ F(t)=t; \qquad else\end{cases} \tag{5.4}$$

Using this theorem, Falk and Michel (2006) proposed four different tests for tail dependence namely Neymann-Pearson, Fisher's, Kolmogorov-Smirnov and Chi square tests. Frick et al. (2007) proposed a generalization of Falk and Michel's test, based on a second-order differential expansion of the spectral decomposition of the non-degenerate distribution function. This test is based on the following theorem Eq. (5.5).

$$P(X+Y>ct \mid X+Y>c)=\begin{cases}F(t)=t^{1+\rho}; tail\ independence\\ F(t)=t; \qquad tail\ dependence\end{cases} \tag{5.5}$$

where: c→0 is a threshold, ρ≥0 is independence measure and $F(t)$ is the standard uniform distribution with t ϵ [0,1]. According to the central limit theorem, the p-values of the optimal test is given by Eq. (5.6).

$$p=\Phi\left(\frac{\sum_{i=1}^{m} log\bar{C}_i+m}{m^{1/2}}\right) \tag{5.6}$$

where $\bar{C}_i$ =(X_i+Y_i)/c, i=1,…,m and Φ is the standard normal density distribution function.

In the LLHR test, threshold in GPD is selected based on the threshold stability plot. Frick et al. (2007) suggested the tail dependence test is quite sensitive to the threshold c. Hence, the threshold is chosen so that the number of exceedance is about 10% to 15% of the total number of observed data.

5.3.4 Selection of marginal distribution

The work by Vittal et al. (2015) suggested that it is important to use both nonparametric and parametric distributions for a selection of the appropriate marginal distributions for each flood characteristic. There are more than one parametric distributions that can be fitted to the

sample data. Hence, identifying the best fitting distribution to the sample need to be tested with several distributions rather than assuming that the particular distribution will be sufficient to provide the necessary insight for flood variables (Lang. et al., 1999, Nguyen et al., 2017, Vittal et al., 2015). The LN, P3, LP3, GPD, Gumbel, and GEV distributions which have been widely used for modelling extreme values (Bezak et al., 2014, Lang. et al., 1999, Saf, 2009a, Salas Jose et al., 2013) are used.

For nonparametric distribution, kernel density estimator with Epanechnikov, Gaussian, triangular and rectangular kernel functions are used in this study. Both parametric and nonparametric distributions are used to find the best marginal distribution for each flood variable in this study.

5.3.5 Extreme value copula and no tail dependence copula functions

A copula is defined as a joint distribution function of standard uniform random variables. If F(x,y) is any continuous bivariate distribution function with marginal distribution $F_1(x)$ and $F_2(y)$, the copula function can be express as Eq. (5.7)

$$F(x,y)=C[F_1(x), F_2(y)] \quad (5.7)$$

If the $F_1(x)$ and $F_2(y)$ are continuous, the copula function C is unique and can be written as Eq. (5.8)

$$C(u,\upsilon)=F[F_1^{-1}(u), F_2^{-2}(\upsilon)] \quad (5.8)$$

where the quantile function F_1^{-1} and $F_2^{-2}$are defined by $F_1^{-1}(u)$=inf[x: $F_1(x)\geq u$] and $F_2^{-1}(\upsilon)$=inf[x: $F_2(y)\geq \upsilon$] respectively.

There are several copula families and among them, the most well-known are the elliptical, Archimedean and extreme values copulas. Elliptical copulas come from elliptical distributions. The most popular copula belonging to this family is Gaussian copula. This copula represents the dependence structure of the data via a correlation matrix in which the elements describe the dependence between pairs of variables. Archimedean copula is widely used because of entailing a broad variety of dependence structure and being easily formulated via generator function. Clayton and Frank copulas are used extensions and recommended for

performing hydrological analysis.

Extreme value copulas are more popular for hydrological application, particularly for extreme events. Indeed, the extreme value copulas with upper tail dependence are considered as appropriate models for the dependence structure for extreme events. Extreme value copulas can be used as a convenient choice in modelling data with positive correlation and arise naturally in the domain of extreme events (Gudendorf and Segers, 2011, Mirabbasi et al., 2012). The families of extreme value copulas considered in this study include: Gumbel-Hougaard, Husler-Reiss and Galambos. Besides, Gaussian, Frank and Clayton copulas are also used in circumstances where diagnostic checks suggest data to be asymptotically independent. More details and descriptions can be found in Gudendorf and Segers (2011), Salvadori et al. (2013) and Poulin et al. (2007). The relevant expression for their dependence function and tail dependence coefficient are presented in Table. 5.1.

Table 5.1: Definition and upper tail dependence coefficient of the copula used in this study

Copula	$C_\Theta(u,\upsilon)$	λ_u
Gumbel	$C(u,\upsilon)=\exp[-(\ln(u))^{\Theta}+-(\ln(\upsilon))^{\Theta}]^{1/\Theta}$	$2-2^{1/\Theta}$
Galambos	$C(u,\upsilon)= u\upsilon\exp[-(\ln(u))^{-\Theta}+-(\ln(\upsilon))^{-\Theta}]^{-1/\Theta}$	$2-2\phi(1/\Theta)$
Husler-Reiss	$C(u,\upsilon)=\exp(\hat{u}\Phi[\frac{1}{\theta}+\frac{1}{2}\Theta\ln(\frac{\hat{u}}{\hat{\upsilon}})]-\hat{\upsilon}\Phi[\frac{1}{\theta}+\frac{1}{2}\Theta\ln(\frac{\hat{\upsilon}}{\hat{u}})])$	$2-2^{1/\Theta}$
Gaussian	$C(u,\upsilon)=\Phi[\phi^{-1}(u),\phi^{-1}(\upsilon)]$	0
Clayton	$C(u,\upsilon)=(u^{-\Theta}+\upsilon^{-\Theta}-1)^{-1/\Theta}$	0
Frank	$C(u,\upsilon)=-\frac{1}{\theta}ln[1+\frac{(e^{-\theta u}-1)((e^{-\theta\upsilon}-1)}{e^{-\theta}}]$	0

where ϕ, Φ are the cumulative density function of a standard normal and multivariate normal distribution with mean 0 and covariance $\sum$ respectively, $u=\ln(\hat{u})$ and $\upsilon=\ln(\hat{\upsilon})$.

Several methods have been proposed to estimate the parameter of the copulas. The Maximum Pseudo-likelihood (MPL) method is a modification of the classical maximum likelihood method where the empirical marginal distribution are used. This method can be applied to both one and multi parameters copulas. The MPL method consists of transforming the marginal variables into uniformly distributed vectors using its empirical distribution

function. Then the copula parameters are estimated using maximization of pseudo log-likelihood function. The form of the log-likelihood function is presented in Eq. (5.9).

$$l_\theta = \sum_{i=1}^{n} log\left[c_\theta\{F_1(X_{i,1}), F_2(X_{i,2})\}\right] \quad (5.9)$$

where $F_1(x) = R_i/(n+1)$ and $F_2(x) = S_i/(n+1)$ are non-parametric marginal probability solely based on rank.

Genest et al. (1995) and Cherubini et al. (2004) suggested the MPL and canonical maximum likelihood approaches in case of an unknown marginal distribution to estimate copula parameters. In order to allow marginal distribution to be free and not restricted by parametric families, the MPL method is suggested because the marginal distribution is considered as the empirical distribution function. Furthermore, Kim et al. (2007), Genest and Favre (2007) and Kojadinovic and Yan (2010a) showed that MPL is the best choice of estimating copula parameters. Therefore, MPL is used in this study.

Selection of appropriate copula is a complex process and need to be considered through several different measures. Only one measure can fail to identify the suitable copulas that can lead to an inappropriate the joint probability of flood characteristics (Fu and Butler, 2014). There are several different methods to select the best copula, including graphical method, GoF tests and model selection criteria. The first two methods are used to measure the discrepancy between the theoretical distribution and empirical distribution, while the model selection criteria such as AIC, which penalize the minimized negative log-likelihood function for the number of parameters estimated, would be more appropriate than repeated tests of significance whose outcomes lose their interpretability (Katz, 2013).

In the graphical method, the theoretical non-exceedance joint probabilities obtained using copula functions are compared to the empirical non-exceedance joint probabilities which can be estimated by Gringorten plotting position formula Eq. (5.10).

$$F_{XY}(x_i, y_i) = \mathrm{P}(\mathrm{X} \le x_i, \mathrm{Y} \le y_i) = \frac{\sum_{m=1}^{i}\sum_{l=1}^{i} n_{ml} - 0.44}{N + 0.12} \quad (5.10)$$

where n_{ml} is the number of pairs (x_j, y_j) counted as $x_j \le x_i$ and $y_j \le y_i$; $i,j = 1,\ldots,N$; $1 \le j \le i$ and N is the sample size. Besides the graphical method, GoF test is also used to test the adequacy of

the hypothesized copulas. Genest et al. (2009) reviewed and compared several GoF tests for copula. They proved that Cramer-von Mises (S_n^I) test comparing the empirical and theoretical copulas is the best GoF test. However, there is no difference between extreme value copulas in this test. In order to overcome this shortcoming, the test based on a Cramer-von Mises (S_n^{II}) statistic, measuring the distance between parametric and non-parametric estimator of the Pickands dependence function, was introduced by Genest et al. (2011). This test is defined in Eq. (5.11).

$$S_n^{II} = \int_0^1 n|A_n(t) - A_{\theta n}(t)|^2 dt \tag{5.11}$$

where $A_n(t)$ and $A_{\theta n}(t)$ are the non-parametric and parametric estimator of Pickands dependence function A. Based on the objective and availability data in this study, S_n^{II} is used to find out the appropriate copula functions.

5.3.6 Joint return period estimation

The return periods of hydrological extreme events are normally associated with a certain exceedance probability. Several theoretical bivariate return periods have been defined in the literature. The joint return period (OR) in which either x or y have exceedance (i.e., X>x or Y>y), which is denoted by $T_{X,Y}^{OR}$. The joint return period (AND) in which both x and y are exceeded (i.e., X>x and Y>y), and is denoted by $T_{X,Y}^{AND}$. These two types of joint return period are given by Eqs. (5.12) and (5.13).

$$T_{X,Y}^{AND} = \frac{\mu_T}{P(X \geq x \; and Y \geq y)} = \frac{\mu_T}{1 - F_X(x) - F_Y(y) + F_{XY}(x,y)} \tag{5.12}$$

$$T_{X,Y}^{OR} = \frac{\mu_T}{P(X \geq x \; or Y \geq y)} = \frac{\mu_T}{1 - F_{XY}(x,y)} \tag{5.13}$$

where, μ_T is the mean inter-arrival time (years).

The above equations are used for both AM and POT approaches. In the case of block maxima, μ_T is equal to 1.0 (Shiau, 2003, Vittal et al., 2015). Since POT is also applied in this study, the mean inter-arrival time is determined based on the observed flood events.

5.4 Results and Discussion

5.4.1 Identification of flood characteristics

The POT approach is used to extract flood characteristics in this study. The threshold is selected based on the three different approaches, namely, the MRL, threshold stability plots and fitting distribution diagnostics. Fig. 5.3a shows the MRL plot for observed daily discharge for Trian. It is clear that after the threshold value of u=950 m^3/s, the MRL is consistent with a straight line. Furthermore, with the threshold of u=950 m^3/s, the shape and modified scale parameters begin to reach a plateau (Fig. 5.3b). Besides, the diagnostic plots (i.e., P-P and Q-Q) for the fitted P3 distribution with the threshold (950 m^3/s) after declustering (r=10 days) are shown in Fig. 5.3c and they show a good agreement between the model and empirical values.

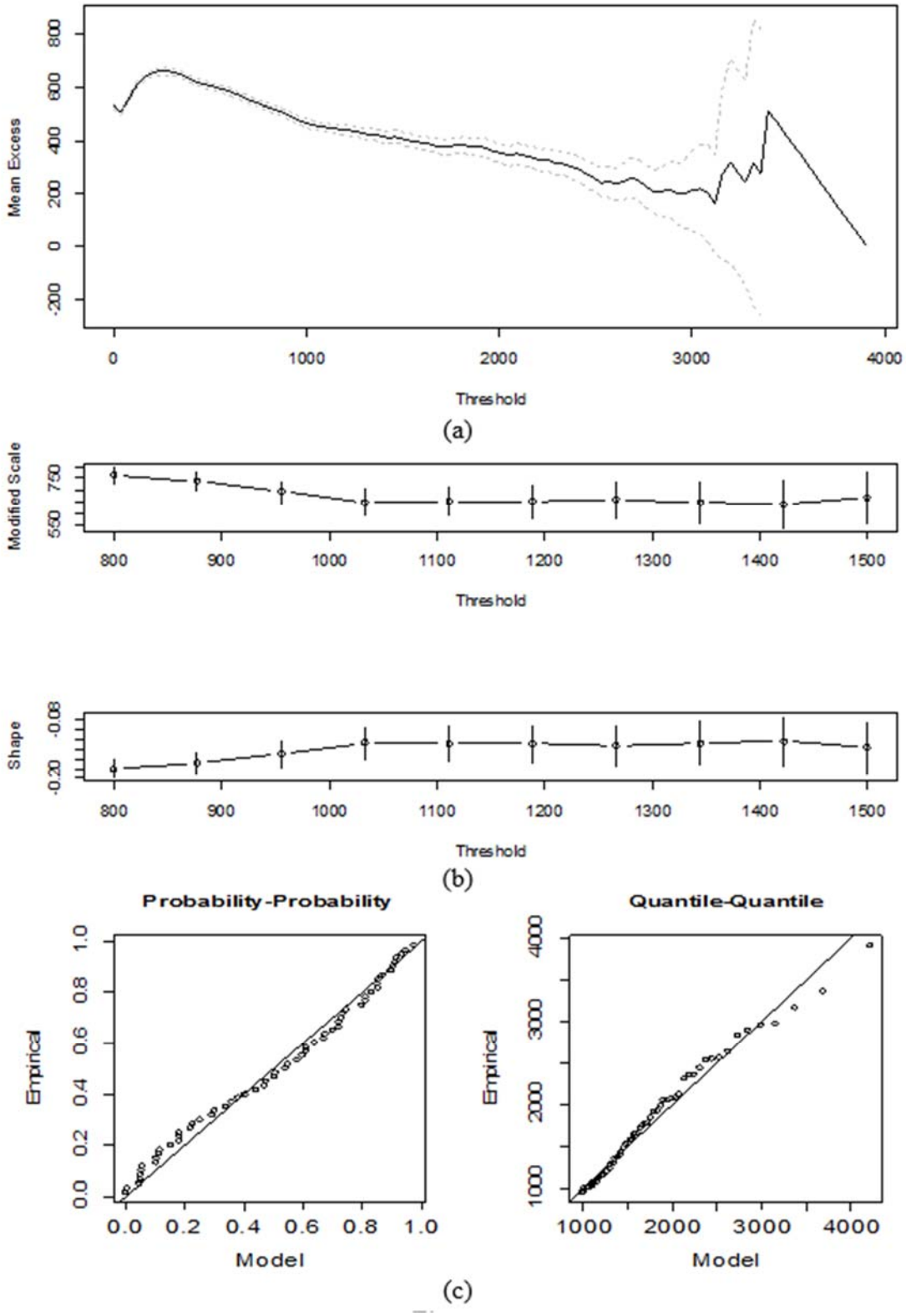


Fig. 5.3: (a) Mean residual life plot; (b) threshold stability plots; (c) diagnostic plots for observed daily flood data

Fig. 5.4 shows that there is insignificant autocorrelation for all flood characteristics. The IID flood variables assumption is still maintained based on this threshold. Therefore, threshold value of u=950 m^3/s is a suitable threshold for Trian. This threshold is used for all future flood characteristics. Flood duration and volume are also determined based on this threshold. The M-K test for peak, volume and duration of observed data showed that there is no significant trend for any of the flood variables observed at the Trian gauge. It indicates that the flood events in the present data are still stationary. Therefore, stationary flood frequency analysis is used to estimate the joint return periods.

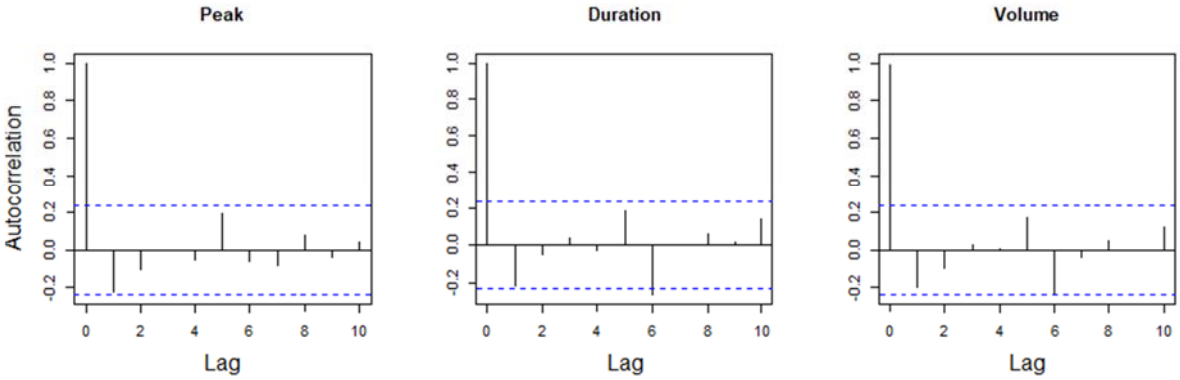


Fig. 5.4: The autocorrelation plot up to lag ten for all the flood variables

5.4.2 Tail independence test

The pair of extremal measures $(\chi, \bar{\chi})$ is used to detect whether the flood variables have asymptotically dependent or not. Nevertheless, in this study, the value of χ (u) is nearly equal to 0.5. It means that the pair of flood characteristics has asymptotic dependence for all u. However, the value of $\bar{\chi}$ shows that the pair of flood characteristics is independent for many cases. For example, Fig. 5.5 shows the Chi and Chi bar plot for the pair of observed flood peak and volume. Therefore, it is difficult to identify between the asymptotically dependence and independence based on these plots.

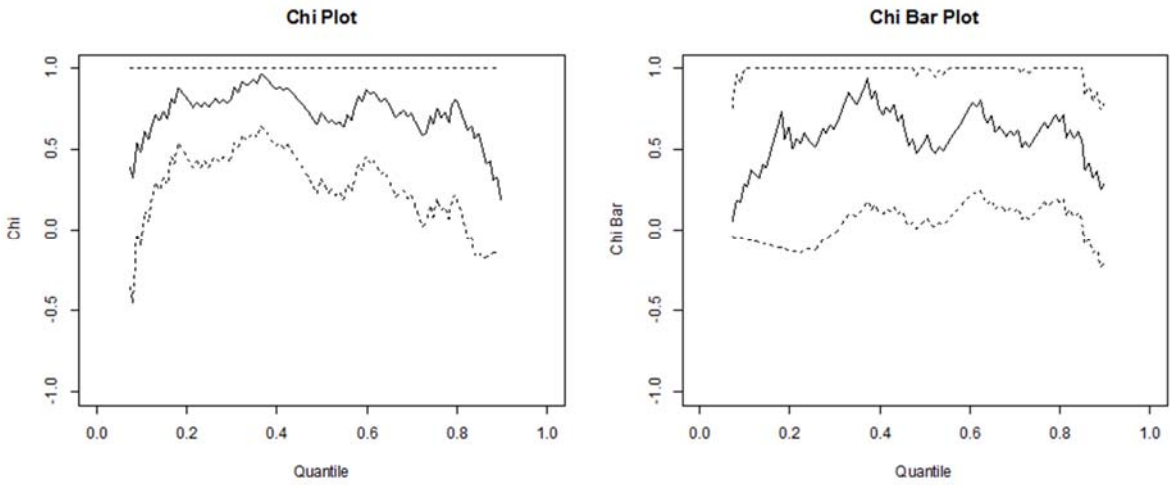


Fig. 5.5: Extremal measures for dependence of observed flood peak and volume

LLHR and TailDep tests are used to decide asymptotically (in)dependent variables in case of the extremal measures do not work. The results from two tests are nearly similar. Table 5.2 shows the p-value of LLHR and tail dependence tests for all pair of observed flood variables. Based on extremal measures and these tests, asymptotically dependence and independence are identified.

Table 5.2: Likelihood ratio and tail dependence test p-value

Tests	p-value		
	DV	DP	PV
LLH	0.04	0.01	0.01
TailDep	0.3	0.02	0.04
Diagnostic	Dep	InDep	InDep

5.4.3 Marginal distribution of flood variables

To determine the most appropriate marginal distribution for all flood characteristics, GEV, Gumbel, LN, P3, GPD and LP3 distributions belonging to the parametric distribution and Epanechnikov, Gaussian, triangular and rectangular kernel functions belonging to nonparametric distribution are used in this study. The maximum likelihood estimation is used to estimate the parameters of the distributions. The selection of the appropriate distribution is based on the AIC value. The selected marginal distributions are presented in Table 5.3, which provides a comparison of performances for all several marginal distributions. The results

indicate that LP3 distribution is most appropriate for modelling flood volume and duration while P3 is found to be the best for flood peak.

Table 5.3: AIC values for all marginal distributions

Flood variable	Parametric					
	LN	Gumbel	GEV	P3	LP3	GPD
V	1,309	1,318	1,317	1,297	1,285	
P	1,066	1,066	1,068	1,061	1,067	1,068
D	654.9	656.0	658.0	641.7	628.9	
Flood variable	Nonparametric					
	Gaussian kernel	Triangular kernel	Rectangular kernel		Epanechnikov kernel	
V	1,310	1,311	1,317		1,313	
P	1,074	1,074	1,092		1,074	
D	634.8	644.2	644.5		647.0	

Note: V is volume (10^6 m^3/s), P is peak (m^3/s) and D is duration (days).

5.4.4 Copula selection

Fig. 5.6 shows the theoretical and empirical joint non-exceedance probabilities of asymptotic tail independence data. It is observed that Frank and Gaussian copulas fit the dataset which is diagnosed as asymptotic independence better than extreme value copulas. Additionally, AIC value and GoF test also indicated that the copula function that has no tail dependence may work well when variables are asymptotically independent.

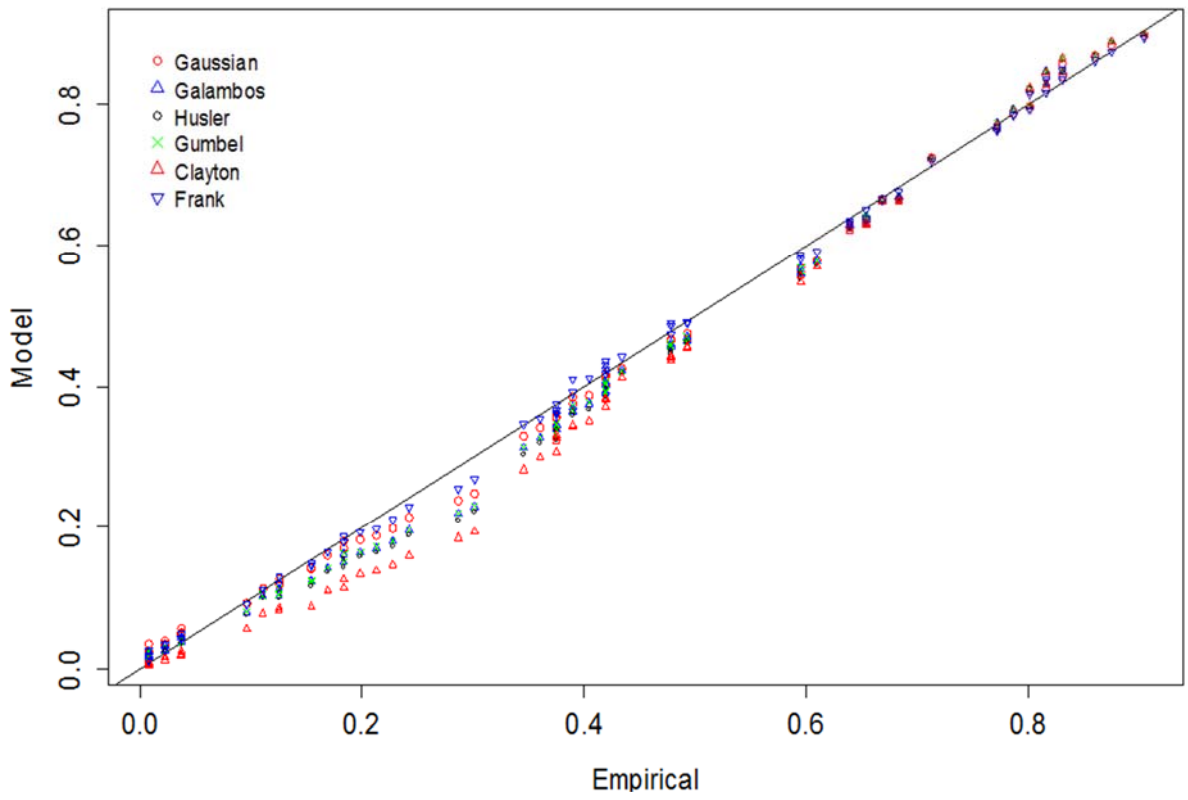


Fig. 5.6: Theoretical and empirical joint non-exceedance probabilities of observed flood duration and volume (asymptotic independence)

The joint return period (AND) of observed flood duration and peak pair are estimated by using the best fitted models of each group copulas. Gumbel-Hougaard copula (extreme value copulas) and Frank copula (the no tail dependence copulas) are selected to estimate the joint return period of the observed flood duration and peak pair. Fig. 5.7 shows the comparison of joint return period curves of the pairs of observed duration and peak which are estimated by Frank copula (black) and Gumbel copula (blue). This plot indicates that there are huge differences between two copulas. For lower return period, the two corresponding curves are very close to each other. However, there are large differences in the central part in the 50-year and 100-year return periods.

Besides, the shape of the joint return period of each copula has significant differences. The bound limits shrink significantly for the Gumbel-Hougaard copula while this situation is not shown by the Frank copula. For example, at 5-year return period, the corresponding bound for Gumbel-Hougaard copula is wider than that of Frank copula. At return period of 10-year,

50-year and 100-year, the phenomenon is the opposite and the curve from Gumbel-Hougaard becomes sharper. This result indicates that choosing inappropriate copula function will lead to be serious difference the joint return period results. In this study suggest that the copula function is selected based on the dependence structure of the variables. The result from tail dependence test may provide useful additional information about the adequacy of the chosen copula functions.

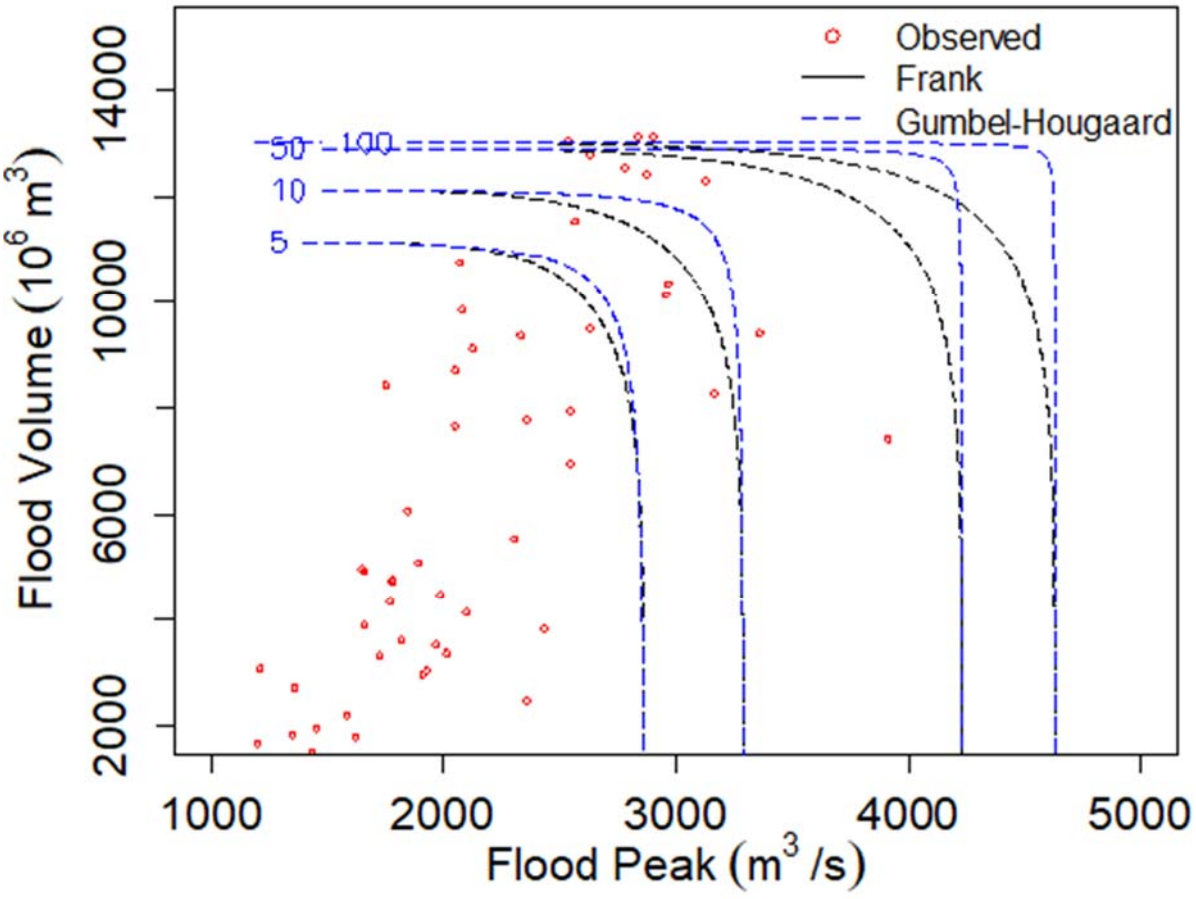


Fig. 5.7: Joint return period of the pair of flood peak and volume modelling by Frank and Gumbel copulas

Based on the above analysis, in this study, three extreme value families of copulas (Gumbel-Hougaard, Galambos and Husler-Reiss) are chosen to model asymptotically dependence pair of flood characteristics. Gaussian, Frank and Clayton copulas are used in modelling asymptotically independence pair of flood characteristics. The dependence parameters of copulas are estimated using MPL method. The copula dependence parameter, AIC and GoF statistics are given in Table 5.4.

Table 5.4: Copula dependence parameters, AIC and corresponding GoF statistics

Copulas	DV			
	Parameter	AIC	S	p-value
Gumbel-Hougaard	6.007	-165.01	0.00579	0.003
Galambos	5.268	-162.4	0.00583	0.002
Husler-Reiss	4.377	-137.98	0.00784	0.007
Copulas	DP			
	Parameter	AIC	S	p-value
Gaussian	0.785	-57.509	0.114	0.065
Clayton	1.774	-47.817	0.504	0.001
Frank	8.455	-67.695	0.063	0.285
Copulas	PV			
	Parameter	AIC	S	p-value
Gaussian	0.835	-73.575	0.119	0.05
Clayton	2.066	-55.25	0.477	0.0002
Frank	10.396	-86.929	0.058	0.335

Fig. 5.8a shows the P-P plot of model and empirical joint non-exceedance probabilities for observed flood duration and volume. This plot indicates that extreme value copulas (Gumbel-Hougaard, Galambos, Husler-Reiss) give the best fit to the dataset. However, identifying the differences among three copula functions is difficult. Therefore, AIC and GoF test are used to choose the best copula function. For example, the AIC value (-165.013) and statistical test value (0.00579) are shown in Table 5.3, which indicate that Gumbel-Hougaard copula provides the best performance for the pair of observed flood duration and volume.

For asymptotically independence case, Fig. 5.8b shows the P-P plot of the model and empirical joint non-exceedance probabilities for the pair of observed flood duration and peak. It is clear that all copulas (Gaussian, Clayton and Frank) give a good fit to the data. However, Frank copula fits better than other copulas. Similarly, the best fit copula using AIC (-67.695) and statistical test values (0.285) is Frank copula (Table 5.3). All measures indicate that Frank copula is the best fit to the data sample (observed flood duration and peak). The best copula based on AIC value and GoF test is used to estimate the joint return period for modelling the pair of flood characteristics.

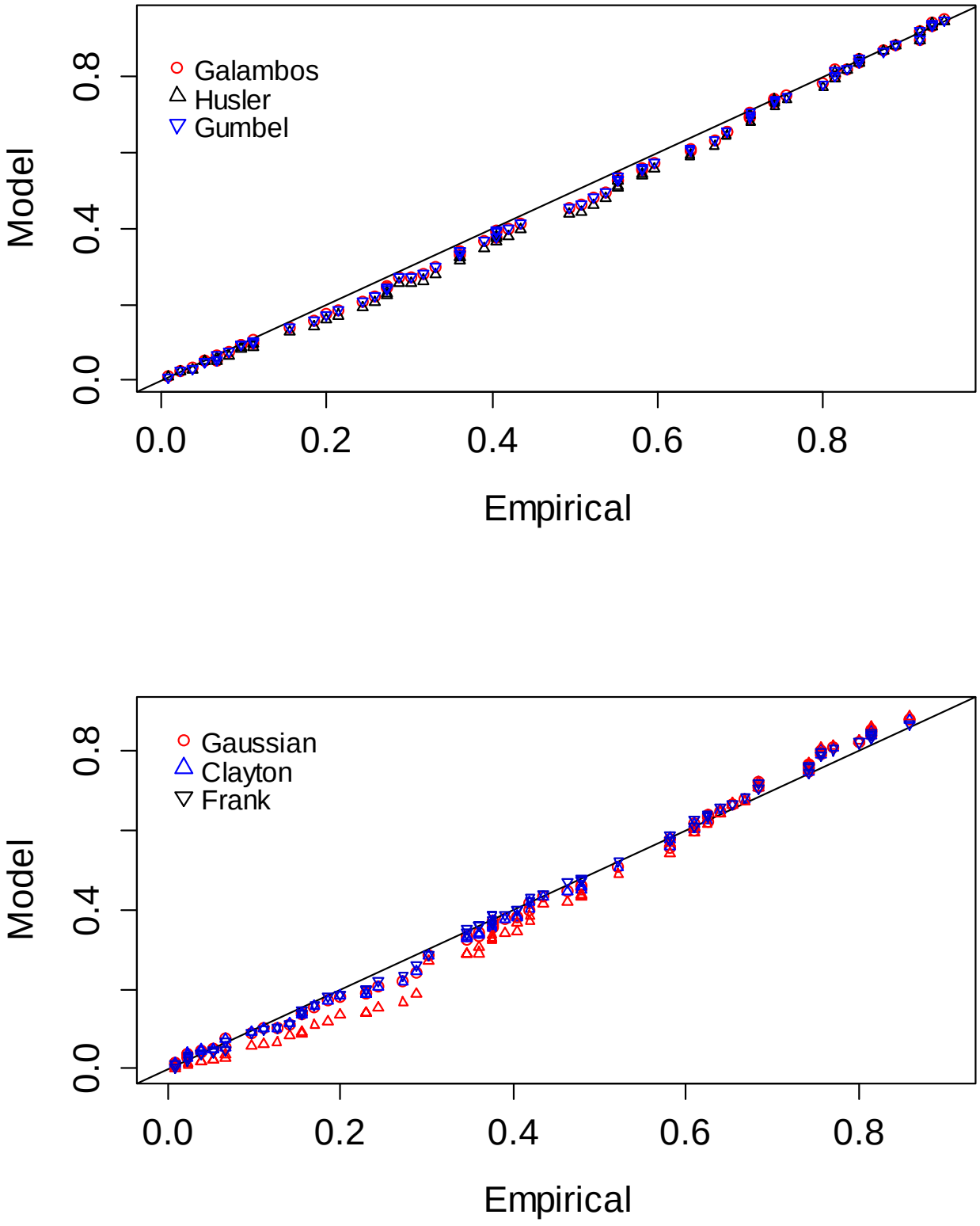


Fig. 5.8: Theoretical and empirical joint non-exceedance probabilities of: (top) Observed duration and volume; (down) Observed duration and peak

5.4.5 Joint return period estimation

The joint return periods (AND and OR) of flood peak and volume for 5, 10, 50, 75 and 100-year return periods are shown in the Fig. 5.9. For example, the flood peak (m^3/s)-volume (10^6 m^3) pairs, (4,011-11,020), (4,119-11,432) and (4,297-11,674) are the joint return period (OR) of 50, 75 and 100-year respectively. The results from this figure also indicate that the joint return periods (AND) provide lower flood variable quantile than joint return periods (OR) for all return periods. Several combinations of flood peak and volume as well as other flood characteristics in the same return period also are obtained through bivariate frequency analysis. These results provide more possible choices for a decision maker to select the flood event for structure designing and water resources planning as well as assessing the variability of the obtained flood map inundation that cannot be achieved through univariate frequency analysis.

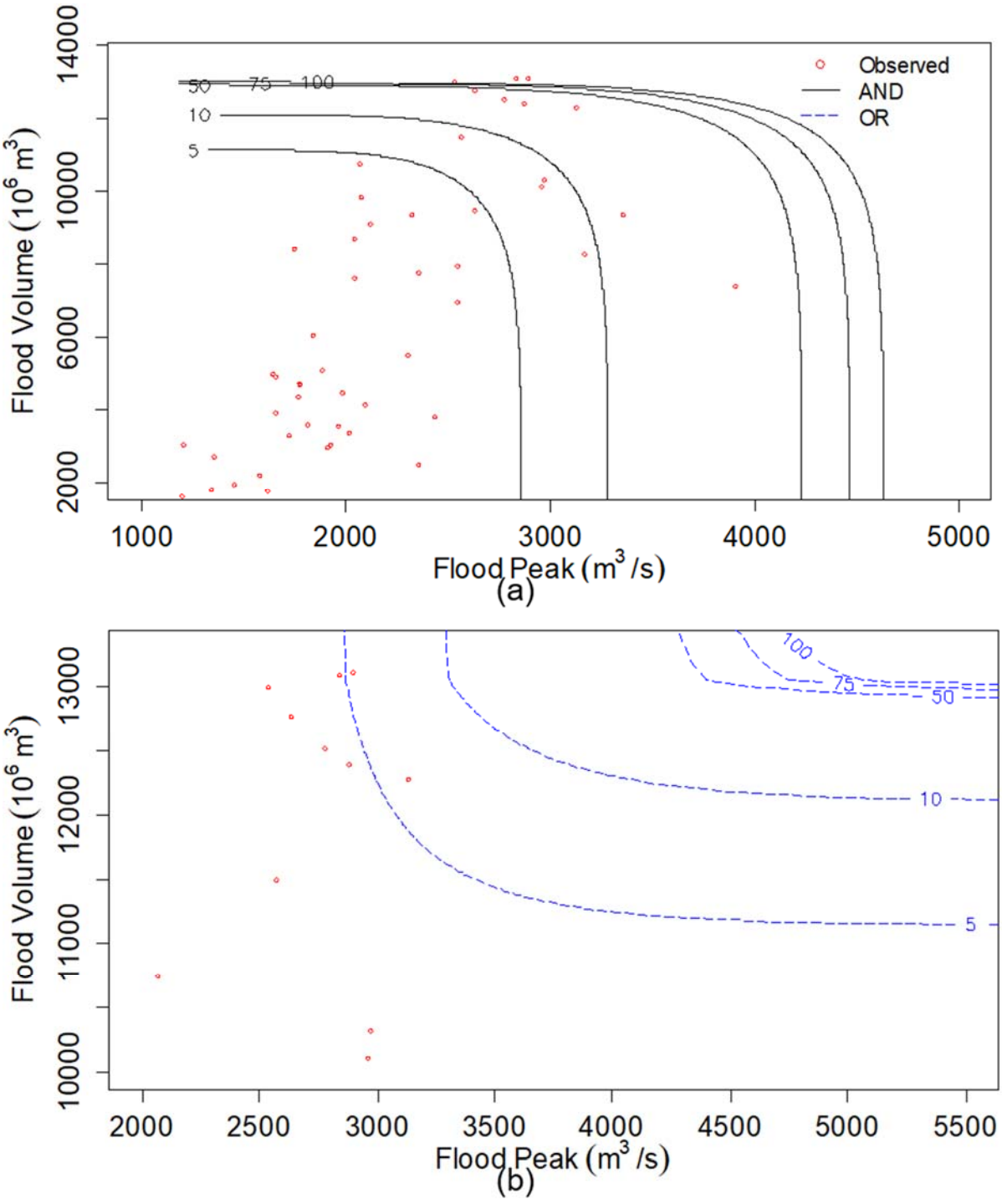


Fig. 5.9: The joint return periods of peak and volume: (a) AND both peak and volume are exceeded; (b) OR either peak or volume are exceeded

5.5 Summary and Conclusions

The main emphasis of this study is on the tail dependence test before the selection of copula function which best fits the data sample. Indeed, extremal measurement is useful approach, but in many cases, it cannot detect whether data is asymptotically dependence or not. LLHR and tail dependence tests are used to identify the asymptotically (in)dependence of observed flood variables. Two pairs of flood characteristics (i.e., peak-volume and duration-peak) have asymptotically independence while flood duration and volume pair has asymptotically dependence in this study. Three extreme value families of copula, namely, Gumbel-Hougaard, Galambos and Husler-Reiss are evaluated to model asymptotically dependence pair of flood characteristics.

The extreme value copulas with upper tail dependence have proved that they are appropriate models for the dependence structure of the flood characteristics. However, identifying the differences among three copula functions is difficult. Therefore, the test based on a Cramer-von Mises (S_n^{II}) statistic measuring the distance between parametric and non-parametric estimator of the Pickands dependence function is used and it is proved that it is highly efficiency for extreme value copula.

Similarly, Gaussian, Frank and Clayton copulas are the appropriate copula models in case of variables which are diagnosed as asymptotic independence. Then, the best fit copula models are used to calculate the joint return periods of flood characteristics. These results provide more possible choices for decision maker to select the flood event for design of hydraulic structures and water resources planning as well as assessing the variability of the obtained flood map inundation in the present situation that cannot achieve through univariate frequency analysis.

Chapter 6

Flood Hazard in Saigon-Dongnai River Basin Under Climate Change Context

6.1 Introduction

A warmer climate is already causing extreme weather events that affect the lives of million people around the word (Schiermeier, 2011). Specifically, extreme climate events seem to occur more frequently in different parts and seasons in Asia in the future, particularly in East Asia (IPCC, 2014). Brunner et al. (2017) indicated that climate change has an impact not only on the peak but also on the volume and hydrograph shape of the flood. Hence, assessment of climate change impacts on flood should consider all flood characteristics (i.e., peak, volume and duration) than only the flood peak. Therefore, in order to make appropriate adaptation strategies, decisions and policies under climate change context, it is important to understand the change of the flood characteristics and potential flood risk in the river basin scale.

Vietnam is one of the countries that is highly influenced by climate change (Rutten et al., 2014, Trinh et al., 2013). The Saigon-Dongnai River basin plays an important role in social and economic development in the South of Vietnam. However, this basin also is facing many problems related to climate change (Noi and Nitivattananon, 2015). Indeed, this basin is one of the most susceptible regions to flood disasters. For example, 971,000 people, which account for nearly 12% of Ho Chi Minh City's (HCMC) population have been affected by flooding and 154 communes and wards in HCMC have been covered with flood waters (ADB, 2010). Several studies indicated that the common causes of flooding in this region are high tides, extreme rainfall, reservoir releases from the upstream of Saigon-Dongnai River (i.e., Trian, Dautieng and Phuochoa) and strong urbanization (ADB, 2010, Lasage et al., 2014, Storch and Downes, 2011, World Bank, 2010). Trian reservoir, located in the upper reach of the Saigon-Dongnai River basin, is one of the largest sub-basins. Therefore, the impact of this reservoir on flood risk in the downstream is greater than other reservoirs. It is, hence, necessary to evaluate the potential flood risk for this region to provide the information on present and future flood hazard (climate change condition) for establishing a flood risk mitigation policy.

There are several studies related the assessing the impact climate change on floods. Most of the flood frequency analysis studies only focus on the flood peak (Camici et al., 2014, Dobler et al., 2012, Qin and Lu, 2014). However, flood is a complex phenomenon defined by the strong correlation of its characteristics such as peak, duration and volume. If univariate frequency analysis is used to assess the behaviour of each flood characteristics, it will lead to over or underestimation of associated flood risk. In addition, in order to develop flood hazard maps, flood peak alone cannot give a reliable evaluation of hazard. Therefore, it is also important to consider simultaneously flood peak along with other flood characteristics (Aronica et al., 2012a, Candela et al., 2014).

The bivariate flood frequency is an important extension for climate change impact studies (Brunner et al., 2017). Copula approach is a very versatile approach for modelling the joint probability in a more realistic way than other approaches. Copulas can model the dependence structure independently of the marginal distribution and allow modelling different marginal distributions (Dupuis, 2007, Favre et al., 2004, Zhang and Singh, 2007). Hence, copula approach is used in this study to assess the impact of climate change on flood characteristics. Besides, Karamouz et al. (2011) suggested that floodplain areas determination

is one of the important parts of flood risk assessment which has a significant implication on the planning of social and economic development. Therefore, the joint return period of flood peak and volume, which are obtained from bivariate frequency analysis, is used as the input data for the hydrodynamic model to estimate flood hazard mapping at river basin scale.

Grimaldi et al. (2013) indicated that there are three main issues in creating the flood hazard maps using hydrodynamic models, namely the resolution of topography data, hydraulic model selection and design flood hydrograph estimating. The difference between this study and other studies related to flood hazards under climate change condition is that above three issues will be applied to improve the accuracy of the flood hazard mapping estimation.

The estimation of flood hydrograph is truly important in defining the inundation areas which are used to determine the flood hazard mapping. Angela and Giuseppe (2017) suggested that using the same joint probabilities of occurrence of flood peak and volume in developing the design flood hydrograph will significantly enhance the reliability of flood hazard mapping (Angela and Giuseppe, 2017). Hence, in order to improve the accuracy of flood hazard mapping, it is necessary to estimate the design flood hydrographs under climate change context.

Deterministic approach has been extensively used to develop flood hazard mappings. A simple form of deterministic approach is that the design flood hydrograph is used as input data for 1D hydraulic model to estimate the water level and then water depth is obtained by extracting the water level with a digital terrain (Merwade et al., 2008). However, flood events have complex spatial dynamics caused by the interaction between channel and floodplain flows and the detailed topography of floodplain areas (Stephens et al., 2012). The advanced deterministic approach, which consists of a combination of a physically-based 1D and 2D hydraulic models, allows simulating accurately the spatial and temporal dynamics of the flood process (Di Baldassarre et al., 2009, Prestininzi et al., 2011, Zhou et al., 2012).

The quality and the accuracy of the inundation maps are highly depended on the quality of topography data which are used to extract cross-section for 1D hydraulic model or surface elevation for 2D hydraulic model. In fact, in most of the studies, cross-sections have been extracted from DEM (Merwade et al., 2008). However, the cross-sections obtained from these sources are often unrealistic because of their lack of adequate horizontal resolution and poor vertical accuracy (Gichamo et al., 2012, Vaze et al., 2010). Similarly, LiDAR data cannot

provide a reliable measurement for river cross-section because LiDAR systems use near infrared laser that cannot penetrate water (Xu et al., 2010). In order to overcome this issue, all the surveyed cross-sections which are used as input in 1D hydraulic model are applied in this study.

In case of 2D hydraulic model, information about mesh resolution and topographic detail can decide the accuracy of model performance result (Dottori et al., 2013). Fine resolution mesh is necessary for capturing intricate flow paths and connectivity in urban areas (Yin et al., 2016). However, it is difficult to model large areas with fine resolution mesh because of the limitation of model preparation and computational intensity (Schubert and Sanders, 2012, Shen et al., 2015, Teng et al., 2017, Zhou et al., 2012). If the selection of mesh resolution is carefully considered, the result of the model will be significantly improved. Hence, the multi-scale modelling approach, where fine resolution is applied for channel or raised embankment areas and coarser resolution, is developed for uniform topographic height, may be an efficient solution for this situation. Surface elevations which are used as input in 2D hydraulic model are obtained using the multi-scale mesh along with detailed topography data. The detailed topography is derived from 1m resolution LiDAR (HCMC area) and 10m resolution DEM (remaining regions).

Inundation duration is an important parameter for flood risk assessment, mainly in the evaluation of transport blockades and access to emergency services. They have not been considered in many flood hazard studies due to the cost involved in its estimation. However, with the help of latest advances in flood modelling techniques such as flexible meshes generation and the advantages of coupled hydrodynamic model, in this study, the inundation duration is also determined. More importantly, with the increases of land use and climate changes that threaten more extreme flooding, developing the computationally efficient model is extremely needed. In fact, the need to developing efficient model flooding is growing in several fields such as flood risk management, real-time flood forecasting system, and flood evacuation planning.

This study aims to address the following issues: (i) estimation of the flood design hydrographs using copula theory of flood variables (i.e., volume and peak) and shapes through historical observed flood hydrograph combined with cluster analysis; (ii) developing the

computationally efficient flood model using advanced deterministic approach based on couple 1D-2D hydraulic model and high quality of topography data; (iii) creation of flood hazard maps, which is quantified by considering the flood depth and velocity in combination; and (iv) assessing flood risk changes under climate change context.

6.2 Study Area and Data

The Saigon-Dongnai river basin is located in the southern part of Vietnam within latitudes 10°30'N-13°00'N and longtitudes105°15'E-109°30'E. This basin has very significant impacts on the socio-economic development of south of Vietnam. The surface area of this basin is 14,800 km^2 (Trian catchment) and 40,680 km^2 (Sea). This basin has four major rivers, namely Dongnai (mainstream), Be, Saigon as major tributaries and Vamcodong rivers that join the Dongnai river before flowing into the sea (Ringler et al., 2012). The lower part of Saigon-Dongnai river basin is a lowland area which is covered between downstream of Dautieng, Phuochoa and Trian reservoirs and coastal line as shown in Fig. 6.1.

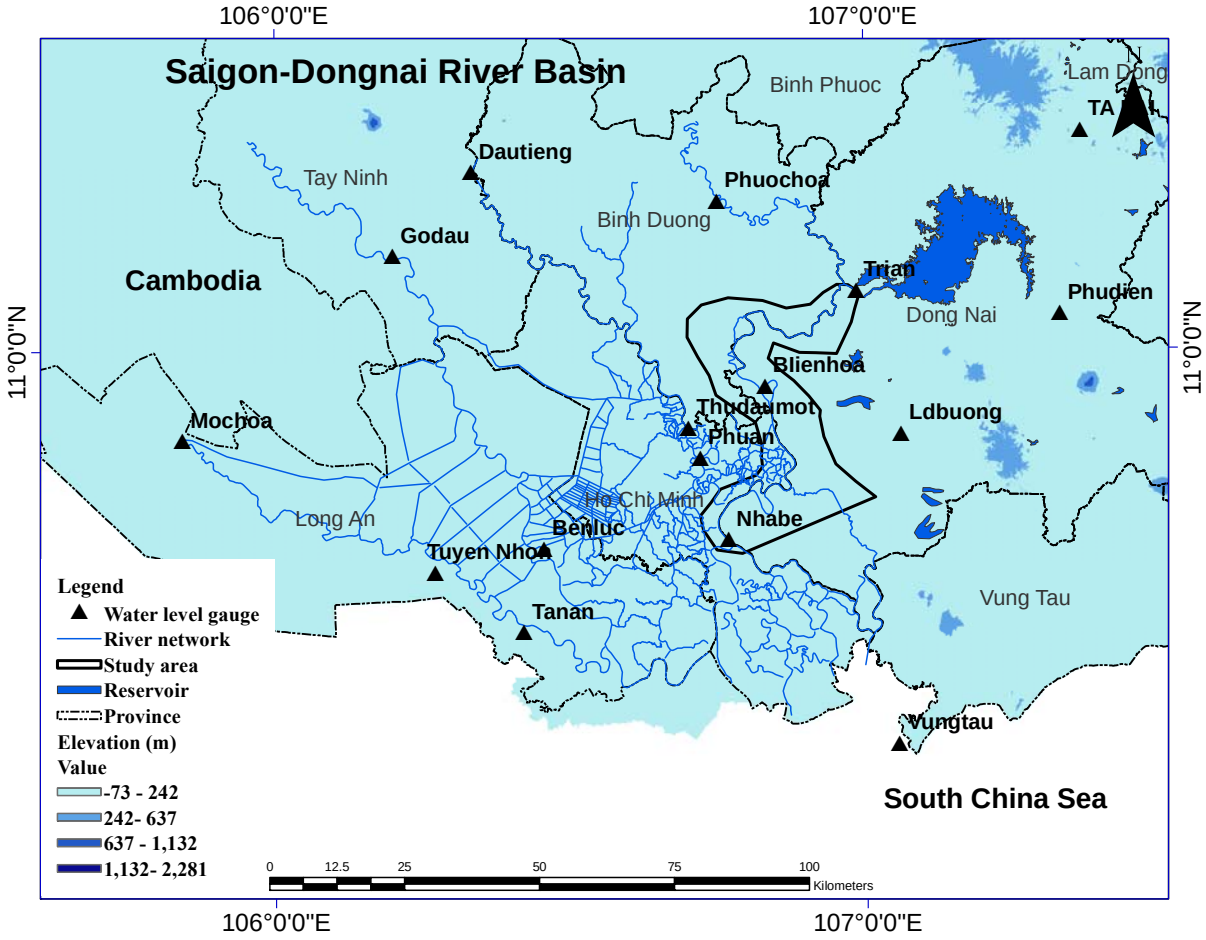


Fig. 6.1 The downstream of Trian catchment

The climate of this basin has been divided into two distinct seasons, namely, rainy (April to November) and dry (December to March) seasons. The climate is controlled by the North-East and South-West monsoons. 90% of rainfall occurs in the seven months of rainy season. The annual average rainfall and temperature are about 2,400 mm and 25.9^{o}C respectively. The hydrological regime of Saigon-Dongnai River basin is influenced by the semi-diurnal tide, precipitation and upstream reservoir releases.

6.3 Methodology

The methodology used in this study is shown in the flowchart (Fig. 6.2). Two scenarios (i.e., RCP 4.5 and 8.5) from five RCMs, namely, HadGEM3-RA, SNU-MM5, SNU-WRF, RegCM4 and YSU-RSM belonging to the CORDEX-EA projects are used. Climate outputs in the future period (2020-2045) obtained from these RCMs are downscaled with respect to the

observed data. And then, it is used as the input for the physically-based hydrological model to simulate the future streamflow data. The present and future discharge time series are used to extract the flood characteristics. Flood design hydrographs are obtained using bivariate frequency analysis and shape through observed historical flood events and cluster analysis. Finally, flood design hydrographs are used as input to the hydraulic models to develop flood hazard maps for present and future scenarios.

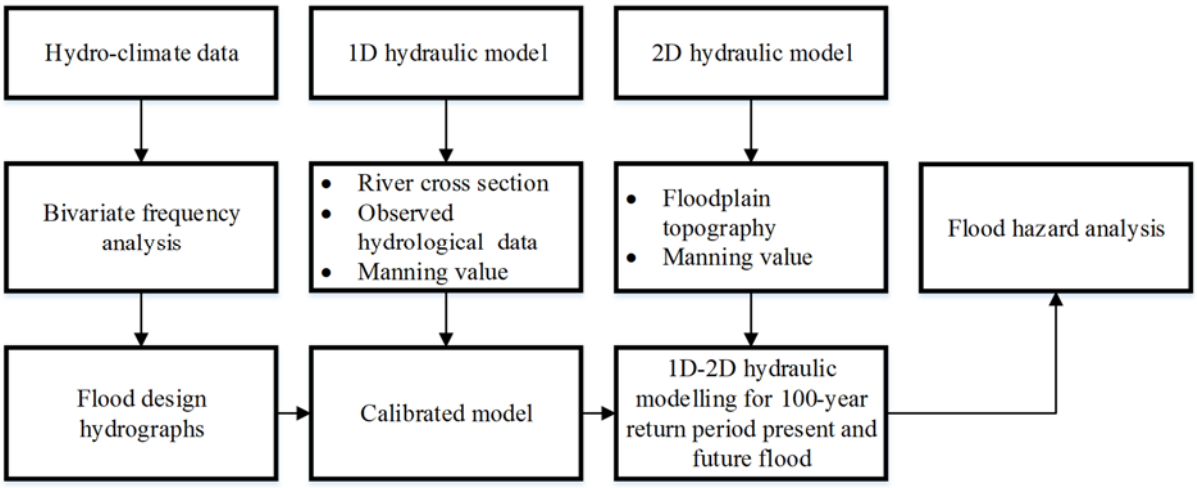


Fig. 6.2: Framework of the methodology

6.3.1 Extracting flood characteristics

AM and POT approaches are widely used to extract flood characteristics. However, block maxima cannot consider multiple occurrences of flood events. Unlike the AM, which extracts only one event per year, POT considers a wider range of events and provides more information than block maxima. Further, several studies suggested the main advantage of using the POT approach for smaller sample sizes (Beguería, 2005, Bezak et al., 2014). Therefore, POT is used to extract the present and future flood characteristics in this study. The lag-autocorrelation plot is also used to check the assumption of IID flood variables (i.e., peak, volume and duration).

6.3.2 Bivariate frequency analysis

Unlike traditional multivariate parametric distribution, in copula approach, the dependence structure is independently modelled with the marginal distributions which allow

for multivariate random events which are modelled using several different marginal distributions. Several researchers have used copula to investigate the bivariate frequency analysis (Dung et al., 2015, Reddy and Ganguli, 2012, Sraj et al., 2015). The most important step in the modelling process using copula approach is the selection of copula function which best fits the data sample (Favre et al., 2004). Dupuis (2007) and Poulin et al. (2007) suggested that copulas should be chosen based on several classes of copulas and several degrees of tail dependence.

To describe dependence in multivariate extreme values, there are two possible situations, namely, asymptotic dependence or asymptotic independence (Coles et al., 1999). Determination of whether the variables have asymptotic dependence or asymptotic independence is very important in multivariate extreme analysis. In fact, in situations where diagnostic checks suggest data to be asymptotically independence, modelling with the extreme value copulas is likely to lead to misleading results (Coles et al., 2001, Ledford and Tawn, 1996). Therefore, copulas are chosen based on the results of the tail dependence tests. As presented in Chapter 4, the LLHR test based on the difference between the log-likelihood of asymptotic dependence and asymptotic independence and tail dependence test introduced by Frick et al. (2007), is used to determine whether the variables have asymptotic dependence or asymptotic independence.

Extreme value copulas can be a suitable selection in modelling data with positive correlation and arise naturally in the domain of extreme events (Gudendorf and Segers, 2011, Mirabbasi et al., 2012). In addition, numerous recent studies indicated that Gumbel-Hougaard copula, belonging to extreme value copulas, can work well when flood variables are diagnosticated as asymptotical dependence (Dung et al., 2015, Karmakar and Simonovic, 2009, Poulin et al., 2007, Zhang and Singh, 2006).

Frank and Clayton copulas, belonging to the Archimedean family, have been widely used in hydrologic analysis because it can model both negative and positive associated variables. Besides, Frank and Clayton copulas, which have zero dependencies in both tails, may be suitable in case of the nonexistence of tail dependence (Dung et al., 2015, Poulin et al., 2007, Sraj et al., 2015). Further, Gaussian copula, which belongs to the family of elliptical, is characterized by correlation matrix and generates a wider range of dependence behaviour

(Bortot et al., 2000). Renard and Lang (2007) showed that Gaussian copula can be useful in modelling the dependence structure in the hydrological application because of its simplicity. Therefore, Clayton, Frank and Gaussian copulas are used to model the dependence structure where pairs of flood characteristics are diagnosed as asymptotic independence.

Several distributions which have been widely used for modelling extreme values (i.e., LN, P3, LP3, Gumbel, GPD and GEV) are used in this study to identify the best fitting distribution to the flood variables (Bezak et al., 2014, Lang. et al., 1999, Saf, 2009a, Salas et al., 2012). In order to choose the appropriate copula, AIC and test introduced by Genest et al. (2011) based on a Cramer-von Mises (S_n^{II}) statistic measuring the distance between parametric and non-parametric estimator of the Pickands dependence function are used in this research.

Another important step in bivariate frequency analysis is the selection of the appropriate return period. The joint return period called OR (X≥x or Y≥y) and AND (X≥x and Y≥y) have been widely used (Shiau, 2003, Vittal et al., 2015). However, in term of flood hazard evaluation, flooding can occur when only flood peak or volume is exceeded (Angela and Giuseppe, 2017). Hence, OR joint return period is used to obtain the flood design hydrograph.

6.3.3 Design flood hydrograph

For developing the design flood hydrograph, in addition to the determination of the value of flood characteristics pair (i.e., flood peak and volume) characterized by the same return period, it also requires the knowledge of shape to assign them (Aronica et al., 2012a). Cluster analysis, determined the similarity between different samples using an algorithm to identify the relationships among attributes, is applied to find the shapes of the hydrograph. Hierarchical Cluster Analysis (HCA) calculated the distances between all samples using Euclidean or Manhattan distance and is the most common approach in which cluster can be formed (Berrueta et al., 2007, Patras et al., 2011). There are several methods to define the distance between two groups, namely single, complete, centroid and Ward's method. However, Ward's method is widely used in the hydrological analysis (Angela and Giuseppe, 2017, Aronica et al., 2012a). In the Ward's method, the distance between two groups is defined as minimizing the increase in sum of squared errors of the distance of any two clusters.

6.3.4 Coupled hydrodynamic modelling

Flood models can be categorized into several types depending upon their data requirements, level of complexity of the underlying equation and the resolution. Using 1D hydraulic model is common practice in determining floodplain area because it is easy to use and require fewer input data and the simulation results can be quickly obtained (Pappenberger et al., 2005). 1D models can solve problems of flood flows in open channels with assumptions that vertical acceleration is not significant and that water level in the channel cross-section is approximately horizontal are valid. However, problems arise when the channel is embanked and water levels are different in the floodplain than in the channel and in such situations, 2D models are needed. 2D numerical models solve full shallow water equations, which are able to simulate timing and duration of inundation with high accuracy. 2D flood inundation models are now important parts of flood risk management practice because they are capable of adequately predicting water depth, velocity and flood risk with high accuracy (Lamb et al., 2009, Teng et al., 2017). However, 2D has some limitations such as taking a long time to set up and run 2D model, particularly for the large area.

To overcome these disadvantages of both 1D and 2D numerical models, coupled 1D and 2D flood model are considered as suitable tools to model flood flows, both in channel and floodplains. The coupled models offer a great advantage for real-time simulation of flood events. There are several coupled models for simulating the hydrodynamic regime in the river. SOBEK 1D and 2D and MIKE FLOOD developed by Delft hydraulics and DHI respectively are widely used in recent years. MIKE FLOOD allows exploiting the best features of both MIKE 11 and MIKE 21 to simulate flood regime. Furthermore, MIKE FLOOD allows avoiding many of the limitations of resolution and accuracy encountered when using MIKE 11 or MIKE 21 separately.

In the 2D hydraulic modelling, the accuracy of the model can be associated with the computation grid resolution. The selection of the resolution of the grid is based on several factors such as the available of topographic data, the requirement of the minimizing errors in the schematization of the physical processes. Dottori et al. (2013) defined the coarse resolution as a mesh of 20m, fine resolution as a mesh between 2 and 5m and very fine resolution as a mesh below 2m in the urban areas. Huthoff et al. (2015) and Chatterjee et al. (2008) suggested

that 50m resolution of topographic data is sufficient to resolve flood response for non-urban areas. However, high resolution of computation mesh cannot be feasible for large scale floodplain areas. Hence, the different mesh resolution corresponding with the quality of topographic data in hydraulic modelling is applied in this study. A flexible mesh is an advanced approach in the model discretization of space, and this is carefully constructed to have a fine resolution around complex areas to reflect huge changes in the topography and coarser resolution over large areas with a little spatial variance to save computational effort. The flexible mesh has a great impact on model performance (Teng et al., 2017).

2D flexible mesh are used as an input data for 2D hydrodynamic models because it easily allows representing the complex geometry associated with the flood ways. It also allows to be represented a small scale features at a finer resolution and the broader floodplain at a coarser resolution without resorting to model grid nesting (Mackay et al., 2015). It also allows fully hydrodynamic with higher order scheme simulations. The floodplain in the lower of Trian catchment is modelled using flexible mesh. This allowed higher definition meshes to be applied over known flood runner and flood ways while much coarser meshes were applied to the broader floodplain where the terrain if more uniform require less detail. The meshes have been developed to represent the flood ways and broader floodplain, the triangular mesh with various sizes depend on the level of resolution required.

6.3.5 Flood hazard analysis

Flood hazard maps provide useful information about the flood severity for the decision maker, planner, and local governor in flood risk management, master planning development and flood emergency response planning. Flood hazard maps can be drawn using hazard curves by combining different flood parameters such as flood extent, water depth and flow velocity with the return period into different hazard classes (Mazzoleni et al., 2014). In the past, flood hazard maps were generated based on only on inundation depths. From the flood hazard maps obtained only from flood depth, it was found that it has the potential to reflect the unsafe areas for people during the extreme events within the floodplain. Ali (2018) used only flood depth to generate flood hazard maps using the modifications standards of Japan International Corporation Agency (JICA). Flood hazard also depends on the duration of inundation, the amount of energy that is contained in the overland flow and the available warning time to

evacuate the population. Therefore, it is important to identify a technique to generate multiple parameters flood hazard maps by incorporating all possible hazardous elements of flood events.

Physical damage to properties during a flood event is closely related to the amount of energy contained in the floodwater. The extent of soil erosion or the displacement of properties also depends on flood velocity (Tennakoon, 2004). Therefore, the flood hazard should be obtained using both flood depth and velocity. Further, Luke et al. (2018) suggested that if multiple flow velocities are incorporated with flood depths, flood hazard maps can be generated which would be suitable for reflecting of risk for others (i.e., people, car and building).

Australian studies suggested flood hazardous zones, describing by velocity flow and water depth, defined by velocity multiple by flood depth thresholds (AEMI, 2014). Similarly, US Department of Land and Soil Conversation classified the flood hazard categories based on maximum energy of floodwater and the hazard zone is divided into high hazard and low hazard zones (OEH, 2005). High hazards correspond to possible danger to people, vehicles and buildings while low hazards correspond to less difficulty in evacuating people and protecting properties. The classification by the US Department of Land and Soil Conservation is realistic for urban applications. However, according to this classification, the threshold of 1.2 m flood depth with 1.4 m/s velocity is classified as high hazard zone. Their hazard category is inadequate for an area experiencing a heavy flood. Generally, it is reported that flood depth ranges from as small as 0.3 m though hazardous to children would be within the self-help range for adults (Mani et al., 2014).

It was, hence, decided to use the flood hazard classification introduced by AEMI (2014) in this study. In this classification, the flood hazard assessment considering the impact on people, vehicle and buildings that may occur when floods meet the following conditions: (i) flood depth is greater than 0.3; (ii) water velocity is greater than 2 m/s and (iii) the product of these two parameters is greater than 0.3 m^2/s. The flood hazard classifications that relate to specific thresholds are presented in Table 6.1, which also contains the limits for classification. The combined effect of two parameters, viz., flood depth and velocity have been considered. The hazard classification is done on a 6-point scale, H1 to H6 scale, H1 being safe for vehicles, people and buildings while H6 being unsafe for vehicles, people and all buildings considered vulnerable to failure.

Table 6.1: Combined hazard curves-vulnerable thresholds and limits for classification

Hazard classification	Description	Limiting still water depth (D) m	Limiting velocity (V) m/s	Classification limit (D×V) m^2/s
H1	Generally safe for a vehicle, people and buildings	0.3	2	D×V≤0.3
H2	Unsafe for small vehicles	0.5	2	D×V≤0.3
H3	Unsafe for vehicles, children and the elderly	1.2	2	D×V≤0.6
H4	Unsafe for vehicles and people	2	2	D×V≤0.6
H5	Unsafe for vehicles and people. All building types are vulnerable to structural damage. Some less robust building types vulnerable to failure	4	4	D×V≤4
H6	Unsafe for vehicles and people. All building types considered vulnerable to failure	-	-	D×V>4

Next to the flood depth and velocity, duration of flooding is also an important parameter in flood risk analysis. The duration of inundation above flood depth threshold is an important baseline information when considering isolation aspects of emergency management. Mani et al. (2014) indicated that the flood inundation of 0.3 m for 1 to 2 days duration can cause severe damage to crop productivity and manufacturing production process. Besides, in order to maintain or enhance the ecological functioning of floodplain wetlands, it is important to understand the timing and duration of connectivity to the river channel vary with streamflow and climate (Karim et al., 2016). Most of the regions in this study area are agricultural land, wetland and industrial zones, it is, hence, very important to assess the impact of the inundation duration on the crop productivity, manufacturing production process as well as floodplain ecosystems.

6.4 Results and Discussion

6.4.1 Bivariate frequency analysis

LLHR and TailDep tests are used to identify the asymptotic dependence and independence of flood characteristics pairs. The results from two tests are nearly similar. Table

6.2 shows the p-value of LLHR and tail dependence tests for all pair of observed and future flood variables. Based on these tests, the asymptotic dependence and independence of present and a future pair of flood characteristics are shown in Table 6.2.

Table 6.2: p-value of LLHR and TailDep tests

Scenarios	Tests	p-value					
		RCP 4.5			RCP 8.5		
		DV	DP	PV	DV	DP	PV
Obs	LLH	0.04	0.01	0.01			
	TailDep	0.30	0.02	0.04			
	Diagnostic	Dep	InDep	InDep			
Had	LLH	0.38	0.24	0.07	0.49	0.25	0.20
	TailDep	1.00	0.12	0.16	1.00	0.11	0.35
	Diagnostic	Dep	Dep	Dep	Dep	Dep	Dep
Reg	LLH	0.65	0.19	0.20	0.43	0.24	0.73
	TailDep	0.56	0.19	0.13	0.51	0.12	1.00
	Diagnostic	Dep	Dep	Dep	Dep	Dep	Dep
YSU	LLH	0.65	0.16	0.30	0.80	0.60	0.55
	TailDep	1.00	0.25	0.33	1.00	0.16	0.31
	Diagnostic	Dep	Dep	Dep	Dep	Dep	Dep
WRF	LLH	0.38	0.35	0.22	0.50	0.13	0.13
	TailDep	1.00	0.88	0.65	1.00	0.16	0.11
	Diagnostic	Dep	Dep	Dep	Dep	Dep	Dep
MM5	LLH	0.37	0.13	0.13			
	TailDep	1.00	0.04	0.04			
	Diagnostic	Dep	InDep	InDep			
GCM	LLH	0.88	0.04	0.05	0.23	0.18	0.39
	TailDep	1.00	0.08	0.09	1.00	1.00	1.00
	Diagnostic	Dep	InDep	Indep	Dep	Dep	Dep

The best marginal distribution for each flood variable is selected based on the minimum AIC values. Table A. 5 shows the value of AIC for all marginal distributions used for peak, volume and duration. The result indicates that most scenarios data follow the LP3 and P3 distributions, namely 36.11% and 30.56 % respectively. 16.67 % and 11.11% of the samples fit the LN and GEV distributions respectively. Only the 5.56 % data follows the Gumbel distribution.

Based on the result of tail dependence tests, three extreme value families of copulas (i.e., Gumbel-Hougaard, Galambos and Husler-Reiss) are chosen to model asymptotical dependence

pair of flood characteristics. Gaussian, Frank and Clayton copulas are used in modelling asymptotically independence pair of flood characteristics. The copula dependence parameter, AIC value and the Cramer-von Mises (S_n^{II}) statistic and p-value are given in Table A. 6 and Table A. 7. Based on the AIC value and S_n^{II} test results, the most appropriate copulas are selected to calculate the joint return periods for all pair of flood characteristics.

The values of the pair of flood peak and volume in the same return period can be obtained through bivariate frequency analysis. For example, the value of flood peak (m^3/s)-volume (10^6 m^3) pairs in the same joint return period (OR) of 100-year are (5,585-12,918), (6,368-21,987) and (4,467-27,641) for the present, mean RCP 4.5 and mean RCP 8.5 scenarios respectively (Table 6.3).

Table 6.3: Flood magnitudes for the same joint return period of 100 years

The same joint return period of 100 years									
Scenarios	Variables	Values			Scenarios	Variables	Values		
Obs	Peak	5,585	4,690	4,295	Had_85	Peak	4,278	4,034	3,927
	Volume	12,918	12,990	13,099		Volume	30,557	31,860	35,670
Reg_45	Peak	6,278	4,036	3,875	GCM_45	Peak	7,824	6,962	6,280
	Volume	29,187	30,502	34,355		Volume	33,808	35,556	41,003
YSU_45	Peak	6,159	4,053	3,754	Reg_85	Peak	4,348	3,484	3,038
	Volume	24,343	24,900	25,517		Volume	21,545	22,008	23,093
MM5_45	Peak	6,278	4,308	4,036	YSU_85	Peak	3,300	3,148	2,972
	Volume	20,259	20,333	20,408		Volume	29,495	31,133	38,228
WRF_85	Peak	3,650	3,367	3,367	WRF_45	Peak	5,755	3,449	3,407
	Volume	33,616	34,411	37,472		Volume	5,897	7,496	10,603
Had_45	Peak	5,913	3,797	3,608	GCM_85	Peak	6,761	5,696	5,468
	Volume	18,429	19,275	20,335		Volume	22,994	23,497	24,675

Note: Peak (m^3/s) and Volume (10^6 m^3)

From the results, it is indicated that the future flood could have more volume than the present flood in both scenarios. The future flood peak of mean RCP 4.5 is larger than observed

flood peak whereas the flood peak of mean RCP 8.5 is smaller than observed flood peak. Besides, the flood volume of mean RCP 8.5 is largest compared with others.

6.4.2 Flood design hydrograph

In order to find the typical shape of the hydrograph, cluster analysis with Ward method is used. A visual representation of the distance at which shapes are combined is shown in the dendrogram (Fig. A. 6). The number of hydrograph clusters should be large enough to describe the hydrograph variation but small enough to identify the characteristic of hydrograph shapes. Therefore, three typical shapes of hydrographs are chosen based on the cluster analysis. The sampled pairs of flood peak and volume are associated with a randomly selected normalized characteristics hydrograph conditioned by their probability of occurrences. Merging the normalized hydrograph and generated peak and volume pair from copulas, flood hydrographs are obtained. Fig. A. 7 shows three synthetic flood hydrographs corresponding with 100-year return period for the present and future floods. And then, these flood hydrographs are used as the upper boundary condition into the hydrodynamic model.

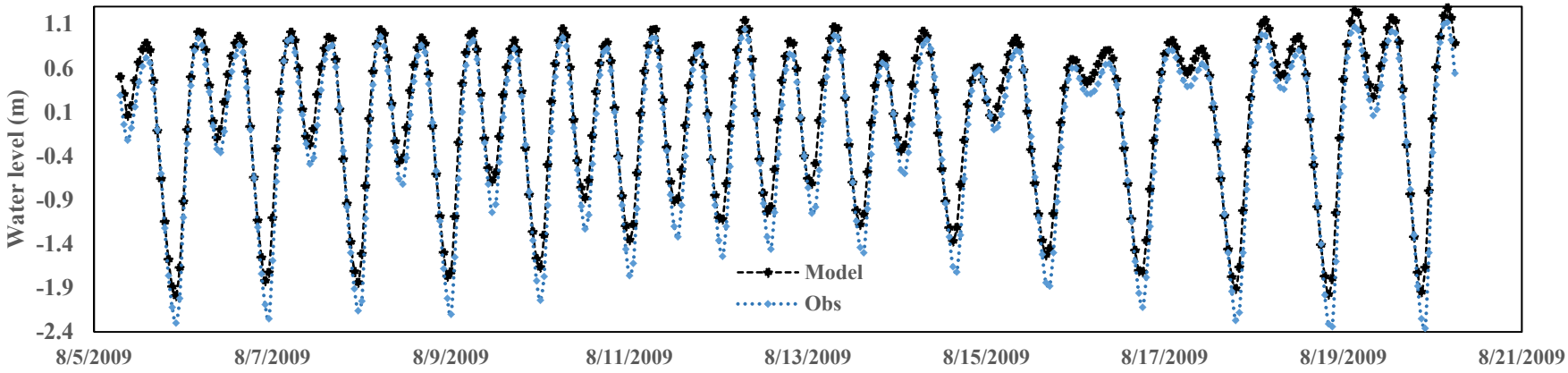
6.4.3 Hydraulic modelling

MIKE 11 is a 1D hydrodynamic model for solving the Saint-Venant equations in the river networks. The river network, the cross-section data and the boundary condition are defined for each river. The lower part of Saigon-Dongnai River basin which is covered within downstream of Dautieng, Phuochoa and Trian reservoirs and coastal line is used in the network for the model. Five main rivers (i.e., Dongnai, Saigon, Be, Vamcotay and Vamcodong) and 251 small streams are used. More than 300 observed cross-sections surveyed in the year 2009 are used in this study. Observed releases from Trian, Phuochoa and Dautieng reservoirs, observed discharges from Godau streamflow gauge located in the VamcoTay river and observed water levels from Mochoa located in the Vamcodong river are used as the upstream boundary conditions. Observed sea level (Vungtau gauge) is considered as the downstream boundary condition. Two hourly observed water levels at Nhabe and Bienhoa, which are located in the Dongnai River are used in calibration and validation for 1D hydrodynamic model.

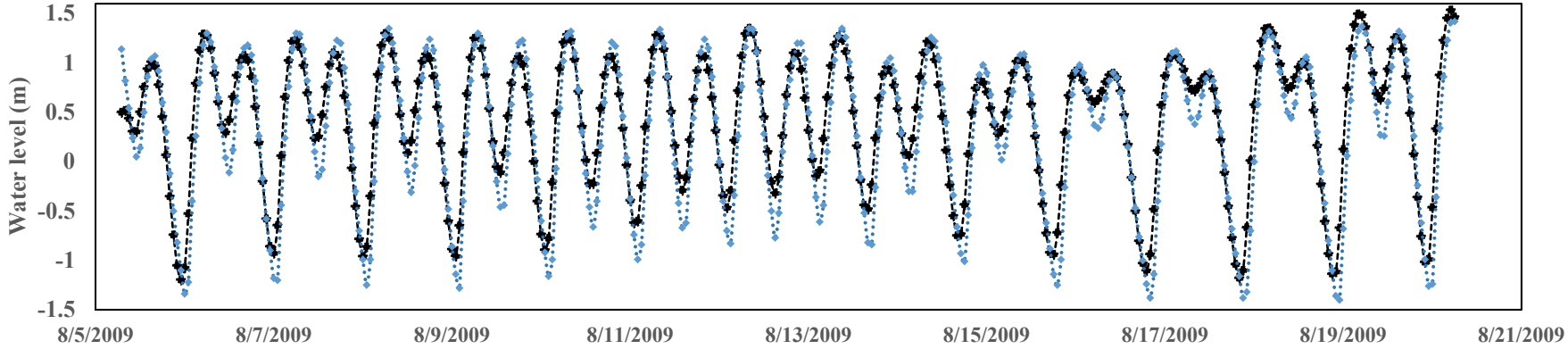
The shuffled complex evolution algorithm is used to optimize the parameters. Manning's coefficient is selected as the model calibration parameter. Based on the objective

and the availability of observed data, the observed water level in Nhabe and Bienhoa are used for calibration and validation of 1D model. The model is calibrated for the using the data of the year 2009 while the model is validated for the year 2012, for the flood season. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used to assess the model performance. There is good agreement between the simulated and observed water levels at most of the time in both locations, as seen from Fig. 6.3. RMSE and MAE values are 0.236 (0.220) m and 0.286 (0.190) m for Nhabe and Bienhoa respectively in the calibration procedure. Similarly, these values are 0.339 (0.280) m and 0.193 (0.140) m for validation step.

A flexible mesh is carefully constructed in 2D hydraulic modelling in this study. The potential inundation areas, complex areas (i.e., channel and raised embankment) and regions in which LiDAR data are available are assigned with finer resolution while the coarser resolution is used for the remaining areas. The flexible mesh with 75,263 elements and 38,202 nodes covering an area of about 738.02 km^2 is built. The surface elevations for the study area are derived from with 1m resolution of LiDAR for HCMC area and 10m resolution DEM for the remaining areas. The Manning's n value for floodplain areas, which are selected from the land use map, can be set for each computational element. The Manning's n values used are 0.2, 0.07 and 0.03 $m^{1/3}/s$ for residential, agriculture and water bodies respectively.

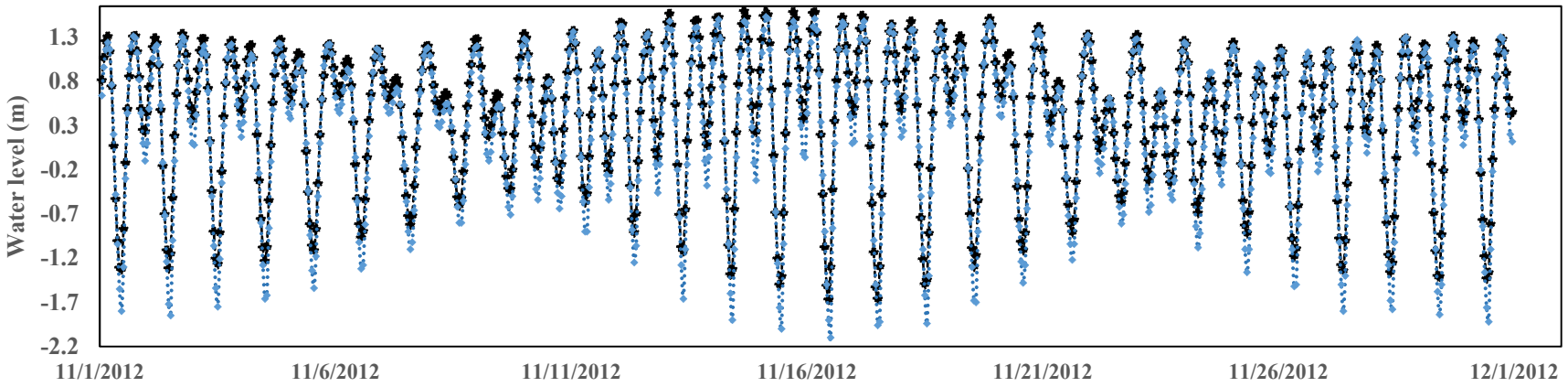


(a)

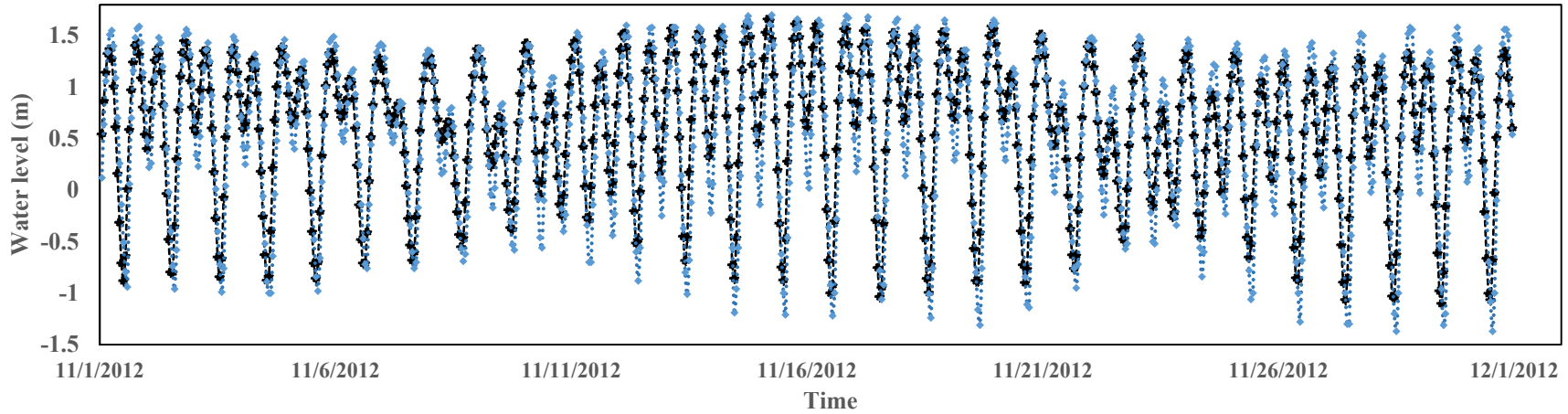


(b)

Fig. 6.3: Simulated and observed water levels for (a) Nhabe and (b) Bienhoa during period of August 2009



(c)



(d)

Fig. 6.3 (continues): Simulated and observed water levels for (c) Nhabe and (d) Bienhoa during period of November 2012

6.4.4 Flood hazard analysis

The changes in 100-year return periods of future flood characteristics in the period (2020-2045) in the Trian watershed were reflected in the changes of the flood inundation in the downstream. In fact, the future flood characteristics are showing increasing values in for the inundation areas and decreasing the flood characteristics will lead to reduce the flooding areas (Fig. 6.4). The total inundation areas under different hazard indexes are given in Table 6.4. The inundation areas of 100-year return period under mean RCP 4.5 is greater than the observed flood. Under mean RCP 8.5 scenario, the inundation area of 100-year return period is smaller than the present observed flood peak.

Table 6.4: Inundation areas under different hazard indices

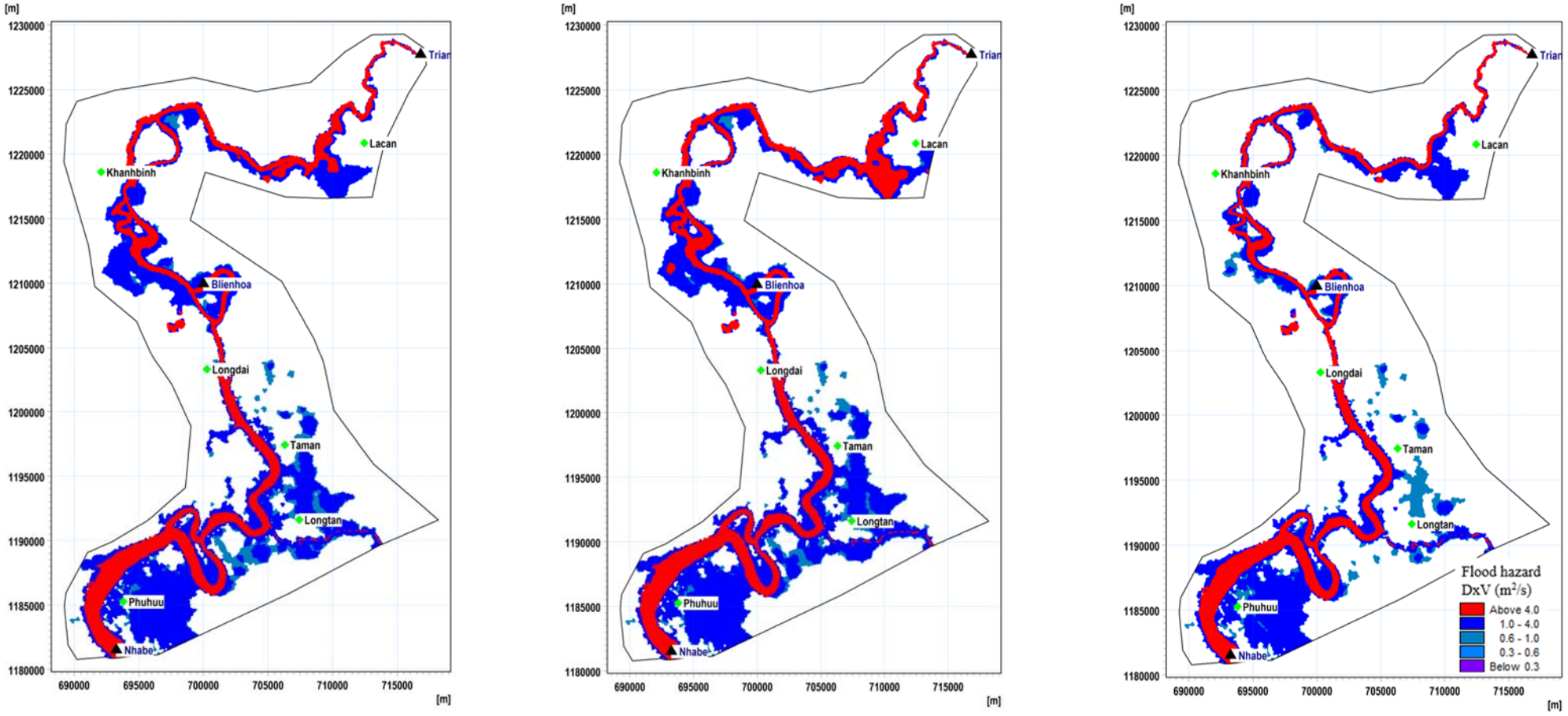
Scenarios	Total of inundation areas (km^2)	Flood hazard index: H1 (km^2)	H2 (km^2)	H3 (km^2)	H4 (km^2)	H5 (km^2)	H6 (km^2)
Obs	199.740	37.367	26.227	45.092	22.359	14.230	54.465
RCP 4.5	317.645	34.406	28.186	53.188	27.894	33.352	140.620
RCP 8.5	153.900	29.576	19.831	33.983	8.538	11.069	50.903

The land area under the H6 hazard index (i.e., unsafe for vehicles and people and all building types considered vulnerable to failure) is greater than any other category. For example, 54.465 km^2 of the study area is subject to H6 hazard index under 100-year return period of the present flood. 140.620 km^2 and 50.903 km^2 are under H6 hazard index for RCP 4.5 and RCP 8.5 respectively.

Based on the spatial distribution of flood hazard in the Saigon-Dongnai River basin, flood inundation can be more intense in the lowlands and agricultural areas (i.e., Longtan, Lacan). The urban areas (i.e., Bienhoa and Longdai) are less affected by flood than others because of ground elevation in these areas may be upgraded for construction purposes.

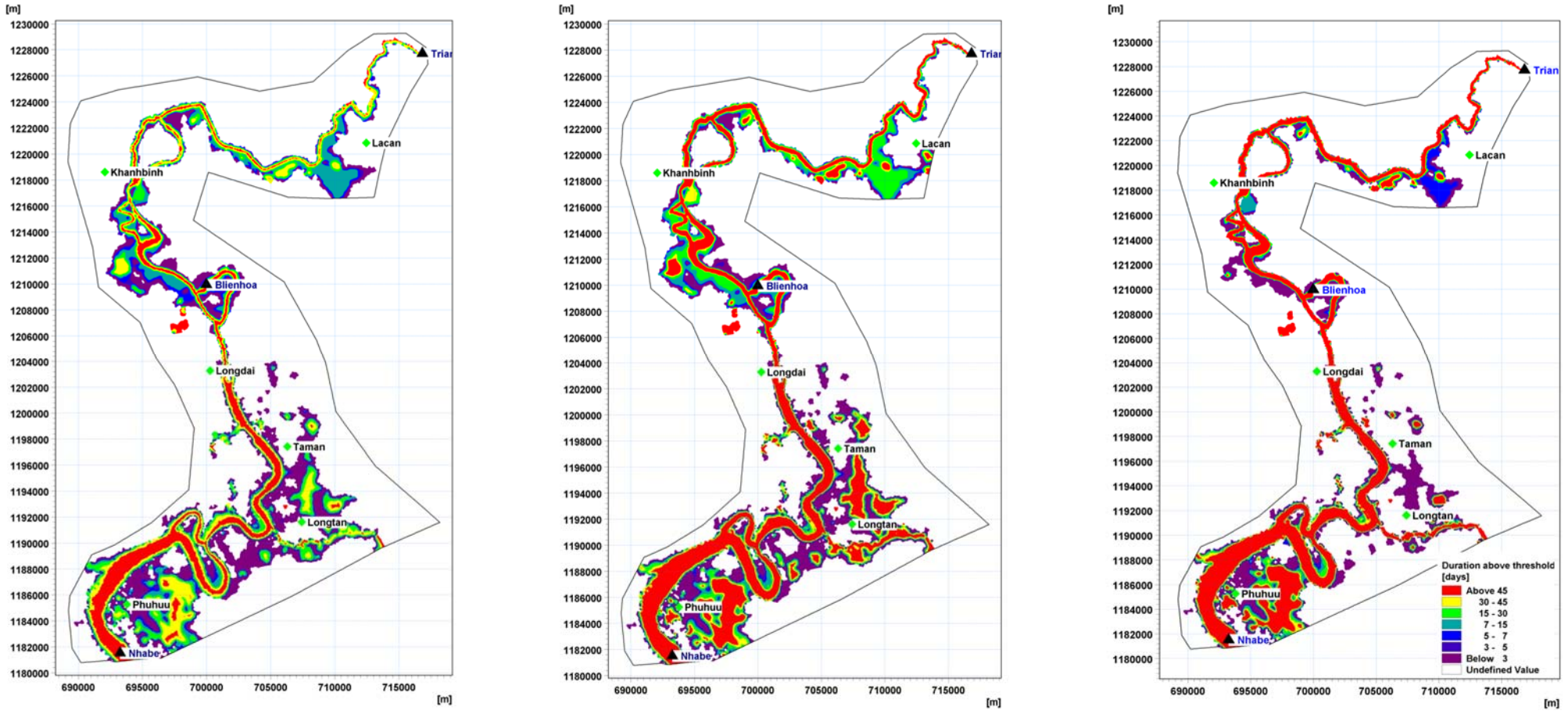
Besides, considering the inundation areas under climate change context, in this study, the duration of the inundation is also investigated. Fig. 6.5 shows the maximum inundation duration above flood depth threshold (0.3 m) for present and future floods. It is clear that the duration of inundation will increase corresponding to its flood magnitude. However, the duration of inundation is also greatly affected by the flood volume. For example, the duration

of inundation in the Phuhuu in the Saigon-Dongnai River basin under RCP 8.5 is greater than the present flood although the present flood peak is larger than RCP 8.5 scenario. Spatial variation in inundation duration is mainly controlled by the land topography. Indeed, the lowland in the lower part of this basin is recorded the inundation duration longer than the upper parts of floodplains. The main reason for this situation is that the geomorphologic processes, generally result in large adjacent areas of flat land in the lower parts of a floodplain than its upper parts.



(a) (b) (c)

Fig. 6.4: Flood hazard maps for flood event with 100-year return period of (a): Observed (b): Mean RCP 4.5 (c): Mean RCP 8.5



(a) (b) (c)

Fig. 6.5: Maps of maximum inundation duration above threshold (a): Observed (b): Mean RCP 4.5 (c): Mean RCP 8.5

6.5 Summary and Conclusions

In this study, a comprehensive approach for developing a flood hazard mapping for the downstream of Trian reservoir under climate change context is presented. The approach combines the bivariate flood frequency and the coupled hydrodynamic model for simulating the inundation in a spatial context.

Flood hydrograph estimation has an important role in accurate assessment of flood hazard. Three flood variables (i.e., peak, volume and shape) are estimated using bivariate frequency and cluster analysis. Indeed, the multivariate statistical analysis provides a better approach than the univariate method in term of extreme events statistic. Therefore, the copula approach is used to model the dependence structure between flood peak and volume. More importantly, the result from tail dependence test provides more useful information to choose the more appropriate copula in multivariate frequency analysis. Extreme value copulas are used when flood variables are diagnosed as asymptotical dependence whereas Clayton, Gaussian and Frank copulas are considered for asymptotical independence cases.

High-resolution DEM based on the LiDAR data along with fully 2D hydrodynamic model provides more detail flood hazard mappings for this study area. Flood hazard is quantified by considering the combination of flood depth and velocity. Spatial variability of flood hazard is also created. This research provides significant potential for better flood inundation estimation in the future at river basin scale.

This study also indicated that the inundation duration not only is controlled by flood magnitude but also by the flood volume. From the results, it is observed that most of the agricultural areas located downstream of Trian catchment will be greatly inundated under climate change context. The lower part of this river basin tends to have longer inundation duration than other parts because of flat land topography.

These results of this study will help the policy makers and stakeholders to plan for the social and economic development in this river basin. In addition, it also provides significant information for the emergency preparedness plan, including aid and relief operations for inundation areas in the future flood events.

It is very important to highlight that the flood protection and adaptation measures for agricultural areas need to be implemented to minimize the consequence of flood damages under climate change conditions. Use of flood-tolerant species may be an effective approach in the areas where flood is frequently recorded. More importantly, a change in the cropping pattern is needed. Annual crops need to be replaced by summer period growing varieties for these areas.

Chapter 7

Summary and Conclusions

7.1 Summary

Floods are one of the worst natural disasters, which cause huge damage annually including loss of human lives. The damage and loss of life caused by floods could be higher in the future also due to climate change. Assessment the changes of flood characteristics under climate change context plays a considerable role in managing the flood risk. Quantifying the vulnerable areas associated with the changes of climate allows local authorities to provide a good the future development planning. Therefore, in this study, quantifying the impact of climate change on flood risk is carried out at river basin scale.

The design and assessment of flood risk of hydraulic structures, water resources planning, reservoir management and flood hazard maps involve the identification of the given flood events with a low probability of exceedance. FFA seeks to connect the magnitude of extreme events with their frequency of occurrence via probability distribution. The objective of FFA is to estimate the return period associated with a flood of given magnitude. The return

period is a standard criterion in the design of the hydraulic structures or the flood control. It becomes necessary, in most studies, to obtain the frequency curve fitting the PDF to the observed data to estimate flood quantiles associated with given return periods.

To investigate the changes in flood magnitude, AM and POT are commonly used to extract the flood events. AM sample is defined by the maximum peak flow of each year. However, AM cannot be used in the case of short data series because it does not provide adequate information. Unlike the AM, which extracts only one event per year, POT considers a broader range of events and provides more information than AM, primarily for estimation of flood magnitude. Many researchers have investigated the choice between AM and POT. Some have emphasized that POT approach is more suitable for extreme value analysis and provides more accurate estimates of flood quantiles than corresponding AM approach.

The frequency of occurrence of the flood variables provides important information for the design of hydraulic structures, water resources planning, reservoir management and flood hazard mapping. Furthermore, flood is a complex phenomenon defined by strongly correlated characteristics such as peak, duration and volume. Therefore, it is necessary to study the simultaneous probabilistic behaviour of flood characteristics.

Traditional multivariate parametric distributions have widely been applied for hydrological applications. However, this approach has some drawbacks such as the dependence structure between the variables, which depends on the marginal distributions or the flood variables that have the same type of marginal distributions. Copulas are widely applied to overcome the restriction of traditional bivariate frequency analysis by choosing the marginals from different families of the probability distribution for flood variables. The most important step in the copula modelling is the selection of copula function which is the best fit to data sample. The choice of copula may significantly impact the bivariate quantiles. This study will investigate the potential of performing the tail dependence tests for the pairs of flood characteristics and evaluating the performance of extreme value copulas for asymptotic dependence variable and Clayton, Frank and Gaussian copulas for asymptotic independence variables.

Flood hazard mapping is one of the critical aspects of flood risk assessment which has a significant implication on the planning of social and economic development. The information

of flood hazard is also essential to provide various strategies for mitigating the flood risk, which in turn, can reduce the losses of human life and damages in urban and rural sectors. Assessing the flood risk at the river basin is not a simple task, because of the complex nature of flood generation caused by a combination of several sources such as precipitation, tidal, river basin characteristics and anthropogenic activities. However, the development of numerical flood modelling methods in recent years and the availability of advanced flood modelling and modern survey techniques for collection of high-quality input data for those models allow to simulate flood behaviour and to study the characteristic of future floods.

Flood hazard maps show the intensity of floods and their associated exceedance probability. To develop reliable flood hazard maps, a methodology, combining the advantageous features of 1D and 2D hydraulic models, bivariate flood frequency analysis and high-resolution topographic data, are typically applied. Flood peak alone cannot give a reliable evaluation of the hazard. It is also essential to consider simultaneously the flood peak along with other flood characteristics in developing flood hazard mappings. Furthermore, rapid and accurate flood modelling at high spatial-temporal resolutions remains a significant challenge in hydrologic and hydraulic studies. Therefore, it is necessary to establish an advanced deterministic approach, including bivariate flood frequency analysis, efficient and flexible hydrodynamic models and high-resolution topographic data in developing the flood hazard maps under climate change context.

The research reported in this thesis contributes towards assessing the flood risk under climate change context at the river basin scale. Initially, the potential impact of climate change on flood frequency is evaluated for the Saigon-Dongnai River basin, Vietnam. The correctness of directly using the asymptotic distribution to model the future POT dataset are also assessed in this study. In addition, the tail dependence tests for the pairs of flood characteristics are carried out to select the appropriate copula functions. The performance of extreme value copulas for asymptotic dependence variables and Clayton, Frank and Gaussian copulas for asymptotic independent variables are assessed. Finally, the last part of this thesis contributes towards developing flood hazard maps obtained using coupled hydrodynamic models, bivariate flood frequency analysis and flood hazard tools.

7.2 Conclusions

Based on this study, the following conclusions can be arrived.

The changes in Trian reservoir belonging to the Saigon-Dongnai River basin, Vietnam are analyzed using GCM/RCMs outputs from CORDEX project. The change factors for two scenarios such as RCP4.5 and 8.5 are calculated by using the outputs of five RCMs and one GCM. Then, these change factors are used as an input to LARS-WG to generate the daily precipitation and daily temperature for both scenarios. The continuous hydrological model (HEC-HMS), which is calibrated and validated using observed discharge data, is applied to simulate future discharge data based on the future climate data. The changes of flood frequency of flood peak extracted using POT approach is compared between historical and future time periods. The analysis reveals that flood magnitudes increase significantly in the future period for the study area.

Since the flood is a complex phenomenon defined by strongly correlated characteristics, univariate frequency analysis approach cannot describe accurately the random correlated flood characteristics. Based on the advantage of copula approach in multivariate frequency analysis, the copula approach is used to model the joint dependence structure of flood characteristics. The most important step in the modelling processing using copula is the selection of copula function that fits the data sample. The chosen copulas should include several classes and degrees of tail dependence. Therefore, the potential of performing the tail dependence tests for the pairs of flood characteristics in selection appropriate copula function are assessed in Chapter 5.

LLHR and tail dependence tests are used to identify the asymptotically dependence of flood characteristics. Three extreme value copulas (i.e., Gumbel-Hougaard, Galambos and Husler-Reiss) are evaluated to model asymptotically dependence of flood variables. In addition, Gaussian, Frank and Clayton copulas are considered as the appropriate copula function in case of which are diagnosed as asymptotic independence. Besides, the extreme value copulas also with upper tail dependence have proven that they are appropriate copula function for the dependence structure of flood variables.

Accurate and reliable flood risk maps are ideal tools for decision makers to reduce social

and economic losses from flood events. These maps provide useful information for organizations dealing with emergency situations to calibrate and adjust warning systems and prepare priority evacuation plans. Modelling the potential flood regime due to river flow from upstream considering the present condition and two future scenarios RCP4.5 and 8.5 are carried out. Hydrodynamic flow modelling is simulated using coupled model in the lower stream of Saigon-Dongnai River basin, Vietnam.

The high-resolution flood hazard maps are obtained in this work through three key components: (i) design flood hydrograph estimated using bivariate flood frequency analysis is used as the input the hydrodynamic model; (ii) high-quality topographic data (i.e., DEM and LiDAR), collected using modern survey, are used as an input for the hydraulic models and (iii) flexible meshes generation, which are the latest advances in flood modelling, are selected for the coupled hydrodynamic model. Two parameters, namely, flood depth and flow velocity, which are obtained from coupled hydrodynamic model, have been used for quantifying the flood hazard.

Flood hazard maps, which were assessed using both flood depth and velocity, are developed in Chapter 6. In addition, safe locations have also been identified for the industries. Besides, flood inundation duration is an important parameter for flood risk assessment also as pointed out in this study. Furthermore, flood hazard maps also provide useful information for the evaluation of transport blockades and access to emergency services.

To summarise, the major conclusions from this study are:

- The changes of flood frequency of flood peak extracted using POT approach is compared between historical and future time periods. The analysis reveals that flood magnitudes increase significantly in the future period for the study area. The results of this study also indicate that directly using the asymptotic distribution to model the POT dataset sometimes provides wrong insights.
- There is a huge difference in the joint return period estimation using the families of extreme value copulas and no upper tail copulas (i.e., Frank, Clayton and Gaussian)

if there exists asymptotic dependence in the flood characteristics. The extreme value copulas with upper tail dependence have proved that they are appropriate models for the dependence structure of the flood characteristics and Frank, Clayton and Gaussian copulas are the appropriate copula models in case of variables which are diagnosed as asymptotic independence.

- The high-resolution flood hazard maps are obtained in this work using both flood depth and velocity, which will help in identifying the potential flood hazard regions as well as safe locations for the setting up industries and planning for social and economic development in the river basin scale.

7.3 Scope for Future Studies

The work presented in this thesis could be further extended if relevant data were available and time is not a constraint.

- Assessing the changes in flood frequency due to the changes of river basin characteristics changes or land use and land cover
- Analysis of the bivariate flood frequency of nonstationary flood characteristics
- Analysis of regional flood frequency analysis of nonstationary flood characteristics
- Quantifying the uncertainty of the flood hazard maps
- Development of flood hazard management system, including a flood evacuation strategy.

APPENDIX A

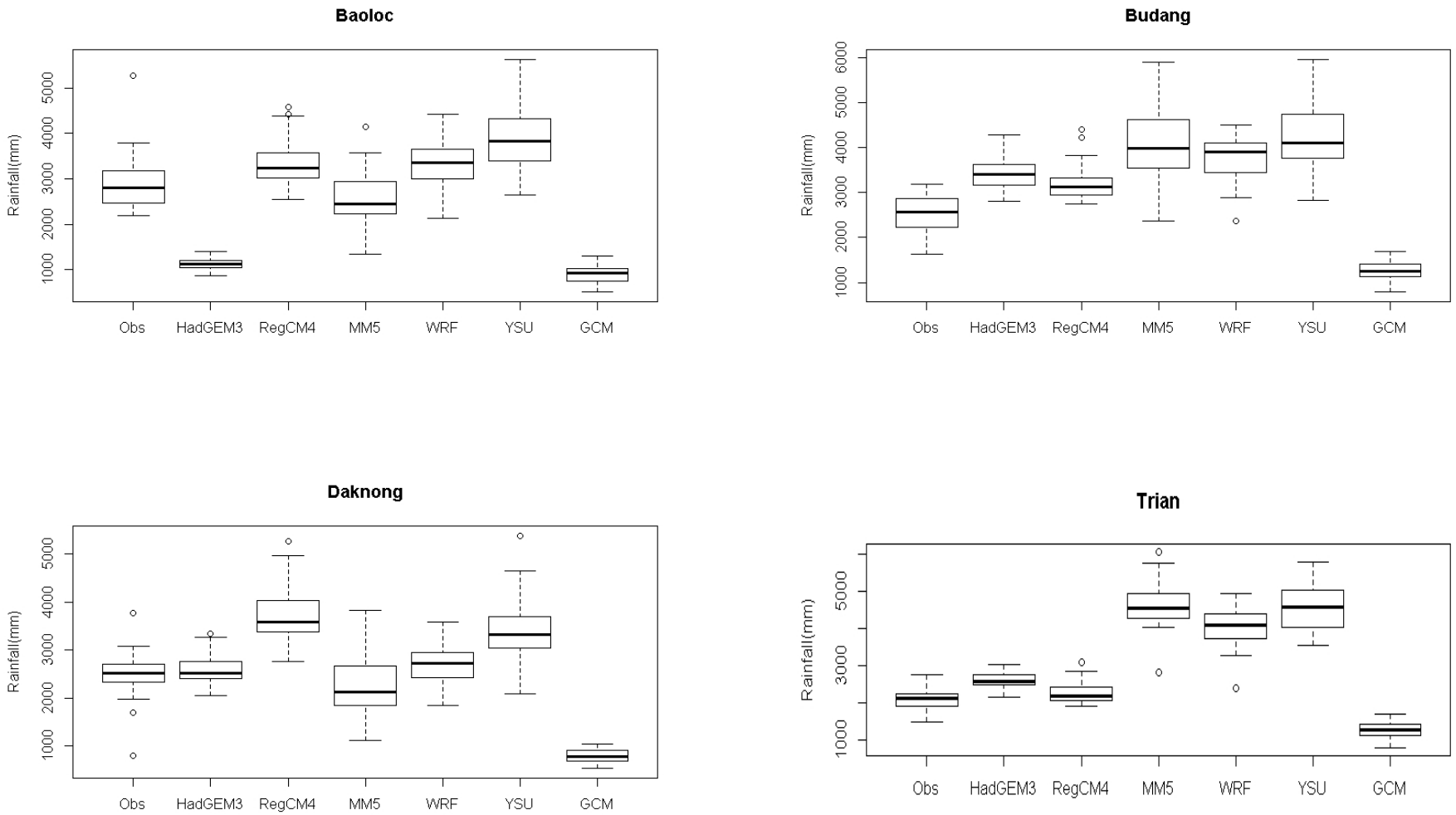


Fig. A.1:Box plots of observed, RCMs and GCM simulated annual precipitation during baseline period (1980-2005)

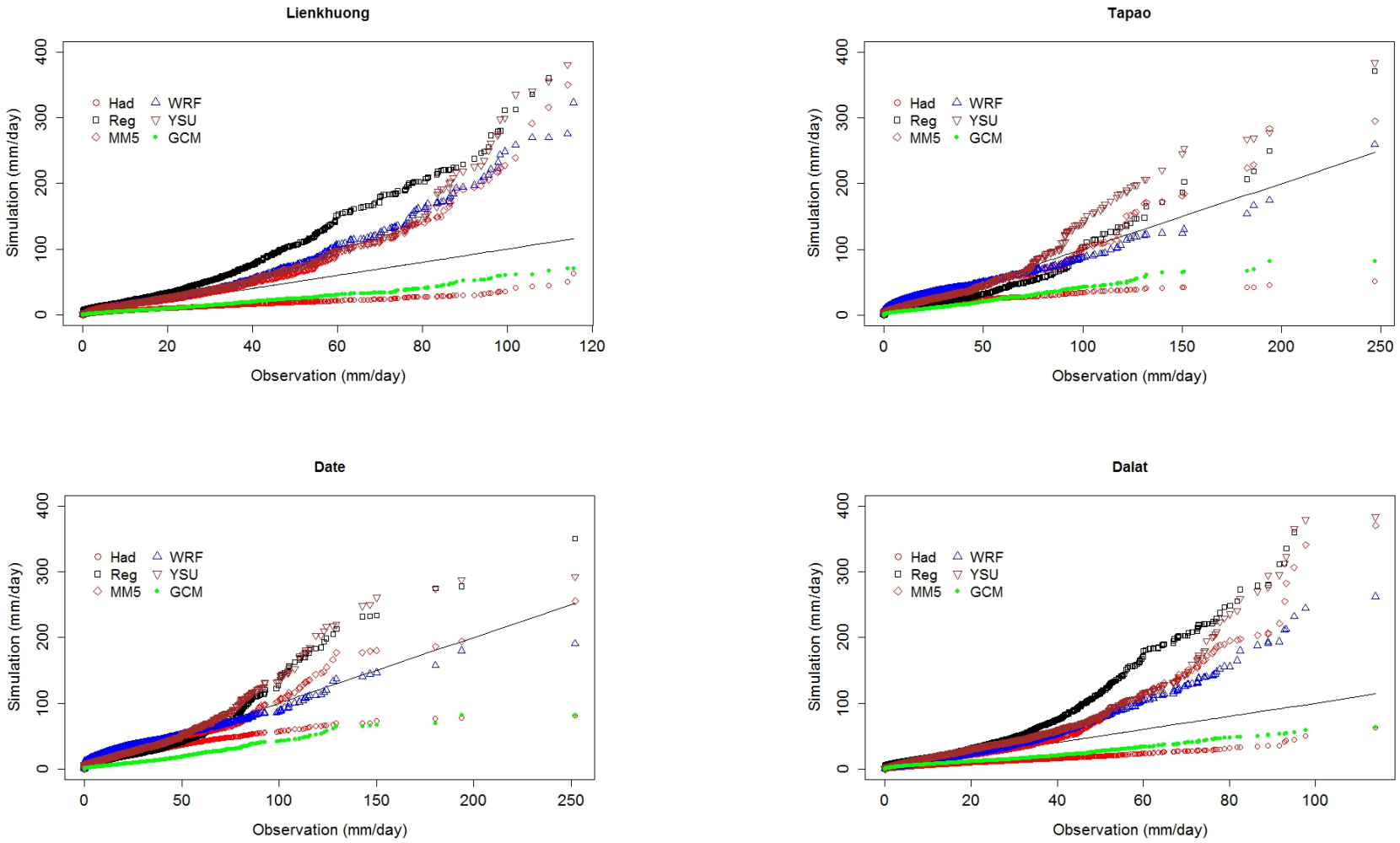


Fig. A.2: Q-Q plots between observed and raw RCMs & GCM simulated daily precipitation

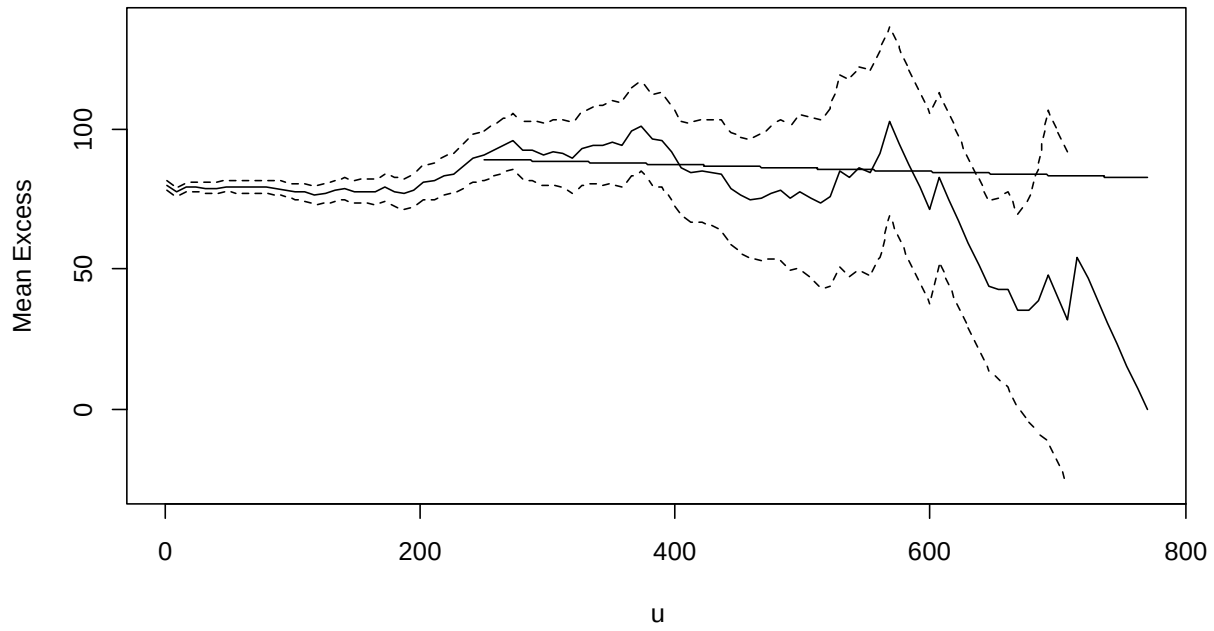


Fig. A.3: Mean residual life plot for daily flood data in Tapao

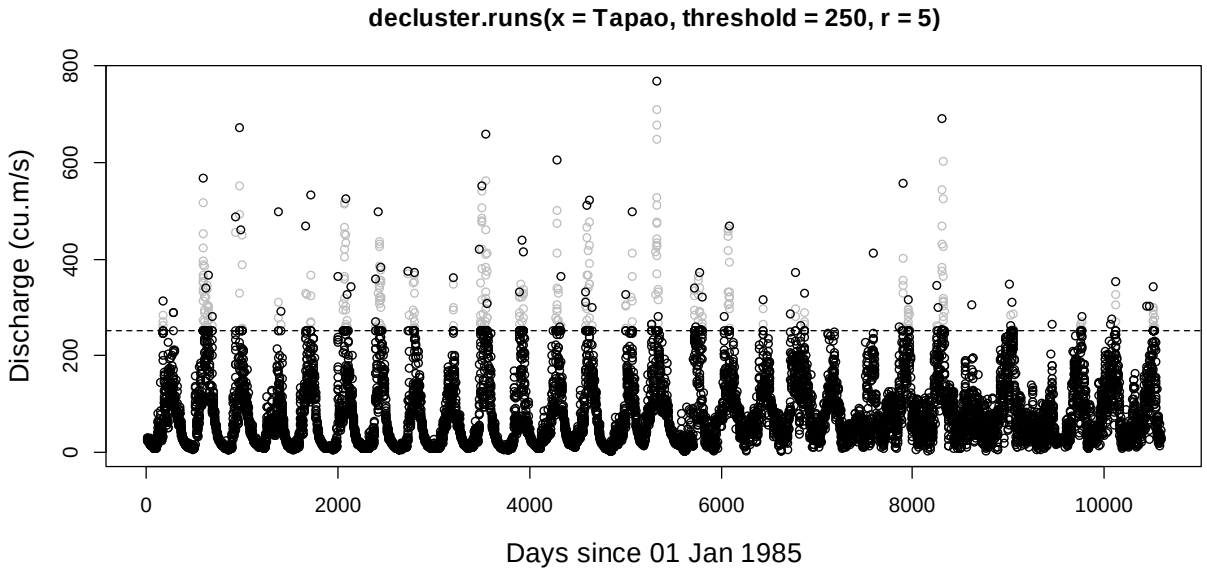


Fig. A.4: POT series of Tapao station based on the threshold value of 250 m^3/s and after declustering with independence criteria of 5 days

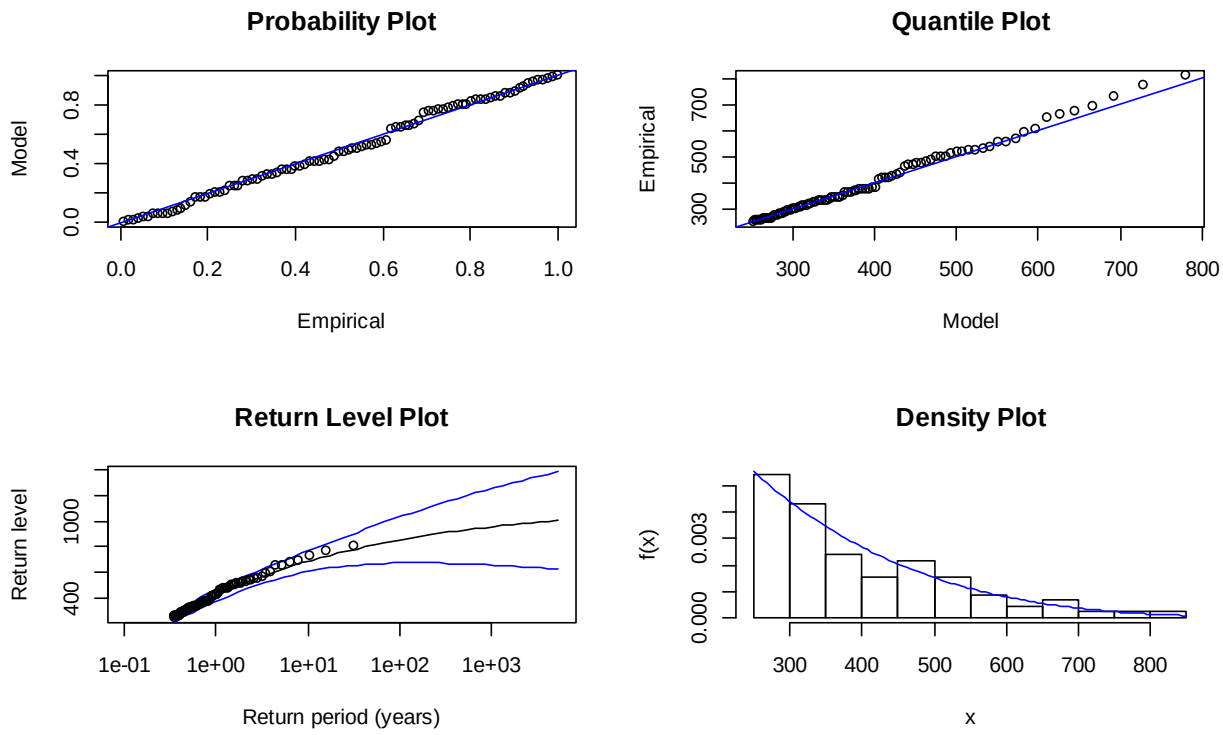


Fig. A.5: Diagnostic plots for threshold (250 m^3/s) excesses of Tapao observed discharge fitted with GPD distribution

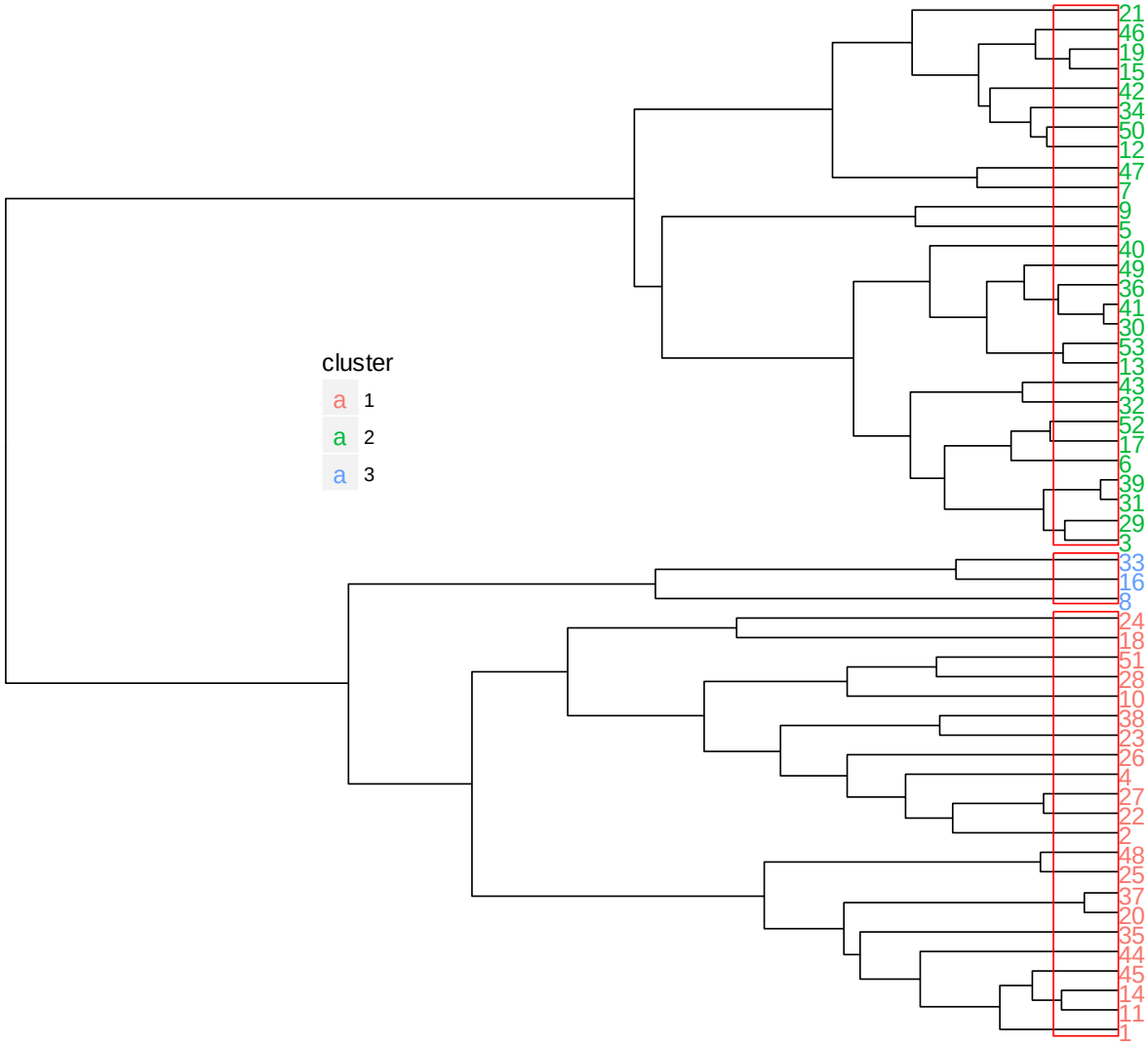


Fig. A.6: Dendrogram of the cluster analysis of hydrographs

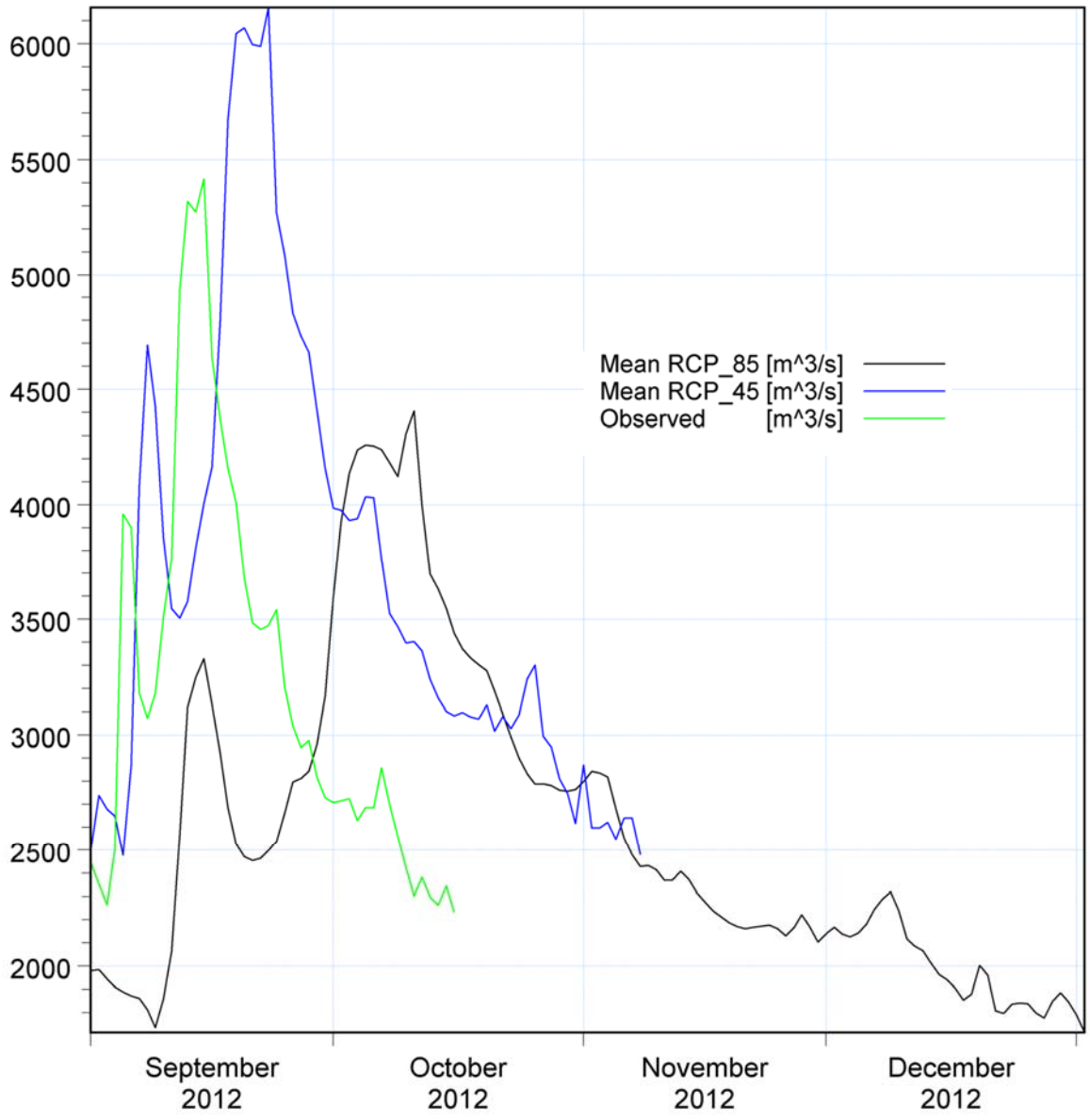


Fig. A.7: Design flood hydrographs for 100-year return period for present and future scenarios

Table A.1: Results of AD test for POT datasets of future (2020-2045) scenario

Scenarios	Talai						Tapao						Trian					
	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3
	Anderson-Darling																	
GCM_45	0.46	2.23	2.10	2.15	0.51	1.07	0.66	6.22	6.04	6.93	1.23	0.77	0.27	1.19	1.07	0.78	0.63	1.15
GCM_85	0.44	2.66	2.58	2.90	0.52	1.08	0.51	1.82	1.80	2.23	0.58	1.12	1.32	2.87	2.69	2.26	0.72	1.44
Had_45	0.22	0.81	0.79	0.97	0.73	0.62	4.80	1.16	1.15	1.09	1.61	1.65	0.99	2.10	2.04	1.93	1.01	1.55
Had_85	0.38	1.78	1.74	2.00	0.52	0.64	0.19	2.39	2.38	3.17	0.33	0.73	1.40	1.59	1.54	1.33	1.01	1.64
Reg_45	0.40	1.86	1.82	2.12	0.68	1.39	0.25	2.89	2.88	3.77	0.43	0.44	1.86	2.25	2.16	2.10	0.80	1.27
Reg_85	0.29	1.39	1.38	1.64	0.73	1.45	0.28	2.29	2.24	3.60	0.40	0.56	1.24	1.30	1.29	1.24	1.20	2.10
MM5_45	0.72	3.08	3.03	3.15	0.63	0.79	0.45	2.25	2.23	2.63	0.54	1.31	1.24	2.80	2.69	2.48	0.87	0.99
WRF_45	0.52	2.80	2.75	2.69	0.71	0.70	0.63	2.45	2.45	3.13	1.23	1.05	0.49	3.44	3.39	3.62	0.92	0.59
WRF_85	0.35	1.07	1.05	1.67	0.68	1.32	1.50	2.78	2.78	3.24	0.58	1.60	0.97	2.51	2.43	2.24	0.84	1.08
YSU_45	0.51	2.16	2.14	2.86	0.57	0.58	0.69	3.96	3.91	4.24	0.62	0.91	1.13	2.58	2.52	2.29	1.56	1.95
YSU_85	0.41	1.47	1.46	2.27	0.57	0.73	0.69	3.13	3.13	3.46	0.70	1.01	0.20	1.07	1.05	1.38	0.50	0.98

Table A.2: Results of KS test for POT datasets of future (2020-2045) scenario

Scenarios	Talai						Tapao						Trian					
	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3
	Kolmogorov-Smirnov																	
GCM_45	0.07	0.09	0.09	0.09	0.05	0.09	0.07	0.16	0.16	0.19	0.10	0.07	0.06	0.09	0.09	0.08	0.08	0.12
GCM_85	0.06	0.11	0.11	0.11	0.05	0.07	0.06	0.11	0.11	0.12	0.06	0.08	0.09	0.12	0.11	0.11	0.08	0.10
Had_45	0.05	0.07	0.07	0.09	0.06	0.07	0.32	0.22	0.22	0.22	0.20	0.20	0.09	0.11	0.11	0.11	0.10	0.11
Had_85	0.06	0.13	0.13	0.11	0.09	0.08	0.05	0.11	0.11	0.11	0.05	0.08	0.13	0.15	0.15	0.14	0.12	0.15
Reg_45	0.05	0.09	0.09	0.11	0.06	0.09	0.04	0.10	0.10	0.13	0.05	0.04	0.10	0.11	0.11	0.11	0.08	0.08
Reg_85	0.04	0.09	0.09	0.11	0.07	0.09	0.05	0.09	0.09	0.12	0.05	0.06	0.11	0.10	0.10	0.11	0.09	0.12
MM5_45	0.09	0.14	0.14	0.14	0.07	0.06	0.05	0.09	0.08	0.10	0.05	0.08	0.13	0.16	0.16	0.17	0.09	0.09
WRF_45	0.09	0.16	0.16	0.16	0.09	0.09	0.11	0.18	0.18	0.17	0.14	0.13	0.07	0.14	0.14	0.16	0.09	0.08
WRF_85	0.06	0.07	0.07	0.09	0.06	0.08	0.08	0.10	0.10	0.12	0.05	0.08	0.11	0.16	0.16	0.16	0.08	0.09
YSU_45	0.06	0.12	0.12	0.12	0.08	0.07	0.08	0.15	0.15	0.15	0.08	0.08	0.10	0.15	0.15	0.16	0.12	0.15
YSU_85	0.07	0.10	0.10	0.14	0.08	0.08	0.09	0.13	0.13	0.12	0.09	0.08	0.04	0.09	0.09	0.10	0.06	0.08

Table A.3: Results of CVM test for POT datasets of future (2020-2045) scenario

Scenarios	Talai						Tapao						Trian					
	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3
	Cramer Von Mises																	
GCM_45	0.07	0.26	0.26	0.31	0.05	0.13	0.12	1.01	1.00	1.25	0.23	0.12	0.04	0.14	0.13	0.10	0.08	0.20
GCM_85	0.07	0.39	0.38	0.46	0.04	0.11	0.07	0.26	0.26	0.33	0.07	0.15	0.21	0.44	0.42	0.36	0.08	0.18
Had_45	0.03	0.12	0.12	0.15	0.07	0.10	0.44	0.19	0.19	0.18	0.13	0.13	0.17	0.34	0.33	0.31	0.15	0.23
Had_85	0.05	0.24	0.24	0.28	0.07	0.09	0.03	0.31	0.31	0.47	0.04	0.10	0.19	0.25	0.25	0.21	0.13	0.23
Reg_45	0.06	0.26	0.26	0.33	0.06	0.17	0.04	0.37	0.37	0.59	0.05	0.04	0.26	0.32	0.31	0.33	0.06	0.12
Reg_85	0.04	0.21	0.21	0.24	0.09	0.20	0.04	0.27	0.27	0.52	0.04	0.07	0.13	0.17	0.17	0.16	0.16	0.31
MM5_45	0.12	0.44	0.44	0.49	0.08	0.10	0.05	0.31	0.31	0.40	0.06	0.19	0.22	0.43	0.42	0.40	0.12	0.11
WRF_45	0.09	0.46	0.45	0.47	0.10	0.07	0.11	0.31	0.31	0.46	0.15	0.13	0.07	0.54	0.54	0.64	0.16	0.07
WRF_85	0.06	0.15	0.15	0.27	0.07	0.16	0.10	0.38	0.38	0.49	0.07	0.23	0.18	0.40	0.39	0.37	0.12	0.13
YSU_45	0.06	0.32	0.32	0.44	0.08	0.09	0.13	0.61	0.60	0.67	0.11	0.15	0.19	0.44	0.43	0.38	0.25	0.30
YSU_85	0.07	0.18	0.18	0.35	0.07	0.09	0.12	0.45	0.45	0.54	0.11	0.16	0.03	0.15	0.14	0.20	0.04	0.11

Table A.4: Results of AIC for POT datasets of future (2020-2045) scenario

Scenarios	Talai						Tapao						Trian					
	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3	GPD	GL	Gumbel	LN	P3	LP3
	Akaike Information Criterion																	
GCM_45	1,361	1,404	1,399	1,392	1,361	1,381	1,404	1,481	1,477	1,500	1,409	1,408	1,073	1,073	1,074	1,072	1,074	1,084
GCM_85	1,422	1,479	1,474	1,469	1,422	1,430	1,239	1,280	1,278	1,290	1,244	1,251	1,098	1,098	1,100	1,097	1,100	1,106
Had_45	1,292	1,317	1,314	1,313	1,292	1,304	207	218	216	217	211	213	1,105	1,104	1,106	1,102	1,105	1,113
Had_85	1,132	1,173	1,170	1,164	1,136	1,139	1,579	1,628	1,626	1,642	1,582	1,587	847	842	843	844	846	821
Reg_45	1,253	1,295	1,291	1,289	1,257	1,268	1,401	1,449	1,447	1,461	1,404	1,406	812	813	813	813	813	823
Reg_85	1,333	1,361	1,358	1,357	1,338	1,354	1,388	1,435	1,432	1,451	1,388	1,389	1,007	1,008	1,009	1,009	1,009	1,039
MM5_45	1,087	1,118	1,114	1,111	1,093	1,104	1,592	1,642	1,640	1,652	1,595	1,605	763	764	764	764	763	768
WRF_45	1,459	1,508	1,504	1,504	1,461	1,464	586	628	626	637	592	589	699	696	692	690	692	710
WRF_85	1,367	1,415	1,412	1,410	1,363	1,374	1,622	1,688	1,686	1,698	1,619	1,631	893	894	895	894	895	912
YSU_45	1,222	1,253	1,250	1,249	1,226	1,236	1,605	1,669	1,666	1,681	1,604	1,608	956	956	958	954	957	981
YSU_85	1,583	1,633	1,629	1,634	1,582	1,590	1,412	1,469	1,467	1,480	1,415	1,419	1,043	1,040	1,042	1,041	1,041	1,062

Table A.5: AIC values for all marginal distributions

Data	Flood variable	AIC						Data	Flood variable	AIC					
		LN	Gumbel	GEV	P3	LP3	GPD			LN	Gumbel	GEV	P3	LP3	GPD
	V	3,188	3,197	3,201	3,176	3,164			V	2,056	2,077	2,071	2,053	2,046	
Obs	P	1,066	1,066	1,068	1,061	1,067	1,068	**WRF_45**	P	699.1	695.6	692.2	690	692.3	709.9
	D	654.9	656	658	641.7	628.9			D	356.4	373.2	349.3	349.7	348.2	
	V	3,363	3,402	3,373	3,349	3,348			V	2,791	2,816	2,808	2,782	2,791	
Had_45	P	1,105	1,104	1,106	1,102	1,105	1,113	**WRF_85**	P	892.7	893.8	894.6	894	894.6	912.0
	D	659.8	692	643.1	645.2	647.5			D	607.4	626.6	611	607.8	608.5	
	V	2,515	2,533	2,534	2,514	2,501			V	2,979	3,005	2,995	2,960	2,980	
Had_85	P	846.8	842.4	843.3	844.1	845.5	821.4	**YSU_45**	P	956.4	955.9	957.9	954	957.4	980.6
	D	546.8	563.5	547.2	538.6	537.2			D	612.8	636	612.4	613.5	612.8	
	V	2,547	2,565	2,565	2540	2543			V	3,257	3,294	3,276	3,261	3,251	
Reg_45	P	811.5	812.9	812.7	813	812.9	823.0	**YSU_85**	P	1,043	1,040	1,042	1,041	1,041	1,062
	D	546.5	561.5	551.8	542.8	542.2			D	633.4	666.5	625.5	636.3	625.9	
	V	3,159	3,188	3,174	3,148	3,154			V	3,068	3,101	3,087	3,057	3,066	
Reg_85	P	1,007	1,008	1,009	1,009	1,009	1,039	**GCM_45**	P	1,073	1,073	1,074	1,072	1,074	1,084
	D	642.1	666.5	639.8	640.7	642.8			D	642.3	666.5	643.2	643.5	643.7	
	V	2,340	2,353	2,357	2,332	2,328			V	3158	3202	3175	3154	3148	
MM5_45	P	762.6	763.8	763.5	763.8	763.2	768.1	**GCM_85**	P	1,098	1,098	1,100	1,097	1,100	1,106
	D	519.5	528.7	526.5	513.4	508.3			D	628.6	667	619.9	629	619.4	

Table A.6: Copula dependence parameters, AIC and GoF statistics for both tail independence and dependence

Copula	DV				Copula	DP				PV			
	Parameter	AIC	S	p-value		Parameter	AIC	S	p-value	Parameter	AIC	S	p-value
					Observed								
Gumbel-Hougaard	6.007	-165.013	0.00579	0.003	Gaussian	0.785	-57.509	0.114	0.065	0.835	-73.575	0.119	0.050
Galambos	5.268	-162.401	0.00583	0.002	Survival Clayton	1.774	-47.817	0.504	0.000	2.066	-55.250	0.477	0.0002
Husler-Reiss	4.377	-137.984	0.00784	0.007	Frank	8.455	-67.695	0.063	0.285	10.396	-86.929	0.058	0.335
					GCM_45								
Gumbel-Hougaard	10.246	-214.737	0.00029	0.082	Gaussian	0.770	-50.723	0.113	0.100	0.828	-49.339	0.125	0.050
Galambos	9.550	-214.675	0.00029	0.067	Survival Clayton	1.572	-39.228	0.148	0.120	2.457	-44.756	0.430	0.000
Husler-Reiss	10.914	-212.448	0.00031	0.102	Frank	6.368	-45.719	0.131	0.040	10.299	-58.872	0.052	0.535
					MM5_45								
Gumbel-Hougaard	24.578	-236.231	0.00006	0.301	Gaussian	0.791	-41.126	0.146	0.035	0.842	-72.335	0.121	0.020
Galambos	23.869	-236.175	0.00007	0.266	Survival Clayton	2.179	-39.994	0.478	0.000	2.387	-62.494	0.134	0.110
Husler-Reiss	25.882	-233.192	0.00011	0.182	Frank	8.969	-49.683	0.071	0.290	7.990	-62.159	0.171	0.005

Table A.7: Copula dependence parameters, AIC and GoF statistics for tail dependence

Copula	DV				DP				PV			
	Parameter	AIC	S	p-value	Parameter	AIC	S	p-value	Parameter	AIC	S	p-value
						Had_45						
Gumbel-Hougaard	11.443	-284.359	0.000127	0.077	1.906	-43.244	0.004544	0.435	2.124	-55.486	0.0033	0.460
Galambos	10.741	-284.327	0.000126	0.097	1.202	-44.276	0.002987	0.624	1.420	-56.376	0.002684	0.520
Husler-Reiss	12.288	-282.523	0.000131	0.147	1.749	-45.605	0.001439	0.813	1.987	-57.567	0.002191	0.475
						Had_85						
Gumbel-Hougaard	20.001	-234.074	0.000253	0.077	3.929	-70.249	0.011193	0.057	4.429	-80.635	0.011463	0.017
Galambos	19.288	-233.980	0.000254	0.117	3.244	-69.799	0.011684	0.027	3.744	-80.073	0.012062	0.017
Husler-Reiss	20.919	-230.559	0.000356	0.062	3.663	-64.134	0.014125	0.027	4.048	-72.450	0.015373	0.007
						GCM_85						
Gumbel-Hougaard	9.644	-221.748	0.000114	0.206	2.271	-45.373	0.003497	0.545	2.773	-67.963	0.000373	0.983
Galambos	8.944	-221.557	0.000114	0.291	1.576	-45.493	0.002865	0.580	2.081	-68.404	0.000242	0.988
Husler-Reiss	10.011	-217.262	0.000136	0.162	2.148	-45.205	0.00252	0.530	2.759	-68.993	0.000285	0.978
						WRF_45						
Gumbel-Hougaard	12.430	-157.319	0.000264	0.127	2.641	-34.075	0.013971	0.221	3.117	-45.893	0.006588	0.251
Galambos	11.718	-157.080	0.000266	0.132	1.939	-33.860	0.013231	0.122	2.421	-45.858	0.005407	0.281
Husler-Reiss	12.228	-151.082	0.000314	0.137	2.527	-33.234	0.014539	0.167	3.128	-45.691	0.004445	0.361
						WRF_85						
Gumbel-Hougaard	16.276	-235.865	0.001539	0.007	2.755	-48.428	0.019469	0.047	3.106	-59.932	0.018623	0.022
Galambos	15.606	-235.402	0.001548	0.007	2.057	-48.388	0.019887	0.042	2.412	-59.965	0.018891	0.022
Husler-Reiss	11.831	-197.939	0.00269	0.002	2.686	-47.984	0.020904	0.052	3.089	-59.470	0.019738	0.007
						YSU_45						
Gumbel-Hougaard	20.02	-294.98	5.20E-05	0.147	2.629	-52.902	0.001056	0.918	2.822	-60.029	0.000767	0.928
Galambos	19.31	-294.91	5.25E-05	0.226	1.935	-53.433	0.001214	0.903	2.126	-60.480	0.000875	0.878
Husler-Reiss	21.37	-292.45	9.21E-05	0.172	2.628	-54.603	0.001777	0.689	2.845	-61.585	0.001343	0.729

YSU_85												
Gumbel-Hougaard	11.956	-267.953	6.26E-05	0.236	2.171	-48.540	0.000705	0.953	2.416	-61.211	0.002446	0.540
Galambos	11.248	-267.525	6.47E-05	0.261	1.467	-48.864	0.000742	0.923	1.713	-61.711	0.002492	0.450
Husler-Reiss	11.584	-256.973	1.38E-04	0.137	2.027	-49.155	0.001341	0.759	2.322	-62.460	0.002704	0.391
Reg_45												
Gumbel-Hougaard	16.449	-216.406	0.00021	0.1269	3.8116	-72.2819	0.009135	0.177	3.976	-77.196	0.01018	0.117
Galambos	15.750	-216.208	0.000213	0.1418	3.1200	-72.6325	0.008907	0.117	3.287	-77.654	0.00973	0.087
Husler-Reiss	15.967	-206.714	0.000319	0.1020	4.0214	-73.8075	0.009631	0.097	4.257	-79.391	0.009921	0.087
Reg_85												
Gumbel-Hougaard	14.896	-279.101	0.000561	0.0075	2.4645	-54.3523	0.001851	0.754	2.704	-64.716	0.005309	0.231
Galambos	14.185	-278.887	0.000569	0.0323	1.7691	-54.8656	0.002046	0.639	2.007	-65.149	0.005423	0.201
Husler-Reiss	15.209	-272.913	0.001136	0.0075	2.4069	-55.6898	0.002745	0.495	2.667	-65.769	0.005865	0.182

References

Abbaspour, Rouholahnejad, Vaghefi, Srinivasan, Yang & Kløve 2015. A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology,* 524**,** 733-752.

Abrahart & See 2007. Neural network modelling of non-linear hydrological relationships. *Hydrol. Earth Syst. Sci.,* 11**,** 1563-1579.

Adamson, Metcalfe & Parmentier 1999. Bivariate extreme value distributions: an application of the Gibbs Sampler to the analysis of floods. *Water Resources Research,* 35**,** 2825-2832.

ADB. 2010. *(Asian Development Bank) Ho Chi Minh City Adaptation to Climate Change: Summary Report* [Online]. Available: https://www.adb.org/publications/ho-chi-minh-city-adaptation-climate-change-summary-report [Accessed 27 Dec 2018].

Aemi 2014. Technical flood risk management guideline: Flood hazard, Australian Emergency Management Institute, Australian Government Attorney General's Department, Melbourne, Australia.

Agarwal, Babel & Maskey 2014. Analysis of future precipitation in the Koshi river basin, Nepal. *Journal of Hydrology,* 513**,** 422-434.

Aghakouchak, Ciach & Habib 2010. Estimation of tail dependence coefficient in rainfall accumulation fields. *Advances in Water Resources,* 33**,** 1142-1149.

Aissia, Chebana, Ouarda, Roy, Desrochers, Chartier & Robichaud 2012. Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada. *Hydrological Processes,* 26**,** 130-142.

Alfieri, Burek, Feyen & Forzieri 2015. Global warming increases the frequency of river floods in Europe. *Hydrol. Earth Syst. Sci.,* 19**,** 2247-2260.

Alfieri, Salamon, Bianchi, Neal, Bates & Feyen 2014. Advances in pan-European flood hazard mapping. *Hydrological Processes,* 28**,** 4067-4077.

Ali 2018. *Flood Inundation Modeling and Hazard Mapping under Uncertainty in the Sungai Johor Basin, Malaysia*, CRC Press.

Alkema 2007. *Simulating floods: On the application of a 2D-hydraulic model for flood hazard and risk assessment*, Utrecht University.

Anandhi, Frei, Pierson, Schneiderman, Zion, Lounsbury & Matonse 2011. Examination of change factor methodologies for climate change impact assessment. *Water Resources Research,* 47.

Andréasson, Bergström, Carlsson, Graham & Lindström 2004. Hydrological Change – Climate Change Impact Simulations for Sweden. *AMBIO: A Journal of the Human Environment,* 33**,** 228-234.

Angela & Giuseppe 2017. Probabilistic Flood Hazard Mapping Using Bivariate Analysis Based on Copulas. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering,* 3.

Angeles, Gonzalez, Erickson & Hernández 2007. Predictions of future climate change in the Caribbean region using global general circulation models. *International Journal of Climatology,* 27**,** 555-569.

Arnell 1999. The effect of climate change on hydrological regimes in Europe: a continental perspective. *Global environmental change,* 9**,** 5-23.

Arnell & Gosling 2016. The impacts of climate change on river flood risk at the global scale. *Climatic Change,* 134**,** 387-401.

Arnell & Gosling 2013. The impacts of climate change on river flow regimes at the global scale. *Journal of Hydrology,* 486**,** 351-364.

Aronica, Candela, Fabio & Santoro 2012a. Estimation of flood inundation probabilities using global hazard indexes based on hydrodynamic variables. *Physics and Chemistry of the Earth, Parts A/B/C,* 42**,** 119-129.

Aronica, F., D. & C. 2012b. Probabilistic evaluation of flood hazard in urban areas using Monte Carlo simulation. *Hydrological Processes,* 26**,** 3962-3972.

Babel, Nguyen Dinh, Mullick & Nanduri 2012. Operation of a hydropower system considering environmental flow requirements: A case study in La Nga river basin, Vietnam. *Journal of Hydro-environment Research,* 6**,** 63-73.

Bacro, Bel & Lantuéjoul 2010. Testing the independence of maxima: from bivariate vectors to spatial extreme fields. *Extremes,* 13**,** 155-175.

Badrzadeh, Sarukkalige & Jayawardena 2015. Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. *Journal of Hydrology,* 529**,** 1633-1643.

Bárdossy & Pegram 2011. Downscaling precipitation using regional climate models and circulation patterns toward hydrology. *Water Resources Research,* 47.

Beguería 2005. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology,* 303**,** 215-230.

Benjankar, Tonina & Mckean 2015. One-dimensional and two-dimensional hydrodynamic modeling derived flow properties: impacts on aquatic habitat quality predictions. *Earth Surface Processes and Landforms,* 40**,** 340-356.

Bergstrom 1976. Development and application of a conceptual runoff model for Scandinavian catchments.

Berrueta, Alonso-Salces & Héberger 2007. Supervised pattern recognition in food analysis. *Journal of Chromatography A,* 1158**,** 196-214.

Bezak, Brilly & Šraj 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal,* 59**,** 959-977.

Bonnin, Martin, Lin, Parzybok, Yekta & Riley 2006. Precipitation-frequency atlas of the United States. *NOAA atlas,* 2.

Booij 2005. Impact of climate change on river flooding assessed with different spatial model resolutions. *Journal of Hydrology,* 303**,** 176-198.

Bortot, Coles & Tawn 2000. The multivariate Gaussian tail model: an application to oceanographic data. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 49**,** 31-049.

Boughton & Droop 2003. Continuous simulation for design flood estimation—a review. *Environmental Modelling & Software,* 18**,** 309-318.

Brunner, Sikorska & Seibert 2017. Bivariate analysis of floods in climate impact assessments. *Science of The Total Environment*.

Cameron 2006. An application of the UKCIP02 climate change scenarios to flood estimation by continuous simulation for a gauged catchment in the northeast of Scotland, UK (with uncertainty). *Journal of Hydrology,* 328**,** 212-226.

Camici, Brocca, Melone & Moramarco 2014. Impact of Climate Change on Flood Frequency Using Different Climate Models and Downscaling Approaches. *Journal of Hydrologic Engineering,* 19**,** 04014002.

Candela, Brigandì & Aronica 2014. Estimation of synthetic flood design hydrographs using a distributed rainfall–runoff model coupled with a copula-based single storm rainfall generator. *Nat. Hazards Earth Syst. Sci.,* 14**,** 1819-1833.

Capéraà, Fougères & Genest 1997. A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika***,** 567-577.

Castellarin, Kohnová, Gaál, Fleig, Salinas, Toumazis, Kjeldsen & Macdonald 2012. Review of applied-statistical methods for flood-frequency analysis in Europe.

Chatterjee, Förster & Bronstert 2008. Comparison of hydrodynamic models of different complexities to model floods with emergency storage areas. *Hydrological Processes,* 22**,** 4695-4709.

Chebana & Ouarda 2011. Multivariate quantiles in hydrological frequency analysis. *Environmetrics,* 22**,** 63-78.

Chen, Brissette & Leconte 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *Journal of Hydrology,* 401**,** 190-202.

Cherubini, Luciano & Vecchiato 2004. *Copula methods in finance*, John Wiley & Sons.

Chowdhary, Escobar & Singh 2011. Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrology Research,* 42**,** 193-216.

Coles, Bawa, Trenner & Dorazio 2001. *An introduction to statistical modeling of extreme values*, Springer.

Coles, Heffernan & Tawn 1999. Dependence measures for extreme value analyses. *Extremes,* 2**,** 339-365.

Crawford & Linsley 1966. Digital Simulation in Hydrology'Stanford Watershed Model 4.

Dariane, Javadianzadeh & James 2016. Developing an Efficient Auto-Calibration Algorithm for HEC-HMS Program. *Water Resources Management,* 30**,** 1923-1937.

Davison & Smith 1990. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)***,** 393-442.

Devia, Ganasri & Dwarakish 2015. A Review on Hydrological Models. *Aquatic Procedia,* 4**,** 1001-1007.

Di Baldassarre, Castellarin, Montanari & Brath 2009. Probability-weighted hazard maps for comparing different flood risk management strategies: a case study. *Natural Hazards,* 50**,** 479-496.

Di Baldassarre, Schumann, Bates, Freer & Beven 2010. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. *Hydrological Sciences Journal,* 55**,** 364-376.

Dobler, Bürger & Stötter 2012. Assessment of climate change impacts on flood hazard potential in the Alpine Lech watershed. *Journal of Hydrology,* 460-461**,** 29-39.

Dottori, Di Baldassarre & Todini 2013. Detailed data is welcome, but with a pinch of salt: Accuracy, precision, and uncertainty in flood inundation modeling. *Water Resources Research,* 49**,** 6079-6085.

Duan & Mei 2014. A comparison study of three statistical downscaling methods and their model-averaging ensemble for precipitation downscaling in China. *Theoretical and applied climatology,* 116**,** 707-719.

Duan, Mei & Zhang 2016. Copula-based bivariate flood frequency analysis in a changing climate—A case study in the Huai River Basin, China. *Journal of Earth Science,* 27**,** 37-46.

Dung, Merz, Bárdossy & Apel 2015. Handling uncertainty in bivariate quantile estimation – An application to flood hazard analysis in the Mekong Delta. *Journal of Hydrology,* 527**,** 704-717.

Dupuis 2007. Using Copulas in Hydrology: Benefits, Cautions, and Issues. *Journal of Hydrologic Engineering,* 12**,** 381-393.

El Adlouni, Bobée & Ouarda 2008. On the tails of extreme event distributions in hydrology. *Journal of Hydrology,* 355**,** 16-33.

Escalante-Sandoval 2007. Application of bivariate extreme value distribution to flood frequency analysis: a case study of Northwestern Mexico. *Natural hazards,* 42**,** 37.

Falk & Michel 2006. Testing for Tail Independence in Extreme Value models. *Annals of the Institute of Statistical Mathematics,* 58**,** 261-290.

Fatichi, Ivanov & Caporali 2011. Simulation of future climate scenarios with a weather generator. *Advances in Water Resources,* 34**,** 448-467.

Favre, Adlouni, Perreault, Thiémonge & Bobée 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research,* 40**,** W01101.

Feldman & Center 2000. *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*, US Army Corps of Engineers, Hydrologic Engineering Center.

Filipova, Lawrence & Klempe 2018. Effect of catchment properties and flood generation regime on copula selection for bivariate flood frequency analysis. *Acta Geophysica*.

Fowler, Blenkinsop & Tebaldi 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology,* 27**,** 1547-1578.

Frahm, Junker & Schmidt 2005. Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance: Mathematics and Economics,* 37**,** 80-100.

Frana. 2012. *Applicability of MIKE SHE to simulate hydrology in heavily tile drained agricultural land and effects of drainage characteristics on hydrology.* Iowa State University.

Frick, Kaufmann & Reiss 2007. Testing the tail–dependence based on the radial component. *Extremes,* 10**,** 109-128.

Fu & Butler 2014. Copula-based frequency analysis of overflow and flooding in urban drainage systems. *Journal of Hydrology,* 510**,** 49-58.

Gaál, Szolgay, Kohnová, Hlavčová, Parajka, Viglione, Merz & Blöschl 2015. Dependence between flood peaks and volumes: a case study on climate and hydrological controls. *Hydrological Sciences Journal,* 60**,** 968-984.

Gain & Hoque 2013. Flood risk assessment and its application in the eastern part of Dhaka City, Bangladesh. *Journal of Flood Risk Management,* 6**,** 219-228.

Garrote, Alvarenga & Díez-Herrero 2016. Quantification of flash flood economic risk using ultra-detailed stage–damage functions and 2-D hydraulic models. *Journal of Hydrology,* 541**,** 611-625.

Genest & Favre 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering,* 12**,** 347-368.

Genest, Favre, Béliveau & Jacques 2007. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research,* 43.

Genest, Ghoudi & Rivest 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika***,** 543-552.

Genest, Kojadinovic, Nešlehová & Yan 2011. A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli,* 17**,** 253-275.

Genest, Rémillard & Beaudoin 2009. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics,* 44**,** 199-213.

Gichamo, Popescu, Jonoski & Solomatine 2012. River cross-section extraction from the ASTER global DEM for flood modeling. *Environmental Modelling & Software,* 31**,** 37-46.

Gilleland & Katz 2016. extRemes 2.0: an extreme value analysis package in R. *Journal of Statistical Software,* 72**,** 1-39.

Golmohammadi, Prasher, Madani & Rudra 2014. Evaluating three hydrological distributed watershed models: MIKE-SHE, APEX, SWAT. *Hydrology,* 1**,** 20-39.

Government. 2016. *The operation of this multipurpose dam system in the Saigon-Dongnai River basin* [Online]. Vietnam. Available: http://www.chinhphu.vn/portal/page/portal/chinhphu/hethongvanban?mode=detail&document_id=184010 [Accessed 27 Dec 2018].

Grimaldi, Petroselli, Arcangeletti & Nardi 2013. Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. *Journal of Hydrology,* 487**,** 39-47.

Gu, Wang, Yu & Mei 2012. Assessing future climate changes and extreme indicators in east and south Asia using the RegCM4 regional climate model. *Climatic Change,* 114**,** 301-317.

Gudendorf & Segers 2011. Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *Journal of multivariate analysis,* 102**,** 37-47.

Gyawali & Watkins David 2013. Continuous Hydrologic Modeling of Snow-Affected Watersheds in the Great Lakes Basin Using HEC-HMS. *Journal of Hydrologic Engineering,* 18**,** 29-39.

Haddad & Rahman 2011. Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment,* 25**,** 415-428.

Hao & Singh 2016. Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography,* 40**,** 549-578.

Hassan, Shamsudin & Harun 2014. Application of SDSM and LARS-WG for simulating and downscaling of rainfall and temperature. *Theoretical and applied climatology,* 116**,** 243-257.

Heffernan, Stephenson & Gilleland 2012. Ismev: an introduction to statistical modeling of extreme values. *R package version,* 1.

Hirabayashi, Mahendran, Koirala, Konoshima, Yamazaki, Watanabe, Kim & Kanae 2013. Global flood risk under climate change. *Nature Clim. Change,* 3**,** 816-821.

Huthoff, Remo & Pinter 2015. Improving flood preparedness using hydrodynamic levee-breach and inundation modelling: Middle Mississippi River, USA. *Journal of Flood Risk Management,* 8**,** 2-18.

Ibrahim-Bathis & Ahmed 2016. Rainfall-runoff modelling of Doddahalla watershed—an application of HEC-HMS and SCN-CN in ungauged agricultural watershed. *Arabian Journal of Geosciences,* 9**,** 170.

IPCC 2014. Climate change 2014: Impacts, Adaptations, and Vulnerability.

Jeong, Sushama, Khaliq & Roy 2014. A copula-based multivariate analysis of Canadian RCM projected changes to flood characteristics for northeastern Canada. *Climate Dynamics,* 42**,** 2045-2066.

Joe, Smith & Weissman 1992. Bivariate Threshold Methods for Extremes. *Journal of the Royal Statistical Society. Series B (Methodological),* 54**,** 171-183.

Jung, Chang & Moradkhani 2011. Quantifying uncertainty in urban flooding analysis considering hydro-climatic projection and urban development effects. *Hydrol. Earth Syst. Sci.,* 15**,** 617-633.

Kabiri, Ramani Bai & Chan 2015. Assessment of hydrologic impacts of climate change on the runoff trend in Klang Watershed, Malaysia. *Environmental Earth Sciences,* 73**,** 27-37.

Kalyanapu, D.R., T.N. & S.J. 2012. Monte Carlo-based flood modelling framework for estimating probability weighted flood risk. *Journal of Flood Risk Management,* 5**,** 37-48.

Karamouz, Noori, Moridi & Ahmadi 2011. Evaluation of floodplain variability considering impacts of climate change. *Hydrological Processes,* 25**,** 90-103.

Karim, Hasan & Marvanek 2017. Evaluating Annual Maximum and Partial Duration Series for Estimating Frequency of Small Magnitude Floods. *Water,* 9**,** 481.

Karim, Petheram, Marvanek, Ticehurst, Wallace & Hasan 2016. Impact of climate change on floodplain inundation and hydrological connectivity between wetlands and rivers in a tropical river catchment. *Hydrological Processes,* 30**,** 1574-1593.

Karlsson, Sonnenborg, Refsgaard, Trolle, Børgesen, Olesen, Jeppesen & Jensen 2016. Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. *Journal of Hydrology,* 535**,** 301-317.

Karmakar & Simonovic 2009. Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management,* 2**,** 32-44.

Karmakar & Simonovic 2008. Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques. *Journal of Flood Risk Management,* 1**,** 190-200.

Katz 2013. Statistical methods for nonstationary extremes. *Extremes in a Changing Climate.* Springer.

Katz, Parlange & Naveau 2002. Statistics of extremes in hydrology. *Advances in Water Resources,* 25**,** 1287-1304.

Kay, Davies, Bell & Jones 2009. Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change,* 92**,** 41-63.

Kay & Jones 2012. Transient changes in flood frequency and timing in Britain under potential projections of climate change. *International Journal of Climatology,* 32**,** 489-502.

Kay, Reynard & Jones 2006. RCM rainfall for UK flood frequency estimation. I. Method and validation. *Journal of hydrology,* 318**,** 151-162.

Kharin, Zwiers, Zhang & Hegerl 2007. Changes in Temperature and Precipitation Extremes in the IPCC Ensemble of Global Coupled Model Simulations. *Journal of Climate,* 20**,** 1419-1444.

Kilsby, Jones, Burton, Ford, Fowler, Harpham, James, Smith & Wilby 2007. A daily weather generator for use in climate change studies. *Environmental Modelling & Software,* 22**,** 1705-1719.

Kim, Silvapulle & Silvapulle 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis,* 51**,** 2836-2850.

Kioutsioukis, Melas & Zerefos 2010. Statistical assessment of changes in climate extremes over Greece (1955–2002). *International Journal of Climatology,* 30**,** 1723-1737.

Klein, Pahlow, Hundecha & Schumann 2010. Probability Analysis of Hydrological Loads for the Design of Flood Control Systems Using Copulas. *Journal of Hydrologic Engineering,* 15**,** 360-369.

Kojadinovic & Yan 2010a. Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics,* 47**,** 52-63.

Kojadinovic & Yan 2010b. Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software,* 34**,** 1-20.

Krause, Boyle & Bäse 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences,* 5**,** 89-97.

Laio, Di Baldassarre & Montanari 2009. Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research,* 45.

Lamb, Crossley & Waller 2009. A fast two-dimensional floodplain inundation model. *Proceedings of the Institution of Civil Engineers - Water Management,* 162**,** 363-370.

Lang., Ouarda & Bobée 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology,* 225**,** 103-117.

Laouacheria & Mansouri 2015. Comparison of WBNM and HEC-HMS for Runoff Hydrograph Prediction in a Small Urban Catchment. *Water Resources Management,* 29**,** 2485-2501.

Larsen, Refsgaard, Drews, Butts, Jensen, Christensen & Christensen 2014. Results from a full coupling of the HIRHAM regional climate model and the MIKE SHE hydrological model for a Danish catchment. *Hydrol. Earth Syst. Sci.,* 18**,** 4733-4749.

Lasage, Veldkamp, De Moel, Van, Phi, Vellinga & Aerts 2014. Assessment of the effectiveness of flood adaptation strategies for HCMC. *Natural Hazards and Earth System Sciences,* 14**,** 1441-1457.

Ledford & Tawn 1996. Statistics for near independence in multivariate extreme values. *Biometrika,* 83**,** 169-187.

Li, Guo, Chen & Guo 2012. Bivariate flood frequency analysis with historical information based on copula. *Journal of Hydrologic Engineering,* 18**,** 1018-1030.

Li, Xu, Chen & Simonovic 2009. Streamflow Forecast and Reservoir Operation Performance Assessment Under Climate Change. *Water Resources Management,* 24**,** 83.

Lima, Lall, Troy & Devineni 2015. A climate informed model for nonstationary flood risk prediction: Application to Negro River at Manaus, Amazonia. *Journal of Hydrology,* 522**,** 594-602.

Lin, Chen, Yao, Chen, Liu, Gao & James 2015. Analyses of landuse change impacts on catchment runoff using different time indicators based on SWAT model. *Ecological Indicators,* 58**,** 55-63.

Lu, Qin & Mandapaka 2015. A combined weather generator and K-nearest-neighbour approach for assessing climate change impact on regional rainfall extremes. *International Journal of Climatology,* 35**,** 4493-4508.

Luke, Sanders, Goodrich, Feldman, Boudreau, Eguiarte, Serrano, Reyes, Schubert & Aghakouchak 2018. Going beyond the flood insurance rate map: insights from flood hazard map co-production. *Natural Hazards and Earth System Sciences,* 18**,** 1097.

Mackay, Suter, Albert, Morton & Yamagata 2015. Large scale flexible mesh 2D modelling of the Lower Namoi Valley.

Malamud & Turcotte 2006. The applicability of power-law frequency statistics to floods. *Journal of Hydrology,* 322**,** 168-180.

Mani, Chatterjee & Kumar 2014. Flood hazard assessment with multiparameter approach derived from coupled 1D and 2D hydrodynamic flow model. *Natural Hazards,* 70**,** 1553-1574.

Maraun, Wetterhall, Ireson, Chandler, Kendon, Widmann, Brienen, Rust, Sauter, Themeßl, Venema, Chun, Goodess, Jones, Onof, Vrac & Thiele-Eich 2010. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics,* 48.

Masood & Takeuchi 2012. Assessment of flood hazard, vulnerability and risk of mid-eastern Dhaka using DEM and 1D hydrodynamic model. *Natural Hazards,* 61**,** 757-770.

Mazzoleni, Bacchi, Barontini, Baldassarre, Pilotti & Ranzi 2014. Flooding Hazard Mapping in Floodplain Areas Affected by Piping Breaches in the Po River, Italy. *Journal of Hydrologic Engineering,* 19**,** 717-731.

Mcleod & Ai 2011. Package 'Kendall'. *Kendall rank correlation and Mann-Kendall trend test. R package version,* 2.

Mehan, Guo, Gitau & Flanagan 2017. Comparative Study of Different Stochastic Weather Generators for Long-Term Climate Data Simulation. *Climate,* 5**,** 26.

Merwade, Olivera, Arabi & Edleman 2008. Uncertainty in flood inundation mapping: current issues and future directions. *Journal of Hydrologic Engineering,* 13**,** 608-620.

Mikhail, Roger, Elaine & Clarence 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate Research,* 10**,** 95-107.

Milly, Wetherald, Dunne & Delworth 2002. Increasing risk of great floods in a changing climate. *Nature,* 415**,** 514.

Mirabbasi, Fakheri-Fard & Dinpashoh 2012. Bivariate drought frequency analysis using the copula method. *Theoretical and Applied Climatology,* 108**,** 191-206.

Moriasi, Arnold, Liew, Bingner, Harmel & Veith 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *ASABE,* 50(3).

Mosquera-Machado & Ahmad 2007. Flood hazard assessment of Atrato River in Colombia. *Water Resources Management,* 21**,** 591-609.

Nguyen, Jayakumar & Agilan 2017. Impact of Climate Change on Flood Frequency of the Trian reservoir in Vietnam using RCMS. *Journal of Hydrologic Engineering (Accepted for publication).*

Noi & Nitivattananon 2015. Assessment of vulnerabilities to climate change for urban water and wastewater infrastructure management: Case study in Dong Nai river basin, Vietnam. *Environmental Development,* 16**,** 119-137.

Oeh 2005. Floodplain Development Manual: The Management of Flood Liable Land. *NSW, Australia: Department of Natural Resources*.

Pappenberger, Beven, Horritt & Blazkova 2005. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *Journal of Hydrology,* 302**,** 46-69.

Pappenberger, Dutra, Wetterhall & Cloke 2012. Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrol. Earth Syst. Sci.,* 16**,** 4143-4156.

Park, Min, Lee, Cha, Suh, Kang, Hong, Lee, Baek, Boo & Kwon 2016. Evaluation of multiple regional climate models for summer climate extremes over East Asia. *Climate Dynamics,* 46**,** 2469-2486.

Pathiraja, Westra & Sharma 2012. Why continuous simulation? The role of antecedent moisture in design flood estimation. *Water Resources Research,* 48.

Patras, Brunton, Downey, Rawson, Warriner & Gernigon 2011. Application of principal component and hierarchical cluster analysis to classify fruits and vegetables commonly consumed in Ireland based on in vitro antioxidant activity. *Journal of Food Composition and Analysis,* 24**,** 250-256.

Perrin, Michel & Andréassian 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology,* 279**,** 275-289.

Petheram, Rustomji, Chiew & Vleeshouwer 2012. Rainfall–runoff modelling in northern Australia: A guide to modelling strategies in the tropics. *Journal of Hydrology,* 462-463**,** 28-41.

Ponce 1994. *Engineering Hydrology: Principles and Practices,* New York, Prentice Hall.

Poulin, Huard, Favre & Pugin 2007. Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering,* 12**,** 394-403.

Prestininzi, Di Baldassarre, Schumann & Bates 2011. Selecting the appropriate hydraulic model structure using low-resolution satellite imagery. *Advances in Water Resources,* 34**,** 38-46.

Prudhomme, Reynard & Crooks 2002. Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrological Processes,* 16**,** 1137-1150.

Qin & Lu 2014. Study of climate change impact on flood frequencies: A combined weather generator and hydrological modeling approach. *Journal of Hydrometeorology,* 15**,** 1205-1219.

Raff, Pruitt & Brekke 2009. A framework for assessing flood frequency based on climate projection information. *Hydrol. Earth Syst. Sci.,* 13**,** 2119-2136.

Rahman, Rahman, Zaman, Haddad, Ahsan & Imteaz 2013. A study on selection of probability distributions for at-site flood frequency analysis in Australia. *Natural Hazards,* 69**,** 1803-1813.

Raj, Elizabeth, Indrajeet, M. & J. 2016. Watershed-scale impacts of bioenergy crops on hydrology and water quality using improved SWAT model. *GCB Bioenergy,* 8**,** 837-848.

Reddy & Ganguli 2012. Bivariate flood frequency analysis of upper Godavari river flows using Archimedean copulas. *Water resources management,* 26**,** 3995-4018.

Renard & Lang 2007. Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources,* 30**,** 897-912.

Requena, Chebana & Mediero 2016. A complete procedure for multivariate index-flood model application. *Journal of Hydrology,* 535**,** 559-580.

Ribatet 2007. POT: Modelling peaks over a threshold. *Threshold,* 5**,** 15-30.

Ringler, Vu Huy & Msangi 2012. Water allocation policies for the Dong Nai River Basin: an integrated perspective.

Roth, Buishand, Jongbloed, Klein Tank & Van Zanten 2012. A regional peaks-over-threshold model in a nonstationary climate. *Water Resources Research,* 48.

Rutten, Van Dijk, Van Rooij & Hilderink 2014. Land Use Dynamics, Climate Change, and Food Security in Vietnam: A Global-to-local Modeling Approach. *World Development,* 59**,** 29-46.

Saf 2009a. Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey. *Water Resources Management,* 23**,** 531-551.

Saf 2009b. Regional flood frequency analysis using L moments for the Buyuk and Kucuk Menderes river basins of Turkey. *Journal of Hydrologic Engineering,* 14**,** 783-794.

Salas, Heo, Lee & Burlando 2012. Quantifying the uncertainty of return period and risk in hydrologic design. *Journal of Hydrologic Engineering,* 18**,** 518-526.

Salas Jose, Heo Jun, Lee Dong & Burlando 2013. Quantifying the Uncertainty of Return Period and Risk in Hydrologic Design. *Journal of Hydrologic Engineering,* 18**,** 518-526.

Salvadori, Durante & De Michele 2013. Multivariate return period calculation via survival functions. *Water Resources Research,* 49**,** 2308-2311.

Sampson, Smith, Bates, Neal, Alfieri & Freer 2015. A high-resolution global flood hazard model. *Water Resources Research,* 51**,** 7358-7381.

Sarhadi, Burn, Concepción Ausín & Wiper 2016. Time varying nonstationary multivariate risk analysis using a dynamic Bayesian copula. *Water Resources Research*.

Sarhadi, Soltani & Modarres 2012. Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis. *Journal of Hydrology,* 458–459**,** 68-86.

Scarrott & Macdonald 2012. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Statistical Journal,* 10**,** 33-60.

Schepsmeier, Stoeber, Brechmann, Graeler, Nagler & Erhardt 2012. VineCopula: Statistical inference of vine copulas. *R package version,* 1.

Schiermeier 2011. Increased flood risk linked to global warming: likelihood of extreme rainfall may have been doubled by rising greenhouse-gas levels. *Nature,* 470**,** 316-317.

Schmidt & Stadtmüller 2006. Non-parametric Estimation of Tail Dependence. *Scandinavian Journal of Statistics,* 33**,** 307-335.

Schubert & Sanders 2012. Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources,* 41**,** 49-64.

Seckin, Haktanir & Yurtal 2011. Flood frequency analysis of Turkey using L-moments method. *Hydrological Processes,* 25**,** 3499-3505.

Semenov 2008. Simulation of extreme weather events by a stochastic weather generator. *Climate Research,* 35**,** 203-212.

Semenov, Brooks, Barrow & Richardson 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate research,* 10**,** 95-107.

Semenov & Stratonovitch 2010. Use of multi-model ensembles from global climate models for assessment of climate change impacts. *Climate research,* 41**,** 1-14.

Serinaldi, Bárdossy & Kilsby 2015. Upper tail dependence in rainfall extremes: would we know it if we saw it? *Stochastic Environmental Research and Risk Assessment,* 29**,** 1211-1233.

Setegn, Rayner, Melesse, Dargahi & Srinivasan 2011. Impact of climate change on the hydroclimatology of Lake Tana Basin, Ethiopia. *Water Resources Research,* 47.

Sharma, Sorooshian & Wheater 2008. Hydrological Modelling in Arid and Semi-Arid Areas. *New York, Cambridge University*.

Shen, Wang, Cheng, Rui & Ye 2015. Integration of 2-D hydraulic model and high-resolution lidar-derived DEM for floodplain flow modeling. *Hydrology and Earth System Sciences,* 19**,** 3605-3616.

Sheng 2001. A bivariate gamma distribution for use in multivariate flood frequency analysis. *Hydrological Processes,* 15**,** 1033-1045.

Shiau 2003. Return period of bivariate distributed extreme hydrological events. *Stochastic Environmental Research and Risk Assessment,* 17**,** 42-57.

Shiau, Wang & Tsai 2006. Bivariate frequency analysis of flood using copulas 1 *JAWRA Journal of the American Water Resources Association,* 42**,** 1549-1564.

Shin, Guillaume, Croke & Jakeman 2015. A review of foundational methods for checking the structural identifiability of models: Results for rainfall-runoff. *Journal of Hydrology,* 520**,** 1-16.

Shin & Kim 2017. Assessment of the suitability of rainfall–runoff models by coupling performance statistics and sensitivity analysis. *Hydrology Research,* 48**,** 1192-1213.

Shortridge, Guikema & Zaitchik 2016. Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrol. Earth Syst. Sci.,* 20**,** 2611-2628.

Solari & Losada 2012. A unified statistical model for hydrological variables including the selection of threshold for the peak over threshold method. *Water Resources Research,* 48.

Solomatine & Ostfeld 2008. Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics,* 10**,** 3-22.

Sraj, Bezak & Brilly 2015. Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrological Processes,* 29**,** 225-238.

Stephens, Bates, Freer & Mason 2012. The impact of uncertainty in satellite data on the assessment of flood inundation models. *Journal of Hydrology,* 414**,** 162-173.

Storch & Downes 2011. A scenario-based approach to assess Ho Chi Minh City's urban development strategies against the impact of climate change. *Cities,* 28**,** 517-526.

Strupczewski, Singh & Feluch 2001. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *Journal of Hydrology,* 248**,** 123-142.

Sun, Lall, Merz & Dung 2015. Hierarchical Bayesian clustering for nonstationary flood frequency analysis: Application to trends of annual maximum flow in Germany. *Water Resources Research,* 51**,** 6586-6601.

Sunyer, Madsen & Ang 2012. A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change. *Atmospheric Research,* 103**,** 119-128.

Svensson, Kundzewicz & Maurer 2005. Trend detection in river flow series: 2. Flood and low-flow index series/Détection de tendance dans des séries de débit fluvial: 2. Séries d'indices de crue et d'étiage. *Hydrological Sciences Journal,* 50.

Taye, Ntegeka, Ogiramoi & Willems 2011. Assessment of climate change impact on hydrological extremes in two source regions of the Nile River Basin. *Hydrology and Earth System Sciences,* 15**,** 209-222.

Teng, Huang & Ginis 2018. Hydrological modeling of storm runoff and snowmelt in Taunton River Basin by applications of HEC-HMS and PRMS models. *Natural Hazards,* 91**,** 179-199.

Teng, Jakeman, Vaze, Croke, Dutta & Kim 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software,* 90**,** 201-216.

Tennakoon. Parameterisation of 2D hydrodynamic models and flood hazard mapping for Naga city, Philippines. 2004. Itc.

Thompson, Cai, Reeve & Stander 2009. Automated threshold selection methods for extreme wave analysis. *Coastal Engineering,* 56**,** 1013-1021.

Thompson & Frazier 2014. Deterministic and probabilistic flood modeling for contemporary and future coastal and inland precipitation inundation. *Applied Geography,* 50**,** 1-14.

Timbadiya, Patel & Porey 2015. A 1D–2D Coupled Hydrodynamic Model for River Flood Prediction in a Coastal Urban Floodplain. *Journal of Hydrologic Engineering,* 20**,** 05014017.

Tramblay, Badi, Driouech, El Adlouni, Neppel & Servat 2012. Climate change impacts on extreme precipitation in Morocco. *Global and Planetary Change,* 82–83**,** 104-114.

Trinh, Vu, Van Der Steen & Lens 2013. Climate Change Adaptation Indicators to Assess Wastewater Management and Reuse Options in the Mekong Delta, Vietnam. *Water Resources Management,* 27**,** 1175-1191.

Trzaska & Schnarr 2014. A review of downscaling methods for climate change projections. *United States Agency for International Development by Tetra Tech ARD***,** 1-42.

Van Roosmalen, Christensen, Butts, Jensen & Refsgaard 2010. An intercomparison of regional climate model data for hydrological impact studies in Denmark. *Journal of Hydrology,* 380**,** 406-419.

Vaze, Teng & Spencer 2010. Impact of DEM accuracy and resolution on topographic indices. *Environmental Modelling & Software,* 25**,** 1086-1098.

Vazquez-Amábile & Engel 2005. Use of SWAT to compute groundwater table depth and streamflow in the Muscatatuck River watershed. *Transactions of the ASAE,* 48**,** 991-1003.

Viglione & Blöschl 2009. On the role of storm duration in the mapping of rainfall to flood return periods. *Hydrol. Earth Syst. Sci.,* 13**,** 205-216.

Viglione, Hosking, Laio, Miller, Gaume, Payrastre, Salinas, N'guyen, Halbert & Viglione 2018. Package 'nsRFA'. *Non-supervised Regional Frequency Analysis. CRAN Repository*.

Villarini, Serinaldi & Krajewski 2008. Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. *Advances in Water Resources,* 31**,** 1674-1686.

Villarini, Smith, Serinaldi, Bales, Bates & Krajewski 2009. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources,* 32**,** 1255-1266.

Vittal, Singh, Kumar & Karmakar 2015. A framework for multivariate data-based at-site flood frequency analysis: Essentiality of the conjugal application of parametric and nonparametric approaches. *Journal of Hydrology,* 525**,** 658-675.

Vu, Aribarg, Supratid, Raghavan & Liong 2016. Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok? *Theoretical and Applied Climatology,* 126**,** 453-467.

Weller, Cooley & Sain 2012. An investigation of the pineapple express phenomenon via bivariate extreme value theory. *Environmetrics,* 23**,** 420-439.

Wilks & Wilby 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography,* 23**,** 329-357.

World Bank. 2010. *Climate risks and adaptation in Asian coastal megacities: a synthesis report* [Online]. Available: http://documents.worldbank.org/curated/en/866821468339644916/Climate-risks-and-adaptation-in-Asian-coastal-megacities-a-synthesis-report [Accessed 27 Dec 2018].

Xu 2002. Hydrologic models. *Textbooks of Uppsala University. Department of Earth Sciences Hydrology*.

Xu, Zhang, Shen & Li 2010. Storm surge simulation along the US East and Gulf Coasts using a multi-scale numerical model approach. *Ocean Dynamics,* 60**,** 1597-1619.

Yaseen, Jaafar, Deo, Kisi, Adamowski, Quilty & El-Shafie 2016. Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq. *Journal of Hydrology,* 542**,** 603-614.

Yilmaz, Imteaz & Perera 2017. Investigation of non-stationarity of extreme rainfalls and spatial variability of rainfall intensity–frequency–duration relationships: a case study of Victoria, Australia. *International Journal of Climatology,* 37**,** 430-442.

Yin, Lin & Yu 2016. Coupled modeling of storm surge and coastal inundation: A case study in New York City during Hurricane Sandy. *Water Resources Research,* 52**,** 8685-8699.

Yue 1999. Applying bivariate normal distribution to flood frequency analysis. *Water International,* 24**,** 248-254.

Yue, Ouarda & Bobee 2001. A review of bivariate gamma distributions for hydrological application. *Journal of hydrology,* 246**,** 1-18.

Yue & Wang 2004. A comparison of two bivariate extreme value distributions. *Stochastic Environmental Research and Risk Assessment,* 18**,** 61-66.

Zaman, Rahman & Haddad 2012. Regional flood frequency analysis in arid regions: A case study for Australia. *Journal of Hydrology,* 475**,** 74-83.

Zhang & Singh 2006. Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng.,* 11**,** 150.

Zhang & Singh 2007. Bivariate rainfall frequency distributions using Archimedean copulas. *Journal of Hydrology,* 332**,** 93-109.

Zhou, Mikkelsen, Halsnæs & Arnbjerg-Nielsen 2012. Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. *Journal of Hydrology,* 414**,** 539-549.

Research papers resulting from the thesis

Publications in Peer-Reviewed Journals

- **Dong Nguyen**, K.V. Jayakumar & Agilan 2018. Impact of Climate Change on Flood Frequency of the Trian Reservoir in Vietnam Using RCMS. ***Journal of Hydrologic Engineering***, 23(2): 05017032.

- **Dong Nguyen** and Jayakumar 2018. Assessing the copula selection for bivariate frequency analysis based on the tail dependence test. ***Journal of Earth System Science,*** https://doi.org/10.1007/s12040-018-0994-4

- **Dong Nguyen,** K.V. Jayakumar & Agilan 2018. Bivariate flood frequency analysis of nonstationary flood characteristics**.** ***Journal of Hydrologic Engineering,*** https://doi.org/10.1061/(ASCE)HE.1943-5584.0001770

- **Dong Nguyen** and K.V. Jayakumar 2018. Flood hazard in Saigon-Dongnai river basin under climate change context. ***Journal of Flood Risk Management (Under review).***

Acknowledgments

I am highly grateful to all those who helped me in completing this thesis possible.

First of all, I would like to specially thank Prof. K.V. Jayakumar for encouraging and guiding me during the doctorate as well as for sharing his knowledge and enthusiasm for collaborating with people. I would also like to thank Prof. N.V. Umamahesh and all the other faculty members of the Water & Environmental Division for their numerous supports. I gratefully acknowledge Prof. P Anand Raj, Prof. D M Vinod Kumar and Prof. Rathish Kumar, the members of the Doctoral Scrutiny Committee for their constant support at various stages of the work. I also thank all the Heads of the Department of Civil Engineering during my stay here for their support.

I thank the India Council for Cultural Relations, Government of India and Government of Vietnam for the financial assistance through a scholarship to successfully complete this study. I also thank the Thuyloi University, Ho Chi Minh City, Vietnam for the support to take up this research work.

National Hydro-Meteorological Service and the Ministry of Natural Resources and Environment, Vietnam provided me with all the necessary data for this study. I place on record my deep sense of gratitude to the officials of these organizations.

I am also thankful to all faculty members and another staff member of Institute of Technology, Warangal for their support during my stay in India.

I thank my colleagues Agilan, Uday, Srilatha, Jew Das and Sruthi who motivated me to continue researching and their assistance in completing this study.

Finally, I would like to express my heartfelt thanks to my family for both financial and emotional support and encouragements.

Dang Dong Nguyen

Date: