Feihong Liu

# Analyzing the influence of compound events on flooding in the downstream reach of the Houston Ship Channel

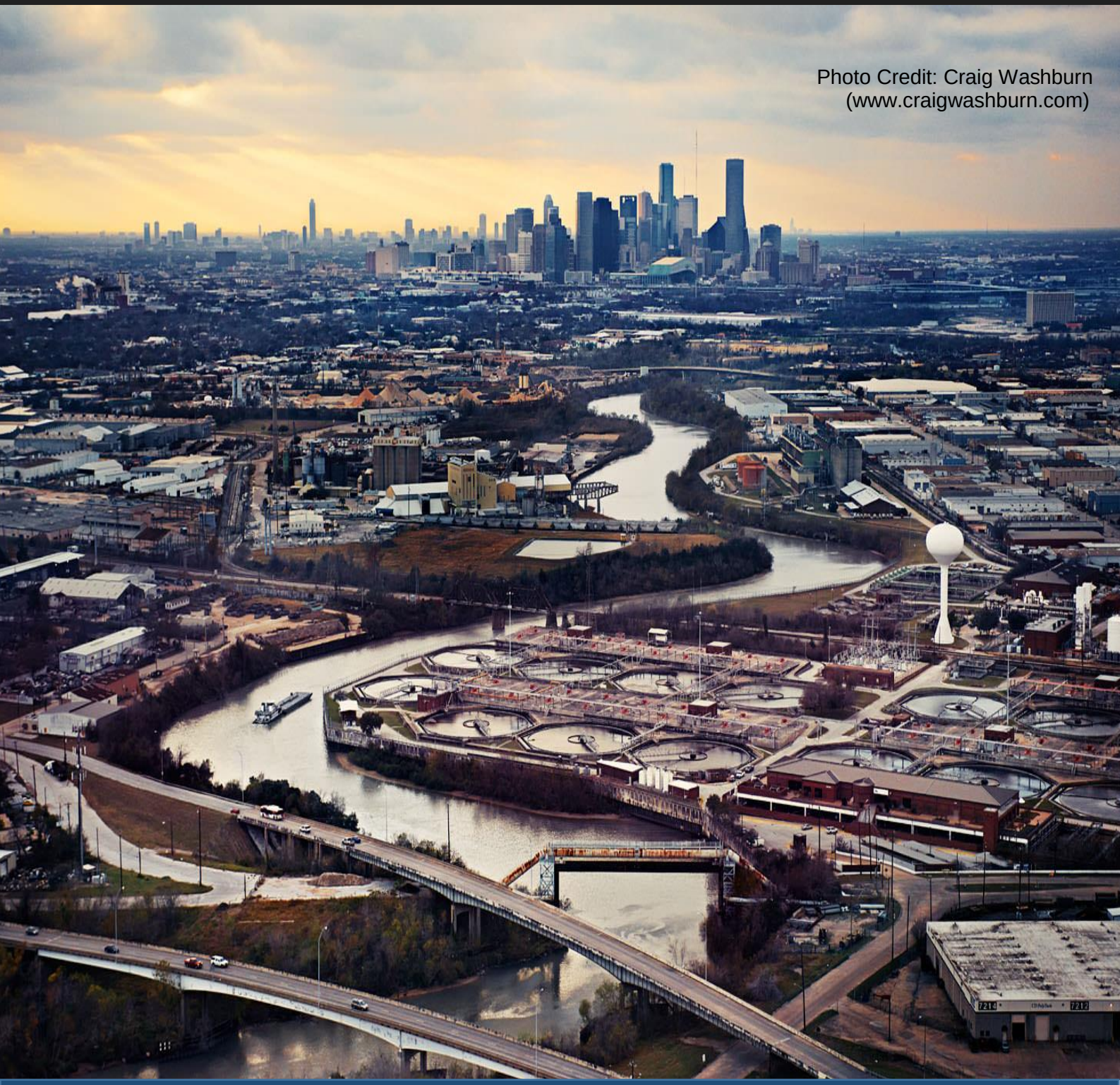


Photo Credit: Craig Washburn
(www.craigwashburn.com)

TU Delft

MSc Thesis

# Analysing the influence of compound events on flooding in the downstream reach of the Houston Ship Channel

By

Feihong Liu

in partial fulfilment of the requirements for the degree of

**Master of Science**
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on Monday, October 30, 2017, at 14:00 PM.

| | | |
|---|---|---|
| Thesis committee: | Prof.dr.ir. Jonkman, S.N. | TU Delft |
| | Dr.ir. Sebastian, A.G. | TU Delft |
| | Dr.ir. Morales-Napoles, O. | TU Delft |
| | Dr.ir. Labeur, R.J. | TU Delft |
| | Ir. Jager, W.S. | TU Delft |



TU Delft

# Preface

This thesis is the last piece of the puzzle to complete my MSc programme in Hydraulic Engineering, Delft University of Technology. In the past ten months, I have gained valuable experience through this thesis which has improved my level of writing, broadened my mind, and let me have a more profound understanding in the field of flood risk, especially the compound flood risk.

I would like to thank all my graduation committee members: Prof. Jonkman, S.N., Dr. Sebastian, A.G., Dr. Morales-Napoles, O., Dr. Labeur, R.J. and Ir. Jager, W.S. who have provided a lot of valuable suggestions and guidance on my thesis and helped me finish this thesis better. In addition, I want to thank my family and friends who have been supporting and encouraging me all the time.

*Feihong Liu*

*October 22, 2017, Delft*

# Abstract

Flooding is one of the most frequent and terrible natural disasters. For the coastal areas connecting inland with the open sea, the occurrence of flooding could be caused by high flows from inland rivers or high-water levels at open sea (e.g., from tsunamis or storm surge). There is also one extreme case that high sea level and high inland flow happen simultaneously, which is also known as the compound flood event. If a compound flood event happens, it has the potential to cause huge losses.

The Houston Ship Channel (HSC) is an important passage connecting “Greater Houston” area to the Gulf of Mexico which plays a critical role in the economic development of Texas and even the whole country. However, because of its special location, the HSC area is prone to flooding caused by both storm surge and heavy rainfall, which has been demonstrated by many historical flood events e.g. Tropical Storm Allison (2001), Hurricane Ike (2008) and Hurricane Harvey (2017). In the wake of Hurricane Ike, the Severe Storm Prediction, Education, and Evacuation from Disasters (SSPEED) Center at Rice University proposed to build a storm surge barrier near the downstream of Fred Hartman Bridge for the protection of the industrial facilities along the HSC from storm surges. One of the major questions related to the design of the storm surge barrier is ‘what are the probable boundary conditions associated with compound flood events (i.e., the combination of storm surge and upstream rainfall-runoff) in the Houston Ship Channel?’. Because once the barrier is closed, the upstream flow cannot flow out to the Galveston Bay, there is the potential to cause the flooding behind the barrier when the closure time is long and upstream discharge is large enough. Therefore, in this thesis, it is focused on the exploration of potential combinations of compound floods in the downstream reach of the HSC. To achieve the goal, this thesis introduces a set of methods to link the return period of a compound flood event with the joint occurrence probability of the hydraulic boundary conditions, which can provide a quick, preliminary judgement when predicting the compound flood events especially in the area lack of sufficient historical data.

To determine the boundary conditions for the downstream reach of the Houston Ship Channel, historical data was collected at eight tidal and stream gages for the period between 02/10/1997 and 31/01/2017. More specifically, the sea level at NOAA 8771450 (GP21) and NOAA 8770613 (Morgan) are collected for the prediction of the sea level at downstream boundary of the study area while daily maximum water level at USGS 08074710 (TB), daily maximum water level at USGS08075500 (SB), daily mean water level at USGS 08076700 (GB), daily mean water level at USGS 08072050 (SJR), daily mean discharge at USGS 08075000 (BB) and daily mean discharge at USGS 08075770 (HB) are collected for the prediction of the total discharge at upstream boundaries. Thereafter, a Non-parametric Bayesian Network (NPBN) model is constructed to build the joint distribution of the selected variables based on which the hydraulic boundary conditions can be generated with the help of uniform flow assumption and the assumption of linear relationship between the discharge and contributing area in a watershed. The final results show that, for the downstream reach of the HSC, the most likely combinations of hydraulic boundary conditions are 1.79m+NAVD88 and 10290$m^3$/s at downstream and upstream boundaries, respectively, for the 100-year compound flood events, 1.93m+NAVD88 and 11634$m^3$/s, for the 500-year compound flood events, and 2.00m+NAVD88 and 12303$m^3$/s, for the 1000-year compound flood events. Furthermore, the corresponding most likely 100-, 500-, and 1000-year return frequency water levels are modelled based on a 1-D steady hydraulic model and the predicted hydraulic boundary conditions. The 1% return frequency water level is 4.28m+NAVD88 at the upstream boundary and 1.79m+NAVD88 at the downstream boundary while the 0.2% return frequency water level is 4.57m+NAVD88 at the upstream boundary and 1.94m+NAVD88 at the downstream boundary; the 0.1% return frequency water level is 4.70m+NAVD88 at the upstream boundary and 2.00m+NAVD88 at the downstream boundary.

After the general analysis, Hurricane Harvey is selected as the case study to test the practicability of the methodology introduced in this thesis (i.e., 1-D steady hydraulic model and the BN). Since the measured highest sea level at the downstream boundary of the study area during Harvey was about 1.445m+NAVD88 on August 27, 2017, that day is considered to be the most significant day and the upstream boundary conditions on that day is estimated to be 15229.36 $m^3/s$ according to the collected data. Based on the hydraulic boundary conditions generated from the collected data on August 27, the return period of Hurricane Harvey is estimated. In addition, when using the BN to predict the upstream boundary condition, the value could be from 4390.34 $m^3/s$ to 31106.48$m^3/s$. Corresponding to the boundary conditions generated from BN, the modelled water levels at upstream boundary could vary from 3m+NAVD88 to 7.9m+NAVD88 while the modelled water level based on the boundary conditions generated from the collected data is about 5m+NAVD88.

# Contents

# List of Figure

# List of Tables

# 1

# Introduction

## 1.1. Motivation

Flooding is one of the most frequent and terrible natural disasters. There are many factors which can cause flooding; for example, a tsunami caused by the earthquake, a river flood caused by heavy rainfall, or storm surge caused by hurricanes. Each type of flood has the potential to cause large social, environmental, and economic damages to human society. For example, in 1998, China suffered an extremely large flood along the Yangtze River which was originally due to the persistent heavy rainfall in the Yangtze River basin. The official death loss was almost 3600 people and the economic loss was predicted higher than US$ 36 billion. It remains one of the worst natural disasters in the history of China (Ye & Glantz, 2005). In addition to Asia, the other regions in the world also face the flood risk. For example, in the Netherlands, the flood is the most frequent hazard because of its low-lying topography. In history, countless flood events happened in the Netherlands, including the North Sea flood of 1953 which was a terrible flood induced by a heavy storm and took 1836 lives ("The flood of 1953," n.d.). Since then, the Netherlands has invested considerably in structural protection against future flooding, e.g. the Delta Works (Bouwer & Vellinga, 2007). The United States is also prone to flooding and many recent hurricane-induced floods have caused tens of billions of USD in losses, e.g. Hurricane Katrina (2005), Hurricane Ike (2008), Superstorm Sandy (2012), and Hurricane Harvey (2017) (Blake, Landsea, & Gibney, 2007). It can be concluded from above examples that flood risk is a global problem and reduction of flood risk is critical to the development of the human society.

From the engineering point of view, risk is usually defined as the probability of an undesired event multiplied by its consequences (Kaplan & Garrick, 1981). Therefore, flood risk can be seen as the combination of the occurrence probability of a flood hazard and its consequence (e.g., damages to the built environment). Thus, flood risk in coastal areas is especially high because the population density in coastal areas is much higher than other areas because of the special position which provides the convenience for transportation and commercial trade as well as the rich resources, beautiful landscape and so on (Neumann et al., 2015). In addition to population, coastal areas are also important to the development of global economy since most of the metropolises locate in the coastal areas e.g. New York, Houston, Shanghai, Tokyo (Neumann et al., 2015). Therefore, it can be expected the potential damage would be extremely large once floods happen in the coastal areas especially when the compound flood events happen (Leonard et al., 2014; Wahl et al., 2015).

In coastal areas, especially deltas and estuaries which connect inland areas with the open sea, the occurrence of a flood hazard could be caused by high flows from inland rivers or high-water levels at open sea (e.g., from tsunamis or storm surge). In addition, there are extreme conditions in which high sea levels and inland flows occur simultaneously, potentially causing more extreme flooding than if they were to occur individually. According to the definition proposed by The Intergovernmental Panel on Climate Change (IPCC), compound events are (IPCC, 2012):

*"(1) two or more extreme events occurring simultaneously or successively, (2) combinations of extreme events with underlying conditions that amplify the impact of the events, or (3) combinations of events that are not themselves*

*extremes but lead to an extreme event or impact when combined. The contributing events can be of similar (clustered multiple events) or different type(s).”*

In this thesis, the joint occurrence of high sea levels and high river flows is called a compound event.

Wahl et al. (2015) have particularly analyzed the risk of compound floods from storm surge and rainfall for major cities in the United States. Based on their work, the interaction between storm surge and rainfall may exacerbate the overall influence of floods in coastal areas through three main mechanism: 1) the water levels in estuarine regions may be elevated to a warning level over which flood will be triggered because of the joint occurrence of storm surge and rainfall; 2) any serious rainfall will worsen the flood event if a severe storm surge already cause a widespread flooding; 3) when a storm surge is not strong enough to trigger flooding but still slow down the stormwater drainage, then a heavy rainfall on top of it can probably cause flooding. This research also points out that the number of compound flood events has increased largely during the last century at many tested coastal cities and the risk of compound flooding is higher for the Atlantic (and Gulf) Coasts than the Pacific Coast (Wahl et al., 2015).

## 1.2. Problem Definition

The Houston Ship Channel (HSC) is part of the Port of Houston which provides an important connection between Houston area and the Gulf of Mexico (Sibley, 2017). Many energy-related companies and economic activities are linked to that area (SSPEED, 2014). However, the HSC and the surrounding area is also prone to flooding because of its special location. On one hand, hurricanes are a common problem in the Gulf of Mexico which can cause severe storm surge at the HSC area; on the other hand, rainfall brought by hurricanes or heavy storms may lead to high flows and give rise to flooding (Christian et al., 2014; Schlepers, 2015; Torres et al., 2015). Currently, flood hazard in the lower downstream reach of the HSC is estimated as the high water level caused by either storm surge or rainfall-runoff (FEMA, 2017); however, it is possible that flooding at that area can be caused by a combination of storm surge and upstream rainfall-runoff.

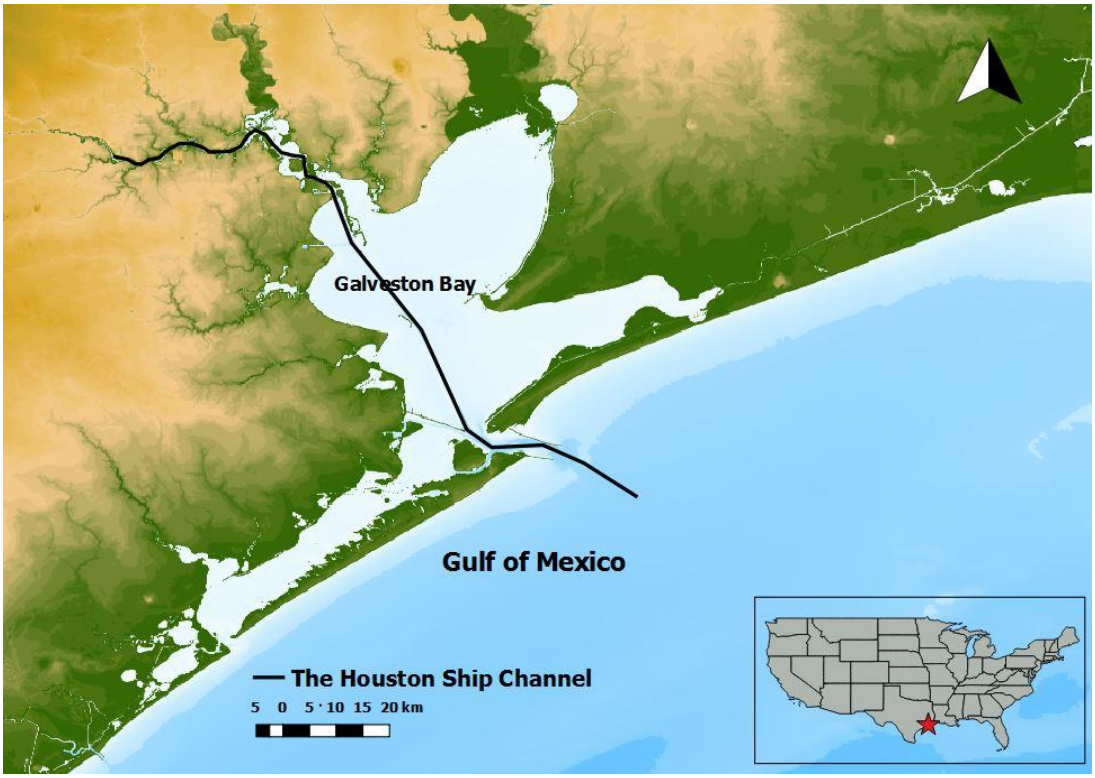



**Figure 1.1: The position of the HSC relative to the Gulf of Mexico. Source: Texas Natural Resources Information System (TNRIS).**

In the wake of recent hurricanes in the Gulf of Mexico, the Severe Storm Prediction, Education and Evacuation from Disasters (SSPEED) Center proposed to build a storm surge barrier at the downstream reach of the HSC near Fred Hartman Bridge which is shown in Figure 1.2 (Christian et al., 2014; SSPEED, 2015). The storm surge barrier would be built with the objective of protecting the industrial facilities along the HSC from storm surges. A number of studies have examined the performance of the proposed storm surge barrier using high resolution hydrologic and hydraulic software to model historical hurricane events (e.g., Ike, Katrina, Rita) (Christian et al., 2014; Torres et al., 2015); however, little is known about the joint probable combinations of rainfall-runoff and storm surge in the HSC area or the return periods of the historical storms. If a storm surge barrier would be constructed in this location, understanding the hazard associated with the compound flood events (i.e. return periods of upstream rainfall-runoff and storm surge) and their potential effects are needed in order to produce an engineering design for the proposed HSC barrier.


**Figure 1.2: Proposed flood risk reduction project. Source: (SSPEED, 2014).**

Moreover, during the writing of this thesis, Hurricane Harvey made landfall as a Category 4 hurricane near Rockport, Texas. The storm stayed in Texas from August 25 to August 30 during which it brought considerable rainfall, especially in the Houston region. Harvey highlighted the potential economic and, especially, environmental devastation that can occur in the HSC and surrounding areas when a hurricane makes landfall on the Texas coast. Moreover, during the event, water levels in Galveston Bay remained high for a number of days, raising the question as to the 'compound' nature of the event and to what extent elevated water levels in Galveston Bay exacerbated flooding in the HSC.

## 1.3. Research Questions and Objectives

Based on the problem described in Section 1.2, this thesis focuses on the following research questions:

1. What are the hazard boundary conditions associated with compound flood events in the downstream reach of the HSC?

2. What are the return frequency water levels (i.e., 1%, 0.2%, 0.1%) in the downstream reach of the HSC under compound flood events?
   2a. To what extent did elevated water levels in Galveston Bay contribute to flooding in the Houston Ship Channel during Hurricane Harvey?
   2b. What is the estimated return period of Hurricane Harvey in downstream reach of the Houston Ship Channel?

According to the above research questions, the primary objectives of this thesis are:

1. Collect and analyse historical water levels and flows in the HSC and contributing watersheds;
2. Build and validate a Non-parametric Bayesian Network (NPBN) to model the hazard boundary conditions that lead to compound floods;
3. Model the return frequency water levels along the downstream reach of the HSC associated with compound flood events;
4. Model the water levels during Hurricane Harvey and determine the return period of Hurricane Harvey.

## 1.4. Research Scope

The main channel of the HSC can be divided into two parts: 1) the upstream reach flowing eastward and following Buffalo Bayou from the Turning Basin to the San Jacinto River; 2) the downstream reach turning southward and flowing into Galveston Bay through San Jacinto River. Figure 1.3 exhibits the general position of the upstream reach and downstream reach of the HSC. In this thesis, the downstream reach of the HSC is selected as the main research area because it encompasses the proposed storm surge barrier.

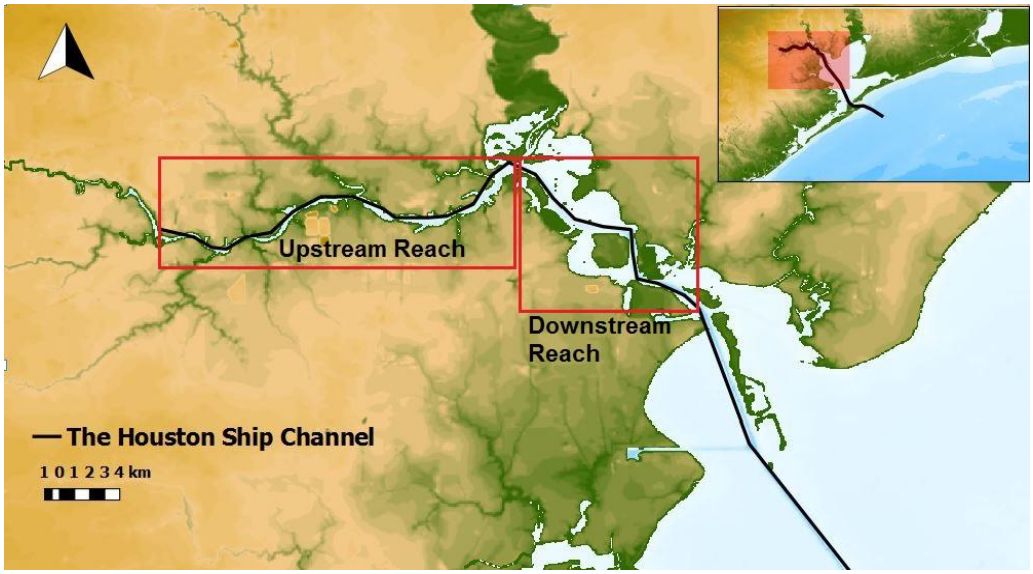



**Figure 1.3: The composition of the HSC. The primary study area in this thesis is the downstream reach of the Houston Ship Channel which includes the rainfall-runoff contribution of the upstream reach and flows from the San Jacinto River south of Lake Houston Dam. Source: Texas Natural Resources Information System (TNRIS).**

## 1.5. Methodology

Figure 1.4 shows the main methodology used in the thesis. First, the historical coastal and riverine data including the sea level data, upstream discharge data and riverine water level data are collected. Then, the Bayesian Network (BN) is chosen and used to build the joint distribution of all the collected variables through which the stochastic combinations of sea levels, riverine discharges and water levels can be generated. Based on the outputs of the

Bayesian Network, the hydraulic boundary conditions i.e. the discharge at the upstream boundary and sea level at the downstream boundary can be predicted so as the joint occurrence probabilities of different boundary conditions. Finally, the water surface profiles of the research area (i.e., the downstream reach of the HSC) are modelled using a 1-D steady hydraulic model where the hydraulic boundary conditions are taken as the input variables. The different models and calculations are described in more detail in the following chapters.

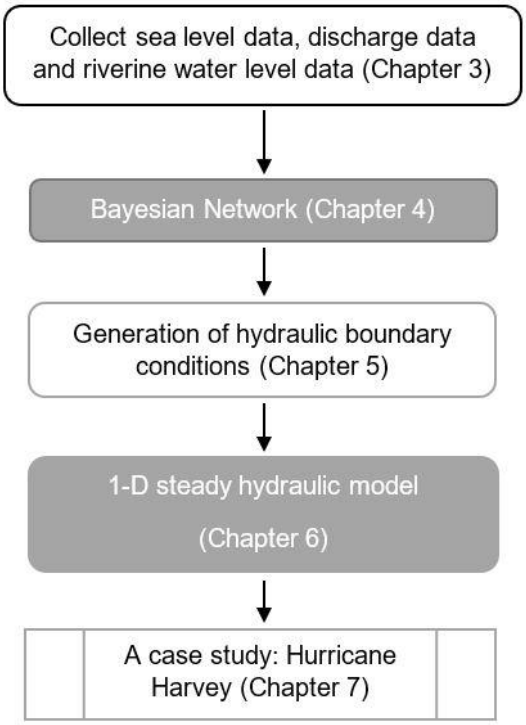



**Figure 1.4: Flowchart of the overall methodology.**

## 1.6. Report Layout

In this thesis, Chapter 2 presents the detailed system description including the geographical position of the study area, tropical cyclone climatology, historical flood events and the review of some related researches. Chapter 3 elaborates the data collection and analysis. Chapter 4 introduces the construction of the Bayesian Network while Chapter 5 describes the generation of the hydraulic boundary conditions and prediction of hazard boundary conditions associated with compound flood events. In Chapter 6, the 1-D steady hydraulic model is introduced and return frequency water levels (i.e., 1%, 0.2%, 0.1%) in the study area under compound flood events are modelled. Thereafter, the case study of Hurricane Harvey is presented in Chapter 7. Finally, in Chapter 8, some recommendations and suggestions for future study are provided.

# 2

# System Description

This Chapter introduces the study area in more detail. In Section 2.1, the location of the HSC is described from the large geographical scope to the small scope. In addition, the additional information about the area, population, economy is introduced as well, which proves the importance of the HSC. Section 2.2 mainly discuss the climatology of the study area which contains the theoretical background of tropical cyclones, the U.S. hurricane season, the potential hurricane frequency, and the precipitation of Houston area. After the introduction of the climatology, some severe historical flood events in Houston area are mentioned in Section 2.3 followed by Section 2.4 presenting the flood risk reduction strategies. In Section 2.5, some researches related to the flood risk in the HSC are reviewed. The summary of the whole chapter is shown in Section 2.6.

## 2.1. Geographical Information

### 2.1.1 Texas

From the perspective of both population and area, Texas is the second largest state in the United States with a total area of 696,241 $km^2$ and a population of nearly 28 million (Plocheck, n.d.). It is in the southwestern U.S. and is neighbored by the states of Louisiana, Arkansas, Oklahoma, and New Mexico. In addition, Texas shares a border with Mexico to the southwest and the Gulf of Mexico to the southeast. Figure 2.1 shows the location of Texas with respect to the rest the U.S. and Mexico. From the economic point of view, Texas is also one of the most important states for U.S. having a gross state product (GDP) of almost $1.62 trillion in 2016, which is in the second place among all states (Bureau of Economic Analysis, n.d.). Thanks to the excellent transportation system including deepwater ports, airports and rail system, export trade become an important part of Texas' economy. Texas has been leading the country in exports for 14 years which exported with a value of more than $232 billion in 2016. In addition, Texas is also attractive to the foreign investment, especially in the energy and chemical industries. Over 50 Fortune 500 companies set offices in Texas (TXEDC, 2016). The above data indicates that Texas plays a critical role in the U.S. economy.


**Figure 2.1: Geographical position of Texas and the "Greater Houston" area. Source: U.S. Census Bureau.**

### 2.1.2 "Greater Houston"

Within Texas, the "Greater Houston" area is the most developed and populous area which is in the southeastern Texas and adjacent to the Gulf of Mexico (see Figure 2.1) and consists of nine counties: Austin County, Waller County, Montgomery County, Liberty County, Chambers County, Galveston County, Brazoria County, Fort Bend County, and Harris County (see Figure 2.2). Among all the nine counties, Harris County is the largest and most developed one which is in the center of Greater Houston and next to the Galveston Bay having an area of 1,778 square miles (nearly 4,605 $km^2$) and a population of around 4.5 million. According to the estimation of U.S. Census Bureau in 2010, Greater Houston is the fifth largest metropolitan area in U.S. covering an area of 9,444 square miles (about 24,459 $km^2$) and home to over 6 million people. In addition, Greater Houston is also one of the most developed and fastest-growing areas of the country which had a GDP of $503.3 billion in 2015 accounting for about 31 percent of the total GDP of Texas (1.61 trillion) at the same year according to the estimation of the U.S. Bureau of Economic Analysis (*Houston Facts*, 2017). The energy industry is the major component of the regional economy. There are approximately 4800 firms accommodated in Greater Houston which cover almost all the segments of the energy industry and about 66 percent of the international integrated oil companies (e.g., ExxonMobil, Shell) have businesses there. The crude oil processed in Greater Houston every calendar day makes up about 40 percent of the total production of Texas and 12 percent of the total nation's capacity (*Houston Facts*, 2017). Greater Houston can be seen as the global headquarter of energy industry which has a significant influence on the domestic and global energy market.

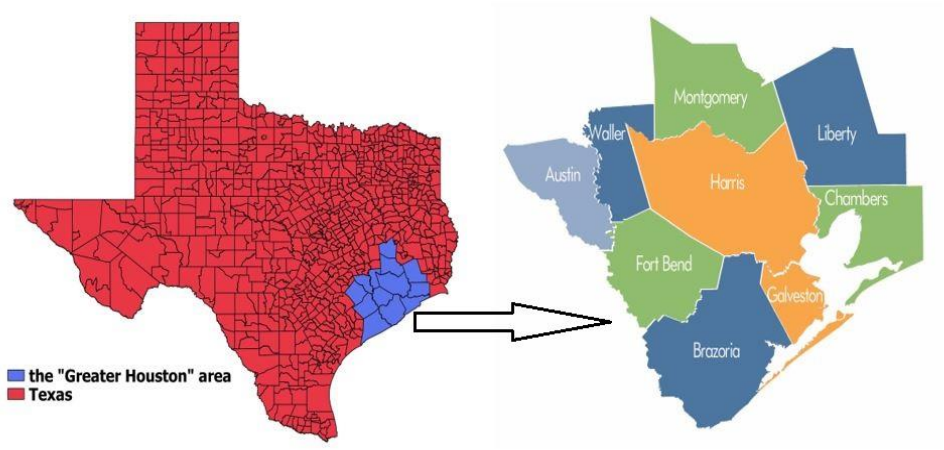



**Figure 2.2: The location and components of the "Greater Houston" area. Source: U.S. Census Bureau, Greater Houston Partnership.**

### 2.1.3 Galveston Bay

Galveston Bay lies in the southeast Texas near Greater Houston which is the largest estuary on the Texas coast and the seventh largest in U.S. with a surface area of approximately 1,554 km$^2$ and a total shoreline length of almost 374 km ("GALVESTON BAY | The Handbook of Texas Online," 2017; Phillips, 2004). From the perspective of classification of estuaries, Galveston Bay belongs to a lagoon-type estuary separated from the Gulf of Mexico by two barrier islands: Galveston Island and Bolivar Peninsula (Phillips, 2004; Schlepers, 2015). Given its special position, Galveston Bay is an important passage for ships navigating between the Gulf of Mexico and Greater Houston.


**Figure 2.3: Geographical position of Galveston Bay. Source: TNRIS, U.S. Census Bureau.**

### 2.1.4 The Houston Ship Channel

The Houston Ship Channel (HSC) connects Greater Houston to the Gulf of Mexico with a length of about 83 km. It flows through the Port of Houston which is the busiest port in U.S. in terms of the foreign tonnage and the second busiest one in terms of the overall tonnage. A report of economic impact of the Port of Houston shows that, in 2014,

the overall cargo activity in the Port of Houston supported 2,695,519 jobs in the U.S. and a total of $629.4 billion of economic activity within which the businesses related to the HSC contribute 1,174,567 jobs throughout Texas and helped generate more than $264.9 billion in statewide economic impact (The Port of Houston Authority, 2015). Along the HSC, there are over 150 companies and 330 terminals which mainly serve the energy-related industry. In 2016, the petroleum products exported from Port of Houston is valued over $21.22 billion, which is the leading export commodities of Port of Houston and almost accounts for 10 percent of the total exports value of Texas (about $232 billion) (*Houston Facts*, 2017; TXEDC, 2016). As Rose (1967) said, because of the presence of the HSC, it is possible to develop the ocean-going trade for Greater Houston which stimulates the economic development of Texas and even the whole country.

The main channel of the HSC, as mentioned in Chapter 1, can be divided into two reaches: 1) the upstream reach flowing eastward and following Buffalo Bayou from Turning Basin to the San Jacinto River; 2) the downstream reach turning southward and flowing into Galveston Bay through San Jacinto River (see Figure 1.2). Buffalo Bayou and San Jacinto river mainly contribute to the HSC. In addition, there are also many other upstream bayous flowing into the HSC. As shown in Figure 2.4, Sims Bayou, Brays Bayou, White Oak Bayou, Hunting Bayou, Greens Bayou, and Buffalo Bayou flow eastward to the upstream reach of the HSC while Carpenters Bayou and San Jacinto River flows southward and combine with the downstream reach of the HSC. Vince Bayou flows into the HSC from the South.

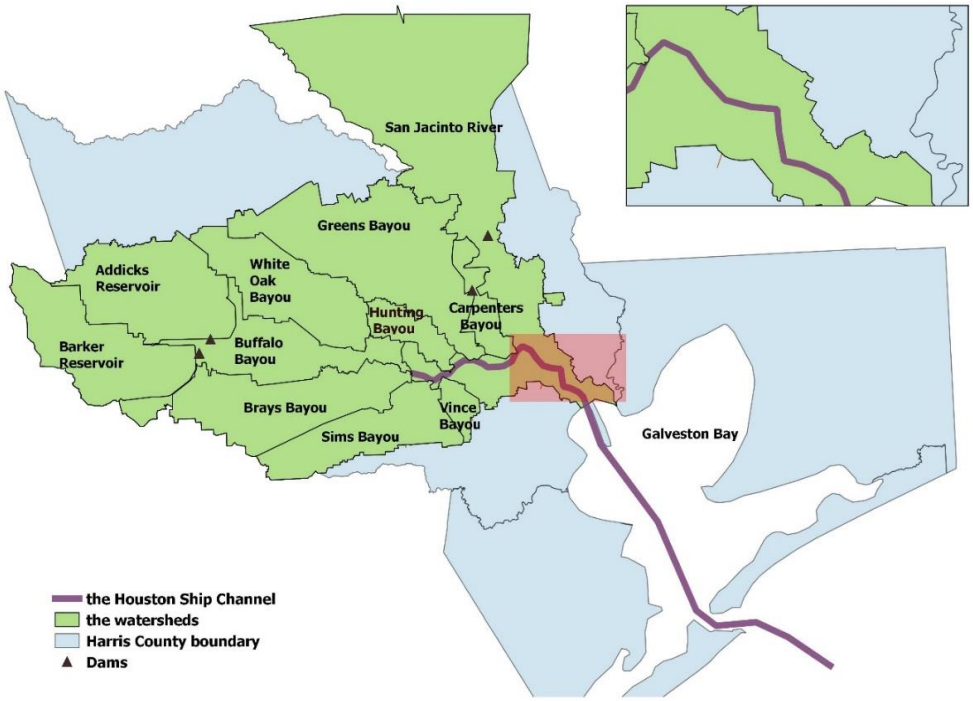



**Figure 2.4: The contributing watersheds of the Houston Ship Channel. Source: Harris County Flood Control District (HCFCD).**

## 2.2. Climatology

### 2.2.1 Tropical Cyclones

A tropical cyclone usually forms over warm ocean waters in tropical areas with a temperature of at least 26.5 Celsius degree (Gray, 1998; Montgomery & Farrell, 1993). Commonly, the tropical cyclones forming in the Atlantic and Northeast Pacific are called hurricane while those forming in the Northwest Pacific are called typhoon and, for the tropical cyclones in the South Pacific and the Indian Ocean, cyclone is their name (Montgomery & Farrell, 1993). In

addition, tropical cyclones rotate counter-clockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere because of the Coriolis force. On the basis of the maximum sustained wind velocity, tropical cyclones in North America can be classified as following (McAdie et al., 2009):

- Tropical Depression: maximum sustained wind speed is less than 38 mph (16.99 m/s).
- Tropical Storm: maximum sustained wind speed is between 39 and 73 mph (between 17.43 m/s and 32.63 m/s).
- Hurricane: maximum sustained wind speed is higher than 74 mph (33.08 m/s).

In more detail, the Hurricane can be further divided into five categories based on the sustained wind speed (Schott et al., 2012):

- Category 1 hurricane: The sustained wind speed is between 119 km/h and 153 km/h;
- Category 2 hurricane: The sustained wind speed is between 154 km/h and 177 km/h;
- Category 3 hurricane: The sustained wind speed is between 178 km/h and 208 km/h;
- Category 4 hurricane: The sustained wind speed is between 209 km/h and 251 km/h;
- Category 5 hurricane: The sustained wind speed is higher than 252 km/h.

The maximum sustained wind is the highest surface winds occurring within the circulation of a storm system where the surface winds are those at the standard meteorological height of 10 m without any block (Chris Landsea, 2006). From the classification, hurricane is the strongest type of tropical cyclones in North America which has horrible destructive power and is also considered as one of the most dangerous natural hazards. The most common hurricane-induced damages include storm surge, heavy rainfall, and strong wind.

The south and southeast of the United States are prone to hurricanes that form in the North Atlantic. The Atlantic hurricane season typically lasts from 1 June to 30 November every year while the period between August and October is the most active (Dorst, n.d.). Historical records of tropical cyclones in the Gulf of Mexico region suggest that the area is most prone to damaging hurricanes during the period between August and October (Blake, Landsea, & Gibney, 2007; Roth, 2010). In addition, Keim et al. (2007) have specifically analyzed the tropical storms and hurricanes occurred from 1901 to 2005 at 45 coastal locations along the Gulf of Mexico and Atlantic Coasts. Their study presents hurricane frequency maps which clearly show the hurricane landfall patterns at different locations based on the return periods of tropical cyclones with different intensities (i.e., tropical cyclone, hurricane and severe hurricane) and the overall tropical hazard index which is the index combined the potential wind, surge damage and the damage to society caused by tropical storms. The overall tropical hazard index can also be used to test the vulnerability of the coastal regions (Keim et al., 2007). Figure 2.6 shows part of the results of Keim et al. (2007) from which the return periods of tropical storms and hurricanes in the Houston-Galveston region are three years and eight years respectively that are relatively small when compared to all the tested locations. It indicates that the Galveston area suffers tropical cyclones with a relatively high frequency. In addition, the hazard index in Galveston area is relatively high which is 104 and ranked 12$^{th}$ of all the locations (Keim et al., 2007).

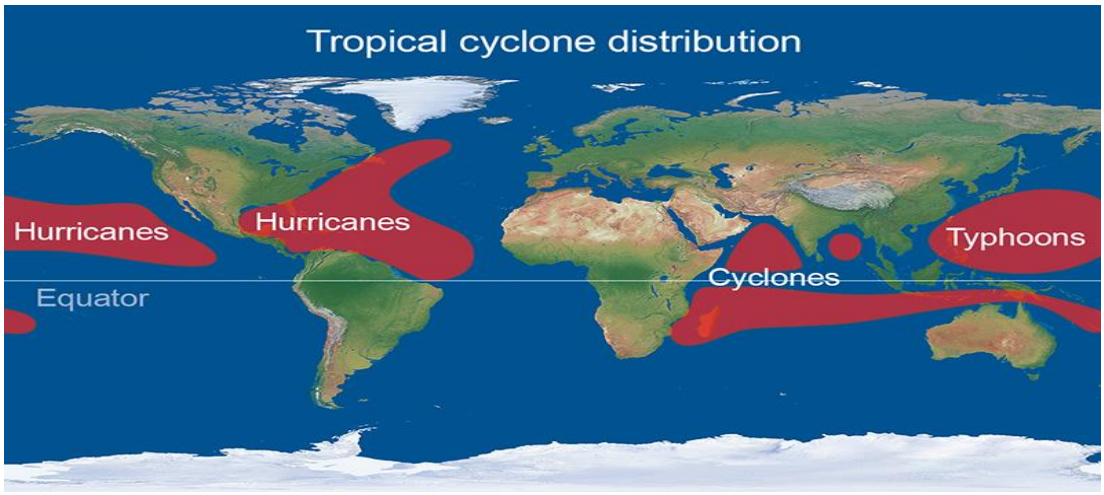



**Figure 2.5: Tropical cyclone distribution around the world. Source: ("Tropical cyclone facts," 2016).**

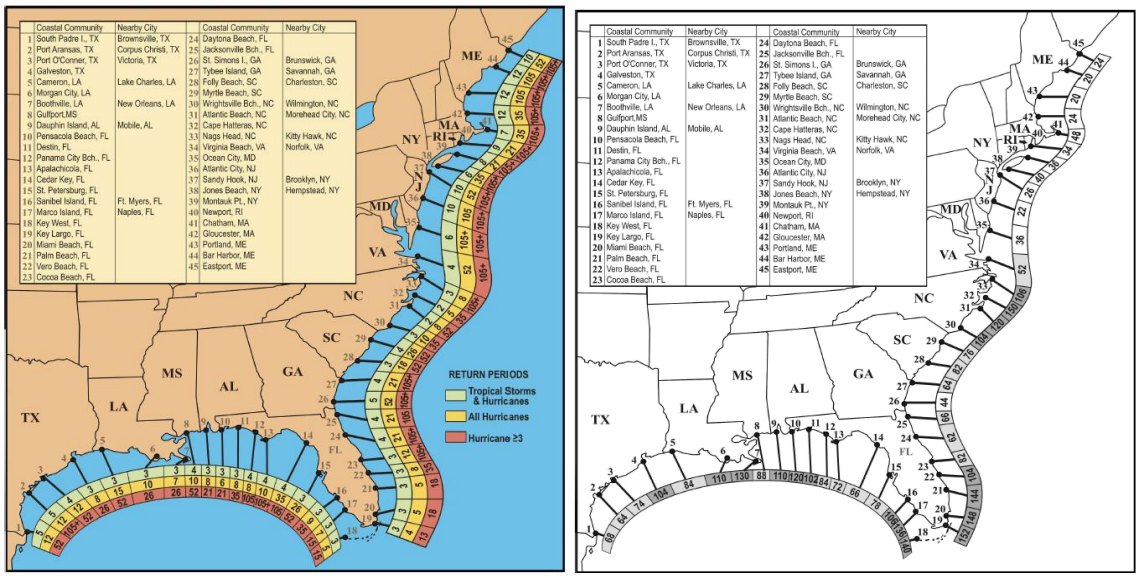



**Figure 2.6: (a) Averaged return period for tropical storms and hurricanes; (b) Hazard index of tropical cyclones. Source: (Keim et al., 2007).**

## 2.2.2 Precipitation

The climate of Greater Houston belongs to humid subtropical and the precipitation mainly composed of rainfall. Figure 2.7 exhibits the monthly precipitation normal of Houston Area and Galveston Area during 1981 and 2010 from which it is observed that, for the Houston Area, the highest precipitation happens in June with a monthly precipitation normal of 5.93 inches (150.62 mm) followed by October with a value of 5.70 inches (144.78 mm) while, for Galveston Area, September and June are the months having largest precipitation normal with 6.03 inches (153.16 mm) and 5.69 inches (144.53 mm) respectively. In addition to the difference of the monthly precipitation normal, there is a common trend in the Houston-Galveston area that the average amount of precipitation normal during the Atlantic hurricane season (from June to November) is larger than the average value of other months, which also proves that compound flood event is likely to happen in the Houston-Galveston area.

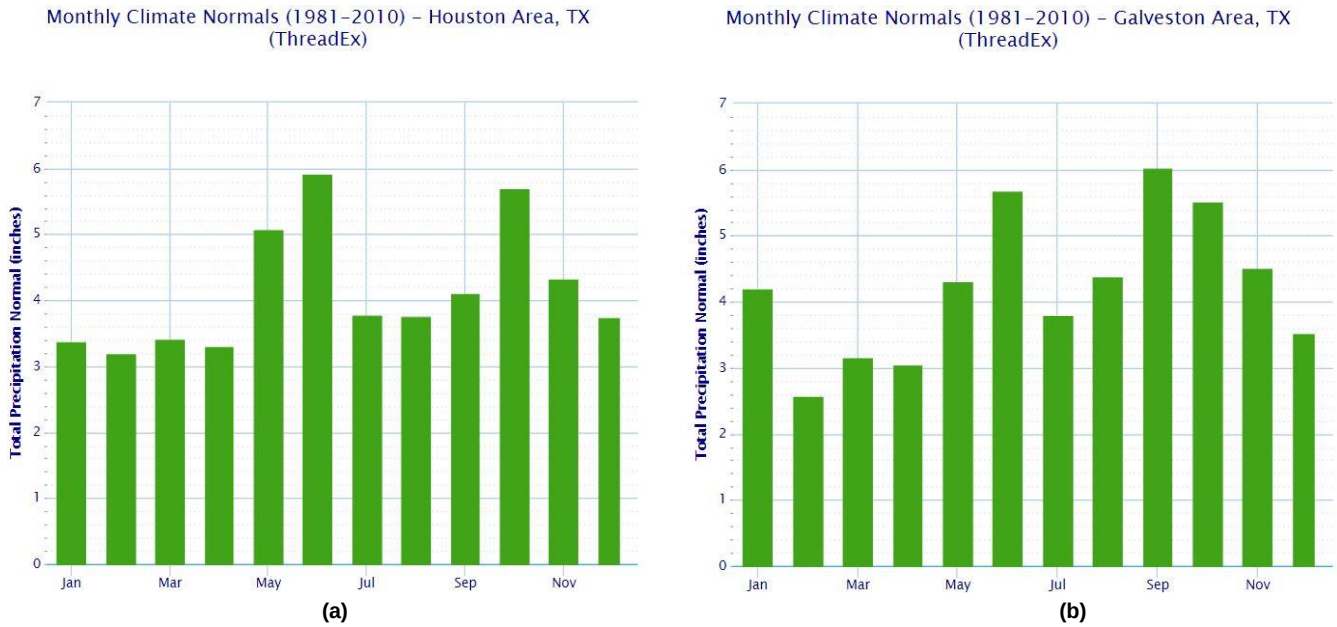



**Figure 2.7: (a) The total precipitation normal in Houston Area (unit: inch); (b) The total precipitation normal in Galveston Area (unit: inch). source: (NOAA, 2014).**

## 2.3. Historical Flood Events

According to the previous introduction, it is already known that floods in Greater Houston could be induced by hurricanes which can cause storm surge and heavy rainfall. In this section, selected historical flood events in the Greater Houston are briefly introduced to provide an example of the flood risk posed by severe storm surge and precipitation events to Greater Houston.

**The Great Galveston Hurricane (1900)**

The Great Galveston Hurricane is classified as a category 4 hurricane which had a maximum wind speed of approximately 233 km/h and made landfall at Galveston, Texas on September 8, 1900. The storm surge triggered by this hurricane was up to 15 feet (4.6m) which almost demolished the whole Galveston island. The death toll in this disaster is about 8000 which is the deadliest U.S. hurricane event in the record according to the work of Blake et al. (2007) (see Figure 2.8). In addition, more than 3600 buildings were destroyed (Hughes, 1998; Roth, 2010; “The Galveston hurricane of 1900,” 2013).

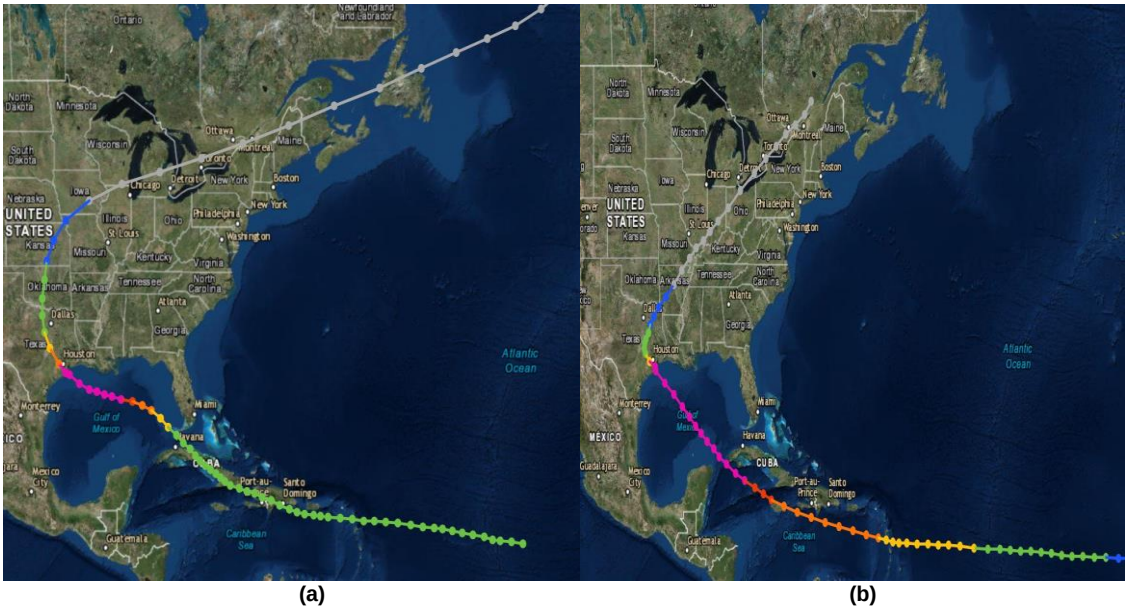
| RANK | HURRICANE | YEAR | CATEGORY | DEATHS |
|---|---|---|---|---|
| 1 | TX (Galveston) | 1900 | 4 | 8000 [a] |
| 2 | FL (SE/Lake Okeechobee) | 1928 | 4 | 2500 [b] |
| 3 | *KATRINA* (SE LA/MS) | 2005 | 3 | 1200 |
| 4 | LA (Cheniere Caminanda) | 1893 | 4 | 1100-1400 [c] |
| 5 | SC/GA (Sea Islands) | 1893 | 3 | 1000-2000 [d] |
| 6 | GA/SC | 1881 | 2 | 700 |
| 7 | *AUDREY* (SW LA/N TX) | 1957 | 4 | 416 [h] |
| 8 | FL (Keys) | 1935 | 5 | 408 |
| 9 | LA (Last Island) | 1856 | 4 | 400 |
| 10 | FL (Miami)/MS/AL/Pensacola | 1926 | 4 | 372 |
| 11 | LA (Grand Isle) | 1909 | 3 | 350 |
| 12 | FL (Keys)/S TX | 1919 | 4 | 287 [e] |
| 13 | LA (New Orleans) | 1915 | 3 | 275 [i] |
| 13 | TX (Galveston) | 1915 | 4 | 275 |
| 15 | New England | 1938 | 3 | 256 [e] |
| 15 | *CAMILLE* (MS/SE LA/VA) | 1969 | 5 | 256 |
| 17 | *DIANE* (NE U.S.) | 1955 | 1 | 184 |
| 18 | GA, SC, NC | 1898 | 4 | 179 |
| 19 | TX | 1875 | 3 | 176 |
| 20 | SE FL | 1906 | 3 | 164 |
| 21 | TX (Indianola) | 1886 | 4 | 150 |
| 22 | MS/AL/Pensacola | 1906 | 2 | 134 |
| 23 | FL, GA, SC | 1896 | 3 | 130 |
| 24 | *AGNES* (FL/NE U.S.) | 1972 | 1 | 122 [f] |
| 25 | *HAZEL* (SC/NC) | 1954 | 4 | 95 |
| 26 | *BETSY* (SE FL/SE LA) | 1965 | 3 | 75 |
| 27 | Northeast U.S. | 1944 | 3 | 64 [g] |
| 28 | *CAROL* (NE U.S.) | 1954 | 3 | 60 |
| 29 | *FLOYD* (Mid Atlantic & NE U.S.) | 1999 | 2 | 56 |
| 30 | NC | 1883 | 2 | 53 |
| 31 | SE FL/SE LA/MS | 1947 | 4 | 51 |
| 32 | NC, SC | 1899 | 3 | 50 [h,j] |
| 32 | GA/SC/NC | 1940 | 2 | 50 |
| 32 | *DONNA* (FL/Eastern U.S.) | 1960 | 4 | 50 |
| 35 | LA | 1860 | 2 | 47 [h] |
| 36 | NC, VA | 1879 | 3 | 46 [h,j] |
| 36 | *CARLA* (N & Central TX) | 1961 | 4 | 46 |
| 38 | TX (Velasco) | 1909 | 3 | 41 |
| 38 | *ALLISON* (SE TX) | 2001 | TS [k] | 41 |
| 40 | Mid-Atlantic | 1889 | TS [l] | 40 [h,j] |
| 40 | TX (Freeport) | 1932 | 4 | 40 |
| 40 | S TX | 1933 | 3 | 40 |

| RANK | HURRICANE | YEAR | CATEGORY | DEATHS |
|---|---|---|---|---|
| 43 | *HILDA* (LA) | 1964 | 3 | 38 |
| 44 | SW LA/Upper TX | 1918 | 3 | 34 |
| 45 | SW FL | 1910 | 3 | 30 |
| 45 | *ALBERTO* (NW FL, GA, AL) | 1994 | TS [k] | 30 |
| 47 | SC, FL | 1893 | 3 | 28 [m] |
| 48 | New England | 1878 | 2 | 27 [h,n] |
| 48 | Texas | 1886 | 2 | 27 [h] |
| 50 | *ANDREW* (S FL, LA) | 1992 | 5 | 26 |
| 50 | *FRAN* (NC) | 1996 | 3 | 26 |
| 52 | LA | 1926 | 3 | 25 |
| 52 | *CONNIE* (NC) | 1955 | 3 | 25 |
| 52 | *IVAN* (NW FL, AL) | 2004 | 3 | 25 |
| ADDENDUM (Not Atlantic/Gulf Coast) | | | | |
| 2 | Puerto Rico (San Ciriaco) | 1899 | 3 | 3369 [i] |
| 6 | P.R., USVI (San Narcico) | 1867 | 3 | 811 [i,j] |
| 6 | Puerto Rico (San Lorenzo) | 1852 | 1 | 800 [f,o] |
| 12 | Puerto Rico (San Felipe) | 1928 | 5 | 312 |
| 17 | USVI, P.R. (San Ciprian) | 1932 | 2 | 225 |
| 25 | *DONNA* (St. Thomas, VI) | 1960 | 4 | 107 |
| 25 | Puerto Rico (San Gil) | 1888 | 1 | 100 [h] |
| 38 | Southern California | 1939 | TS [k] | 45 |
| 38 | *ELOISE* (Puerto Rico) | 1975 | TS [k] | 44 |
| 48 | USVI (Santa Juana") | 1871 | 3 | 27 [h] |
| 52 | Puerto Rico (San Liborio) | 1926 | 2 | 25 |

Notes:

| | |
|---|---|
| a | Could be as high as 12,000 |
| b | Could be as high as 3000 |
| c | Total including offshore losses near 2000 |
| d | August |
| e | Total including offshore losses is 600 |
| f | No more than |
| g | Total including offshore losses is 390 |
| h | At least |
| i | Puerto Rico 1899 and NC, SC 1899 are the same storm |
| j | Could include some offshore losses |
| k | Only of Tropical Storm intensity. |
| l | Remained offshore |
| m | Mid-October |
| n | Four deaths at shoreline or just offshore |
| o | Possibly a total from two hurricanes |

**Figure 2.8: The hurricanes leading to 25 or more death from 1851 to 2010 in the United States. Source:(Blake et al., 2007).**




**(a)** **(b)**

**Figure 2.9: (a) The track of the Great Galveston Hurricane in 1900; (b) The track of the 1915 Galveston Hurricane. Source: ("Historical Hurricane Tracks," n.d.).**

**The 1915 Galveston Hurricane (1915)**

After the Great Galveston Hurricane in 1900, the Galveston area suffered a huge loss again from the hurricane during August 16 and 19, 1915 which made landfall about 48 km southwest of Galveston and was a hurricane of category 3. A 12 ft. (about 3.65m) storm surge inundated the island and about 275 people lost their lives (Roth, 2010; "Upper Texas Coast Tropical Cyclones in the 1910s," n.d.).

**Tropical Storm Allison (2001)**

Tropical Storm Allison formed over the northwestern Gulf of Mexico and moved to the inland over the upper Texas coast (Stewart, 2011). Allison was not a strong storm which however brought extremely heavy rainfall in Texas which was mainly due to its slow movement and long stay in Texas (Roth, 2010). In Figure 2.11, the total rainfall in Harris County from June 5 to June 9, 2001, is depicted, it can be seen that the highest amount of rainfall totally noted were 38.8 inches (985.52 mm) ("Tropical Storm Allison," n.d.). Its heavy rainfall led to significant flooding over the upper Texas coastal area which caused a total loss of about $5 billion while nearly 96 percent of the loss occurred in Houston area. About 22 people in Texas died in this disaster which is one of the worst tropical storms in the history of America (Roth, 2010; Stewart, 2011).

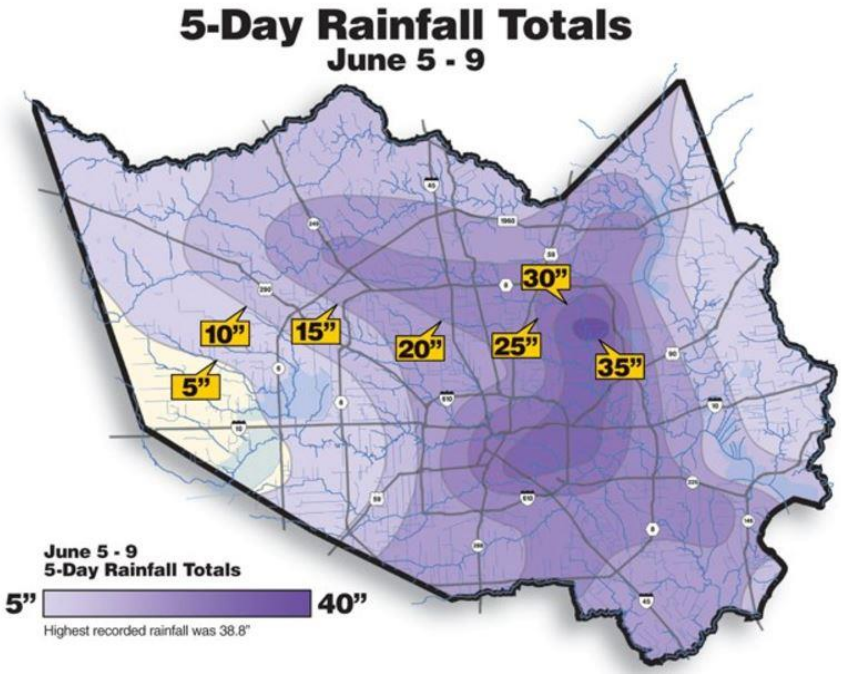



**Figure 2.10: Five-day rainfall total of Harris County during Tropical Strom Allison. Source: ("Tropical Storm Allison," n.d.).**

**Hurricane Ike (2008)**

Hurricane Ike was produced from a tropical wave in North Atlantic on September 1, 2008, and made landfall on Galveston Island on September 13 with the intensity of category 2 based on the Saffir-Simpson Hurricane Scale (Berg, 2009; Roth, 2010). After the landfall, severe storm surge flooding and heavy rainfall flooding caused considerable damage which made Hurricane Ike become the second costliest hurricane from 1900 to 2010 (Blake et al., 2007). The storm surge height across Harris County ranged from 12 feet (3.66m) to 15 feet (4.57m) and the averaged storm surge level was close to the level with 100-year return period for Harris County ("Hurricane Ike," n.d.). At the same time, hurricane Ike caused two heavy rainfall the first of which resulted in 6-10 inches (152.4-254 mm) across Harris County while the second one resulted in 3-8 inches (76.2-203.2 mm). Nearly 92,000 homes and 7,100 businesses were

damaged in the disaster which caused 2,400 injuries (“Hurricane Ike,” n.d.). In Figure 2.10, the inundation map of Hurricane Ike is exhibited which shows that the maximum inundation depth along the Houston Ship Channel was over 10 ft. (about 3.05 m).

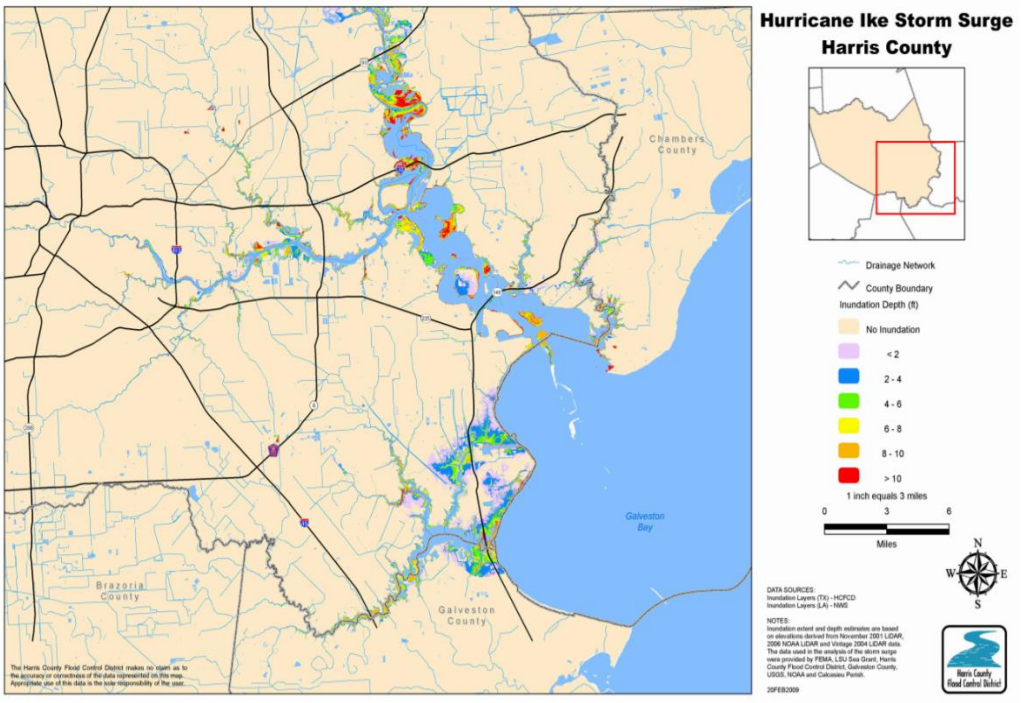



**Figure 2.11: The inundation map of Hurricane Ike. Source: (“Hurricane Ike,” n.d.).**

### Hurricane Harvey (2017)

Hurricane Harvey made landfall as a Category 4 hurricane near Rockport, Texas on August 25, 2017. During this event, the storm brought considerable rainfall to the Houston region and water levels in Galveston Bay remained high for several days at the same time. Although there is no official estimation of the loss caused by Hurricane Harvey, it is still believed that Hurricane Harvey could be the most expensive or second-most expensive natural disaster in the United States (Quealy, 2017). A detailed study of Hurricane Harvey will be presented in Chapter 7.

## 2.4. Flood Risk Reduction Strategies in Houston-Galveston Bay Region

Since Hurricane Ike, increasing attention has been put into the flood risk reduction strategies for Houston-Galveston Region. For example, the Severe Storm Prediction, Education and Evacuation from Disasters (SSPEED) Center has proposed a comprehensive solution i.e., the Houston-Galveston Area protection System (H-GAPS) to protect that region from storm surge. (Christian et al., 2014; Jonkman et al., 2013; Penland & Moore, n.d.; Torres et al., 2015; SSPEED, 2015). Figure 2.12 exhibits the sketch map of H-GAPS from which multiple defensive lines are combined to protect Houston-Galveston area. At the outer line, it is suggested to raise the existing roads i.e. U.S. Highway 87 (F) at Bolivar Island and FM 3005 (G) at Galveston Island which combined with the Galveston levee (H), Galveston seawall (1) and the proposed Bolivar Roads Gate (L) to form a continuous coastal barrier protecting the storm surge from overflowing Galveston Island and Bolivar Island. For the surge created in the Galveston Bay, it is proposed to extend the existing dredged spoils (E) or build the oyster reefs (D) to reduce the wind setup inside Bay which would be combined with a gate across the HSC (see M). In addition, a storm surge barrier is also proposed to build at the outlet of the HSC to Galveston Bay (U) to protect the area behind where Port of Houston locates (Bedient, 2015; SSPEED, 2015).


**Figure 2.12: The sketch map of H-GAPS. Source: (SSPEED, 2015).**

In more detail, to protect the HSC, SSPEED proposed the Centennial Gate Project and gave the alternative solutions (SSPEED, 2014). Figure 2.13 shows the potential alignments of these two solutions and both two alignments include a gate and a levee. The combination of the gate-levee system connects two sides of the outlet of the HSC. The difference of alternatives is the location, in solution A, the gate-levee system is near the downstream of Fred Hartman Bridge while the alignments in solution B is suggested to near the Morgan's Point (SSPEED, 2014). In this thesis, it is focused on the solution A and the location of the proposed storm surge barrier i.e. near the downstream of Fred Hartman Bridge will be selected as the downstream boundary in case of analyzing the compound flood events later.


(a) (b)

**Figure 2.13: (a) Solution A of Centennial Gate Project; (b) Solution B of Centennial Gate Project. Source: (SSPEED, 2014).**

## 2.5. A Review of the Literature on Flood Risk in the HSC

Constructing a storm surge barrier may also raises a problem that the flooding of the HSC may be caused by the combination of storm surge and upstream rainfall-runoff. Once the barrier is closed, the upstream flow cannot flow out to the Galveston Bay that may cause the flooding behind the barrier when the closure time is long and upstream discharge is large enough. To analyze the flood risk in the HSC resulting from compound events and the potential influence of the proposed storm surge barrier on flood frequency, many efforts have been made. Christian et al., (2014) did some study to understand the interaction between storm surge and rainfall-runoff. In their research, the proposed storm surge barrier is modelled as a 4.8-m barrier which connects the topographic ridges on both sides of Lower San Jacinto River with elevations of 7.6 m above MSL and the gate is modelled as a sequence of individual binary gates. The authors used ADCIRC which can provide the prediction of storm surge and flooding in two or three dimensions to simulate the storm surge as the downstream boundary condition and selected *Vflo*, a hydrologic model that is able to simulate runoff, to compute the upstream flow hydrographs through combing the historical radar rainfall for Hurricane Ike and topography. The simulated boundary conditions are inputted in HEC-RAS to get the time-series water surface elevations. Three scenarios (i.e., Hurricane Ike, Ike plus 15% wind speed, and Ike plus 30% wind speed) were considered in this research and the result shows that the proposed storm surge can reduce the water surface level behind the gate well for all scenarios (see Figure 2.14). Although the result is positive, the lack of scenarios makes the research uncompleted since only Hurricane Ike rainfall and storm surge was used. Based on the effort of Christian et al. (2014), Torres et al. (2015) did some improvement. The similar method used by Christian et al. (2014) was used again. The difference is that they combined the ADCIRC and SWAN to build different storm surge scenarios with different landfall locations. And Hurricane Katrina (2005), Ike (2008) and Isaac (2012) with three shifted landfall locations are conducted in this research. Although a lot of research has been done, the simulation of more hurricane events is still necessary which is important to understand more about joint probable combinations of rainfall-runoff and storm surge in the HSC and influence of a storm surge barrier on flood frequencies during compound events.

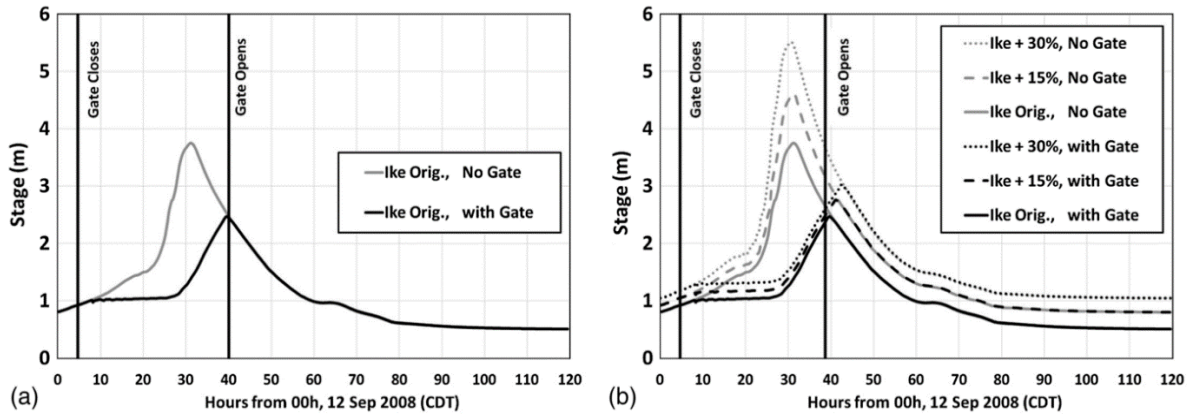



**Figure 2.14: Simulated stage hydrograph upstream of the proposed storm surge barrier for (a) Hurricane Ike surge; (b) Hurricane surge of Hurricane Ike plus 15% and 30% wind speed. Source: (Christian et al., 2014).**

In addition, Couasnon (2017) made effort on the compound flood hazard at the upstream reach of the HSC. In her research, the water level at Lynchburg Landing and discharges at the outlets of Brays Bayou, Buffalo Bayou, White Oak Bayou, Sims Bayou, Hunting Bayou, Greens Bayou, and Vince Bayou are taken as the main variables which are considered as the main sources triggering the compound flood events. A Bayesian Network was constructed to build a joint distribution of the above variables through which stochastic scenarios were generated. To study the water surface profile of upstream reach of the HSC, Couasnon (2017) built a process-based 1-D hydraulic model to estimate the water levels under the compound flood events. According to work of Couasnon (2017), the Bayesian Network is proven to be a reliable addition to the flood risk analysis. However, in the research of Couasnon (2017), neither the downstream reach of the HSC under compound flood events nor the discharge from San Jacinto River were considered. Therefore, in this thesis, the downstream reach of the HSC is analyzed and a more completed Bayesian Network including San Jacinto River is constructed.

## 2.6. Summary

In this Chapter, through the review of Texas, Greater Houston and the HSC, the importance of the HSC is shown which serves as critical for the economy of the entire U.S. However, Greater Houston is also prone to flooding caused by both storm surge and heavy rainfall and the potential damage of flooding would be large, which has been proved by many notable historical flood events including Hurricane Ike (2008) and Hurricane Harvey (2017). Especially Hurricane Ike (2008) highlighted the vulnerability of the Houston Ship Channel after which many researches and government-organized groups in the region have proposed a few possible mitigation strategies to protect the Houston-Galveston region from storm surge. In this thesis, it is focused on the proposal by the SSPEED Center which suggests building a movable storm surge barrier near the downstream of Fred Hartman Bridge. However, this proposal could lead to the flooding behind the barrier in case of compound flood events. Meanwhile, many researches have been done focusing on the flood risk in the upstream reach of the HSC, but little is known about the situation in the downstream reach of the HSC. Therefore, this thesis will focus on the analysis of compound flood events in the downstream reach of the HSC from a perspective of probability. And the research questions mentioned in Section 1.3 will be answered in the following chapters.

# 3

# Data Collection and Analysis

As introduced in the previous chapter, the compound flood event in the HSC area mainly refer to the joint occurrence of storm surge and inland rainfall-runoff. Naturally, storm surge and rainfall-runoff are the most direct variables related to the compound flood event. However, it is very difficult to collect large amounts of storm surge and rainfall-runoff data directly. Therefore, to simplify the problem, it is assumed that the characteristics of storm surge and rainfall-runoff can be reflected on the sea level and riverine discharge respectively and the compound flood event can then be transferred to the joint occurrence of high sea level and high river flow. In Section 3.1, the boundaries of the research area are selected in order to determine where the sea level and riverine discharge data are most needed in this thesis. Section 3.2 describes the data collection and selection. It is worth noting here that there is no gauging station deployed at the selected boundaries; therefore, some assumptions are made in Section 3.2 in order to link the sea level and discharge at the boundaries to the data that can be obtained. Thereafter, the detailed data analysis is operated in Section 3.3. Some discussions about the data selection are shown in Section 3.4 and the work of this chapter is summarized in Section 3.5.

## 3.1. Boundary Selection

**Upstream Boundary**

According to the introduction of the HSC in Section 2.1.4, the upstream discharges flowing into the downstream reach of the HSC mainly come from the Buffalo Bayou and San Jacinto River. Therefore, it is also reasonable to divide the upstream boundary into two parts:

(1) For the discharge from Buffalo Bayou and its tributaries i.e. Brays Bayou (BB), Sims Bayou (SB), Hunting Bayou (HB), Vince Bayou (VB) and Greens Bayou (GB), the mouth of Buffalo Bayou to the downstream reach of the HSC (29.761408°N 95.086903°W) is defined as upstream boundary one (U1).

(2) For the discharge from San Jacinto River, the point just upstream of the confluence of the San Jacinto River and the HSC (29.763150°N 95.080690°W) is defined as upstream boundary two (U2).

The detailed positions of upstream boundaries are shown in Figure 3.1.

**Downstream Boundary**

For the downstream boundary (D), the location of the proposed storm surge barrier (near the downstream of Fred Hartman Bridge) is selected as the downstream boundary (see Figure 3.1).

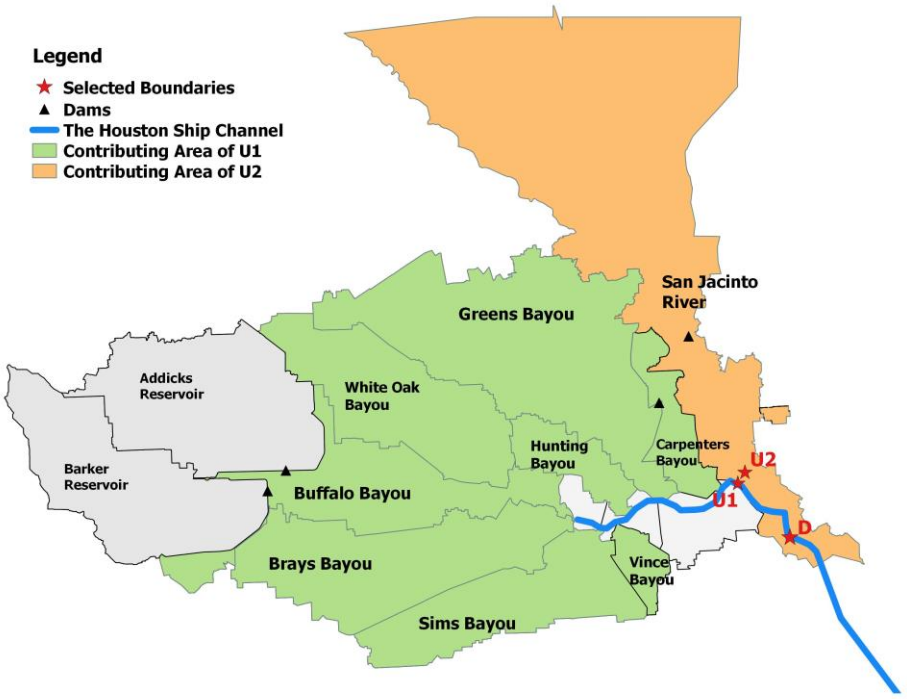



**Figure 3.1: A simplified map of the study area. Source: HCFCD.**

## 3.2. Data Collection

After confirming the boundaries of the research area, it is found that the there is no gauging station deployed at the boundaries; so, the sea level data at downstream boundary and discharge data at upstream boundaries cannot be collected directly. To solve this problem, the discharge at the upstream boundary U1 is assumed as the summation of the discharges at the outlets of Buffalo Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, Greens Bayou, Carpenters Bayou, and Vince Bayou; and the discharge at the upstream boundary U2 is assumed totally from the discharge of San Jacinto River (see Figure 3.1). In addition, the sea level at the downstream boundary is assumed to be equal to the sea level measured at the nearest gauging station. Based on the above assumptions, the collection of boundary data is converted into collecting the discharge data or water level data in Buffalo Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, Greens Bayou, Carpenters Bayou, Vince Bayou and San Jacinto River and the nearest sea level data to the downstream boundary. Here, the riverine water level data is also considered since it can be converted into discharge, which will be introduced later.

### 3.2.1 Gauging Station Selection

The coastal and riverine data are collected from National Oceanic and Atmospheric Administration (NOAA) and U.S. Geological Survey (USGS) respectively. The tide station at Morgan's Point (NOAA station 8770613) is the nearest station to the downstream boundary which is therefore selected and the sea level at Morgan's Point is taken as the downstream boundary condition. In addition, another tide station at Galveston Pier 21 (NOAA station 8771450) is selected since approximately one-century sea level data are captured at this station. The sufficient data is valuable for studying the pattern of sea level change, especially under the extreme conditions. There is one point needs to be noted that, the downstream sea level is defined here as the sum of the tide level and residual water level as shown in Eq. 3.1 where the residual water level includes all the influential factors of sea level except tide (e.g., storm surge, wind setup). Naturally, the characteristics of hurricane-induced storm surge can be mainly reflected on the residual water level. Therefore, in the following study, the residual water level will be analyzed emphatically. This way of dealing with the hurricane-induced flood event is also suggested by Couasnon (2017).

$$H_{downstream} = H_{tide} + H_{residual\ water\ level} \quad \text{[Eq. 3.1]}$$

For the upstream boundary U1, since gauging stations at the outlets of Buffalo Bayou, White Oak Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, Greens Bayou, Carpenters Bayou, and Vince Bayou are non-existent. Therefore, those stations nearest to the outlets are selected. In more detail, USGS 08074710 (TB) is selected representing the condition of Buffalo Bayou and White Oak Bayou, USGS 08075000 (BB) is selected representing the condition of Brays Bayou, USGS 08075500 (SB) is selected representing the condition of Sims Bayou, USGS 08075730 (VB) is selected representing the condition of Vince Bayou, USGS 08075770 (HB) is selected representing the condition of Hunting Bayou, USGS 08076700 (GB) is selected representing the condition of Greens Bayou. Through the selection of riverine stations, almost all the watersheds influencing on the discharge of U1 are included except Carpenters Bayou since there is no gauging station in Carpenters Bayou. Fortunately, Carpenters Bayou is a relatively small watershed with a contributing area of nearly 81 $km^2$ while the largest watershed, Greens Bayou, having a contributing area of almost 550 $km^2$. Therefore, the influence of the discharge from Carpenters Bayou on the U1 is small. For the upstream boundary U2, USGS 08072050 (SJR), the nearest station to the U2, is selected. The detailed information of selected stations is listed in Table 3.1 and Table 3.2. The locations of all the selected stations are shown in Figure 3.2.

**Table 3.1: Original selected coastal data. Data source: NOAA.**

| Stations | Location | Abbreviation of the station | Data type | Data Period |
|---|---|---|---|---|
| NOAA 8771450 | Galveston Pier 21, TX | GP21 | Hourly water level | 01/01/1904-31/01/2017 |
| NOAA 8770613 | Morgan's Point, TX | Morgan | Hourly water level | 17/03/1993-31/01/2017 |

**Table 3.2: Original selected riverine data. Data source: USGS.**

| Stations | Location | Abbreviation of the station | Data type | Data Period |
|---|---|---|---|---|
| USGS 08074710 | Turning Basin, TX | TB | Daily maximum water level | 01/10/1987-01/03/2017 |
| USGS 08075000 | Brays Bayou, Houston, TX | BB | Daily mean discharge | 25/05/1936-01/03/2017 |
| USGS 08075000 | Brays Bayou, Houston, TX | BB | Daily mean gage height | 08/06/1988-01/03/2017 |
| USGS 08075500 | Sims Bayou, Houston, TX | SB | Daily mean discharge | 01/10/1952-30/09/1995 |
| USGS 08075500 | Sims Bayou, Houston, TX | SB | Daily maximum gage height | 01/10/1997-01/03/2017 |
| USGS 08075730 | Vince Bayou, Pasadena, TX | VB | Daily mean discharge | 01/10/1971-01/03/2017 |
| USGS 08075730 | Vince Bayou, Pasadena, TX | VB | Daily mean gage height | 10/05/1997-01/03/2017 |
| USGS 08075770 | Hunting Bayou, IH610, TX | HB | Daily mean discharge | 14/04/1964-01/03/2017 |
| USGS 08075770 | Hunting Bayou, IH610, TX | HB | Daily mean gage height | 05/09/1996-01/03/2017 |
| USGS 08076700 | Greens Bayou, Ley Rd, TX | GB | Daily mean discharge | 02/12/1971-05/12/2016 |
| USGS 08076700 | Greens Bayou, Ley Rd, TX | GB | Daily mean gage height | 02/10/1997-01/03/2017 |
| USGS 08072050 | San Jacinto Rv, Sheldon, TX | SJR | Daily mean gage height | 17/08/1996-01/03/2017 |

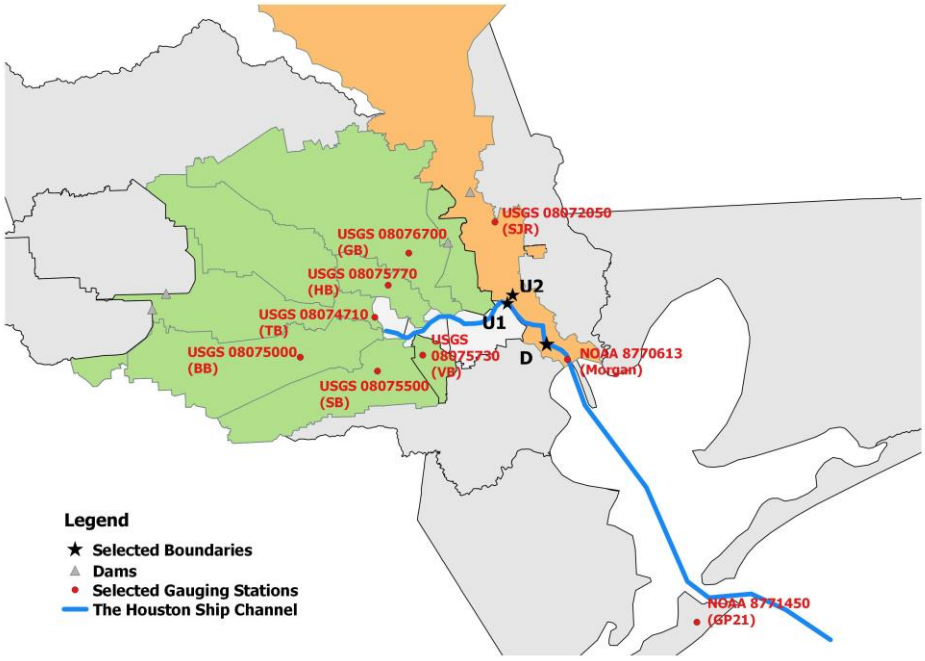



**Figure 3.2: The relative locations of all the selected stations. Source: HCFCD.**

Table 3.3 exhibits the contributing areas of all selected stations. It is worth noting here that the contributing area of USGS 08074710 (TB) consists of a partial contributing area of Buffalo Bayou and whole contributing area of White Oak Bayou. In addition, the contributing area of USGS 08072050 (SJR) only takes into account the part between the U2 and the Lake Houston Dam. All the contributing areas accounted here exclude that of the upstream reservoirs since it is assumed there is no water flowing from the reservoirs into the downstream areas in cases of compound flood events. From Table 3.3, the contributing area of USGS 08074710 (TB) covers almost 94 percent of the total contributing areas of the corresponding watersheds while the contributing areas of USGS 08075730 (VB) and USGS 08075770 (HB) only account for half of total contributing areas of Vince Bayou and Hunting Bayou respectively. On average, combining all the selected riverine stations, nearly 80 percent of the total contributing areas of U1 and U2 has been captured. In addition, the contributing area of Vince Bayou watershed is about 41 $km^2$ which is the smallest watershed less than one-tenth of Greens Bayou which is the largest watershed with a contributing area of 471.4 $km^2$.

**Table 3.3: Contributing area of the selected stations. Data source: USGS, HCFCD.**

| Gage station | Contributing area of the stations (unit: $km^2$) | Contributing area of corresponding watersheds (unit: $km^2$) | Coverage percentage of contributing areas of each station |
|---|---|---|---|
| USGS Station: 08074710 | 520.7 | 551.7 | 94.4% |
| USGS Station: 08075000 | 245.8 | 328.9 | 74.7% |
| USGS Station: 08075500 | 163.2 | 243.5 | 67.0% |
| USGS Station: 08075730 | 21.4 | 41.4 | 51.7% |
| USGS Station: 08075770 | 41.7 | 77.7 | 53.7% |
| USGS Station: 08076700 | 471.4 | 549.1 | 85.8% |
| Total contributing area of U1 | 1464.2 | 1792.3 | 81.7% |
| USGS Station: 08072050 | 132.1 | 202.02 | 65.4% |
| Total contributing area of U1 and U2 | 1596.3 | 1994.32 | 80.0% |

### 3.2.2 Riverine Data

In Table 3.2, one challenging thing can be found that the available data at USGS stations: 08075000 (BB), 08075500 (SB), 08075730 (VB), 08075770 (HB) and 08076700 (GB) include both discharge data and water level data while only water level data can be collected at the USGS stations 08074710 (TB) and 08072050 (SJR). As mentioned previously, the discharge is the optimal variable to replace rainfall-runoff. However, there are two problems coming out once the discharge data are selected for the stations having both discharge and water level data. The first problem is that there is no common data period between the discharge data of USGS 08075500 (SB) which ends at 30/09/1995 and water level data of USGS 08072050 (SJR) which starts from 17/08/1996. Because of this problem, it is impossible to calculate the dependence between the data at these two stations later, which will cause the difficulty of building Bayesian Network. Therefore, the water level data is chosen for USGS 08075500 (SB) which has an almost the same data period as that of USGS 08082050 (SJR). The second problem is that only a few discharge data is available at USGS 08076700 (GB). The amount of valid discharge data of USGS 08076700 (GB) is only around one thousand which is much less than other stations which may give rise to the inaccuracy of calculating dependencies between the data of USGS 08076700 (GB) and other stations. So, it is reasonable to use water level data for USGS 08086700 (GB).

### 3.2.3 Coastal Data

Another challenging thing is found by comparing the data types of NOAA stations to USGS stations that the data collected from NOAA stations is hourly data while the data of USGS stations is daily maximum data or daily mean data. For the consistency of data, the daily maximum residual water levels at NOAA 8771450 (GP21) and NOAA 8770613 (Morgan) are extracted which are calculated based on Eq. 3.1. The selected data information at all the stations is listed in Table 3.4. All the data are converted to S.I. and the same datum (NAVD88).

**Table 3.4: The data selected after the first data filtering. Data source: NOAA, USGS.**

| Stations | Data type | Observed data Period | Total number of days | Number of available data |
|---|---|---|---|---|
| NOAA 8771450 (GP21) | Daily maximum residual water level | 01/01/1904-31/01/2017 | 41305 | 41305 |
| NOAA 8770613 (Morgan) | Daily maximum residual water level | 17/03/1993-31/01/2017 | 8722 | 8722 |
| USGS 08074710 (TB) | Daily maximum water level | 01/10/1987-01/03/2017 | 10745 | 10288 |
| USGS 08075000 (BB) | Daily mean discharge | 25/05/1936-01/03/2017 | 29501 | 29501 |
| USGS 08075500 (SB) | Daily maximum water level | 01/10/1997-01/03/2017 | 7092 | 6631 |
| USGS 08075730 (VB) | Daily mean discharge | 01/10/1971-01/03/2017 | 16589 | 16584 |
| USGS 08075770 (HB) | Daily mean discharge | 14/04/1964-01/03/2017 | 19315 | 18947 |
| USGS 08076700 (GB) | Daily mean water level | 02/10/1997-01/03/2017 | 7091 | 6958 |
| USGS 08072050 (SJR) | Daily mean water level | 17/08/1996-01/03/2017 | 7502 | 7455 |

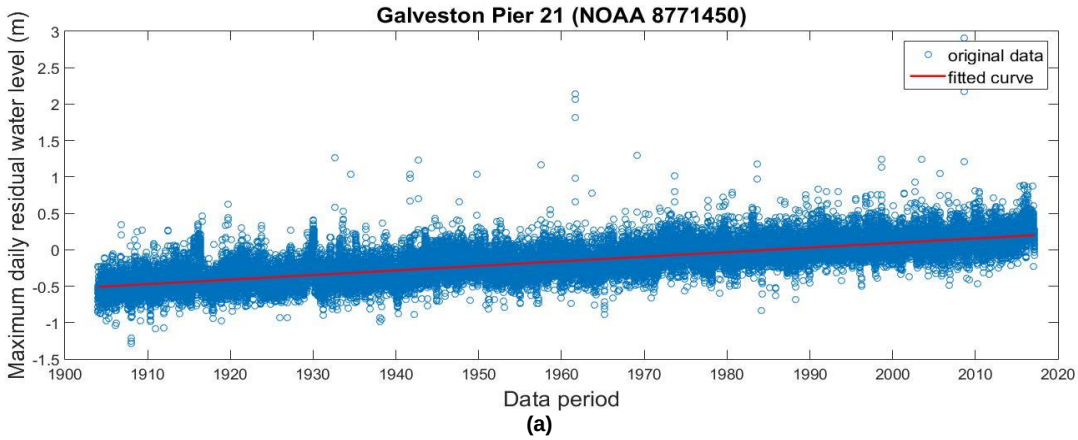
## 3.3. Data Analysis

After the first data filtering, it is still necessary to check whether all the data are suitable for this project or not. Therefore, the detailed data analysis is presented in this section to extract the most appropriate data for the project, which is mainly divided into two parts: data trend, Spearman's rank correlation.

### 3.3.1 Data Trend

Statistically, the trend existing in a data set may interference to the actual correlations with other data sets. For example, the trend of mean sea level rise may lead to an inaccurate measurement of the storm surge height, which may further influence the correlation between the storm surge height and rainfall-runoff in case of the compound flood events. Therefore, in this section, the collected data are plotted to check if any data trend exists.

In Appendix A.1, all the data trends are shown. For the data at NOAA station 8770613 (Morgan), USGS station 08074710 (TB), USGS station 08075500 (SB), USGS station 08075730 (VB), USGS station 08075770 (HB), USGS station 08076700 (GB) and USGS station 08072050 (SJR), there is no obvious trend observed while an almost linear increased trend is found for the daily maximum residual water level at NOAA station 8771450 (GP21) (see Figure 3.3a). The annual average daily maximum residual water level increases from -0.5m+NAVD88 in 1904 to approximately 0.2m+NAVD88 in 2016 with a rising rate of 6.25mm/yr which is close to the trend of mean sea level rise observed by NOAA (i.e. 6.47+/- 0.23mm/yr). Therefore, it is reasonable to consider that the increased trend of the data at NOAA station 8771450 (GP21) is mainly caused by the mean sea level rise. In this thesis, to analyze the influence of environmental hazards (i.e. hurricanes) on the residual water level, the trend of the mean sea level rise needs to be eliminated. Since the vertical datum NAVD88 is opted, the annual averaged daily maximum residual water level in 1988 is selected as the reference and all the data are shifted around it. A comparison between the original daily maximum residual water level and the detrend daily maximum residual water level at NOAA station 8771450 (GP21) is shown in 3.3b. In addition, it is observed that the discharge at USGS station 08075000 (BB) increased significantly before 1985, but stabilizes thereafter (see Figure 3.3c). Previous research attributes this trend to a rapid period of urbanization, which occurred during the period between 1950-1980 (Bass et al., 2016; Couasnon, 2017). Since the purpose of this thesis is dealing with the future extreme situations, the data at USGS station 08075000 (BB) before 01/01/1985 are excluded from the following analysis.




**(a)**

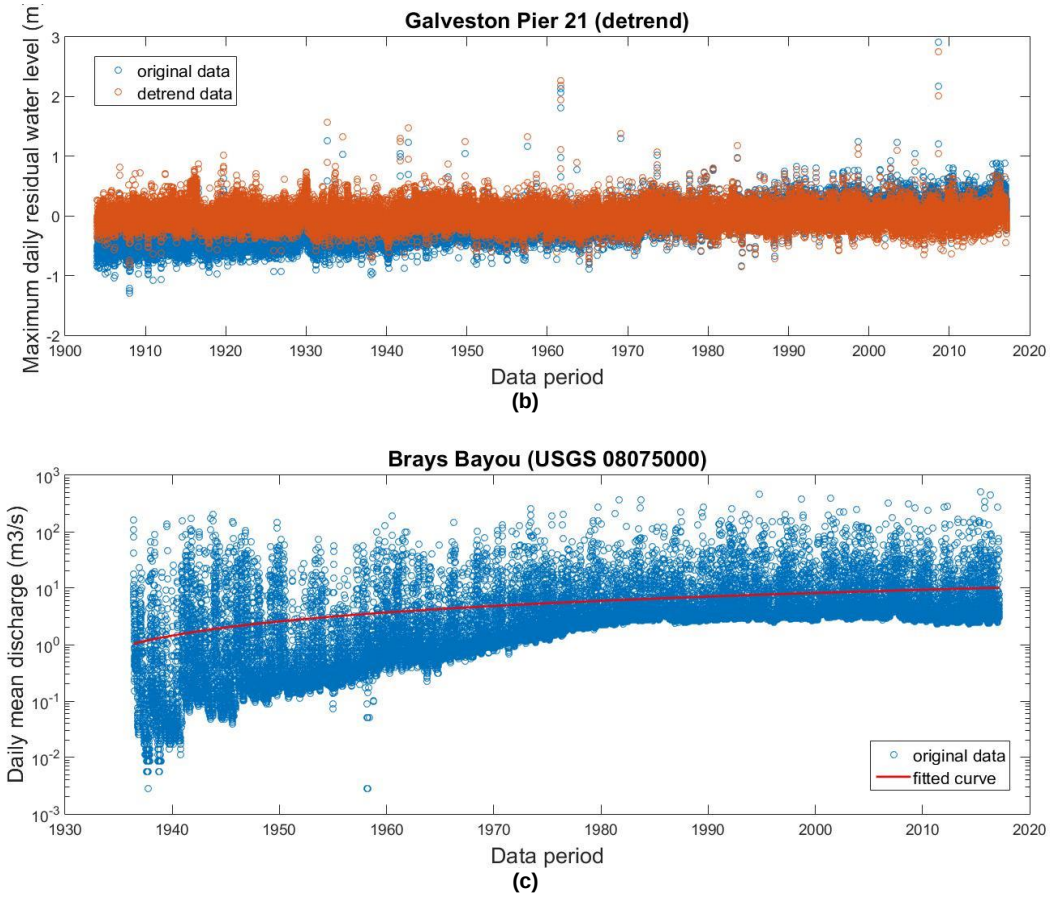



**Figure 3.3: (a) The data trend of the original residual water level at NOAA station 8771450; (b) The detrend residual water level at NOAA station 8771450; (c) The data trend of the discharge at USGS station 08075000. Data source: NOAA, USGS.**

### 3.3.2 Spearman's Rank Correlation

Spearman's rank correlation is the correlation of the ranks of data sets. For instance, the corresponding pair of ranks of a data pair $(y_{i1}, y_{i2})$, $i = 1,...,n$ is $(r_{i1,}\ r_{i2})$, where $r_{i1} = m$ when the $y_{i1}$ is the $m^{\text{th}}$ smallest data among the data set $(y_{11},...,y_{n1})$, and $r_{i2} = n$ when the $y_{i2}$ is the $n^{\text{th}}$ smallest data among the data set $(y_{12},...,y_{n2})$. Then, the Spearman's rank correlation of $(r_{i1}, r_{i2})$ can be expressed as Eq. 3.2 (Joe, 2015).

$$r = \frac{\mathrm{cov}(r_{i1}, r_{i2})}{\sigma_{r_{i1}} \sigma_{r_{i2}}} \qquad \text{[Eq. 3.2]}$$

Spearman's rank correlation is used here since it can detect the monotone dependence of variable pairs based on which the unique copula of the variable pairs can be determined (Genest & Favre, 2007; Joe, 2015). And the copula is also critical to build the Bayesian Network which will be introduced in Chapter 4. In Table 3.5, the Spearman's rank correlations of all the possible variable pairs are shown and more detailed information is listed in Appendix A.2. The meaning of each abbreviation shown in Table 3.5 can be referred to Table 3.4.

**Table 3.5: Rank correlations of the selected variable pairs.**

| | GP21 | Morgan | TB | BB | SB | VB | HB | GB | SJR |
|---|---|---|---|---|---|---|---|---|---|
| GP21 | 1.0000 | 0.8624 | 0.7092 | 0.2267 | 0.6867 | 0.2758 | 0.2925 | 0.6772 | 0.6804 |
| Morgan | 0.8624 | 1.0000 | 0.7062 | 0.2378 | 0.6727 | 0.2441 | 0.2484 | 0.6458 | 0.6716 |
| TB | 0.7092 | 0.7062 | 1.0000 | 0.2413 | 0.9059 | 0.2172 | 0.1780 | 0.7405 | 0.8139 |
| BB | 0.2267 | 0.2378 | 0.2413 | 1.0000 | 0.3170 | 0.5805 | 0.7171 | 0.5025 | 0.2507 |
| SB | 0.6867 | 0.6727 | 0.9059 | 0.3170 | 1.0000 | 0.2891 | 0.2446 | 0.7637 | 0.7863 |
| VB | 0.2758 | 0.2441 | 0.2172 | 0.5805 | 0.2891 | 1.0000 | 0.5739 | 0.4418 | 0.2248 |
| HB | 0.2925 | 0.2484 | 0.1780 | 0.7171 | 0.2446 | 0.5739 | 1.0000 | 0.4552 | 0.2217 |
| GB | 0.6772 | 0.6458 | 0.7405 | 0.5025 | 0.7637 | 0.4418 | 0.4552 | 1.0000 | 0.7983 |
| SJR | 0.6804 | 0.6716 | 0.8139 | 0.2507 | 0.7863 | 0.2248 | 0.2217 | 0.7983 | 1.0000 |

As shown in Table 3.5, on one hand, the rank correlations between the daily maximum residual water level at NOAA 8771450 (GP21), NOAA 8770613 (Morgan) and riverine water level at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08076700 (GB), USGS 08072050 (SJR) are strong, all of which are higher than 0.65. In addition, strong rank correlations can also be found between the discharge at USGS 08075000 (BB), USGS 08075730 (VB) and USGS 08075770 (HB) which are around 0.6. On the other hand, weak correlations exist between the discharge at USGS 08075000 (BB), USGS 08075730 (VB), USGS 08075770 (HB) and daily maximum water level at NOAA 8771450 (GP21), NOAA 8770613 (Morgan) and the riverine water level at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08072050 (SJR) that are less than 0.3. From the results of the rank correlations, predicting reliable discharge value at BB, HB, and VB through the data at other stations will be difficult because of the weak dependencies. In addition, through the previous analysis of contributing area shown in Section 3.2.1, it is already known that Vince Bayou is the smallest watershed among all the watersheds and its discharges are also small. Therefore, it is reasonable to consider the discharge of Vince Bayou as a negligible influential factor to the discharge of the U1. Thus, the discharge data at USGS 08075730 (VB) is abandoned, which can release some calculation burden of the Bayesian Network as well in the following study. It is worth mentioning in advance that the data period of variables used to construct the Bayesian Network shall be the same otherwise it is impossible to calculate the rank correlation of a random variable pair in the Bayesian Network. The maximum overlapped data period of all the variables is from 02/10/1997 to 31/01/2017 which includes the data of almost twenty years. Combining all the information above, the overall information of data selected for constructing the BN is shown in Table 3.6.

**Table 3.6: The data selected for building the Bayesian Network. Data source: NOAA, USGS.**

| | Stations | Data type |
|---|---|---|
| Selected stations and corresponding data types for BN model | NOAA 8771450 (GP21) | Daily maximum residual water level |
| | NOAA 8770613 (Morgan) | Daily maximum residual water level |
| | USGS 08074710 (TB) | Daily maximum water level |
| | USGS 08075000 (BB) | Daily mean discharge |
| | USGS 08075500 (SB) | Daily maximum gage height |
| | USGS 08075770 (HB) | Daily mean discharge |
| | USGS 08076700 (GB) | Daily mean gage height |
| | USGS 08072050 (SJR) | Daily mean gage height |
| Overlapped data Period | 02/10/1997-31/01/2017 | |
| Total number of overlapped days | 7062 | |
| Number of available data | 5664 | |

## 3.4. Discussion

### 3.4.1 Influence of Reducing Data

Comparing the final data period opted for constructing the BN to the original collected data period at each station, it is found that a lot of data is abandoned at NOAA 8771450 (GP21), NOAA 8770613 (Morgan), USGS 08074710 (TB), USGS 08075000 (BB), USGS 08075770 (HB) and USGS 08072050 (SJR). The NOAA 8771450 (GP21) lost 35641 data which is the worst among all the stations followed by USGS 08074710 (BB) and USGS 08074710 (HB) losing 23837 data and 13283 data respectively. Reducing data may cause the inaccuracy of the research. Therefore, it is necessary to verify the influence of reducing data.

**Rank correlation**

One of the potential influence of reducing data is reflected on the rank correlations. A comparison between the rank correlations of all the variable pairs with originally collected data and final selected data respectively are shown in Table 3.7. It is observed that the largest change of rank correlation happened for the variable pair (GP21-BB) which increases from 0.2267 to 0.2864 with a rise of 0.0597 while the largest decrease occurred for the variable pair (TB-BB) with a drop of 0.0345. According to the comparison, reducing data does not influence the rank correlation significantly from a statistical point of view.

**Table 3.7: The rank correlation matrix of selected variable pairs with original collected data and final selected data.**

**(a) Empirical rank correlations of selected variable pairs with original collected data**

| | GP21 | Morgan | TB | BB | SB | HB | GB | SJR |
|---|---|---|---|---|---|---|---|---|
| GP21 | 1.00 | 0.8624 | 0.7092 | 0.2267 | 0.6867 | 0.2925 | 0.6772 | 0.6804 |
| Morgan | 0.8624 | 1.00 | 0.7062 | 0.2378 | 0.6727 | 0.2484 | 0.6458 | 0.6716 |
| TB | 0.7092 | 0.7062 | 1.00 | 0.2413 | 0.9059 | 0.1780 | 0.7405 | 0.8139 |
| BB | 0.2267 | 0.2378 | 0.2413 | 1.00 | 0.3170 | 0.7171 | 0.5025 | 0.2507 |
| SB | 0.6867 | 0.6727 | 0.9059 | 0.3170 | 1.00 | 0.2446 | 0.7637 | 0.7863 |
| HB | 0.2925 | 0.2484 | 0.1780 | 0.7171 | 0.2446 | 1.00 | 0.4552 | 0.2217 |
| GB | 0.6772 | 0.6458 | 0.7405 | 0.5025 | 0.7637 | 0.4552 | 1.00 | 0.7983 |
| SJR | 0.6804 | 0.6716 | 0.8139 | 0.2507 | 0.7863 | 0.2217 | 0.7983 | 1.00 |

**(b) Empirical rank correlations of selected variable pairs with the final selected data**

| | GP21 | Morgan | TB | BB | SB | HB | GB | SJR |
|---|---|---|---|---|---|---|---|---|
| GP21 | 1.00 | 0.8879 | 0.7218 | 0.2864 | 0.7023 | 0.2598 | 0.6839 | 0.6900 |
| Morgan | 0.8879 | 1.00 | 0.7230 | 0.2268 | 0.6816 | 0.2229 | 0.6476 | 0.6899 |
| TB | 0.7218 | 0.7230 | 1.00 | 0.2068 | 0.9087 | 0.1512 | 0.7397 | 0.8128 |
| BB | 0.2864 | 0.2268 | 0.2068 | 1.00 | 0.3137 | 0.7233 | 0.4850 | 0.2554 |
| SB | 0.7023 | 0.6816 | 0.9087 | 0.3137 | 1.00 | 0.2407 | 0.7749 | 0.7931 |
| HB | 0.2598 | 0.2229 | 0.1512 | 0.7233 | 0.2407 | 1.00 | 0.4492 | 0.2215 |
| GB | 0.6839 | 0.6476 | 0.7397 | 0.4850 | 0.7749 | 0.4492 | 1.00 | 0.8062 |
| SJR | 0.6900 | 0.6899 | 0.8128 | 0.2554 | 0.7931 | 0.2215 | 0.8062 | 1.00 |

### Historical events

Another potential influence of reducing data is losing the historical flood information. In Table 3.8, the top ten highest historical sea levels at NOAA 8771450 (GP21) from 1908 to 2016 and top ten highest historical discharges at USGS 08075000 (BB) since 1934 are listed. For GP21, the highest water level occurred at 13/09/2008 which is captured by the final selected data period (02/10/1997-31/01/2017). However, there are seven events out of the range from 1997 to 2017. For BB which includes the discharge data of almost 80 years, the largest discharge is 506.57$m^3$/s occurring at 26/05/2015 and only three events in the top ten events are out of the final selected data period. In conclusion, using the final selected data period will lose some historical extreme data especially for NOAA 8771450 (GP21) which may influence its marginal distribution and cause the incomplete analysis for the historical flood events.

**Table 3.8: (a) The top ten highest water levels at NOAA 8771450 (GP21); (b) The top ten highest discharges at USGS 08075000 (BB). Data source: NOAA, USGS.**

**(a)**

| The Top Ten Highest Water Levels since 1908<br>NOAA Station: 8771450 (GP21) | |
|---|---|
| Date | Value (Unit: m, Datum: NAVD88) |
| 13/09/2008 | 3.527 |
| 17/08/1915 | 2.920 |
| 11/09/1961 | 2.493 |
| 15/09/1919 | 2.280 |
| 18/08/1983 | 1.652 |
| 27/06/1957 | 1.609 |
| 11/09/1998 | 1.564 |
| 15/07/2003 | 1.558 |
| 29/08/1942 | 1.457 |
| 25/07/1934 | 1.457 |

**(b)**

| The Top Ten Highest Discharge since 1934<br>USGS Station: 08075000 (BB) | |
|---|---|
| Date | Value (Unit: $m^3$/s, Datum: NAVD88) |
| 26/05/2015 | 506.57 |
| 18/10/1994 | 461.29 |
| 18/04/2016 | 449.97 |
| 09/06/2001 | 396.20 |
| 11/09/1998 | 384.88 |
| 31/08/1981 | 370.73 |
| 19/09/1983 | 365.07 |
| 31/10/2015 | 339.60 |
| 16/10/2006 | 316.96 |
| 13/09/2008 | 311.30 |

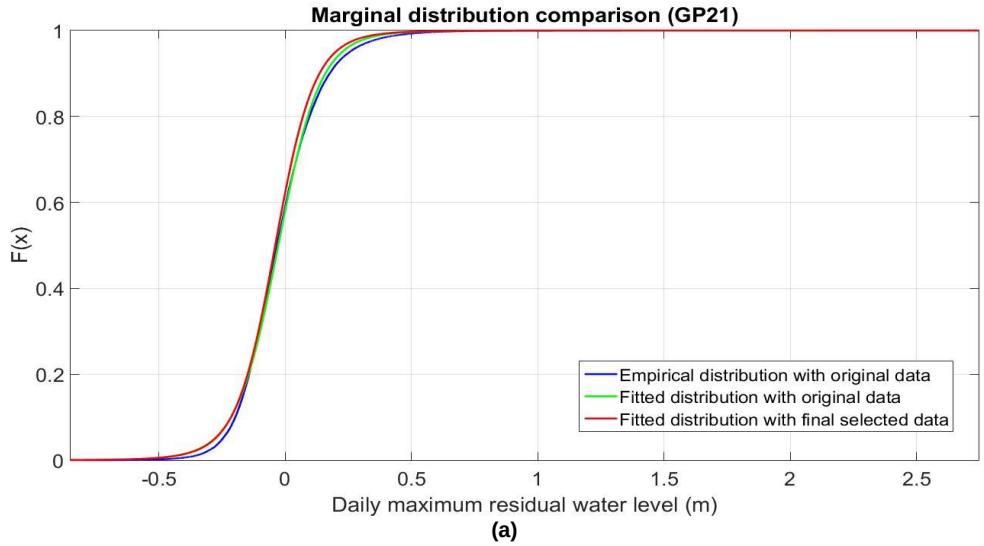
### Marginal distribution

As mentioned above, reducing data will lead to the loss of some critical historical data, which will influence the marginal distribution of the variables especially for NOAA 8771450 (GP21) and USGS 08075000 (BB). To test this effect, the continuous marginal distributions of daily maximum residual water level at GP21 fitting the originally collected data from 1904 to 2017 and final selected data from 1997 to 2017 are estimated and compared to the empirical marginal distribution. Similarly, the marginal distributions of discharge at USGS 08075000 (BB) fitting the originally collected data and final selected data are also estimated and compared to the empirical marginal distribution. The Pearson's chi-squared test is used here to ascertain the best fitted marginal distribution which is introduced specifically in Appendix B.1. A brief conclusion of Pearson's Chi-squared test is shown in Table 3.9 while the detailed results are in Appendix B.2.

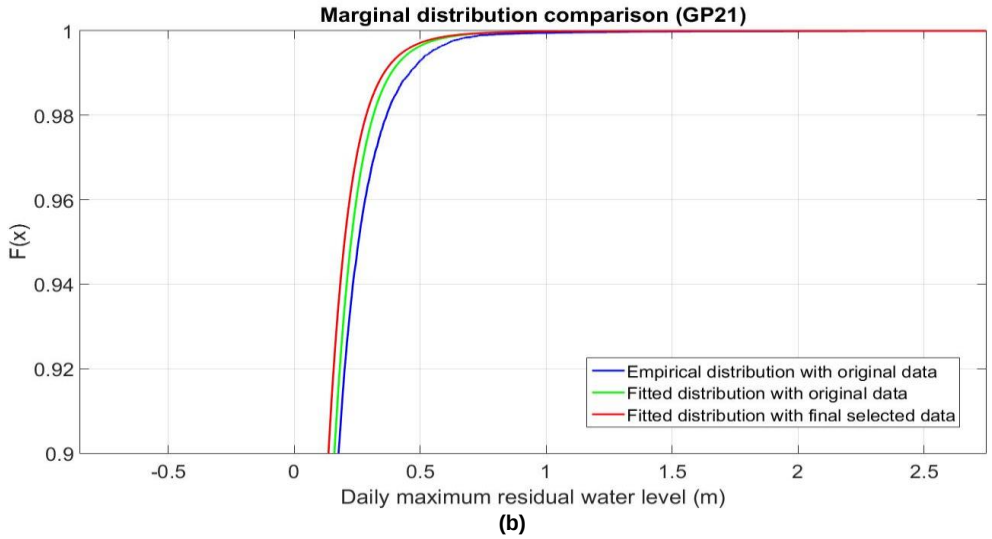
**Table 3.9: Information of fitted continuous marginal distributions of daily maximum residual water level at NOAA 8771450 (GP21) and discharge at USGS 08075000 (BB).**

| Station | Data period | Best fitted distribution type | Parameters | Result of Chi-squared test |
|---|---|---|---|---|
| NOAA 8771450 (GP21) | 01/01/1904-31/01/2017 | Tlocationscale | Mu: -0.0273, Sigma:0.1289, Nu:5.7070 | 0.014 |
| | 02/10/1997-31/01/2017 | Tlocationscale | Mu: -0.0391, Sigma:0.1194, Nu:5.2075 | 0.015 |
| USGS 08075000 (BB) | 01/01/1985-01/03/2017 | Generalized Extreme Value (GEV) | K: 0.9311, Sigma: 1.0536, Mu: 3.3640 | 0.0127 |
| | 02/10/1997-31/01/2017 | Generalized Extreme Value (GEV) | K: 0.9424, Sigma: 1.1487, Mu: 3.4022 | 0.0144 |

Figure 3.4 shows the comparisons of empirical cumulative distribution functions (ECDFs) and the fitted continuous cumulative distribution functions of the daily maximum residual water level at GP21 and discharge at BB. In general, the differences between the distributions are not significant (see Figure 3.4a and 3.4c). However, for the extreme situations with small exceedance probabilities (see Figure 3.4b and 3.4d), the marginal distribution fitting the final selected data underestimates the daily maximum residual water level at GP21 compared to the empirical distribution and the marginal distribution fitting the originally collected data while it overestimates the discharge at BB.




**(a)**




**(b)**

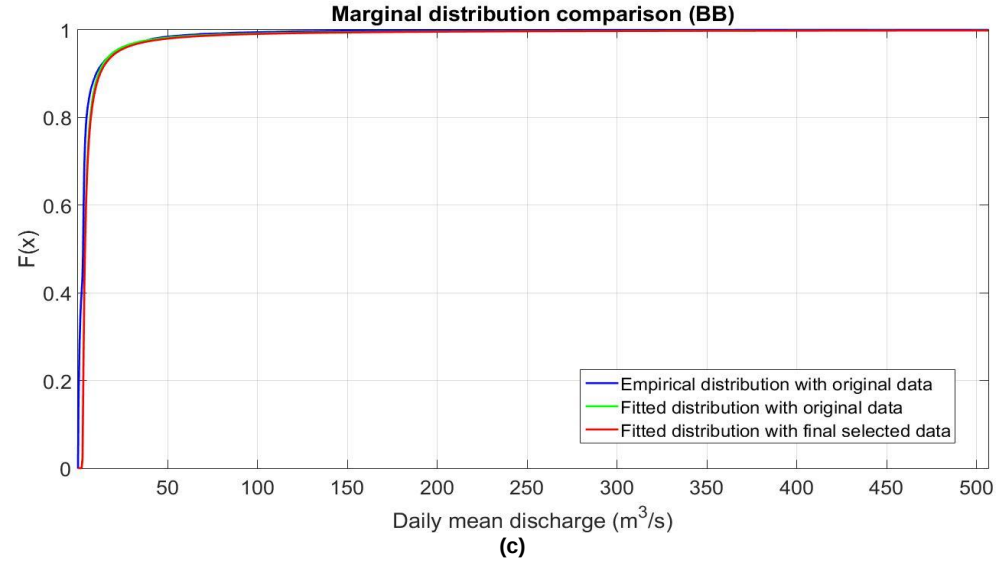



**Figure 3.4: (a) Comparisons of empirical cumulative distribution functions (ECDFs) and the marginal cumulative distribution functions fitting the originally collected data and final selected data of the daily maximum residual water level at NOAA 8771450 (GP21); (b) Comparisons of empirical cumulative distribution functions (ECDFs) and the marginal cumulative distribution functions fitting the original collected data and final selected data of the daily maximum residual water level at NOAA 8771450 (GP21) in the case of small exceedance probabilities; (c) Comparisons of empirical cumulative distribution functions (ECDFs) and the marginal cumulative distribution functions fitting the original collected data and final selected data of the discharge at USGS 08075000 (BB); (d) Comparisons of empirical cumulative distribution functions (ECDFs) and the marginal cumulative distribution functions fitting the original collected data and final selected data of the discharge at USGS 08075000 (BB) in the case of small exceedance probabilities.**

### 3.4.2 Data Period Selection

Based on the introduction of tropical cyclones in Section 2.2.1, the hurricane season in North America is known from June to November. However, with the consideration of covering all the potential flood events, the final selected data period covers the continuous daily data from 02/10/1997 to 31/01/2017 which include the data out of the hurricane season. To test the influence of the additional data on the prediction of compound floods, the data during the hurricane season from 1997 to 2017 are extracted and an additional Bayesian Network is built based on the extracted data. By comparing the different Bayesian Networks, the values of the selected variables generated from the Bayesian Network based on the data during hurricane season are slightly larger than the values generated from the Bayesian Network based on the final selected data. Considering the small influence, the final data period selected i.e. the continuous daily data from 02/10/1997 to 31/01/2017 is acceptable.

## 3.5. Conclusion

This chapter elaborates the process of data selection which lays the foundation for the following work. In Section 3.1, the boundaries of the study area are selected where the upstream boundaries are divided into two parts: U1 at the mouth of Buffalo Bayou to the downstream reach of the HSC and U2 at the point just upstream of the confluence of the San Jacinto River and the HSC and the downstream boundary is selected close to the location of the proposed storm surge barrier (see Figure 3.1). After the confirmation of the boundaries, the required hydraulic boundary conditions related to compound flood events are also fixed i.e. the sea level at downstream boundary and discharges at upstream boundaries. Based on the geographical information shown in Figure 3.1, the discharge at the U1 is further converted into the summation of the discharges at the outlets of Buffalo Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, Greens Bayou and the discharge at the U2 is assumed totally from the discharge of San Jacinto River, which provides the guidance for selecting the USGS stations and the stations nearest to the outlets of the watersheds mentioned above are selected. In addition, the NOAA station at Morgan's Point is selected which is the nearest station to the downstream boundary and the sea level at Morgan's Point is assumed to equal to the downstream boundary sea level. The NOAA station at Galveston Pier 21 is also selected as an addition because of its sufficient data reserve. In Section 3.2 and 3.3, the detailed data selection and analysis are performed for extracting the most suitable data for constructing Bayesian Network. In conclusion, the final selected data and the additional information is listed in Table 3.10.

**Table 3.10: Overall information about the final selected data. Data source: NOAA, USGS.**

| Selected stations | Location | Contributing area of the stations (unit: $km^2$) | Coverage percentage of contributing areas of each station | Selected data type |
|---|---|---|---|---|
| NOAA 8771450 (GP21) | - | - | - | Daily maximum residual water level |
| NOAA 8770613 (Morgan) | - | - | - | Daily maximum residual water level |
| USGS 08074710 (TB) | Buffalo Bayou | 520.7 | 94.4% | Daily maximum water level |
| USGS 08075000 (BB) | Brays Bayou | 245.8 | 74.7% | Daily mean discharge |
| USGS 08075500 (SB) | Sims Bayou | 163.2 | 67.0% | Daily maximum gage height |
| USGS 08075730 (VB) | Vince Bayou | 21.4 | 51.7% | Daily mean discharge |
| USGS 08075770 (HB) | Hunting Bayou | 41.7 | 53.7% | Daily mean discharge |
| USGS 08076700 (GB) | Greens Bayou | 471.4 | 85.8% | Daily mean gage height |
| USGS 08072050 (SJR) | San Jacinto River | 132.1 | 80.0% | Daily mean gage height |
| Final selected data Period | 02/10/1997-31/01/2017 | | | |
| Total number of overlapped days | 7062 | | | |
| Number of available data | 5664 | | | |
| Data datum | NAVD88 | | | |

# 4

# Bayesian Network

Based on the work in Chapter 3, the compound flood events are interpreted as the joint occurrence of high sea level at the downstream boundary and high discharges at upstream boundaries. Additionally, the variables and historical data related to the hydraulic boundary conditions are picked out. In this chapter, all the selected variables are jointed by the Non-Parametric Bayesian Network (NPBN) and a large number of stochastic scenarios are generated through the NPBN for studying the joint probable combinations of the hydraulic boundary conditions in the next chapter. In Section 4.1, background knowledge about Bayesian Networks and the software "UniNet" is introduced. In Section 4.2, it is elaborated how to build and validate the Non-Parametric Bayesian Network (NPBN). A summary is presented in Section 4.3.

## 4.1. Theoretical Background

### 4.1.1 A General Introduction of Bayesian Network

Bayesian Networks (BNs) are graphical models which represent the probability distribution of different variables (Heckerman, 1995; Jensen, 1996). Generally, A BN model is made up by a directed acyclic graph (DAG) and a set of (conditional) distributions (A. Hanea, Morales Napoles, & Ababei, 2015a). For the directed acyclic graph consisting of nodes and arcs, each node represents a random variable while each arc represents the dependence relationship of two different nodes. According to the direction of the arcs, the nodes can be divided into parents which are at the starting point of the arcs and children which are at the end of the arcs. Each parent node can be seen as the predecessor of the corresponding child node. In addition, every node has a marginal distribution and a conditional distribution exists for each child node when the dependencies between different nodes are considered which provide the quantitative information about the strength of the dependencies. Combining the conditional distributions, a directed acyclic graph (DAG) shows the joint distribution of the random variables which can be expressed as Eq. 4.1 (A. Hanea et al., 2015a; Heckerman, 1995):

$$P(x_1,...,x_n) = \prod_i P(x_i \mid pa_i) \qquad \text{[Eq. 4.1]}$$

Where $x_i$ is the value of a random variable $X_i$ and $pa_i$ is the value of the parents of the variable $X_i$. Figure 4.1 shows some basic structures of BNs. In structure (a) and (b) of Figure 4.1, X1 is conditionally independent of X3 given X2 while, in structure (c), X1 and X3 are independent but not conditional independent given X2. These basic structures lay the foundation for constructing the BNs.

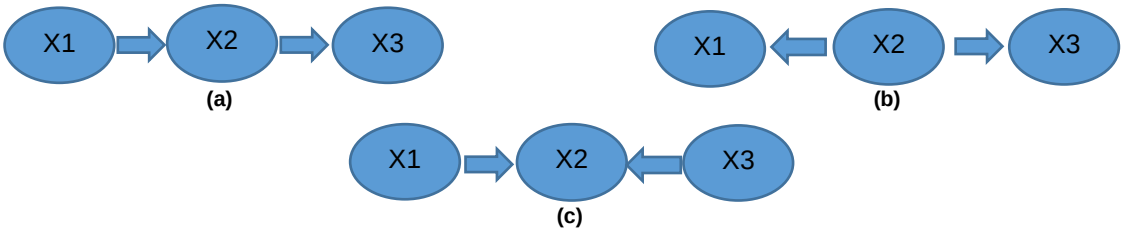



**Figure 4.1: Basic structures of Bayesian Networks.**

Nowadays, BNs have been widely used in many different fields for predicting the uncertain factors. For example, Yu et al. (1999) predicted the reliability of power systems by the BN model and Morales-Nápoles et al. (2012) assessed the failure risk of earth dams through a BN model.

### 4.1.2 Non-Parametric Bayesian Networks

Theoretically, BNs can be further divided into discrete BNs whose nodes represent discrete random variables and Hybrid & Non-Parametric BNs whose nodes can present both discrete and continuous random variables. When comparing these two different types of BNs, discrete BNs have some serious limitations (Hanea et al., 2006; Morales Napoles et al., 2013):

(1) Discrete BN cannot handle the excessive assessment burden in the applications with high complexity in data-sparse situations;

(2) Discrete BN includes only discrete variables, which is insufficient for many problems. Many applications require the assessment of joint behaviour between discrete and continuous variables or only continuous variables.

Therefore, the Non-Parametric Bayesian Network (NPBN) is used here because of considerations regarding the number of random variables (i.e., 8) and availability of data. In NPBNs, nodes represent random variables for which there is no assumption of parametric marginal distribution and the marginal distributions of all the variables are derived from the collected data in this project. In addition, arcs showing the relationship between two variables are associated with one-parameter conditional copulas where the conditional copulas are parameterized by Spearman's rank correlations (A. M. Hanea et al., 2006; Morales, Kurowicka, & Roelen, 2008). Here, the Spearman's rank correlations $r$ are calculated as following:

$$r(X,Y) = \rho(F_X(X), F_Y(Y)) \qquad \text{[Eq. 4.2]}$$

In Eq.4.2, $\rho$ is Pearson's coefficient, $X$ , $Y$ are random variables and $F_X(X)$ , $F_Y(Y)$ are their respective marginal cumulative distribution functions (CDFs) which are derived from the collected data set $(x_1, x_2, x_3, \ldots x_n), (y_1, y_2, y_3, \ldots y_n), \ldots$ and can be expressed as:

$$F_X(x) = \#(x_i \le x)/(n+1) \qquad \text{[Eq. 4.3]}$$

$$F_Y(y) = \#(y_i \le y)/(n+1) \qquad \text{[Eq. 4.4]}$$

Where $\#(x_i \le x)/\#(y_i \le y)$ means the total number of all the observed data $x_i / y_i$ that smaller than a random value $x / y$ . In addition, in Eq. 4.2, $\rho$ is Pearson's coefficient which is defined as:

$$\rho(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} \quad \text{[Eq. 4.5]}$$

Where $E$ is the expectation of the variable and $\sigma$ is the standard deviation. In general, the copula discussed here can be seen as the expression of a multivariate joint distribution corresponding to the ranks of the variables (Joe, 1997). More knowledge about copula is introduced in Appendix C. In conclusion, the NPBNs can describe the joint distribution uniquely by combining the marginal distributions of all the variables, conditional copulas and the conditional independence (Hanea et al., 2006).

### 4.1.3 "UniNet"

In this project, "UniNet" is chosen to build the NPBN which is a standalone uncertainty analysis software and constructs NPBNs under the normal copula assumption which means the bivariate distribution of an arbitrary variable pair in the NPBN obey the normal copula ("Uninet," n.d.).

Visualization is one of the advantages of "UniNet" through which it is easy to see the structures of BNs and the influences of every node on the others. Another significant advantage of "UniNet" is the convenience of testing the adequacy of the constructed BN including the validation of the suitability of the normal copula assumption for the BN and adequacy of the BN structure. As proposed by Hanea et al. (2015), the validation of the normal copula assumption can be determined by the comparison between the empirical rank determinant (DER) and the empirical normal determinant (DNR) while the comparison between the empirical normal determinant (DNR) and determinant of the rank correlation matrix of the BN model using the normal copula (DBBNR) is used to determine the adequacy of the BN structure. If the DER is within the 90% confidence bounds of the DNR, then the normal copula assumption is considered acceptable. Similarly, If the DNR is within the 90% confidence bounds of the DBBNR, the BN structure is adequate. Here, the DER is computed by transforming the marginals to uniforms and then calculating the product moment correlation of the transformed variables while the DNR is obtained by transforming the marginals to standard normal and then transforming the product moment correlations to rank correlations (Hanea et al., 2015).

## 4.2. Determination of Bayesian Network

In this section, the NPBN of the selected variables i.e. daily maximum residual water level at NOAA 8771450 (GP21) and NOAA 8770613 (Morgan); daily maximum water level at USGS 08074710 (TB) and USGS 08075500 (SB); daily mean water level at USGS 08076700 (GB) and USGS 08072050 (SJR); daily mean discharge at USGS 08075000 (BB) and USGS 08075770 (HB) is set up. In Figure 4.2, the geographical positions of the variables marked by the red circles are shown.

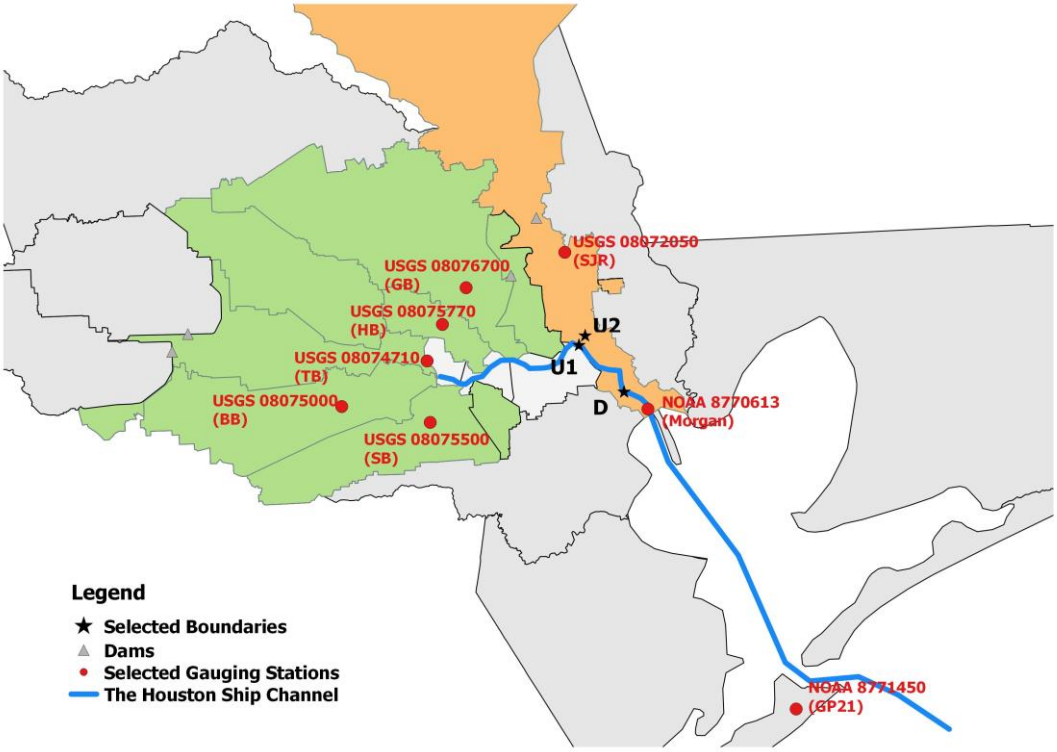



**Figure 4.2: Locations of the gauging stations selected for building the BN. Source: HCFCD.**

### 4.2.1 Model Setup

Since there is no common rule for building a BN, the BN structure could be various. In this case, to build a logical BN structure, the following additional factors are considered:

(1) Basic principle for building a BN model

According to the definition of BNs which are directed acyclic graphs, the BN model is constructed without any loop included (A. M. Hanea et al., 2006).

(2) Purpose of the project

One of the purposes of this project is analyzing the compound flood events at the downstream reach of the HSC. Therefore, it is critical to find out the dependency between the upstream discharge and downstream sea level, which can be reflected in the BN model by appointing the downstream residual water levels at GP21 and Morgan as the parent nodes of the upstream discharge at BB and HB and upstream water level at SJR, GB, TB, and SB (see Figure 4.3).

(3) Rank correlation

According to the rank correlation analysis in Section 3.3.2, it is already known that the most rank correlations of the variable pairs are higher than 0.65 while the rank correlations of variable pairs: BB-GP21, BB-Morgan, BB-TB, BB-SJR, HB-GP21, HB-Morgan, HB-TB, HB-SB and HB-SJR are relatively low which are less than 0.3. Therefore, the dependencies between the variable pairs with rank correlations less than 0.3 can be considered weak and will not be shown in the BN model.

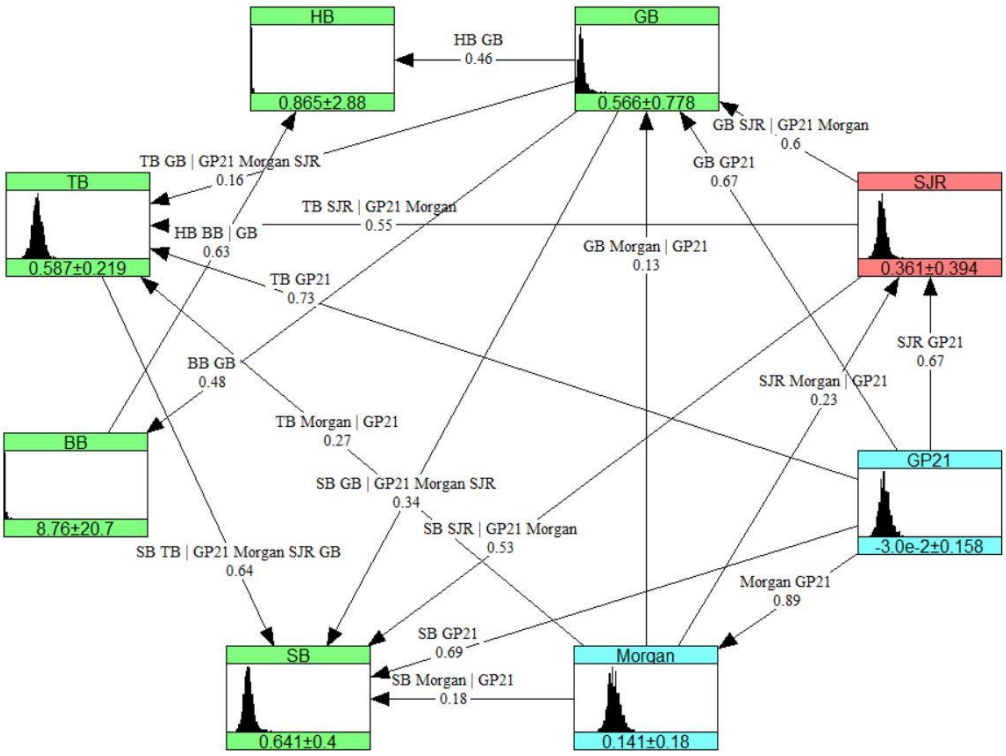



**Figure 4.3: The structure of the NPBN.**

Based on the above considerations, the BN model is built as shown in Figure 4.3 from which each rectangular box represents a node where the upper part of the box is the abbreviation of the variable, the middle part shows the histogram of the inputted historical data of the variable and the lower part shows the mean value and standard deviation corresponding to the histogram. The detailed information about the nodes is in Table 4.1. The conditional rank correlations are presented on the directed arcs. In addition, the variables marked by green are related to the discharge at upstream boundary U1 while the variable marked by pink is related to the discharge at the upstream boundary U2 and the variables marked by blue are related to the residual water level at the downstream boundary. Through the BN model, the stochastic values of daily maximum residual water levels at NOAA 8771450 (GP21) and NOAA 8770613 (Morgan), daily mean water levels at USGS 08072050 (SJR) and USGS 08076700 (GB), daily maximum water levels at USGS 08074710 (TB) and USGS 08075500 (SB), daily mean discharges at USGS 08075000 (BB) and USGS 08075770 (HB) can be generated. In this case, a total of ten-thousand samples is generated from the BN.

**Table 4.1: Detailed information about the nodes in the BN.**

| Abbreviation | Station | Data type | Data Unit | Datum |
|---|---|---|---|---|
| GP21 | NOAA 8771450 | Daily maximum residual water level | m | NAVD88 |
| Morgan | NOAA 8770613 | Daily maximum residual water level | m | NAVD88 |
| SJR | USGS 08072050 | Daily mean water level | m | NAVD88 |
| GB | USGS 08076700 | Daily mean water level | m | NAVD88 |
| HB | USGS 08075770 | Daily mean discharge | $m^3/s$ | NAVD88 |
| TB | USGS 08074710 | Daily maximum water level | m | NAVD88 |
| BB | USGS 08075000 | Daily mean discharge | $m^3/s$ | NAVD88 |
| SB | USGS 08075500 | Daily maximum water level | m | NAVD88 |

### 4.2.2 Validation of the Bayesian Network

As stated in Section 4.1.3, the adequacy of the constructed BN can be checked through the validation of the suitability of the normal copula assumption and the validation of adequacy of the BN structure. Apart from the method mentioned in Section 4.1.3, the semi-correlation test is adopted as a supplement to test the suitability of the normal copula for all the variable pairs in the BN. The detailed introduction about the semi-correlation test is in Appendix D.1.

**Validation of Normal Copula assumption**

(1) Semi-correlation test

Semi-correlation test is used to check whether the Gaussian Copula fits all the variable pairs existing in the BN model best compared to Gumbel Copula and Clayton Copula which are commonly used in the hydrologic analysis (Favre et al., 2004). The detailed results of Semi-correlation test are listed in Appendix D.2. In general, the rank correlations of the variable pairs jointed by Gaussian Copula is closest to the rank correlations calculated with empirical data. Concretely, for the upper right (NE) tail which is the most interesting part where both valuables have large values, the results show that Gaussian Copula can fit the upper right (NE) tail well for the variable pairs having only water level variables. Gumbel Copula fits the upper right (NE) tail better for the variable pairs having both water level variable and discharge variable than Gaussian Copula, but the difference is small.

(2) Comparison of determinants of the correlation matrices

The determinant of the empirical rank correlation matrix (DER) is 4.16E-4 and the determinant of the normal rank correlation matrix (DNR) is 6.05E-4 for a sample size of 2000 and DER falls outside the 90% central confidence band of DNR (i.e. from 5.17E-4 to 7.01E-4).

In conclusion, the result shows that the normal copula assumption is not optimal for all the variable pairs. However, it is still acceptable to use the normal copula assumption especially considering the complexity of the BN since the BN under normal copula assumption can provide a faster calculation for generating stochastic samples.

**Validation of the BN structure**

The determinant of the model rank correlation matrix (DBBNR) is 6.88E-4 and the determinant of the normal rank correlation matrix (DNR) is 6.05E-4 for a sample size of 100 and DNR falls within the 90% central confidence band of DBBNR (i.e. from 5.86E-4 to 7.99E-4). Therefore, the current BN structure is adequate under the normal copula assumption.

## 4.3. Conclusion

This Chapter elaborates the process of building and validating the Non-Parametric Bayesian Network (NPBN). In this thesis, the NPBN is built based on the normal copula assumption (i.e., all the selected variables are jointed by Gaussian Copula). Although the validation shows that the normal copula assumption is not optimal for all the variable pairs, the normal copula assumption is still used when considering the complexity of the BN and the calculation capacity of generating stochastic samples. The final Bayesian Network is shown in Figure 4.3 and ten-thousand stochastic samples of the selected variables are generated which will be used to generate the hydraulic boundary conditions in the next chapter.

# 5

# Prediction of Hazard Boundary Conditions

In this chapter, the first research question will be answered: “What are the hazard boundary conditions associated with compound flood events in the downstream reach of the HSC?”. Based on the Bayesian Network built in Chapter 4, ten-thousand stochastic samples of the selected variables are generated. However, these outputs are not the required hydraulic boundary conditions (i.e. the discharge at upstream boundaries and the sea level at the downstream boundary). The conversion between the BN outputs and hydraulic boundary conditions are still needed. In Section 5.1, the detailed conversion process including the assumptions and equations are introduced. Thereafter, the analysis of the resulting hydraulic boundary conditions including the marginal distribution and dependencies is presented in Section 5.2. In Section 5.3, the hazard boundary conditions associated with 100-, 500-, and 1000-year compound flood events are predicted. Section 5.4 shows the discussion of some problems about the conversion process and the estimated hazard boundary conditions. The whole work is summarized in Section 5.5.

## 5.1. Conversion of Hydraulic Boundary Conditions

It is worth mentioning again that in this thesis, the compound flood events are considered as the joint occurrence of high sea level at the downstream boundary and high discharges at upstream boundaries. Therefore, it is critical to obtain the data of hydraulic boundary conditions for studying the compound flood events in the research area. According to the introduction in the previous chapters, collecting the hydraulic boundary conditions directly is impossible which however can be converted through the variables in the BN. This section introduces the conversion process.

### 5.1.1 Conversion of Upstream Boundary Conditions

For the upstream part, the discharge at the U1 is assumed as the summation of the discharges at the outlets of Buffalo Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, and Greens Bayou while the discharge at the U2 is assumed totally from the San Jacinto River. Based on the assumptions, Figure 5.1 shows the detailed process of generating the discharge at upstream boundaries which mainly contains three steps: (1) convert the water levels at USGS 08072050 (SJR), USGS 08076700 (GB), USGS 08074710 (TB), USGS 08075500 (SB) to discharges at these stations through the uniform flow assumption; (2) convert the discharges at all the USGS stations to the discharges at the outlets of the corresponding watersheds based on the assumption of Linear relationship between the discharge and contributing area; (3) estimate the discharge at upstream boundaries. In the following, these three steps are elaborated.

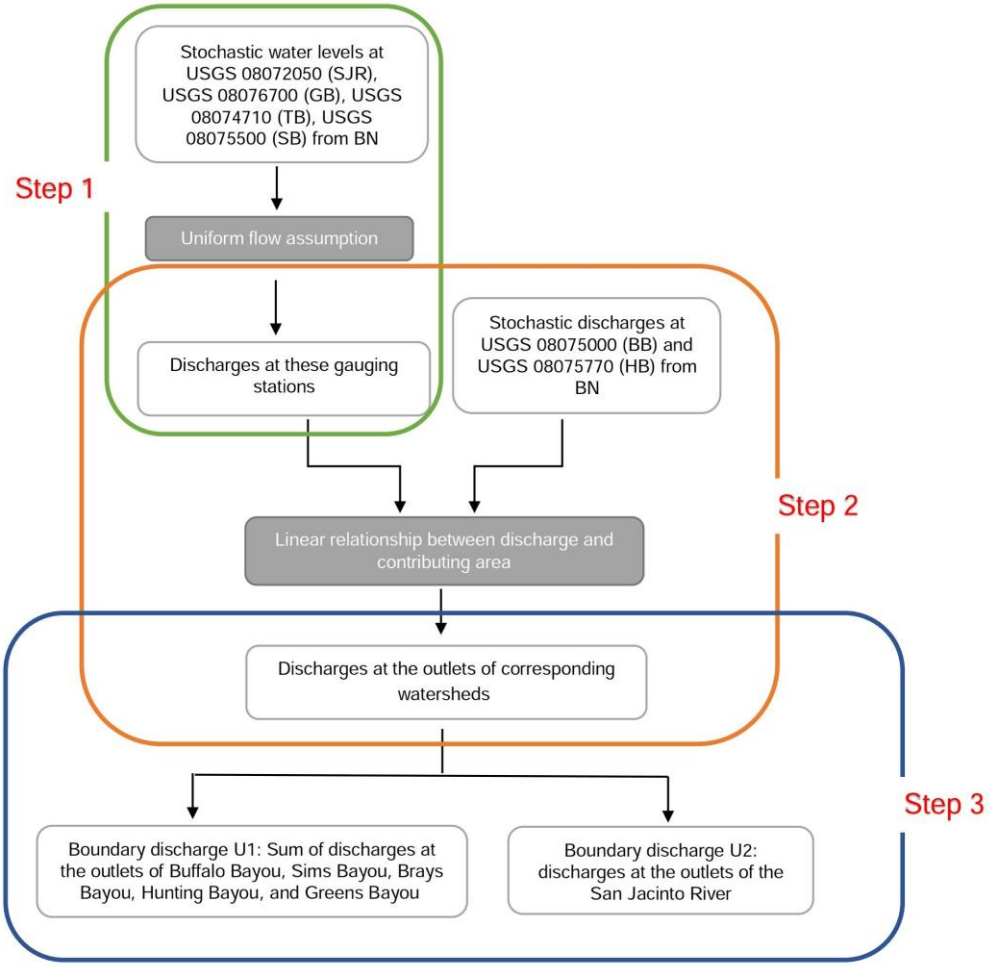



**Figure 5.1: The sketch of generating the upstream boundary conditions.**

### Step 1: Convert water level to discharge

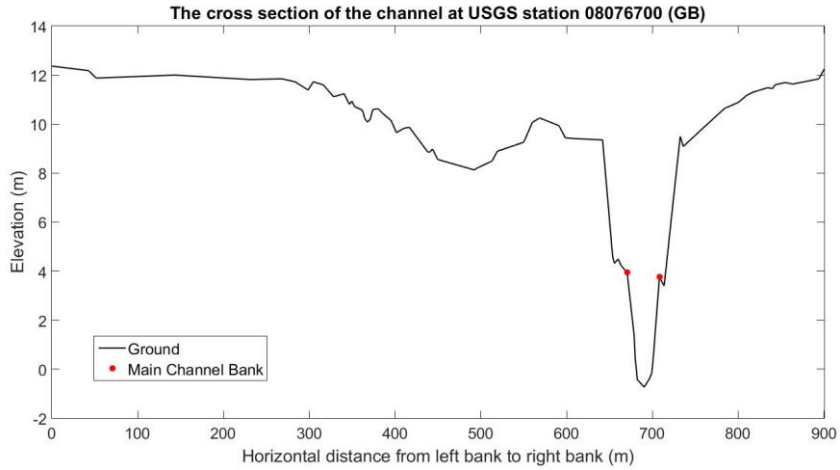
In Chapter 3, the daily maximum water level at USGS 08074710 (TB) and USGS 08075500 (SB); daily mean water level at USGS 08076700 (GB) and USGS 08072050 (SJR) are selected as the input variables in the BN and ten-thousand stochastic water levels at these stations are generated. However, these variables do not directly link with the discharge at upstream boundaries, the conversion from water level to discharge at these stations is needed.

In reality, most USGS stream gages estimate discharge by using the developed stage-discharge relation (i.e. rating curve). This method transmits water level data by satellite to USGS computers and then do the computation, which, however, is out of the scope of this project. Here, the uniform flow assumption is made and Manning's Equation with S.I. unit (see Eq. 5.1) is used to convert water level to discharge:

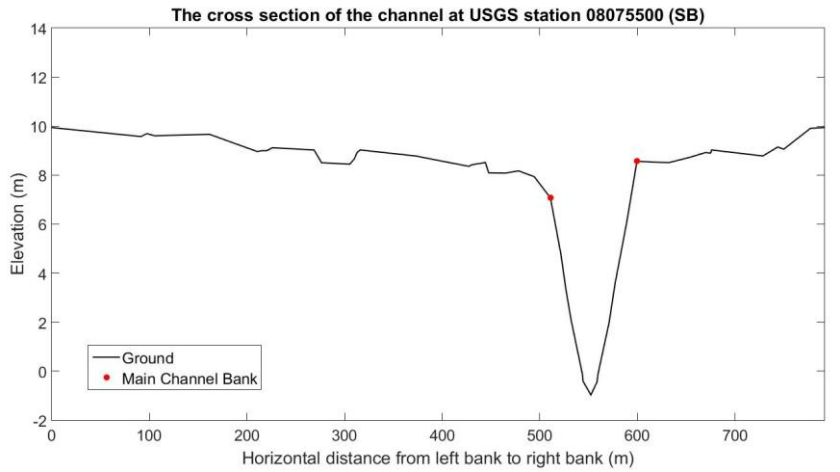
$$Q = \frac{1}{n} A R^{\frac{2}{3}} S^{\frac{1}{2}} \qquad \text{[Eq. 5.1]}$$

Where $Q$ is uniform flow discharge, $A$ is the cross-sectional area of flow, $S$ is the bottom slope of the channel, $n$ is the Manning roughness coefficient and $R$ is hydraulic radius (i.e., area divided by wetted perimeter). In Eq. 5.1, the Manning roughness coefficient $n$ and the bottom slope of the channel $S$ can be seen as constants which mainly depend on the geological conditions of water channels while cross-sectional area of flow $A$ and hydraulic radius $R$ depend on the shapes of water channels and water level.

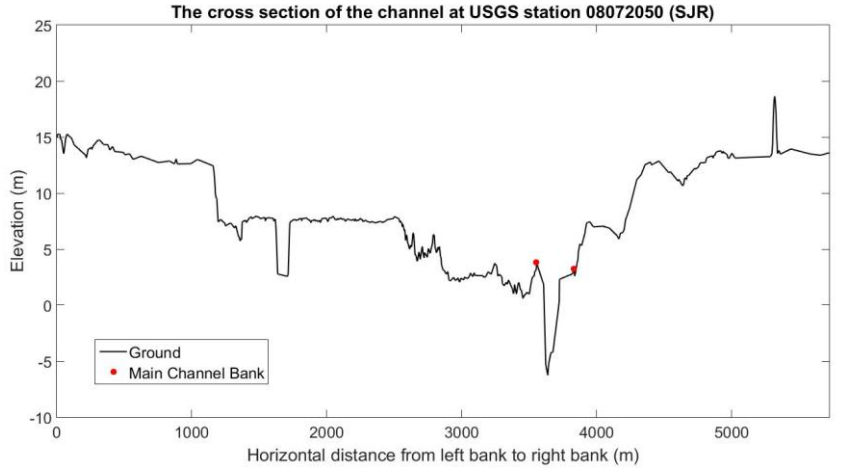
In this project, the geological conditions, and shapes of water channels at the USGS stations can be collected from the Model and Map Management (M3) System of Harris County. Therefore, the only uncertain variable in Eq. 5.1 is water level and accordingly, the uniform discharges at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08076700 (GB) and USGS 08072050 (SJR) can be estimated by the stochastic water levels generated from BN. In Figure 5.2, the cross-sectional shapes of water channels at the above USGS stations are shown.



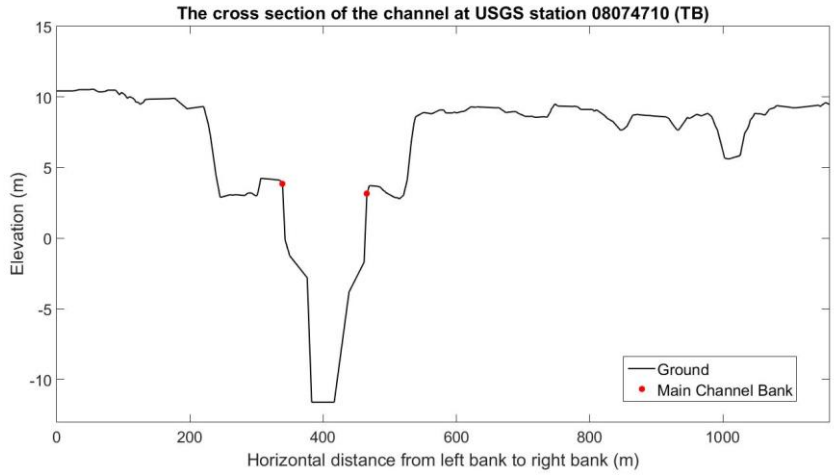

**(a)**




**(b)**



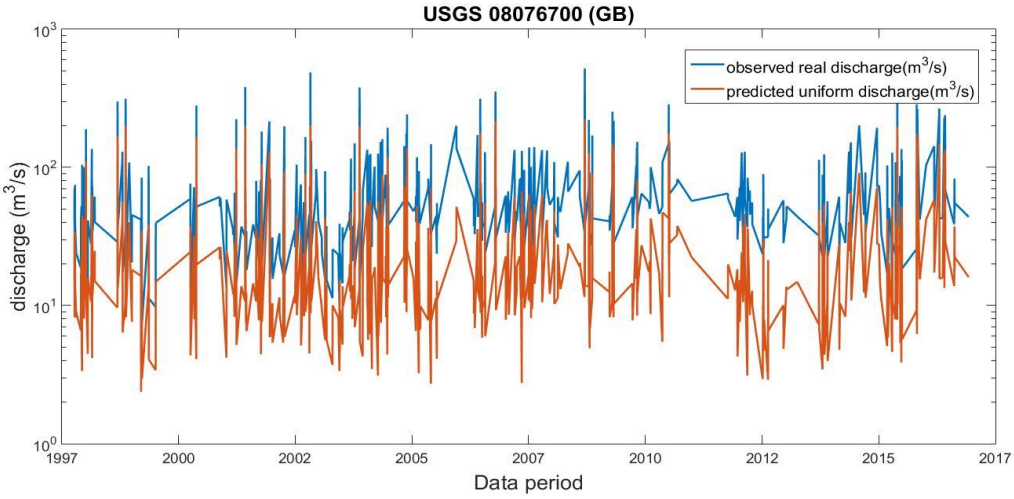

**(c)**




**(d)**

**Figure 5.2: (a) The cross-section of Greens Bayou at USGS station 08076700 (GB); (b) The cross-section of Sims Bayou at USGS station 08075500 (SB); (c) The cross-section of San Jacinto River at USGS station 08072050 (SJR); (d) The cross-section of Buffalo Bayou at USGS station 08074710 (TB). Source: HCFCD.**

After obtaining the uniform discharges, it is necessary to validate the accuracy of the uniform flow assumption. For this validation, the largest two hundred historical discharge values and water level values of the same days at USGS stations: 08076700 (GB) (data period: from 1997 to 2017) and 08075000 (BB) (data period: from 1987 to 2017) are collected respectively and then the uniform discharges are calculated based on the historical water levels. The reason selecting USGS 08076700 (GB) and 08075000 (BB) is that only these two stations can provide sufficient discharge data and water level data of the same days. In addition, the reason choosing the largest two hundred historical discharge values is because it is interesting to know the relationship between observed discharges and estimated uniform discharges under extreme conditions. In Figure 5.3, the historical discharges and estimated uniform discharges of the same days are compared where the orange line represents the estimated uniform discharges and the blue line represents the observed discharges.



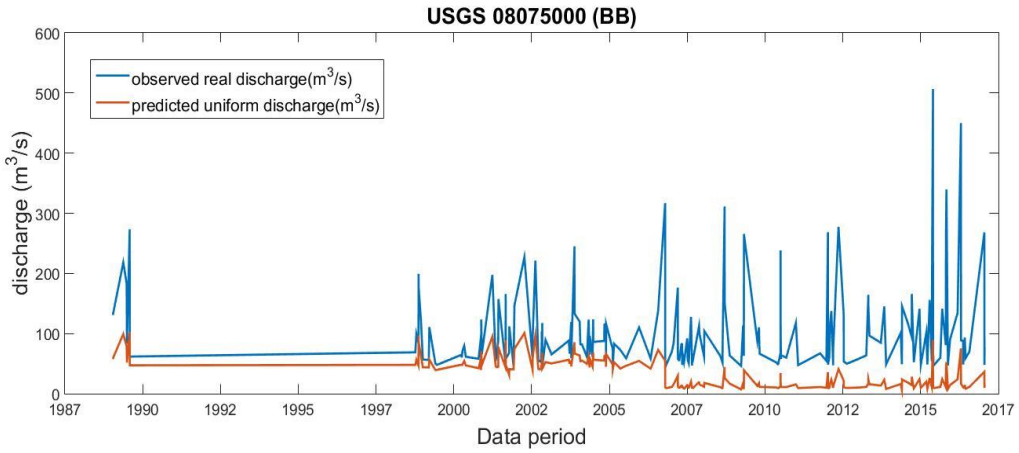



**Figure 5.3: The comparisons between the estimated uniform discharges and observed discharges of the same days at the USGS station 08076700 (GB) and 08075000 (BB), respectively. Source: USGS.**

The comparison results show that, for USGS 08076700 (GB), the uniform discharges are smaller than the observed discharge and the average ratio between the uniform discharges and observed discharges is about 0.36. For the USGS 08075000 (BB), the similar relationship exists with an average ratio of 0.37. Because of lacking information, it is hard to know the actual relationships between the uniform discharges and real discharges at USGS 08072050 (SJR), USGS 08074710 (TB), USGS 08075500 (SB); therefore, it is assumed that the similar relationships exist at these stations as well. After the validation, the final discharge at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08076700 (GB) and USGS 08072050 (SJR) can be estimated as:

$$Q_{USGS\_TB} = Q_{USGS\_TB\_uniform} / 0.35 \quad \text{[Eq. 5.2]}$$

$$Q_{USGS\_SB} = Q_{USGS\_SB\_uniform} / 0.35 \quad \text{[Eq. 5.3]}$$

$$Q_{USGS\_GB} = Q_{USGS\_GB\_uniform} / 0.35 \quad \text{[Eq. 5.4]}$$

$$Q_{USGS\_SJR} = Q_{USGS\_SJR\_uniform} / 0.35 \quad \text{[Eq. 5.5]}$$

Where, $Q_{USGS\_TB}$ is the discharge at USGS station 08074710 (TB), $Q_{USGS\_TB\_uniform}$ is the uniform discharge at USGS station 08074710 (TB); $Q_{USGS\_SB}$ is the discharge at USGS station 08075500 (SB), $Q_{USGS\_SB\_uniform}$ is the uniform discharge at USGS station 08075500 (SB); $Q_{USGS\_GB}$ is the discharge at USGS station 08076700 (GB), $Q_{USGS\_GB\_uniform}$ is the uniform discharge at USGS station 08076700 (GB); $Q_{USGS\_SJR}$ is the discharge at USGS station 08072050 (SJR), $Q_{USGS\_SJR\_uniform}$ is the discharge at USGS station 08072050 (SJR). In addition, 0.35 is chosen as the correction coefficient to make the estimated discharge closer to the actual situation in extreme cases.

**Step 2: Convert the discharges at the gauging stations to those at the outlets**

Based on step 1, the discharges at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08076700 (GB) and USGS 08072050 (SJR) can be estimated; meanwhile, the discharges at USGS 08075000 (BB), USGS 08075770 (HB) have been generated from the BN. According to the assumption of calculating the discharge at upstream boundaries (i.e., the discharge at the U1 is assumed as the summation of the discharges at the outlets of Buffalo Bayou, Sims Bayou, Brays Bayou, Hunting Bayou, and Greens Bayou while the discharge at the U2 is assumed totally from the San Jacinto River), the conversion from the discharges at the selected gauging stations to those at the outlets of the

corresponding watersheds is still needed. In Figure 5.4, the positions of the outlets of Greens Bayou, Hunting Bayou, Buffalo Bayou, Brays Bayou, and Sims Bayou relative to the gauging stations are shown.

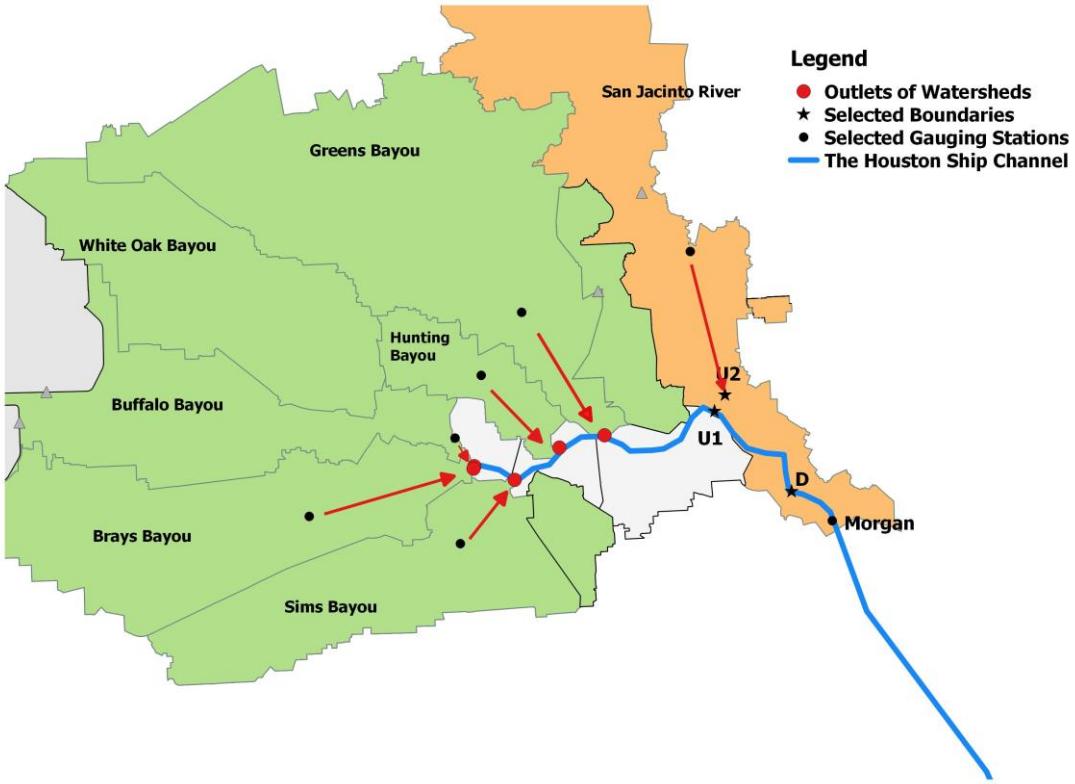



**Figure 5.4: The positions of the outlets of the watersheds relative to the selected gauging stations. Source: USGS.**

In this thesis, a linear relationship is assumed between the discharge and contributing area. Therefore, refer to the information of contributing areas shown in Table 3.3, the discharges at the outlets of Greens Bayou, Hunting Bayou, Buffalo Bayou, Brays Bayou, and Sims Bayou can be calculated as following:

$$Q_{Outlet\_BB} = Q_{USGS\_BB} / 0.747 \quad \text{[Eq. 5.6]}$$

$$Q_{Outlet\_TB} = Q_{USGS\_TB} / 0.944 \quad \text{[Eq. 5.7]}$$

$$Q_{Outlet\_SB} = Q_{USGS\_SB} / 0.67 \quad \text{[Eq. 5.8]}$$

$$Q_{Outlet\_GB} = Q_{USGS\_GB} / 0.858 \quad \text{[Eq. 5.9]}$$

$$Q_{Outlet\_HB} = Q_{USGS\_HB} / 0.537 \quad \text{[Eq. 5.10]}$$

Where, $Q_{Outlet\_BB}$ is the discharge at the outlet of Brays Bayou, $Q_{USGS\_BB}$ is the discharge at USGS station 08075000 (BB); $Q_{Outlet\_TB}$ is the discharge at the outlet of Buffalo Bayou; $Q_{Outlet\_SB}$ is the discharge at the outlet of Sims Bayou; $Q_{Outlet\_GB}$ is the discharge at the outlet of Greens Bayou; $Q_{Outlet\_HB}$ is the discharge at the outlet of Hunting Bayou; $Q_{USGS\_HB}$ is the discharge at USGS station 08075770 (HB).

In addition, from Figure 5.4, the discharge at the U2 can be directly estimated by the discharge at USGS station 08072050 (SJR) based on the relationship of contributing area:

$$Q_{U2} = Q_{USGS\_SJR} / 0.654 \quad \text{[Eq. 5.11]}$$

Where, $Q_{U2}$ is the discharge at the upstream boundary U2.

**Step 3: Generation of the upstream boundary conditions**

On the basis of the previous two steps, the discharge at upstream boundary U1 and the total discharges can be generated as following:

$$Q_{U1} = Q_{Outlet\_BB} + Q_{Outlet\_TB} + Q_{Outlet\_SB} + Q_{Outlet\_GB} + Q_{Outlet\_HB} \quad \text{[Eq. 5.12]}$$

$$Q_{Total} = Q_{U1} + Q_{U2} \quad \text{[Eq. 5.13]}$$

Where, $Q_{U1}$ is the discharge at the upstream boundary U1, $Q_{Total}$ is the total discharge at upstream boundaries which is also selected as the upstream boundary condition.

### 5.1.2 Conversion of Downstream Boundary Conditions

For the downstream part, it is already assumed that the sea level at NOAA 8770613 (Morgan) is the required downstream boundary condition. According to the Eq. 3.1 in Section 3.2.1, the downstream sea level is equal to the sum of tide level and residual water level. Through the BN, the stochastic daily maximum residual water level at NOAA 8770613 (Morgan) have been generated. In addition, the predicted tide levels at Morgan from 01/01/1993 to 01/03/2017 are plotted in Figure 5.5.

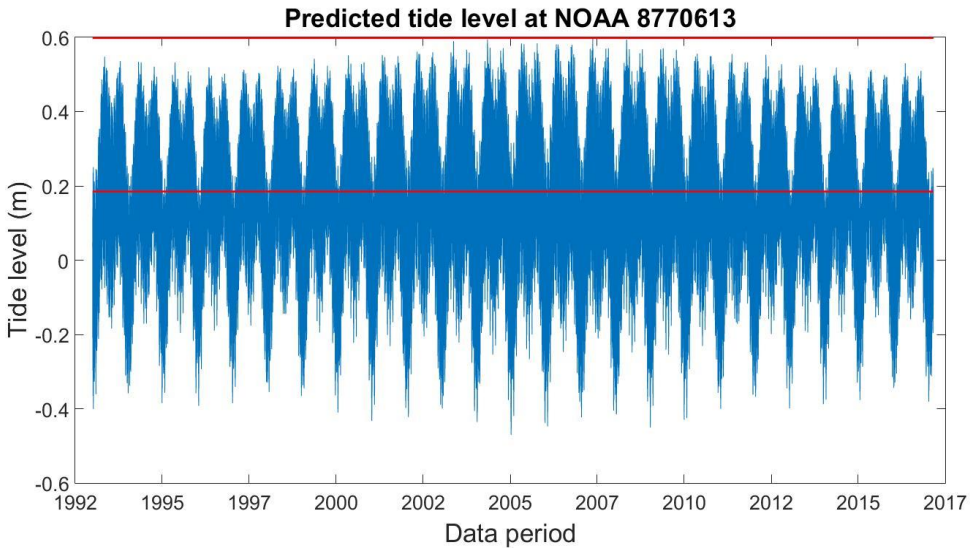



**Figure 5.5: Predicted tide level at NOAA station 8770613 (Morgan) from 1993 to 2017. Source: NOAA.**

From Figure 5.5, the mean predicted tide level from 1993 to 2017 is 0.1848m+NAVD88 while the maximum predicted tide level is 0.5983m+NAVD88. To consider the most dangerous situation, it is assumed that the tide level is a constant value and equal to the maximum. Therefore, the downstream boundary condition can be expressed as:

$$H_{Downstream} = H_{Morgan\_residual} + 0.5983 \quad \text{[Eq. 5.14]}$$

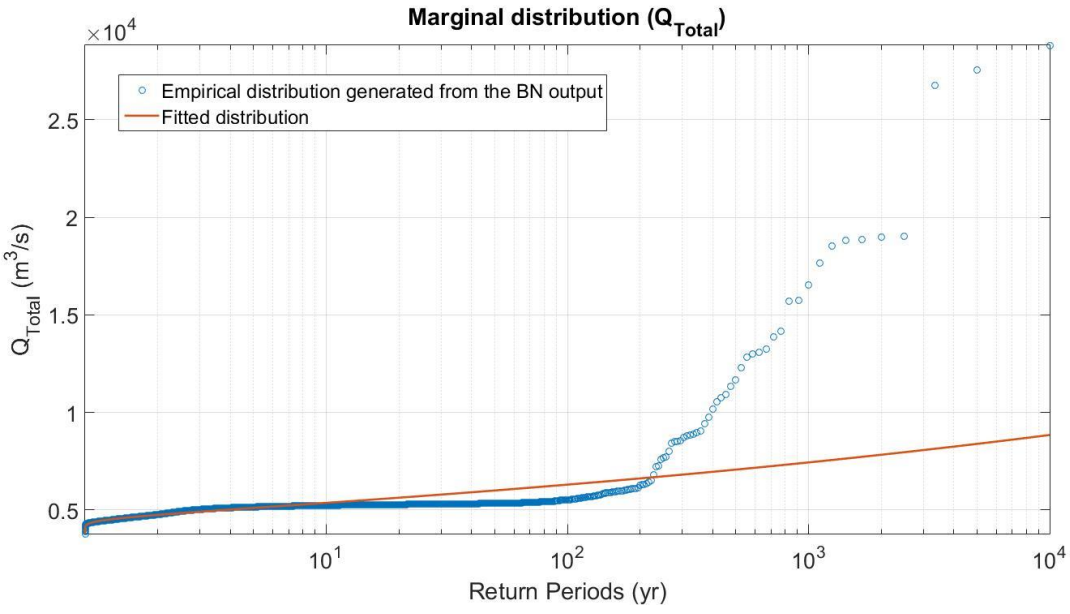
Where, $H_{Downstream}$ is the required sea level at the downstream boundary, $H_{Morgan\_residual}$ is the residual water level at Morgan's Point.

## 5.2. Analysis of Generated Hydraulic Boundary Conditions

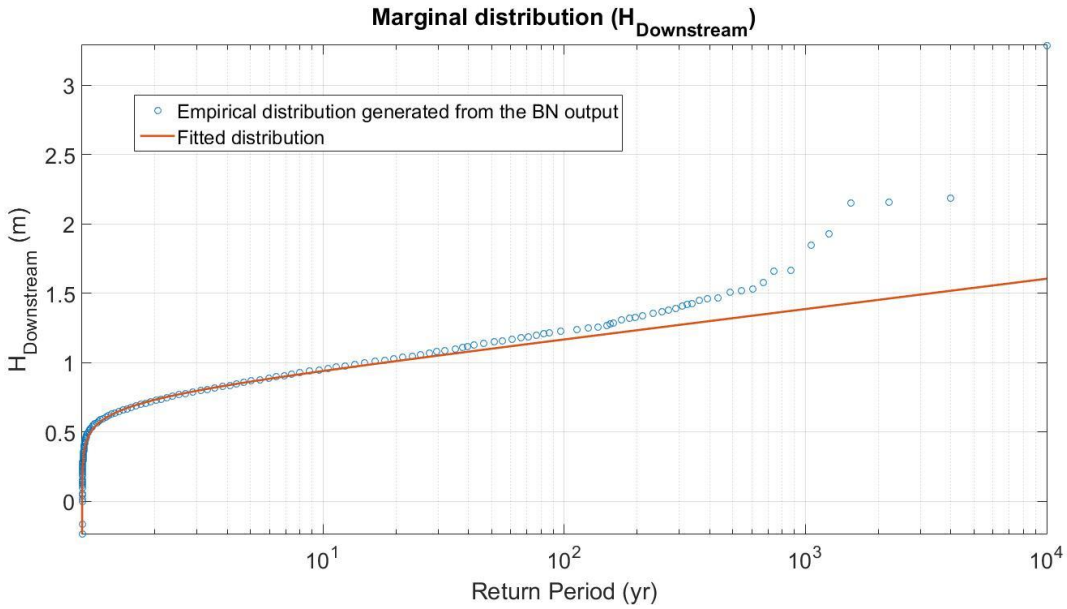
Through the introduced conversion process, the variables: $H_{downstream}$, $Q_{Outlet\_BB}$, $Q_{Outlet\_TB}$, $Q_{Outlet\_SB}$, $Q_{Outlet\_GB}$, $Q_{Outlet\_HB}$, $Q_{U1}$, $Q_{U2}$ and $Q_{Total}$ are generated based on the outputs of the BN with a sample size of ten thousand. Among these variables, $Q_{Total}$ and $H_{downstream}$ are the required variables representing upstream and downstream hydraulic boundary condition respectively. In this section, these generated variables are studied in more detail including their marginal distributions and dependences.

### 5.2.1 Marginal Distribution

The Pearson's chi-squared test is used here to judge the most appropriate theoretical continuous marginal distributions of the variables. The detailed results of Pearson's chi-squared test as well as the graphs of the marginal distributions are shown in Appendix F.1. For the variables representing the boundary conditions, the Generalized Extreme Value distribution (GEV) fits the upstream boundary condition $Q_{Total}$ best ( $\chi^2$ =9.2E-5) with a shape parameter k of 0.0923, scale parameter σ of 289.5650 and location parameter μ of 4637.2 while the logistic distribution can describe the downstream boundary condition $H_{Downstream}$ best ( $\chi^2$ =0.0172) with a scale parameter σ of 0.0948 and mean value μ of 0.0948. In Figure 5.6, the fitted continuous marginal distributions of $H_{downstream}$ and $Q_{Total}$ compared to the empirical distributions are shown.




**(a)**




**(b)**

**Figure 5.6: (a) A comparison between the continuous marginal distribution fitting the total discharge at the upstream boundaries and corresponding empirical distribution; (b) A comparison the continuous marginal distribution fitting the sea level at NOAA 8770613 (Morgan) and corresponding empirical distribution.**

By comparing to the empirical distributions, the fitted continuous marginal distribution of $Q_{Total}$ can provide a reliable prediction of upstream boundary condition when the return period is smaller than 200 years while the significant underestimation exists in the cases of return periods greater than 200 years. The similar pattern can also be observed in the fitted continuous marginal distribution of $H_{downstream}$. The underestimation for hydraulic boundary conditions in extreme events may also influence the prediction of hazard boundary conditions later.

### 5.2.2 Dependence

In this section, the dependencies between $H_{downstream}$ and $Q_{Outlet\_BB}$, $Q_{Outlet\_TB}$, $Q_{Outlet\_SB}$, $Q_{Outlet\_GB}$, $Q_{Outlet\_HB}$, $Q_{U1}$, $Q_{U2}$ and $Q_{Total}$ are tested. If the dependencies between these variables are positive and strong, then the compound flood events are more likely to happen from a perspective of statistics; on the contrary, If the dependencies between these variables are negative and/or weak, the compound flood events are not easy to happen. Here, the rank correlations and semi-correlations are used to test the dependencies.

(1) Rank Correlation

As stated earlier, ten-thousand variables: $H_{downstream}$, $Q_{Outlet\_BB}$, $Q_{Outlet\_TB}$, $Q_{Outlet\_SB}$, $Q_{Outlet\_GB}$, $Q_{Outlet\_HB}$, $Q_{U1}$, $Q_{U2}$ and $Q_{Total}$ are generated on the basis of outputs of the BN, based on which the rank correlations between these variables are calculated (see Table 5.1a). As a comparison, the rank correlations between the original input variables in the BN are shown as well (see Table 5.1b).

**Table 5.1: (a) Rank correlations between the generated variables; (b) Rank correlations between the original input variables in the BN model.**

**(a)**

| | $H_{downstream}$ | $Q_{Outlet\_BB}$ | $Q_{Outlet\_TB}$ | $Q_{Outlet\_SB}$ | $Q_{Outlet\_GB}$ | $Q_{Outlet\_HB}$ | $Q_{U2}$ | $Q_{U1}$ | $Q_{Total}$ |
|---|---|---|---|---|---|---|---|---|---|
| $H_{downstream}$ | 1.0000 | 0.3095 | 0.7346 | 0.6708 | 0.6392 | 0.2892 | -0.2606 | 0.7359 | -0.1069 |

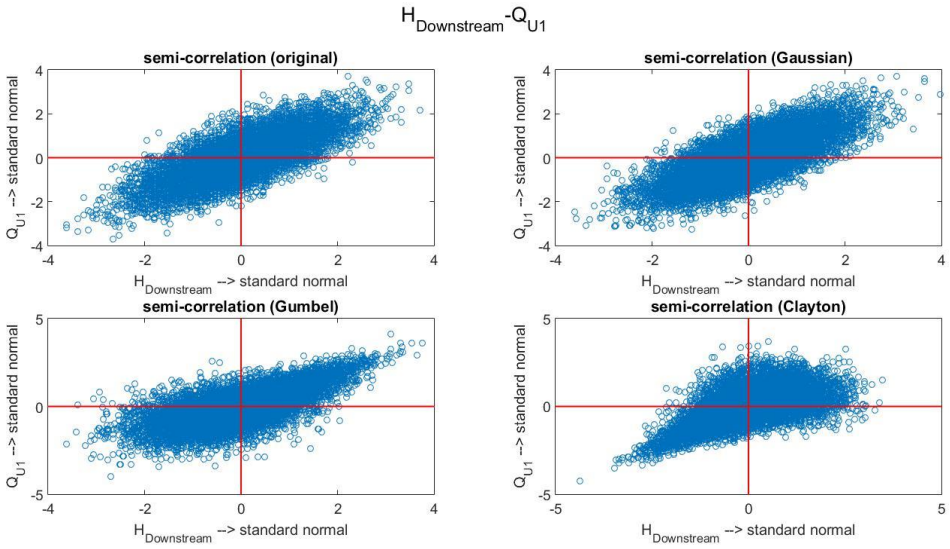
**(b)**

| | Morgan | BB | TB | SB | GB | HB | SJR | |
|---|---|---|---|---|---|---|---|---|
| Morgan | 1.0000 | 0.2268 | 0.7230 | 0.6816 | 0.6476 | 0.2229 | 0.6899 | |

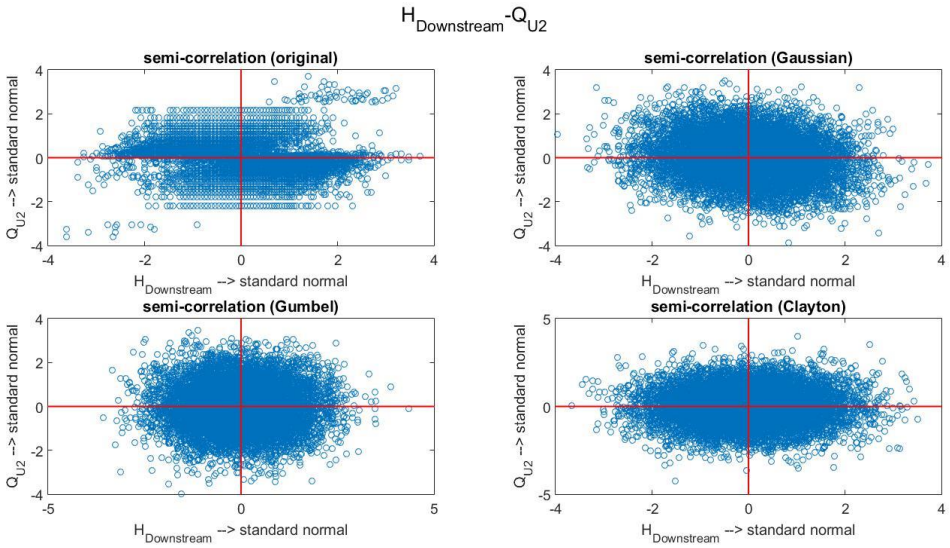
From Table 5.1a, the rank correlation between the sea level at downstream boundary and the discharge at the outlet of Brays Bayou $(H_{downstream}, Q_{Outlet\_BB})$ is 0.3095 which is close to the rank correlation between the daily maximum residual water level at NOAA 8770613 (Morgan) and daily mean discharge at USGS 08075000 (BB) (i.e., 0.2268). In addition, for the variable pairs: $(H_{downstream}, Q_{Outlet-SB})$ , $(H_{downstream}, Q_{Outlet\_GB})$ , $(H_{downstream}, Q_{Outlet\_HB})$ and $(H_{downstream}, Q_{Outlet\_TB})$, the rank correlations of which are also close to the rank correlations between the original input variables of the BN. In addition, a positive and strong dependence can be found between the sea level at the downstream boundary and discharge at the upstream boundary U1 $(H_{downstream}, Q_{U1})$ (i.e., 0.7359) while a weak and negative dependence exists between the sea level at the downstream boundary and discharge at the upstream boundary U2: $(H_{downstream}, Q_{U2})$ (i.e., -0.2606). For the dependence between the hydraulic boundary conditions $(H_{downstream}, Q_{Total})$, the rank correlation is -0.1069. Based on the results of the dependencies, it can be observed that the discharge at the U2 has a larger influence on the total discharge at upstream boundaries than the discharge at the U1. The results of the dependencies will be further discussed in Section 5.4.

(2) Semi-correlation

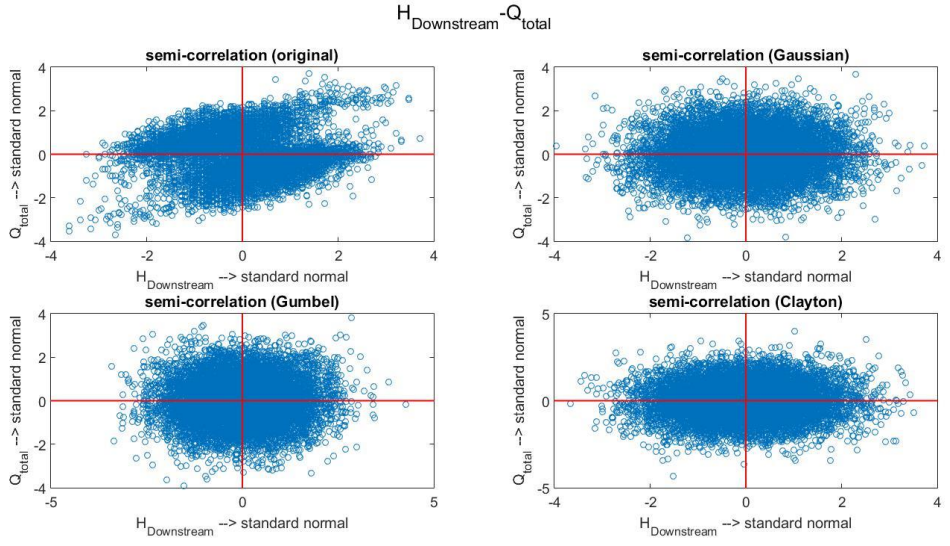
The semi-correlation test is used here to find the appropriate copula describing the joint distributions of the variable pairs. The detailed information about the results of the semi-correlation test is shown in Appendix F.2. In Figure 5.7, the semi-correlations of the variable pairs $(H_{downstream}, Q_{U1})$, $(H_{downstream}, Q_{U2})$, and $(H_{downstream}, Q_{Total})$ are exhibited for reference. It can be seen graphically that a strong, positive dependence exists in the variable pair $(H_{downstream}, Q_{U1})$ while the weak dependencies exist in the variable pairs $(H_{downstream}, Q_{U2})$ and $(H_{downstream}, Q_{Total})$.


H_Downstream-Q_U1
semi-correlation (original)
semi-correlation (Gaussian)
semi-correlation (Gumbel)
semi-correlation (Clayton)
Q_U1 --> standard normal
H_Downstream --> standard normal

**(a)**


H_Downstream-Q_U2
semi-correlation (original)
semi-correlation (Gaussian)
semi-correlation (Gumbel)
semi-correlation (Clayton)
Q_U2 --> standard normal
H_Downstream --> standard normal

**(b)**


H_Downstream-Q_total
semi-correlation (original)
semi-correlation (Gaussian)
semi-correlation (Gumbel)
semi-correlation (Clayton)
Q_total --> standard normal
H_Downstream --> standard normal

**(c)**

**Figure 5.7: (a) Graphical semi-correlation between the downstream boundary condition and discharge at the U1; (b) Graphical semi-correlation between the downstream boundary condition and discharge at the U2; (c) Graphical semi-correlation between the downstream boundary condition and the upstream boundary condition.**

According to the semi-correlation test, Gaussian Copula describes the dependencies of the variable pairs: $(H_{downstream}, Q_{Outlet\_BB})$ , $(H_{downstream}, Q_{Outlet-SB})$ , $(H_{downstream}, Q_{Outlet\_GB})$ , $(H_{downstream}, Q_{Outlet\_HB})$ , $(H_{downstream}, Q_{Outlet\_TB})$ and $(H_{downstream}, Q_{U1})$ well. For the variable pairs: $(H_{downstream}, Q_{U2})$ and $(H_{downstream}, Q_{Total})$, all the Gaussian Copula, Gumbel Copula and Clayton Copula cannot express the dependencies well among which Gaussian Copula is the better choice.

## 5.3. Prediction of Hazard Boundary Conditions

Back to the main question of this chapter (i.e., what are the hazard boundary conditions associated with compound flood events in the downstream reach of the HSC?), it is critical to estimate the most likely hydraulic boundary conditions given certain joint occurrence probabilities. Based on the definition of compound event as stated in Section 1.1, the compound flood event in this thesis can be divided into two situations: (1) the upstream and downstream hydraulic boundary condition are extremes simultaneously; (2) the hydraulic boundary conditions are not extremes simultaneously but lead to an extreme event or impact when combined. Accordingly, the occurrence probability of compound flood events can be divided into two different types: (1) $P_{OR}(U > u \cup V > v)$ i.e. at least one of the hydraulic boundary conditions are higher than an extreme value; (2) $P_{And}(U > u \cap V > v)$ i.e. both upstream and downstream hydraulic boundary conditions are higher than an extreme value. The return period corresponding to these two different probabilities (i.e., $T_{OR}$, $T_{And}$ ) can then be expressed as following (Vandenberghe, Verhoest, Onof, & De Baets, 2011):

$$T_{OR} = \frac{\mu}{P(U > u \cup V > v)} = \frac{\mu}{P_{OR}} = \frac{\mu}{1 - F_{UV}(U,V)} \quad \text{[Eq. 5.15]}$$

$$T_{And} = \frac{\mu}{P(U > u \cap V > v)} = \frac{\mu}{P_{And}} = \frac{\mu}{1 - F_U(u) - F_V(v) + F_{UV}(u,v)} \quad \text{[Eq. 5.16]}$$

Where $\mu$ is the interarrival time of two successive events which is one day (i.e. 1/365 year) in this case, $F_U$ and $F_V$ are the marginal cumulative functions of the hydraulic boundary conditions, $F_{UV}$ is the joint cumulative function of the hydraulic boundary conditions.

As introduced earlier, the joint distribution of the hydraulic boundary conditions is described by the copula, Eq. 5.15 and Eq. 5.16 can then be rewritten as (Couasnon, 2017; Salvadori, 2004; Vandenberghe et al., 2011):

$$T_{OR} = \frac{\mu}{1 - C\{F_U(u), F_V(v)\}} \quad \text{[Eq. 5.17]}$$

$$T_{And} = \frac{\mu}{1 - F_U(u) - F_V(v) + C\{F_U(u), F_V(v)\}} \quad \text{[Eq. 5.18]}$$

Where, $C$ represents the copula. When comparing $P_{And}(U>u \cap V>v)$ and $P_{OR}(U>u \cup V>v)$, it can be found that the event $(U>u \cap V>v)$ is included in the event $(U>u \cup V>v)$. Therefore, $P_{OR}(U>u \cup V>v)$ is considered as the occurrence probability of compound flood events in this thesis since it includes more potential combinations of the hydraulic boundary conditions causing compound flood events. Consequently, given a certain return period $T_{OR}$, the corresponding joint occurrence probability of the hydraulic boundary conditions can be calculated:

$$C\{F_U(u), F_V(v)\} = 1 - \frac{\mu}{T_{OR}} \qquad \text{[Eq. 5.19]}$$

In Table 5.2, the joint occurrence probabilities given the return periods of 100 years, 500 years and 1000 years are listed.

**Table 5.2: Joint occurrence probabilities given the return periods.**

| Return Period (year) | Corresponding joint probability |
|---|---|
| 100 | 0.999973 |
| 500 | 0.999994 |
| 1000 | 0.999997 |

In Figure 5.8, the potential combinations of the hydraulic boundary conditions $(H_{downstream}, Q_{Total})$ are plotted based on Gaussian Copula given the return periods $T_{OR}$ of 100-, 500-, and 1000-year. The red circles represent the most likely combinations of the hydraulic boundary conditions (i.e., having the highest probability density among all the combinations).

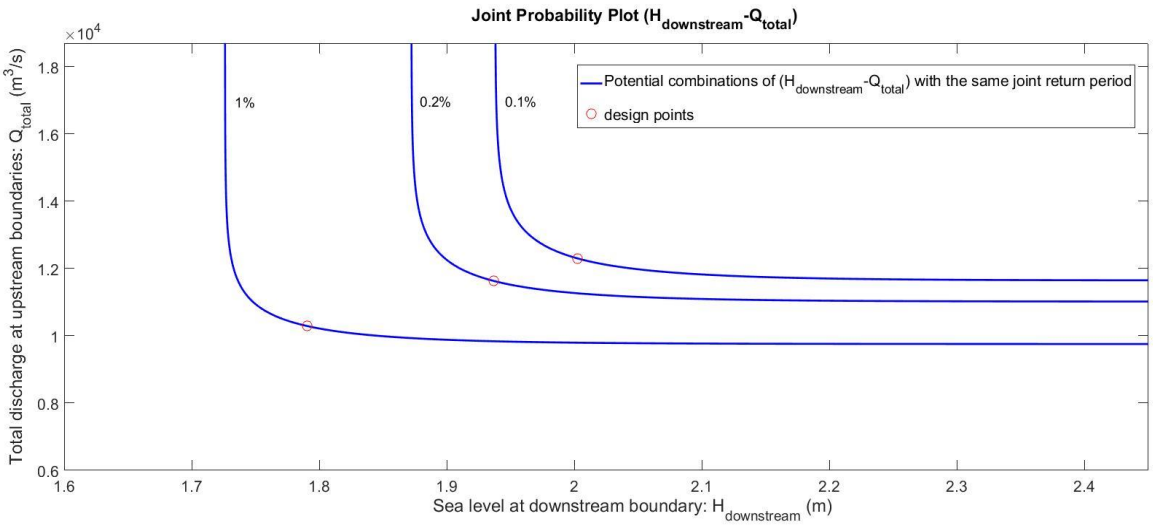



**Figure 5.8: Joint occurrence probability plot for hydraulic boundary conditions. The 100-, 500-, and 1000-year contours are shown as the blue solid lines.**

From Figure 5.8. in case of 100-year compound flood event, the most likely combination of the hydraulic boundary conditions is (1.7905m+NAVD88, 10290m$^3$/s) while for the cases of 500-year and 1000-year compound flood events, the most likely combinations of the hydraulic boundary conditions are (1.9365m+NAVD88, 11634m$^3$/s) and (2.0023m+NAVD88, 12303m$^3$/s), respectively. One thing worth noting is that the differences between the most likely boundary conditions of 100-, 500-, and 1000-year compound flood events are not large, and the reason can be found in the marginal distributions of the boundary conditions. It is already known that the fitted marginal distributions of

boundary conditions underestimate the values in the extreme cases. Additionally, the slopes of the marginal distributions (i.e., the ratio of the boundary condition to the return period) are mild in the extreme conditions (see Figure 5.6), which can explain the small differences between the predicted hazard boundary conditions.

As a comparison, when assuming the independence between the upstream and downstream hydraulic boundary conditions, the sea level at the downstream boundary of 100-year return period is 1.1679m+NAVD88 while the total discharge at upstream boundaries of the same return period is 6297m$^3$/s. Accordingly, the hydraulic boundary conditions of 500- and 1000-year return periods based on the independence assumption are (1.3213m+NAVD88, 7067m$^3$/s) and (1.3871m+NAVD88, 7435m$^3$/s), respectively. The comparison is shown in Figure 5.3 from which the hydraulic boundary conditions of compound flood events are much more severe than the boundary conditions of single flood events of the same return periods.

**Table 5.3: The comparison between compound flood event and single flood event with the same return periods.**

| | Return period (year) | $H_{downstream}$ (unit: m, datum: NAVD88) | $Q_{Total}$ (unit: m$^3$/s) |
|---|---|---|---|
| Compound floods | 100 | 1.7905 | 10290 |
| | 500 | 1.9365 | 11634 |
| | 1000 | 2.0023 | 12303 |
| | | | |
| Independent floods | 100 | 1.1679 | 6297 |
| | 500 | 1.3213 | 7067 |
| | 1000 | 1.3871 | 7435 |

## 5.4. Discussion

### 5.4.1 The Influence of the Conversion Process on the Prediction of Hydraulic Boundary Conditions

To discuss the influence of the conversion process on the prediction of hydraulic boundary conditions, it is important to review the dependencies shown in Table 5.1. By comparing the Table 5.1a and 5.1b, a significant change of the rank correlations can be found between the variable pair: $(H_{downstream}, Q_{U2})$ which shows a weak, negative correlation (-0.2606) and the variable pair: $(Morgan, SJR)$ (i.e., the daily maximum residual water level at NOAA 8770613 and water level at USGS 08072050) which shows a positive, strong correlation (0.6899). According to Eq. 5.14, the downstream sea level $H_{downstream}$ is equal to the residual water level at NOAA 8770613 (Morgan) plus a constant tide level and this conversion process does not affect the calculation of rank correlation. In addition, as introduced in Section 5.1.1, the conversion from water level at USGS 08072050 (SJR) to $Q_{U2}$ includes two steps: (1) convert from water level at USGS 08072050 (SJR) to the discharge at this station $Q_{USGS\_SJR}$ through uniform flow assumption; (2) convert from $Q_{USGS\_SJR}$ to the discharge at the U2 $Q_{U2}$ through linear relationship between the discharge and contributing area. The latter process is a linear process which does not affect the calculation of rank correlation as well. Therefore, the large decrease of the rank correlations from $(Morgan, SJR)$ to $(H_{downstream}, Q_{U2})$ is mainly caused by the uniform flow assumption. As shown in Figure 5.2c, the cross-sectional shape of the water channel at USGS 08072050 (SJR) is very irregular, which leads to the non-linear relationship between the water level and uniform discharge at USGS 08072050 (SJR) (see Figure 5.9a).

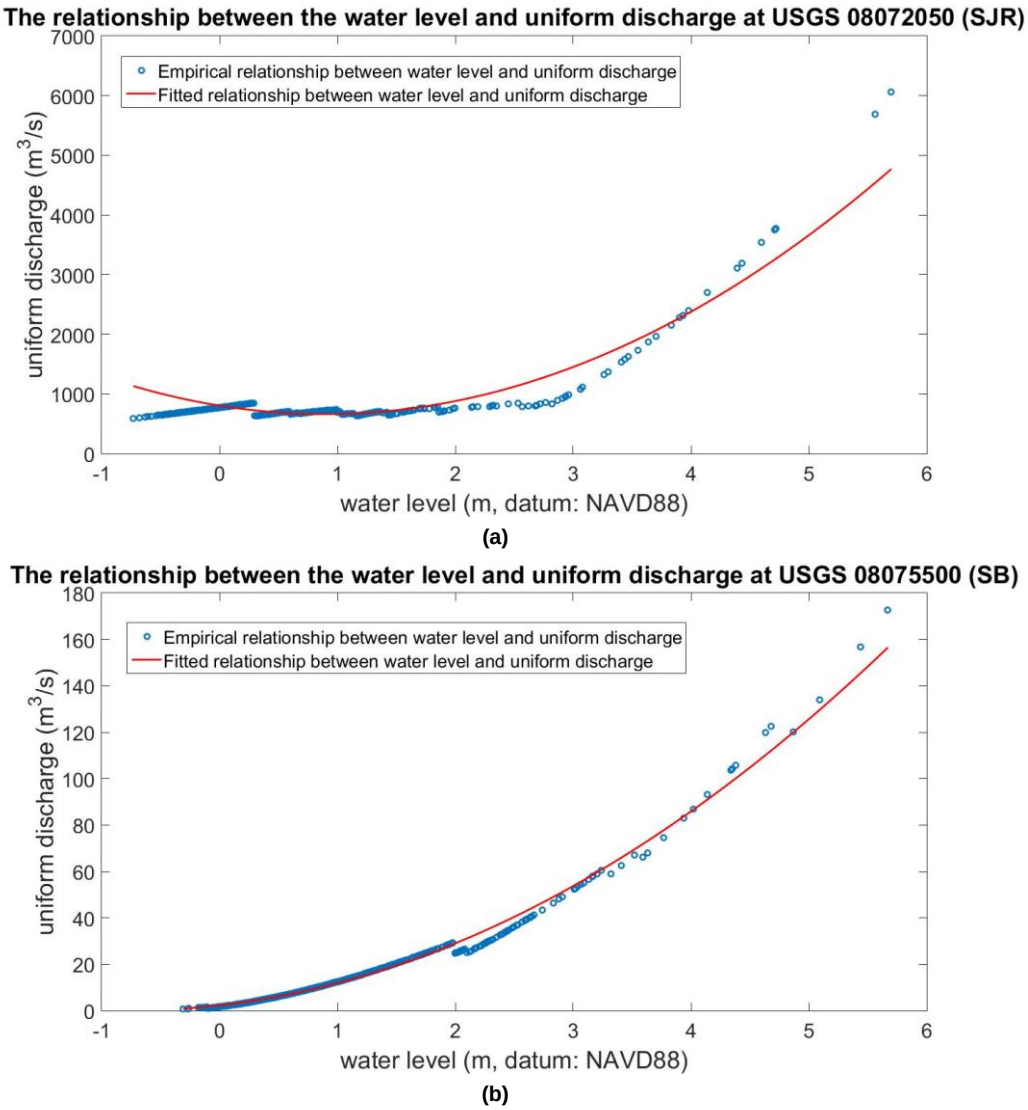



**Figure 5.9: (a) The relationship between the water level and uniform discharge at USGS 08072050 (SJR); (b) The relationship between the water level and uniform discharge at USGS 08075500 (SB). Source: USGS.**

As a comparison, the relationship between the water level and uniform discharge at USGS 08075500 (SB) is shown in Figure 5.9b. the cross-sectional shape of water channels at USGS 08075500 (SB) is quite regular (see Figure 5.2b), which causes a more linear relationship between water level and uniform discharges and therefore, the rank correlations between $(H_{downstream}, Q_{Outlet\_SB})$ and $(Morgan, SB)$ are close (see Table 5.1).

In conclusion, it is found that the uniform flow assumption will influence the dependencies between the discharge at the U2 and sea level at downstream boundary significantly, which may further cause the inaccuracy for estimating the marginal distribution and dependence of hydraulic boundary conditions and influence the prediction of hazard boundary conditions. Therefore, in the future study, using rating curve to convert water level into discharge is still a better choice.

### 5.4.2 The Influence of the Order of Building BN and Converting Hydraulic Boundary Conditions

In addition to the influence of conversion process itself, another concern comes from the order of building the BN and the conversion of the hydraulic boundary conditions. In this thesis, the BN is built first, and the hydraulic boundary conditions are converted based on the outputs of the BN. However, it is also possible to reverse the order. To test the

influence of changing the order, the originally collected data at selected upstream gauging stations are converted into the discharges at the outlets of the corresponding watersheds first, and then build the BN. However, the marginal distributions and dependencies of the hydraulic boundary conditions generated from the new order are still similar to those used in this thesis. Therefore, it is considered that changing the order will not influence the prediction of hydraulic boundary conditions significantly.

## 5.5. Conclusion

This chapter focuses on the first research question: “What are the hazard boundary conditions associated with compound flood events in the downstream reach of the HSC?”. According to the introduction in the previous chapters, the first research question is equivalent to estimate of the most likely hydraulic boundary conditions given certain joint occurrence probabilities. Based on the uniform flow assumption and linear relationship assumption between the discharge and contributing area, the outputs of the BN model in Chapter 4 are converted into the required hydraulic boundary conditions (i.e., the sea level at the downstream boundary $H_{downstream}$ and total discharge at the upstream boundaries $Q_{Total}$) at first. Thereafter, according to the fitted continuous marginal distributions and the dependence between the boundary conditions, Gaussian Copula is used to describe the joint distribution of the hydraulic boundary conditions. Finally, in Section 5.3, the return period of compound flood event is linked with joint occurrence probability of the hydraulic boundary conditions and the hazard boundary conditions associated with compound flood events are predicted. In conclusion, the predicted most likely hazard boundary conditions are (1.7905m+NAVD88, 10290$m^3$/s) corresponding to the 100-year compound flood event, (1.9365m+NAVD88, 11634$m^3$/s) corresponding to the 500-year compound flood event and (2.0023m+NAVD88, 12303$m^3$/s) corresponding to the 1000-year compound flood event. In the next chapter, these predicted hazard boundary conditions will be inputted in the 1-D Steady Hydraulic Model for modelling the return frequency water levels in the study area.

# 6

# 1-D Steady Hydraulic Model

After the work in Chapter 5, the most likely hazard boundary conditions associated with 100-, 500-, and 1000-year compound flood events in the study area are estimated. Based on that, the second research question will be answered in this chapter (i.e., “What are the return frequency water levels (i.e., 1%, 0.2%, 0.1%) in the downstream reach of the HSC under compound flood events?”). The water level in the study area is modelled by a 1-D steady hydraulic model and the hydraulic boundary conditions are used as the input variables to the 1-D model. In Section 6.1, the 1-D steady hydraulic model is introduced including the theoretical background, assumptions, and limitations while the establishment of the model for the study area is presented in Section 6.2. The return frequency water levels (i.e., 1%, 0.2%, 0.1%) are modelled in Section 6.3 and the modelled results is also discussed in the same section. Finally, a summary is shown in Section 6.4.

## 6.1. General Introduction

### 6.1.1 Theoretical Background

To quantitively analyze the potential flood risk to the research area in case of compound flood events, it is important to estimate the water surface profile. In this thesis, a simple steady state, 1-D hydraulic model is built to model the water surface profile and the hydraulic boundary conditions are selected as the main input variables. In reality, the hydraulic boundary conditions are changing in time, leading to a time-dependent water surface profile. However, analyzing the water surface profile under time-dependent situations requires a very high computation capacity and usually is time-consuming. Moreover, the BN developed in the previous chapter only provides peak estimates. Therefore, to simplify the problem, it is assumed that the boundary conditions do not change over time i.e., steady state. Based on this assumption, the 1-D steady model is built based on the Energy equation i.e. Bernoulli equation (see Eq. 6.1) which is similar to the method used in “HEC-RAS” (Brunner, 2016).

$$Z_2 + Y_2 + \frac{a_2 V_2^2}{2g} = Z_1 + Y_1 + \frac{a_1 V_1^2}{2g} + h_e \qquad \text{[Eq. 6.1]}$$

In Eq. 6.1, $Z_1, Z_2$ are elevations of the channel bottom, $Y_1, Y_2$ are the water depths, $V_1, V_2$ are average velocities (total discharge/total flow area), $a_1, a_2$ are velocity weighting coefficients, $g$ is gravitational acceleration and $h_e$ is energy head loss. A graphical explanation of the energy equation is shown in Figure 6.1.

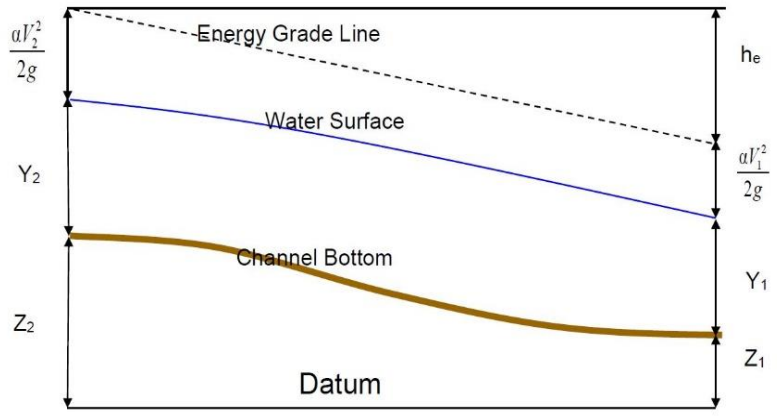



**Figure 6.1: Graphically explanation of the energy equation. Source: (Brunner, 2016).**

In more detail, the energy head loss $h_e$ can be further divided into two parts: friction loss and contraction/expansion loss which give rise to the equation for energy loss (see Eq. 6.2) (Brunner, 2016).

$$h_e = L\overline{S}_f + C\left|\frac{a_2V_2^2}{2g} - \frac{a_1V_1^2}{2g}\right| \qquad \text{[Eq. 6.2]}$$

Where, $L$ is discharge-weighted reach length, $\overline{S}_f$ is representative friction slope between two sections and $C$ is expansion/contraction loss coefficient. For the discharge-weighted reach length, the general calculation is as following:

$$L = \frac{L_{lob}\overline{Q}_{lob} + L_{ch}\overline{Q}_{ch} + L_{rob}\overline{Q}_{rob}}{\overline{Q}_{lob} + \overline{Q}_{ch} + \overline{Q}_{rob}} \qquad \text{[Eq. 6.3]}$$

Where, $L_{lob}, L_{ch}, L_{rob}$ are cross-section reach lengths for the flow in the left overbank, main channel, and right overbank and $\overline{Q}_{lob} + \overline{Q}_{ch} + \overline{Q}_{rob}$ is the arithmetic average of the discharges of the left overbank, main channel, and right overbank. In addition, the friction slope can be calculated from Manning's equation (S.I. unit):

$$S_f = \frac{Q^2n^2}{A^2R^{4/3}} \qquad \text{[Eq. 6.4]}$$

Where, $Q$ is discharge, $n$ is Manning's roughness coefficient, $A$ is cross-section area and $R$ is hydraulic radius (area/wetted perimeter). Finally, to solve the mean Kinetic Energy Head, the general Discharge-Weighted Velocity Head is shown in Eq. 6.5.

$$\frac{a\overline{V}^2}{2g} = \frac{Q_1\frac{V_1^2}{2g} + Q_2\frac{V_2^2}{2g} + \ldots + Q_N\frac{V_N^2}{2g}}{Q_1 + Q_2 + \ldots + Q_N} \qquad \text{[Eq. 6.5]}$$

Where, $Q_1, Q_2, ...Q_N$ are the discharges of subareas of flow area and $V_1, V_2, ...V_N$ are the mean velocities of the subareas. One example of mean kinetic energy head with two subareas is shown in Figure 6.2.

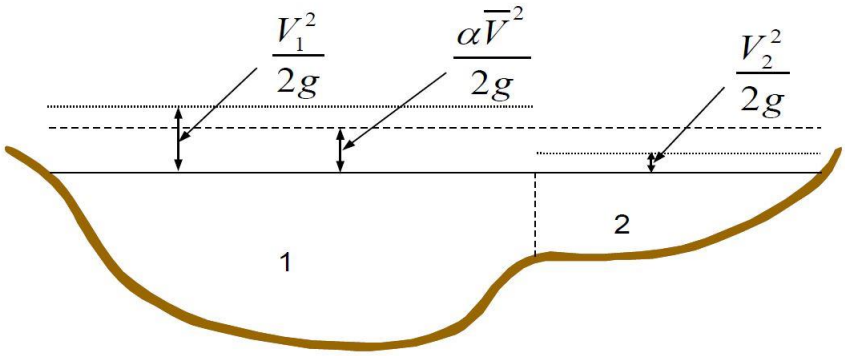



**Figure 6.2: An example of the mean kinetic energy. Source: (Brunner, 2016).**

In this project, because of the limitation of obtaining the discharges of different subareas, the flow area is considered as a whole, and the corresponding mean kinetic energy head is therefore calculated as Eq. 6.6.

$$E_{kinetic} = \frac{Q^2}{2gA^2} \quad \text{[Eq. 6.6]}$$

### 6.1.2 Assumptions and Limitations

Once the 1-D steady model is chosen, there are some implicit assumptions as following (Brunner, 2016):

- Flow is steady;
- Flow is gradually varied since the pressure head in Eq. 6.1 obey hydrostatic distribution;
- Flow is one dimensional which means that the lateral flow is not considered here and the total energy head is unvaried within one cross-section;
- The bottom slope is less than 1:10.

## 6.2. Model Setup

### 6.2.1 Division of the Cross Sections

Since there is no official divided cross section information in the research area the division of cross-sections need to be done before modelling. Based on the one-meter contours obtained from Texas Natural Resources Information System (TNRIS), a total of 18 cross sections are selected along the research area which is shown in Figure 6.3.

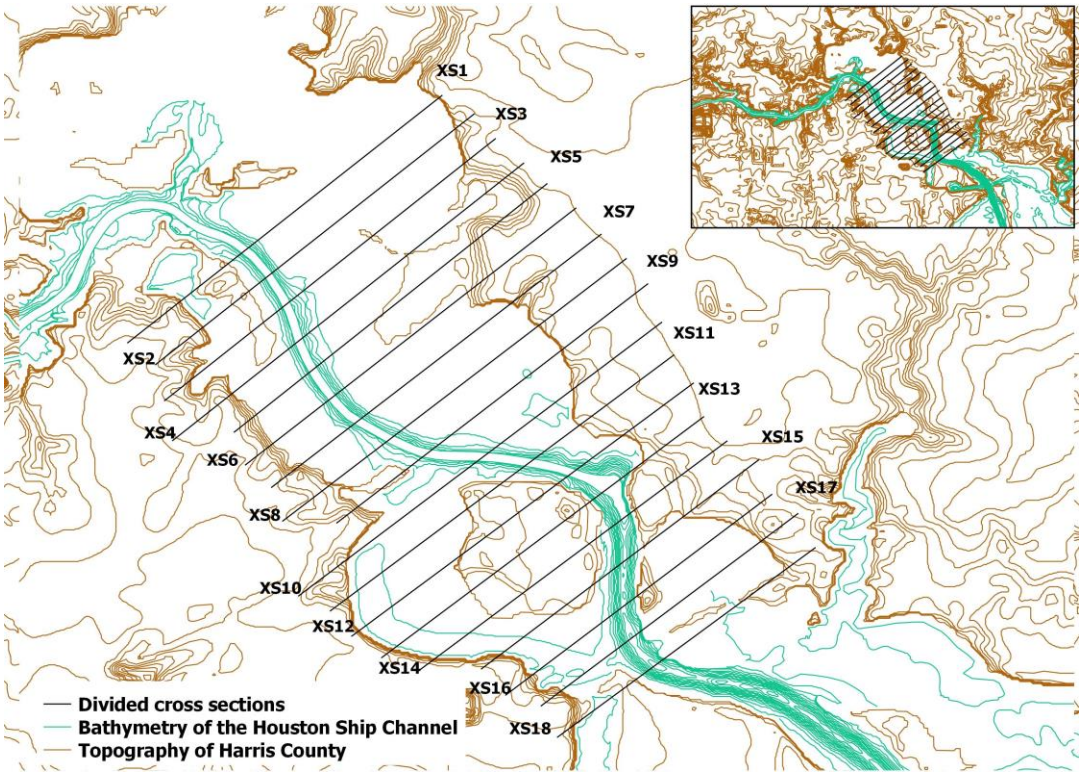
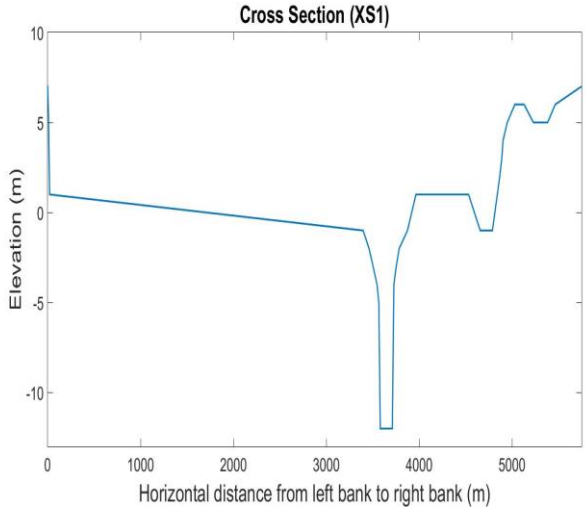
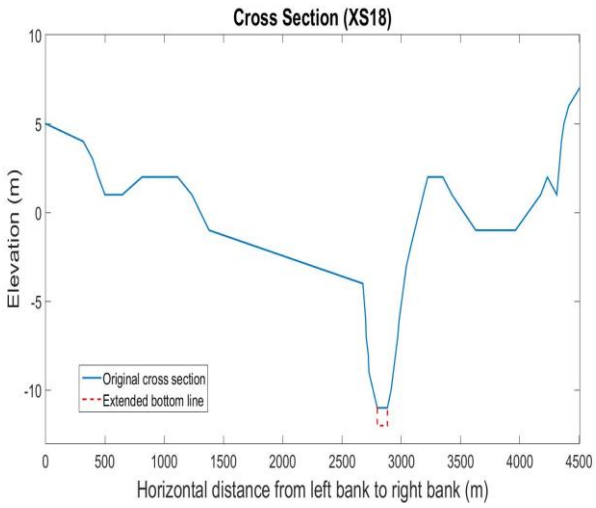



**Figure 6.3: The division of the cross sections along the research area. Source: Texas Natural Resources Information System (TNRIS).**

The division of the cross sections starts from the upstream boundary and ends at the downstream boundary while all the cross sections are parallel each other and interval between two adjacent cross sections is about 500 meters. This was considered to be sufficient resolution because there are no significant changes in the channel between the cross-sections especially the dredged portions. The HSC is considered as the main channel while the surrounding area is assumed as the floodplain. According to the information of contour lines, the deepest bathymetry contour line of the HSC is -12m + NAVD88 which is close to the dredged depth of the HSC i.e. 40 feet (12.19m). However, for the cross-section XS15, XS16, XS17 and XS18 locating near the downstream boundary, the captured deepest contour lines are -10m + NAVD88, -10m + NAVD88, -9m + NAVD88 and -11m + NAVD88 respectively which are much shallower than the real bottom of the HSC. To make the divided cross sections closer to the actual cross-section, the bottom of all the cross sections are extended to -12m + NAVD88. In addition, based on the predicted 100-year flood depth from FEMA, the highest water level along the research area is 21 feet (6.4m). Therefore, the highest contour lines of both left and right floodplain are selected at 7m + NAVD 88 which are high enough to deal with the potential high-water level with 100-year return period. In Figure 5.4, the cross sections XS1 and XS18 are shown for reference. The complete cross sections are in Appendix G. It is worth mentioning that the selected floodplain is not the actual floodplain since there are many harbour facilities included in the selected floodplain.




**(a) Cross section XS1**




**(b) Cross section XS18**

**Figure 6.4: Cross section XS1 and XS18. Source: HCFCD.**

### 6.2.2 Determination of Parameters

**Manning roughness coefficient n**

Manning roughness coefficient $n$ is an important parameter in the model which represents the bed roughness of the main channel and floodplain and is mainly related to properties of the bed material (Limerinos, 1970). In this case, it is hard to find the accurate Manning roughness coefficient of the research area directly. Thus, in this thesis, the $n$ is estimated using the measured Manning roughness coefficient of the upstream reach of the HSC. According to the Federal Emergency Management Agency's (FEMA) effective floodplain models for Harris County, the averaged the Manning roughness coefficient of the upstream reach of the HSC including the main channel and floodplain is nearly 0.06. Therefore, the Manning roughness coefficient of the research area is assumed to be 0.06 as well for the entire cross sections.

**Expansion/contraction coefficient C**

As introduced in Section 6.1, the expansion/contraction coefficient is used to calculate the energy loss due to the change of cross-sectional shape. Since the HSC is a well-dredged channel the cross sections of which do not differ significantly and there is no obstacle along the HSC, it is assumed that the expansion/contraction coefficient is zero.

## 6.3. Return Frequency Water Levels

Based on the 1-D steady hydraulic model introduced above, the return frequency water levels in the downstream reach of the HSC under the most likely 100-, 500-, and 1000-year hydraulic boundary conditions are modelled (see Figure 6.5). In Figure 6.5, the 1% return frequency water level is represented by the pink line while the blue line and the green line represent the 0.2% and 0.1% return frequency water levels, respectively. In addition, as a comparison, the Base Flood Elevation (BFE) in the study area is also shown in Figure 6.5. Here, the BFE refers to the water elevation corresponding to the 100-year flood event.

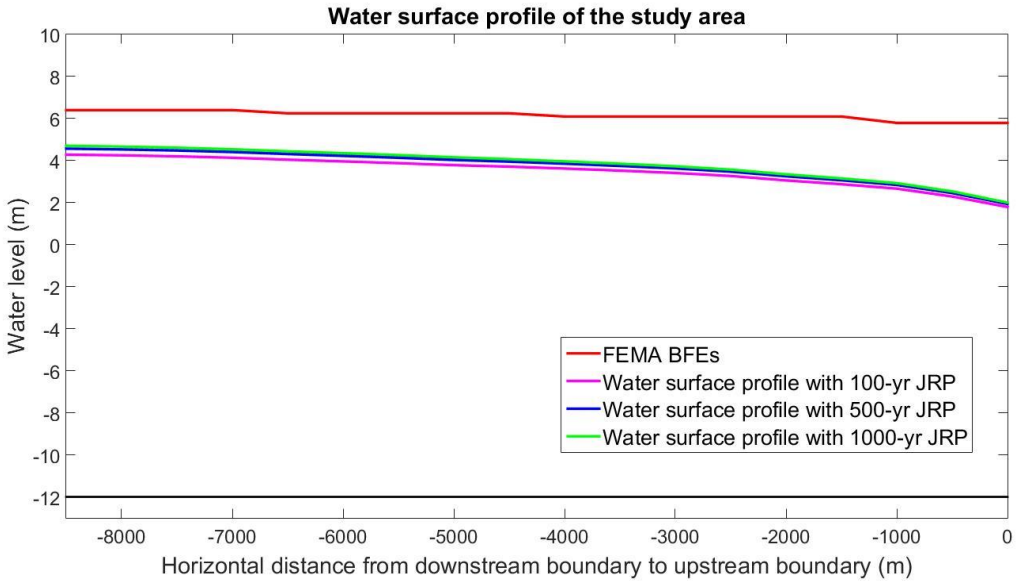



**Figure 6.5: Modelled return frequency water levels (i.e., 1%, 0.2%, 0.1%) in the study area.**

In general, the 1% return frequency water level is 4.28m+NAVD88 at the upstream boundary and 1.79m+NAVD88 at the downstream boundary while the 0.2% return frequency water level is 4.57m+NAVD88 at the upstream boundary and 1.94m+NAVD88 at the downstream boundary; the 0.1% return frequency water level is 4.70m+NAVD88 at the upstream boundary and 2.00m+NAVD88 at the downstream boundary. The water levels close to the downstream boundary are more affected by the downstream boundary conditions while the water levels in the mid-to-upper reach are higher because of the influence from the upstream discharge.

There are two problems can be seen from Figure 6.5. The first one is that the modelled return frequency water levels are much lower than BFE especially at the downstream boundary. The BFE is about 21 ft. (6.4m) at the upstream boundary of the study area and about 19 ft. (5.8m) at the downstream boundary. This is mainly because of the underestimation of hazard boundary conditions.

The second problem is that the differences between different return frequency water levels are small. The main reason causing this problem is the small differences between the inputted hydraulic boundary conditions. In Section 5.3. It is already discussed that the differences between the predicted most likely boundary conditions of 100-, 500-, and 1000-year compound flood events are not large and the inaccuracy of prediction of the hazard boundary conditions directly affects the results of return frequency water levels.

## 6.4. Conclusion

In this chapter, a 1-D steady hydraulic model of the study area is used to estimate the return frequency water levels at the downstream reach associated with the most likely 100-, 500-, and 1000-year hydraulic boundary conditions. The modelled return frequency water levels are shown in Figure 6.5. In conclusion, the 1% return frequency water level is 4.28m+NAVD88 at the upstream boundary and 1.79m+NAVD88 at the downstream boundary while the 0.2% return frequency water level is 4.57m+NAVD88 at the upstream boundary and 1.94m+NAVD88 at the downstream boundary; the 0.1% return frequency water level is 4.70m+NAVD88 at the upstream boundary and 2.00m+NAVD88 at the downstream boundary.

# 7

# A Case Study: Hurricane Harvey

As introduced in Section 2.3, Hurricane Harvey made landfall as a Category 4 hurricane near Rockport, Texas. During this event, the storm brought considerable rainfall to the Houston region and water levels in Galveston Bay remained high for a number of days as well, raising the question as to the 'compound' nature of the event and to what extent elevated water levels in Galveston Bay exacerbated flooding in the study area. Therefore, a case study of Hurricane Harvey is presented in this section. In Section 7.1, some general information about Hurricane Harvey is introduced. In addition, to explore the 'compound' nature, the related sea level data, discharge data and riverine water level data is collected in Section 7.2 while the estimation of the hydraulic boundary condition during Harvey is shown in Section 7.3. Thereafter, In Section 7.4, the water levels in the study area under the estimated hydraulic boundary conditions are modelled based on the 1-D Steady Hydraulic Model built in Chapter 6. The summary is in Section 7.5. This case study is also a good example to test the practicability of the methodology used in this thesis.

## 7.1 General Introduction

Tropical Storm Harvey was the 8$^{th}$ named storm of the Atlantic hurricane season in 2017. It formed approximately 400 kilometers east of Barbados on August 17, 2017, and moved toward the west with a wind speed of about 30 km/h. As it moved across the eastern Caribbean Sea, Tropical Storm Harvey weakened and it degenerated into an open sea. However, the remnant of Harvey moved toward the Gulf of Mexico and because of the warm water, redeveloped into a hurricane on August 24, 2017. The regained hurricane moved northward and made a landfall on the central Texas coast near Rockport at 10 pm, August 25 with a maximum sustained wind speed of near 215 km/h based on which Harvey is defined as a Category 4 hurricane at landfall.

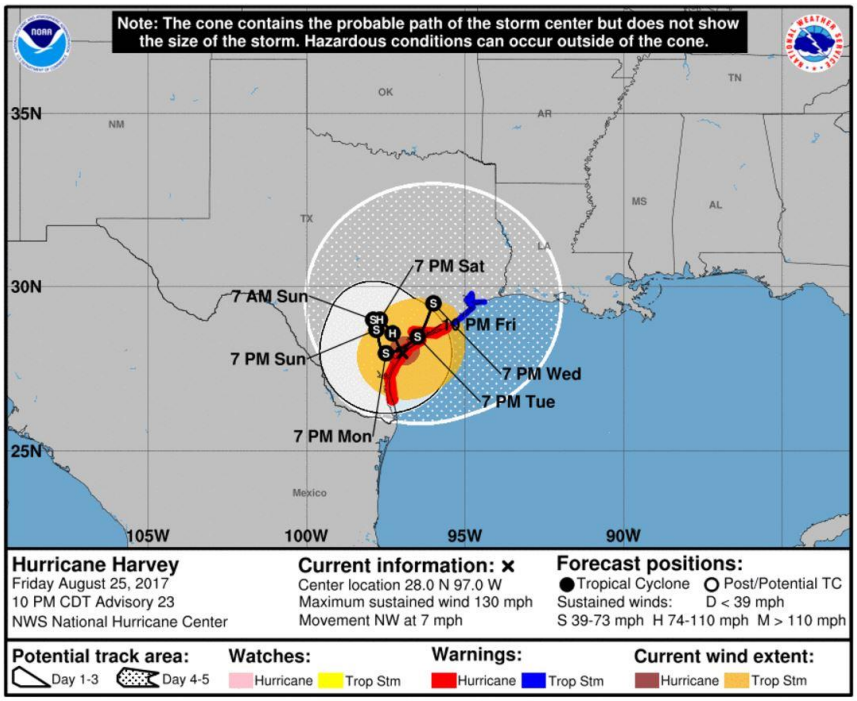



**Figure 7.1: Hurricane Harvey forested track and warnings. Source: The National Hurricane Center, NOAA.**

Hurricane Harvey generated some storm surge and significant winds which damaged coastal towns in the central portions of the Texas coast. However, most notably, Hurricane Harvey remained nearly stationary over coastal Texas until August 30 during which it caused unprecedented heavy rainfall especially in Greater Houston that led to substantial flooding. As shown in Figure 7.2, the total rainfall from August 25 to August 29 in Greater Houston region was over 50 inches (1270mm).

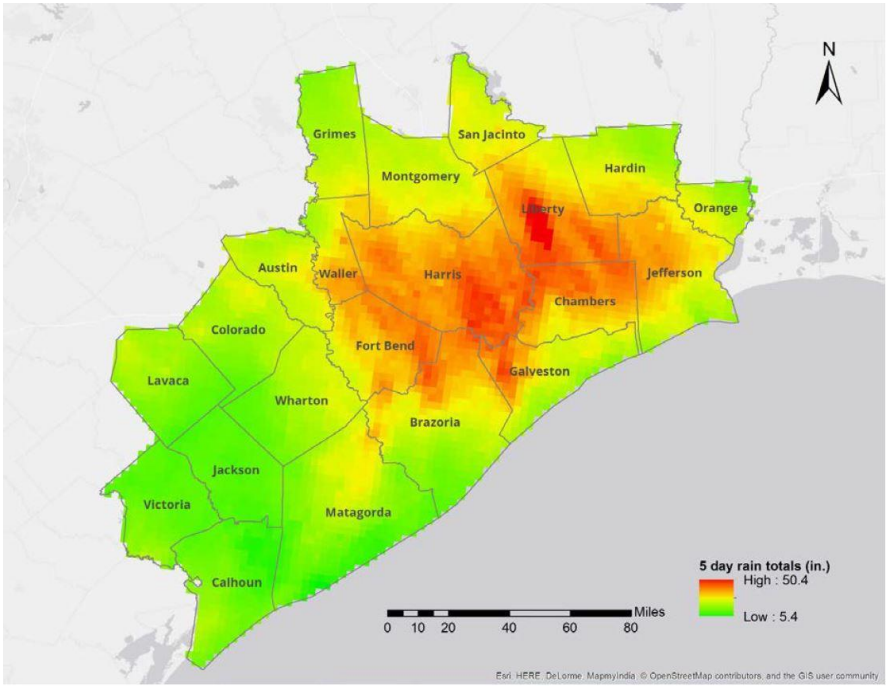


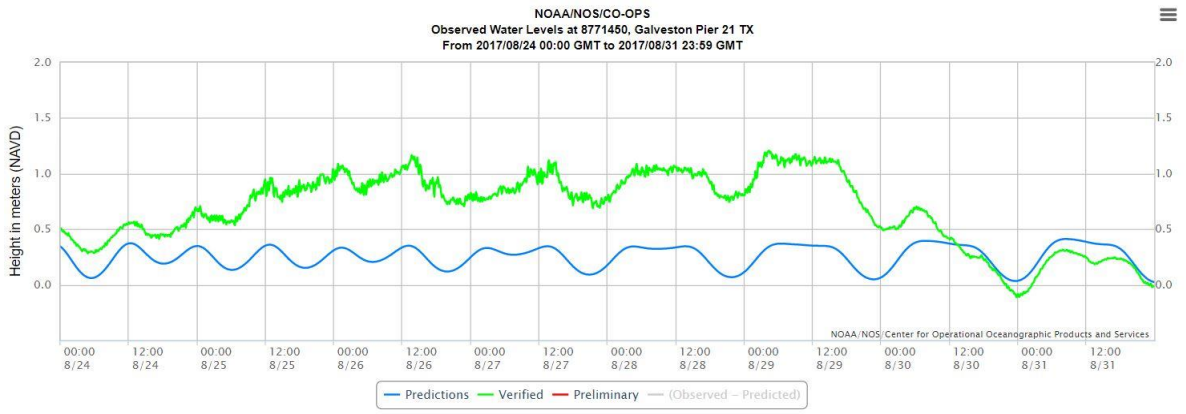

**Figure 7.2: Harvey rainfall totals for South Texas from August 25 to August 29. Source: (SSPEED, 2017).**

Although there is no official estimation of the loss caused by Hurricane Harvey, it is still believed that Hurricane Harvey could be the most expensive or second-most expensive natural disaster in the United States (Quealy, 2017). According to the statement of Texas state governor Greg Abbott, it is estimated that the damages will be between $150 billion and $180 billion, surpassing the $108 billion that caused by Hurricane Katrina in 2005 (Blake et al., 2007; Parraga & McWilliams, 2017). In addition, more than 185,000 homes were damaged and 9,000 destroyed until September 1 based on the statement of Texas Department of Public Safety and the number is still rising (Carroll & Dart, 2017).

## 7.2. Data Collection from Hurricane Harvey

The discharge and water level data at the selected upstream USGS stations during Hurricane Harvey are collected for the period between August 23 and September 3, which corresponds to the rainfall event and possible lag time for watershed response. During this period, it is found that the highest discharge at USGS 08075000 (BB) and USGS 08075770 (HB) were 925.41 $m^3/s$ on August 28 and 279.89 $m^3/s$ on August 27, respectively, both of which are the historical highest recorded discharges. In addition, the highest water levels at USGS 08074710 (TB), USGS 08075500 (SB), USGS 08076700 (GB) and USGS 08072050 (SJR) were 3.67m+NAVD88 on August 29, 8.10m+NAVD88 on August 27, 5.07m+NAVD88 on August 26 and 6.72m+NAVD88 on August 28, respectively, all of which are also the worst cases in the history. It is worth mentioning that, USGS station 08076700 (GB) failed between August 27 and August 31, which means the real situation could be worse.

Apart from the upstream situation, the measured highest daily maximum water level during Hurricane Harvey at NOAA 8771450 (GP21) was 1.166m+NAVD88 on August 26 while the highest daily maximum residual water level during the same period was 0.874m+NAVD88 on August 29. In addition, the measured highest daily maximum water level at the NOAA 8770613 (Morgan) was 1.445m+NAVD88 on August 27 and the highest daily maximum residual water level was 1.238m+NAVD88 on August 29. The detailed sea level variations are shown in Figure 7.3.




**(a)**


**(b)**

**Figure 7.3: (a) The measured water level marked by the green line and tide level marked by the blue line at NOAA 8771450 (GP21); (b) The measured water level marked by the green line and tide level marked by the blue line at NOAA 8770613 (Morgan). Source: NOAA.**

Here, the highest daily maximum residual water level at NOAA 8771450 (GP21) is analyzed specifically for making an intuitive judgment on the impact of Hurricane Harvey on the downstream sea level. The marginal distribution of daily maximum residual water levels at NOAA 8771450 (GP21) from 01/01/1904 to 01/03/2017 was estimated in Section 3.4.1 (see Table 3.9 and Figure 3.4a). Based on the marginal distribution, the return period of the highest daily maximum residual water level during Hurricane Harvey (i.e., 0.874m+NAVD88) is greater than 1000 years compared to both empirical distribution and fitted continuous distribution (see Figure 7.4). In addition, when comparing Hurricane Harvey to other events occurred recently (i.e., Tropical Allison (2001) and Hurricane Ike (2008)), it is found that the highest daily maximum residual water level at NOAA 8771450 (GP21) during Hurricane Harvey is slightly greater than the value during Tropical Storm Allison (2001) (i.e., 0.65m+NAVD88) and far less than the value during Hurricane Ike (2008) (i.e., 2.74m+NAVD88) (see Figure 7.4).

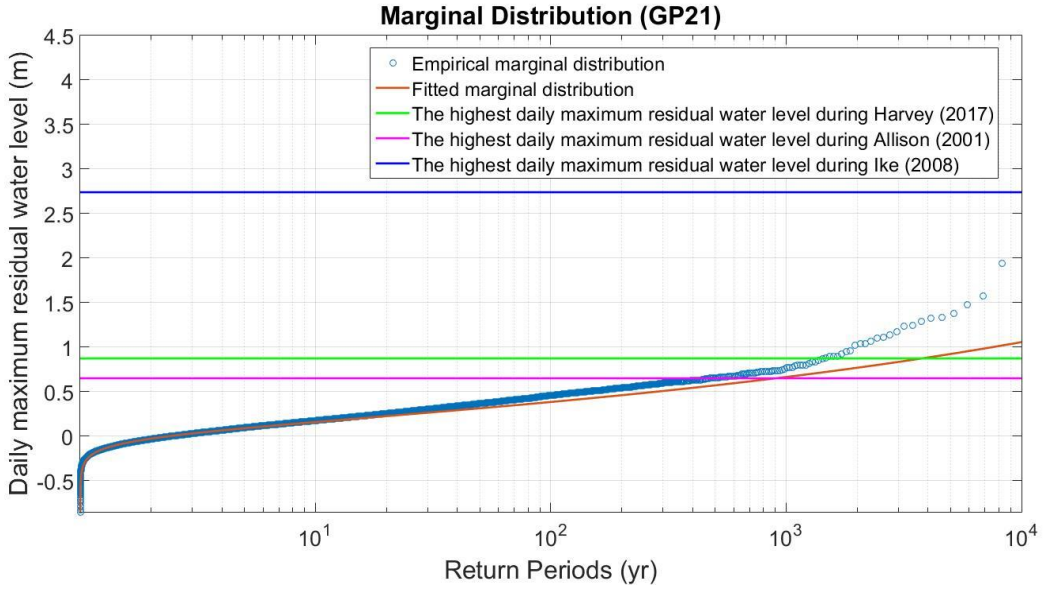



**Figure 7.4: Marginal distribution plot for the daily maximum residual water level at NOAA 8771450 (GP21). The fitted continuous marginal distribution is marked by the red solid line and empirical marginal distribution is marked by blue circles.**

## 7.3. Hydraulic Boundary Conditions during Hurricane Harvey

Given the collected data above, the hydraulic boundary conditions of the downstream reach of the HSC during Hurricane Harvey can be estimated based on the assumptions mentioned in Section 5.1. Thereafter, water levels of downstream reach of the HSC under the estimated boundary conditions can be modelled. In this section, the hydraulic boundary conditions are estimated based on the collected data and the BN model respectively. Since the highest sea level collected at the NOAA 8770613 (Morgan) gage was on August 27, 2017, that day is considered to be the most significant day and the hydraulic boundary conditions on that day are chosen for analysis. The hydraulic boundary conditions estimated from the collected data are listed in Table 7.1 where the total discharge at upstream boundaries is 15229.36 $m^3/s$ while the sea level at the downstream boundary is 1.445m+NAVD88.

**Table 7.1: Estimated hydraulic boundary condition on August 27 based on the collected data. Data source: NOAA; USGS.**

| | Station | Collected highest water level (m, NAVD88) | Collected highest residual water level (m, NAVD88) | |
|---|---|---|---|---|
| Downstream | NOAA 8770613 (Morgan) | 1.445 | 1.182 | |
| | Station | Collected water level (m, NAVD88) | Estimated discharge at station ($m^3/s$) | Estimated discharge at outlet ($m^3/s$) |
| Upstream | USGS 08072050 (SJR) | 3.92 | 6627.14 | 10133.24 |
| | USGS 08076700 (GB) | 5.07 | 189.33 | 220.66 |
| | USGS 08075500 (SB) | 8.10 | 781.07 | 1165.77 |
| | USGS 08074710 (TB) | 3.34 | 1860.46 | 1949.64 |
| | Station | Collected discharge ($m^3/s$) | Estimated discharge at station ($m^3/s$) | Estimated discharge at outlet ($m^3/s$) |
| | USGS 08075000 (BB) | 925.41 | 925.41 | 1238.84 |
| | USGS 08075770 (HB) | 279.89 | 279.89 | 521.21 |
| | Estimated total discharge | | | 15229.36 |

Note: For the USGS 08076700 (GB) and USGS 08075000 (BB), the stations failed on August 27. Therefore, the water level at USGS 08076700 (GB) on August 27 is estimated by the highest collected water level from August 25 to August 31 while the discharge at USGS 08075000 (BB) on August 27 is estimated by the highest collected discharge.

As a comparison, the hydraulic boundary conditions are also estimated through the BN model. To test the influence of the downstream high residual water level on the upstream discharges or riverine water levels, the BN model is conditionalized given the highest daily maximum residual water level at NOAA 8770613 (Morgan) on August 27 (i.e., 1.182m+NAVD88).

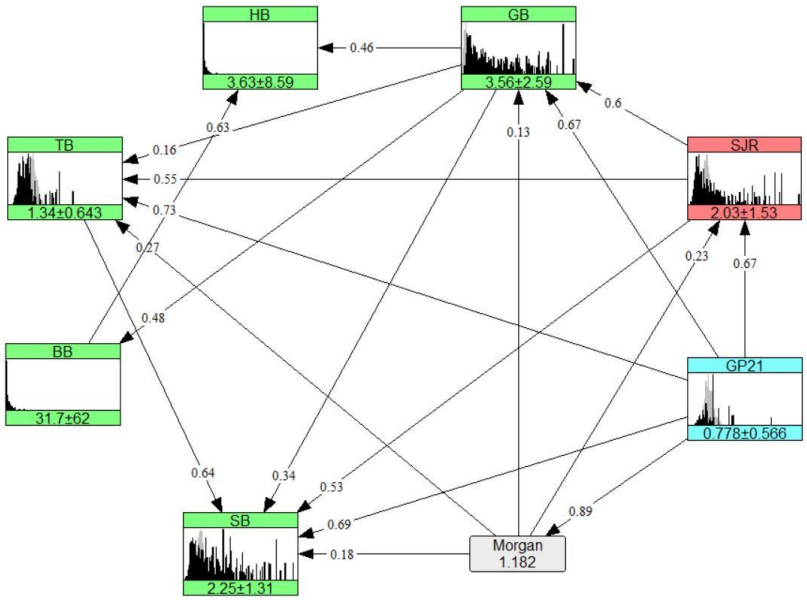



**Figure 7.5: The BN given the daily maximum residual water level at NOAA 8770613 (Morgan) on August 27.**

Based on the conditionalized BN model and the conversion process introduced in Section 5.1, ten-thousand samples of upstream boundary conditions $Q_{Total}$ are generated. The values of the samples are from 4390.34 $m^3/s$ to 31106.48$m^3/s$ while the most likely value is about 5000 $m^3/s$. A histogram representing the empirical probability density function of the upstream boundary condition generated from the BN is shown in Figure 7.6. Comparing the upstream boundary conditions estimated from two different ways, the value estimated from the measured data is included in the values estimated from the BN although the most likely value estimated from the BN is much lower than the former. In the following section, these estimated hydraulic boundary conditions will be used to predict the return period of Hurricane Harvey and the corresponding water levels in the study area.

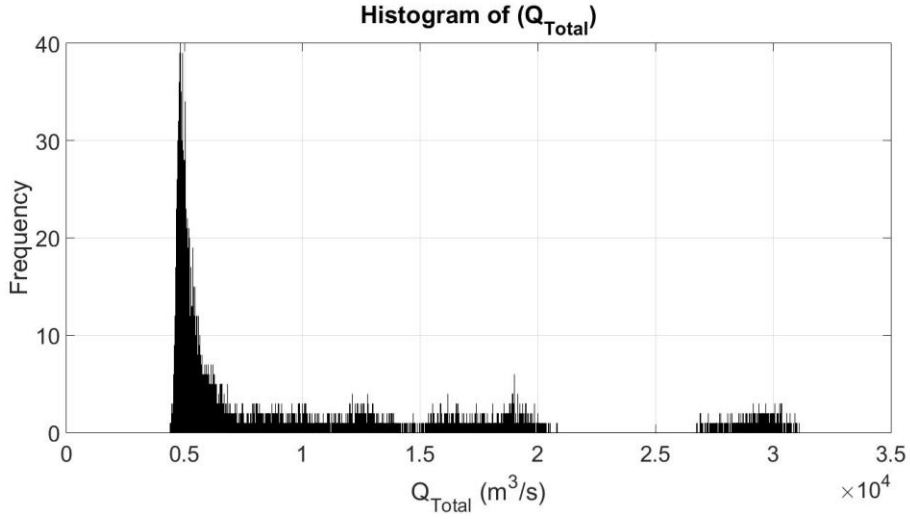



**Figure 7.6: The histogram of generated upstream boundary conditions based on the BN.**

## 7.4. Result Analysis

### 7.4.1 The Return Period of Hurricane Harvey

In Section 5.3, it is already introduced that the return periods of compound flood events are determined by the joint occurrence probabilities of upstream and downstream hydraulic boundary conditions in this thesis. Therefore, based on the estimated hydraulic boundary conditions on August 27, 2017, the return period of Hurricane Harvey can be predicted. Since the hydraulic boundary conditions estimated from the collected data is closer to the real situation, these boundary conditions are used to calculate the return period (see Table 7.2).

**Table 7.2: The return periods of Hurricane Harvey.**

| Upstream Boundary Condition ($m^3/s$) | Exceedance Probability based on Marginal distribution | Joint Occurrence Probability based on Gaussian Copula | Return Periods (yr) |
|---|---|---|---|
| 15229.36 | 1.1327E-7 | 5.4307E-4 | 5.0449 |
| Downstream Boundary Condition (m) | Exceedance Probability based on Marginal distribution | | |
| 1.445 | 5.4296E-4 | | |

The estimated return period of Harvey is about 5 years, not a very high return period, despite Hurricane Harvey has brought the unprecedented precipitation. From Table 7.2, it is observed that the estimation of return period mainly depend on the downstream boundary condition in this case since the joint occurrence probability of boundary conditions is close to the marginal exceedance probability of the downstream boundary condition when the marginal exceedance probability of the upstream boundary condition is extremely low. Therefore, although the upstream discharge is extremely high, the downstream boundary condition mainly influences the return period. In Figure 7.7, the hydraulic boundary conditions estimated from the collected data are plotted in the joint probability graph.

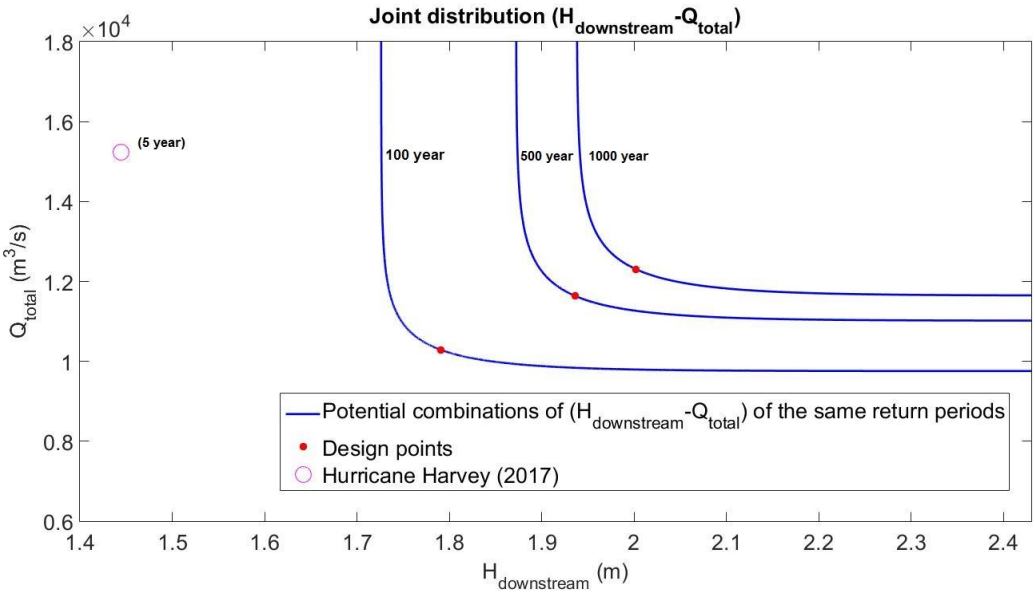



**Figure 7.7: The position of the hydraulic boundary conditions during Hurricane Harvey in the Joint occurrence probability graph.**

### 7.4.2 Modelled Water Level in the Study Area

In addition to the return period of Harvey, the water levels in the downstream reach of the HSC during Harvey is also interested to know. Based on the generated hydraulic boundary conditions in Section 7.3, the water levels on August 27, 2017, are modelled (see Figure 7.8).

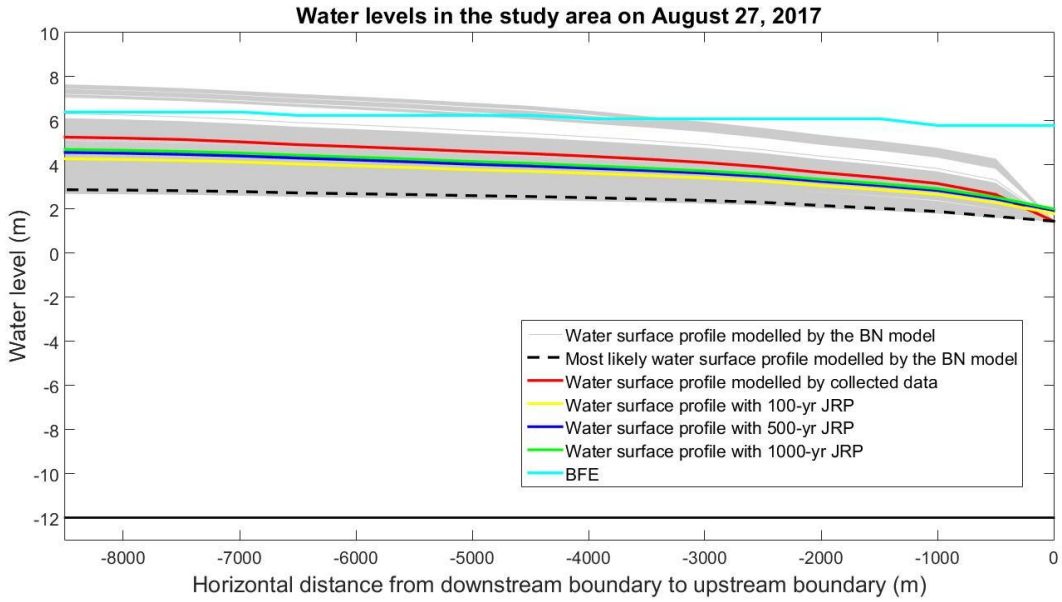



**Figure 7.8: Water surface profile plot for the downstream reach of the HSC on August 27, 2017.**

In Figure 7.8, the water levels corresponding to the hydraulic boundary conditions resulting from the conditionalized BN are shown as the grey solid lines; and the modelled water levels at the upstream boundary could be from about 3m+NAVD88 to nearly 7.9m+NAVD88 with a difference of approximate 4.9m when the sea level at the downstream boundary is selected as 1.445m+NAVD88. It can be seen that the possible highest water level (i.e., 7.9m+NAVD88) at the upstream boundary is higher than the BFE (i.e., the cyan solid line in Figure 7.8). In addition, the most likely water level (i.e., the black dashed line) at the upstream boundary is about 2.85m+NAVD88 which is much lower than the 100-year return frequency water level estimated in Section 6.3 (i.e., the yellow solid line in Figure 7.8).

The water level corresponding to the hydraulic boundary conditions resulting from the collected data is presented by the red solid line; and the water level at the upstream boundary is around 5m+NAVD88 which is higher than the 1000-year return frequency water level estimated in Section 6.3 (i.e., the green solid line in Figure 7.8), but lower than the BFE.

## 7.5. Conclusion

In this chapter, Hurricane Harvey is studied in more detail. According to the measured discharge and water level data at the selected USGS stations, the highest sea level during Hurricane Harvey at Morgan' Point (NOAA 8770613) is 1.445m+NAVD88 on August 27, 2017, while the collected highest discharge data and riverine water level data at the upstream stations are also around that day, August 27 is therefore considered to be the most significant day during Hurricane Harvey. According to the collected data at the selected upstream stream gages, the upstream boundary conditions on that day is estimated to be 15229.36 $m^3/s$ and the downstream boundary condition is selected as 1.445m+NAVD88. Based on the hydraulic boundary conditions generated from the collected data, the return period of

Hurricane Harvey is estimated to be 5 years based on the method introduced in Section 5.3. As a comparison, the upstream boundary conditions are also estimated through the BN built in Chapter 4. The result shows that upstream boundary condition could be from 4390.34 $m^3/s$ to 31106.48$m^3/s$ while the downstream boundary condition is fixed at 1.445m+NAVD88. Corresponding to the boundary conditions resulting from the conditionalized BN, the modelled water levels at upstream boundary could be from 3m+NAVD88 to 7.9m+NAVD88 while the water level corresponding to the boundary conditions generated from the collected data is about 5m+NAVD88 at the upstream boundary.

# 8

# Conclusions and Recommendations

In this chapter, some conclusions of this thesis are provided in Section 8.1. Meanwhile, based on the limitations of the methodology. some recommendations are given in Section 8.2. Finally, in Section 8.3, some suggestions for the future study are introduced.

## 8.1. Conclusion

After an overview of this thesis, it provides a methodology to analyze the influence of compound flood event on flooding in the downstream reach of the HSC. Based on this methodology, there are following findings:

- In the study area, the most likely combinations of the hydraulic boundary conditions i.e., downstream sea level and upstream discharge are 1.79m+NAVD88 and 10290$m^3$/s, respectively, for the 100-year compound flood events; 1.93m+NAVD88 and 11634$m^3$/s, for the 500-year compound flood events; and 2.00m+NAVD88 and 12303$m^3$/s, for the 1000-year compound flood events.

- The 1%, 0.2%, and 0.1% return frequency water levels in the study area are estimated. In short, the 1% return frequency water level is 4.28m+NAVD88 at the upstream boundary and 1.79m+NAVD88 at the downstream boundary while the 0.2% return frequency water level is 4.57m+NAVD88 at the upstream boundary and 1.94m+NAVD88 at the downstream boundary; the 0.1% return frequency water level is 4.70m+NAVD88 at the upstream boundary and 2.00m+NAVD88 at the downstream boundary.

- According to the analysis of Hurricane Harvey, the return period of Hurricane Harvey is predicted as 5 years which is incredibly low. It is found that the inconsistent interarrival times when calculating the return periods of a compound flood event and a single flood event lead to this unreliable value. For the return period of a compound flood event, the interarrival time between two successive events is selected as 1 day while the interarrival time is selected as 1 year for the return period of a single flood event. Based on this mistake, the real return period of Hurricane Harvey is estimated to be over 1000 year.

- Combining all the analysis results, it is found that the methodology provided in this thesis underestimates the influence of compound flood event on flooding in the downstream reach of the HSC, which is mainly caused by the bad fitness of marginal distributions of hydraulic boundary conditions. In addition, the oversimplified variable conversion process and 1-D hydraulic model also lead to the inaccuracy of the methodology.

In general, although the methodology did not give a reliable assessment on the influence of compound flood event in the study area it is still can be used in the future once the fitness of marginal distributions, the conversion process and the 1-D hydraulic model have been improved. And because of Hurricane Harvey, considering the hazard associated with compound flood events in the HSC area is proved important especially in case of designing the hydraulic structures.

## 8.2. Recommendations for Current Work

(1) The Bayesian Network

In this thesis, the compound flood events are predicted in terms of joint probability while the Bayesian Network is the main method to generate the stochastic scenarios. However, during the process of building the BN model, several assumptions are made to simplify the problem and release some calculation burden which causes the inaccuracy of the final results. Therefore, some improvements can be added to enhance the reliability of the results.

- Data collection: In this thesis, the coastal data and riverine data are collected from NOAA and USGS respectively. And because of the availability of the data, the final selected data period is from 02/10/1997 to 31/01/2017, which leads to the reduction of the valuable data. In future, it is suggested to collect the data from more reliable resources and expand the available data period.

- Variable Selection: To make the BN model simple, the discharges from Carpenters Bayou and Vince Bayou are neglected since there is no USGS station along Carpenters Bayou and discharge from Vince Bayou is small enough compared to the discharges from other watersheds. However, these two variables still make contributions to the discharge at the upstream boundary. Therefore, to make the BN model more complete, the discharges from Carpenters Bayou and Vince Bayou should be measured and analyzed in future.

- Construction of BN: The BN model here is built based on the normal copula assumption which means any correlated variable pair in the BN is jointed by Gaussian Copula. This normal copula assumption can reduce the calculation burden of the BN significantly and provide a fast response but, on the other hand, when the Gaussian Copula cannot express the dependence of a variable pair, it probably provides an inaccurate result. In Section 4.2.2, it has been proved that the normal copula assumption is not optimal for all the variable pairs. Therefore, it is worth building a BN without the normal copula assumption.

- Marginal Distributions: In this thesis, predicting the return period of a compound flood event is based on the joint distribution of the hydraulic boundary conditions where the joint distribution is linked with the fitted continuous marginal distributions. As discussed in section 5.2.1, the fitted continuous marginal distributions underestimate the hydraulic boundary conditions when the return period is greater than 200 years compared to the empirical distribution generated from the output of the BN. And this underestimation of the fitted marginal distributions affects the prediction of hazard boundary conditions and their return periods significantly. Therefore, it is recommended to improve the marginal distributions and make them fit the extreme situations better.

(2) The generation of the boundary conditions

For the generation of the required hydraulic boundary conditions, the uniform flow assumption and linear relationship assumption between the discharge and contributing area are used in this thesis to convert the outputs of the BN model to the boundary conditions. However, the assumptions above especially the uniform flow assumption are not the optimal choices. For example, after converting the water level to uniform discharge using uniform flow assumption, it is still necessary to predict the real discharge by the uniform discharge, but this prediction is difficult and inaccurate. To improve the accuracy of generated boundary conditions, the following suggestions are made:

- It is suggested to use the rating curve to transfer the water level to discharge. Although it is difficult to get information about rating curves directly as stated in Section 4.3.1, converting the water level to discharge by the rating curve is still the better way than uniform flow assumption.

- It is suggested to find the historical discharge data at the outlets of the watersheds and make the comparison between the real discharge and the assumed discharge at the outlets of the watersheds to verify the accuracy of linear relationship assumption between the discharge and contributing area.

(3) Calculation of the return period

As stated in Section 8.1, it is found that the inconsistent choice of interarrival time between the return period of a compound flood event and a single flood event lead to the unreliable prediction. The interarrival time between two successive compound events of is selected as 1 day while the interarrival time is selected as 1 year for the single flood events. Since the daily situations are considered in this thesis, it is recommendation to unify the interarrival time for both compound flood events and single flood events as 1 day.

(4) 1-D Steady Hydraulic model

To simulate the water surface profiles at the downstream reach of the HSC better, the following suggestions are made to improve the performance of the 1-D Steady Hydraulic model:

- Make a more precise division of the cross-sections.
- Estimate the Manning roughness coefficient and the energy loss between two adjacent cross sections more accurate.

## 8.3. Recommendations for Future Study

In general, this project provides a rough, preliminary judgement about the influence of potential compound floods on the downstream reach of the HSC. In future, a more advanced methodology could be used for the more reliable analysis. For instance, the 1-D Steady Hydraulic model could be replaced by 1-D Unsteady Hydraulic model or even 2-D, 3-D model. In addition to the methodology, there are still many other topics related to the compound floods that could be extended. Some suggested researches are listed below.

- Analysis of the influence of the proposed storm surge barrier on the downstream reach of the HSC in the cases of compound flood events.
- Analysis of the influence of the climate change (e.g. mean sea level rise) on the generation of compound floods.
- Analysis of the influence of the potential compound floods in the other regions of the world (e.g. Shanghai).

# Bibliography


Bass, B., Juan, A., Gori, A., Fang, Z., & Bedient, P. B. (2016). 2015 Memorial Day Storm Flood Impacts for Changing Watershed Conditions in Houston, TX. *Natural Hazards Review*, *18*(3), 1–11. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000241.

Bedient, P. B. (2015). Design of a Houston-Galveston area protection system (H-GAPS). In *DELFT DELTA DESIGN: HOUSTON GALVESTON BAY REGION TEXAS, USA* (p. 136). Delft University of Techonology.

Berg, R. (2009). *Hurricane Ike Report*. *National Hurricane Centre*. https://doi.org/10.1017/CBO9781107415324.004

Blake, E. S., Landsea, C. W., & Gibney, E. J. (2007). The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2006 (and other frequently requested hurricane facts). *National Oceanic \& Atmospheric Administration (NOAA)—National Hurricane Center (NHC), US Department of Commerce-Technical Memorandum NWS TPC-5*, *2006*(April), 1–45. https://doi.org/NWS TPC-5

Bouwer, L., & Vellinga, P. (2007). On the Flood Risk in the Netherlands. In B. S., S. M.J.F., & H. J.W. (Eds.), *Flood Risk Management in Europe* (pp. 469–484). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-4200-3_24

Brunner, G. W. (2016). *HEC-RAS River Analysis System - Hydraulic Reference Manual*.

Bureau of Economic Analysis, U. S. (n.d.). Gross domestic product (GDP) by state (millions of current dollars). Retrieved July 30, 2017, from https://www.bea.gov/iTable/.

Carroll, R., & Dart, T. (2017). Houston residents begin “massive” cleanup as Harvey death toll hits 45 | US news | The Guardian. Retrieved September 6, 2017, from https://www.theguardian.com/us-news/2017/sep/01/hurricane-harvey-death-toll-rises-houston-residents-return

Chris Landsea. (2006). TCFAQ D4) What does "maximum sustained wind" mean ? Retrieved August 5, 2017, from http://www.aoml.noaa.gov/hrd/tcfaq/D4.html

Christian, J., Fang, Z., Torres, J., Deitz, R., & Bedient, P. B. (2014). Modeling the Hydraulic Effectiveness of a Proposed Storm Surge Barrier System for the Houston Ship Channel during Hurricane Events. *Natural Hazards Review*, *77*(6), 137–147. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000150.

Couasnon, A. A. O. (2017). *Characterizing flood hazard at two spatial scales with the use of stochastic models*. Delft University of Technology.

Dahiya, R. C., & Gurland, J. (1972). Pearson chi-squared test of fit with random intervals. *Biometrika*, *59*(1), 147–153. https://doi.org/10.1093/biomet/59.1.147

Delgado-Hernández, D.-J., Morales-Nápoles, O., De-León-Escobedo, D., & Arteaga-Arcos, J.-C. (2012). A continuous Bayesian network for earth dams’ risk assessment: an application. *Structure and Infrastructure Engineering*. Taylor & Francis. https://doi.org/10.1080/15732479.2012.731416

Dorst, N. (n.d.). When is hurricane season. Retrieved June 29, 2017, from http://www.aoml.noaa.gov/hrd/tcfaq/G1.html

Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, *40*(1), n/a-n/a. https://doi.org/10.1029/2003WR002456

FEMA. (2017). *Flood Insurance Study* (Vol. 1). Technical report, FEMA.

GALVESTON BAY | The Handbook of Texas Online. (2017). Retrieved April 26, 2017, from https://www.tshaonline.org/handbook/online/articles/rrg01

Genest, C., & Favre, A.-C. (2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *Journal of Hydrologic Engineering*, *12*(4), 347–368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Gray, W. M. (1998). The Formation of Tropical Cyclones. *Meteorology and Atmospheric Physics*, *67*, 37–69. https://doi.org/10.1007/BF01277501

Greater Houston Partnership. (2016). *Houston Facts*.

Greater Houston Partnership. (2017). *Houston Facts*.

Hanea, A. M., Kurowicka, D., & Cooke, R. M. (2006). Hybrid Method for Quantifying and Analyzing Bayesian Belief Nets. *Quality and Reliability Engineering International*, *22*, 709–729. https://doi.org/10.1002/qre.808

Hanea, A., Morales Napoles, O., & Ababei, D. (2015a). Non-parametric Bayesian networks: Improving theory and reviewing applications. *Reliability Engineering and System Safety*, *144*(August), 265–284. https://doi.org/10.1016/j.ress.2015.07.027

Hanea, A., Morales Napoles, O., & Ababei, D. (2015b). Non-parametric Bayesian networks: Improving theory and reviewing applications. *Reliability Engineering and System Safety*, *144*, 265–284. https://doi.org/10.1016/j.ress.2015.07.027

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. *Innovations in Bayesian Networks*, (November). Retrieved from /home/obada/Desktop/MAS 1 - ISW/Robot Learning/A tutorial on learning with Bayesian networks.pdf

Historical Hurricane Tracks. (n.d.). Retrieved June 30, 2017, from https://coast.noaa.gov/hurricanes/

Hughes, P. (1998). The Great Galveston Hurricane. *Weatherwise*, *51*(1), 21–24. https://doi.org/10.1080/00431672.1998.9926103

Hurricane Ike. (n.d.). Retrieved June 30, 2017, from https://www.hcfcd.org/storm-center/hurricane-ike-2008/

IPCC. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation*. *Ipcc*. https://doi.org/10.1596/978-0-8213-8845-7

Jensen, F. V. (1996). Bayesian networks basics. *AISB Quarterly*, 9–22. https://doi.org/10.1145/2347736.2347755

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. CRC Press. https://doi.org/10.1201/b13150

Joe, H. (2015). *Dependence Modeling with Copulas*. *CRC Press*.

Jonkman, S. N., van Ledden, M., Lendering, K. T., Mooyaart, L., Stoeten, K. J., de Vries, P., … de Kort, R. (2013). *Report Galveston Bay: Bolivar Roads Flood Risk Reduction Barrier: Sketch Design*.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*, *1*, 11–27.

Keim, B. D., Muller, R. A., & Stone, G. W. (2007). Spatiotemporal patterns and return periods of tropical storm and hurricane strikes from Texas to Maine. *Journal of Climate*, *20*(14), 3498–3509. https://doi.org/10.1175/JCLI4187.1

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., Mcinnes, K., … Stafford-Smith, M. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(1), 113–128. https://doi.org/10.1002/wcc.252

Limerinos, J. T. (1970). *Determination of the Manning Coefficient From Measured Bed Roughness in Natural Channels*. *USGS Water-Supply Paper 1898-B*.

McAdie, C. J., Landsea, C. W., Neumann, C. J., David, J. E., Blake, E. S., & Hammer, G. R. (2009). *Tropical Cyclones of the North Atlantic Ocean 1851-2006*. *NOAA*.

Montgomery, M. T., & Farrell, B. F. (1993). Tropical Cyclone Formation. *Journal of the Atmospheric Sciences*, *50*, 285–310. https://doi.org/10.1175/1520-0469(1993)050<0285:TCF>2.0.CO;2

Morales, O., Kurowicka, D., & Roelen, A. (2008). Eliciting conditional and unconditional rank correlations from conditional probabilities. *Reliability Engineering and System Safety*, *93*(5), 699–710. https://doi.org/10.1016/j.ress.2007.03.020

Morales Napoles, O., Worm, D., van den Haak, P., Hanea, A., Courage, W., & Miraglia, S. (2013). *Introduction to Bayesian Networks*.

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding - A global assessment. *PLoS ONE*, *10*(3). https://doi.org/10.1371/journal.pone.0118571

NOAA. (2014). NowData - NOAA Online Weather Data. Retrieved June 29, 2017, from http://w2.weather.gov/climate/xmacis.php?wfo=hgx

Parraga, M., & McWilliams, G. (2017). Funding battle looms as Texas sees Harvey damage at up to $180 billion. Retrieved September 6, 2017, from https://www.reuters.com/article/us-storm-harvey/funding-battle-looms-as-texas-sees-harvey-damage-at-up-to-180-billion-idUSKCN1BE0TL

Penland, C., & Moore, W. P. (n.d.). *Houston Ship Channel Gate Design*.

Phillips, J. D. (2004). *a Sediment Budget for Galveston Bay*.

Plocheck, R. (n.d.). Facts | Texas Almanac. Retrieved June 28, 2017, from http://texasalmanac.com/topics/facts-profile

Quealy, K. (2017). The Cost of Hurricane Harvey: Only One Recent Storm Comes Close - The New York Times. Retrieved September 6, 2017, from https://www.nytimes.com/interactive/2017/09/01/upshot/cost-of-hurricane-harvey-only-one-storm-comes-close.html?mcubz=3

Rose, W. (1967). Catalyst of an Economy : The Houston Ship Channel. *Land Economics*, *43*, 32–43. https://doi.org/10.2307/3145562

Roth, D. (2010). Texas Hurricane History Table of Contents, 83.

Salvadori, G. (2004). Bivariate return periods via 2-Copulas. *Statistical Methodology*, *1*(1–2), 129–144. https://doi.org/10.1016/j.stamet.2004.07.002

Schlepers, M. H. (2015). *The Houston Ship Channel Barrier*. Master thesis, Delft University of Technology.

Schott, T., Landsea, C., Hafele, G., Lorens, J., Thurm, H., Ward, B., Zaleski, W. (2012). *The Saffir-Simpson Hurricane Wind Scale*. *National Hurricane Center*. Retrieved from http://www.nhc.noaa.gov/pdf/sshws.pdf

Sebastian, A., Dupuits, E. J. C., & Morales-Nápoles, O. (2017). Applying a Bayesian network based on Gaussian copulas to model the hydraulic boundary conditions for hurricane flood risk analysis in a coastal watershed. *Coastal Engineering*, *125*, 42–50. https://doi.org/10.1016/j.coastaleng.2017.03.008

Sibley, M. M. S. (2017). HOUSTON SHIP CHANNEL | The Handbook of Texas Online| Texas State Historical Association (TSHA). Retrieved April 26, 2017, from https://www.tshaonline.org/handbook/online/articles/rhh11

SSPEED. (2014). *SSPEED Center 2014 report*. Technical report, SSPEED.

SSPEED. (2015). *2015 Annual Report Houston-Galveston Area Protection System H-GAPS*. Technical report, SSPEED.

SSPEED. (2017). *Tropical Storm Harvey Summary Report – No . 1*. Technical report, SSPEED.

Stewart, S. R. (2011). *2001-Tropical Storm Allison*.

Texas Economic Development Corporation. (2016). *Texas Trade and Foreign Direct Investment 2016*.

The flood of 1953. (n.d.). Retrieved April 26, 2017, from http://www.deltawerken.com/Rescue-and-consequences/309.html

The Galveston hurricane of 1900. (2013). Retrieved June 30, 2017, from http://oceanservice.noaa.gov/news/features/sep13/galveston.html

The Port of Houston Authority. (2015). *THE 2014 ECONOMIC IMPACT OF MARINE CARGO ACTIVITY AT THE PORT OF HOUSTON ON THE STATE OF* (Vol. 17603).

Torres, J. M., Bass, B., Irza, N., Fang, Z., Proft, J., Dawson, C., Bedient, P. (2015). Characterizing the hydraulic interactions of hurricane storm surge and rainfall-runoff for the Houston-Galveston region. *Coastal Engineering*, *106*, 7–19. https://doi.org/10.1016/j.coastaleng.2015.09.004

Tropical cyclone facts. (2016). Retrieved June 29, 2017, from http://www.metoffice.gov.uk/weather/tropicalcyclone/facts

Tropical Storm Allison. (n.d.). Retrieved June 30, 2017, from https://www.hcfcd.org/storm-center/tropical-storm-allison-2001/

Uninet. (n.d.). Retrieved August 1, 2017, from http://www.lighttwist.net/wp/uninet

Upper Texas Coast Tropical Cyclones in the 1910s. (n.d.). Retrieved June 30, 2017, from https://web.archive.org/web/20080928030349/http://www.srh.noaa.gov/hgx/hurricanes/1910s.htm

Vandenberghe, S., Verhoest, N. E. C., Onof, C., & De Baets, B. (2011). A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall. *Water Resources Research*, *47*(7), 1–16. https://doi.org/10.1029/2009WR008388

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, *5*(July), 1–6. https://doi.org/10.1038/nclimate2736

Ye, Q., & Glantz, M. H. (2005). The 1998 Yangtze floods: The use of short-term forecasts in the context of seasonal to interannual water resource management. *Mitigation and Adaptation Strategies for Global Change*, *10*(1), 159–182. https://doi.org/10.1007/s11027-005-7838-7

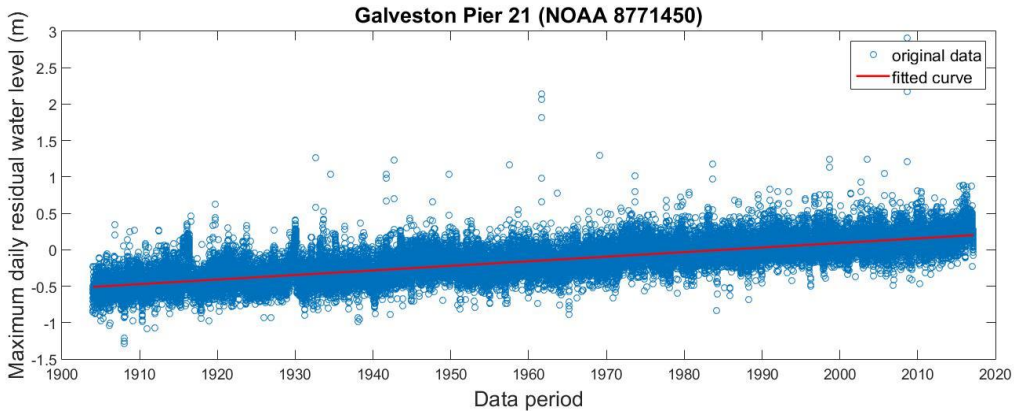
Yu, D. C., Nguyen, T. C., & Haddawy, P. (1999). Bayesian network model for reliability assessment of power systems. *IEEE Transactions on Power Systems*, *14*(2), 426–432. https://doi.org/10.1109/59.761860

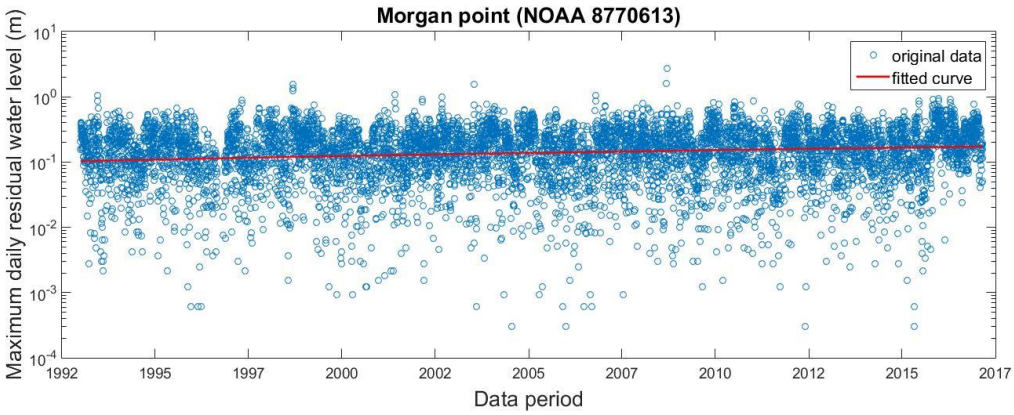
# Appendix A

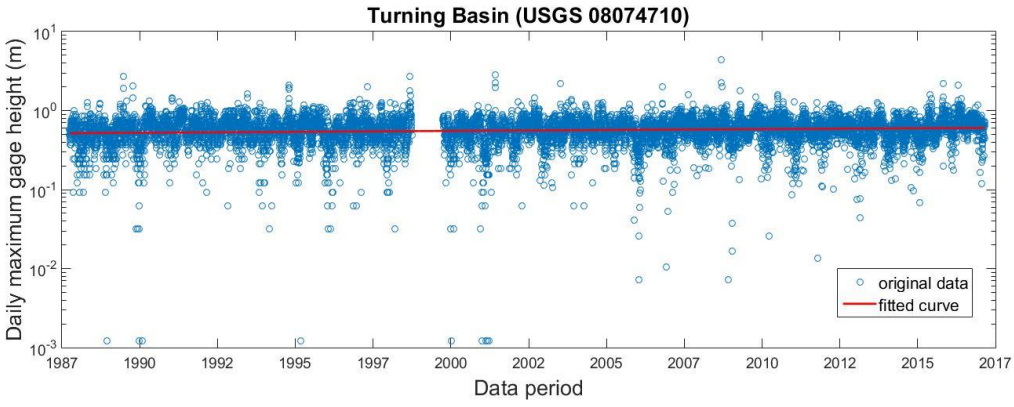
## Detailed Original Data Analysis

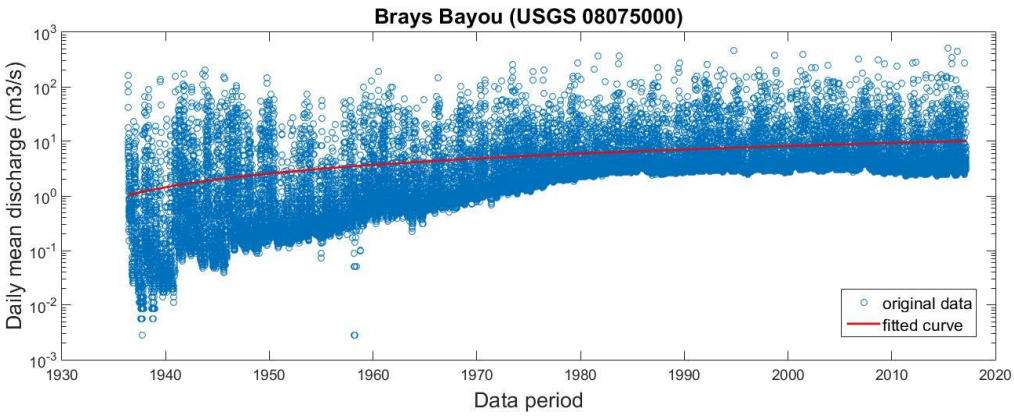
### A.1. Data Trend

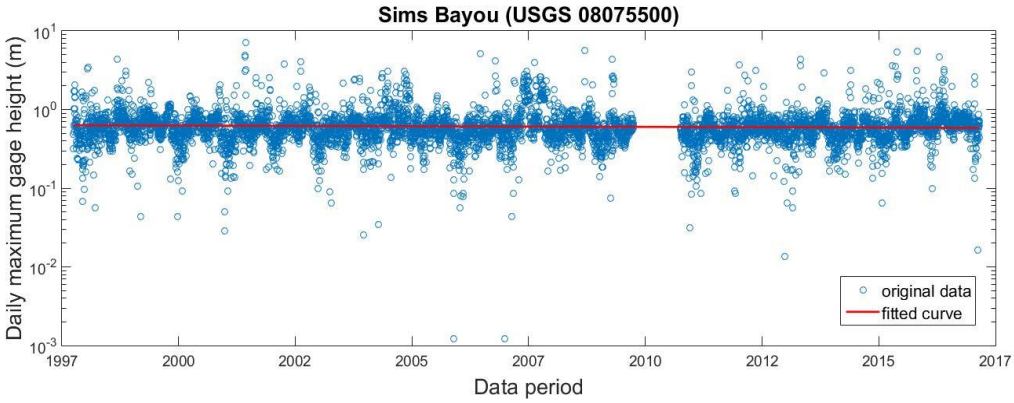
#### A.1.1 Daily Data Trend

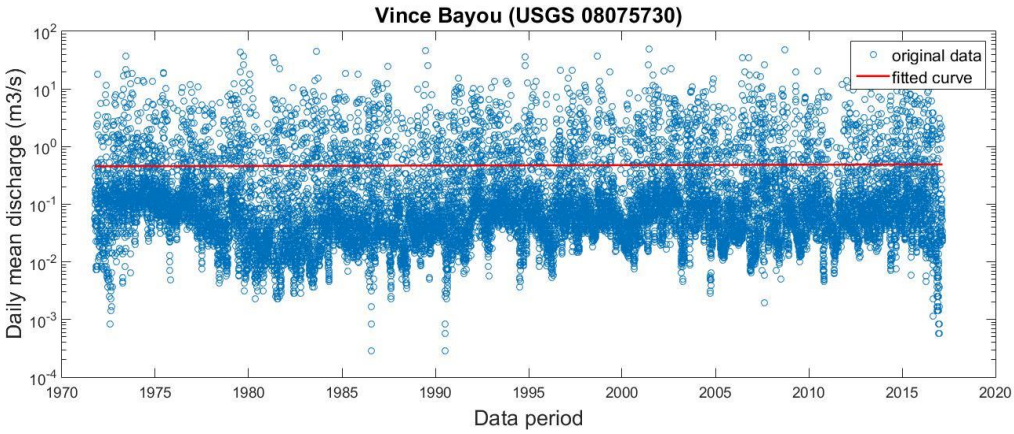


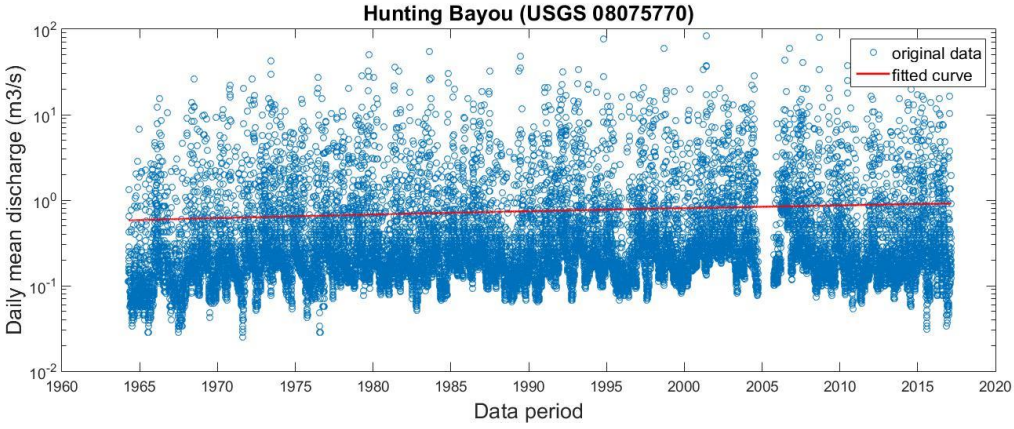




Brays Bayou (USGS 08075000)
Daily mean discharge (m3/s)
Data period
original data
fitted curve


Sims Bayou (USGS 08075500)
Daily maximum gage height (m)
Data period
original data
fitted curve


Vince Bayou (USGS 08075730)
Daily mean discharge (m3/s)
Data period
original data
fitted curve


Hunting Bayou (USGS 08075770)
Daily mean discharge (m3/s)
Data period
original data
fitted curve

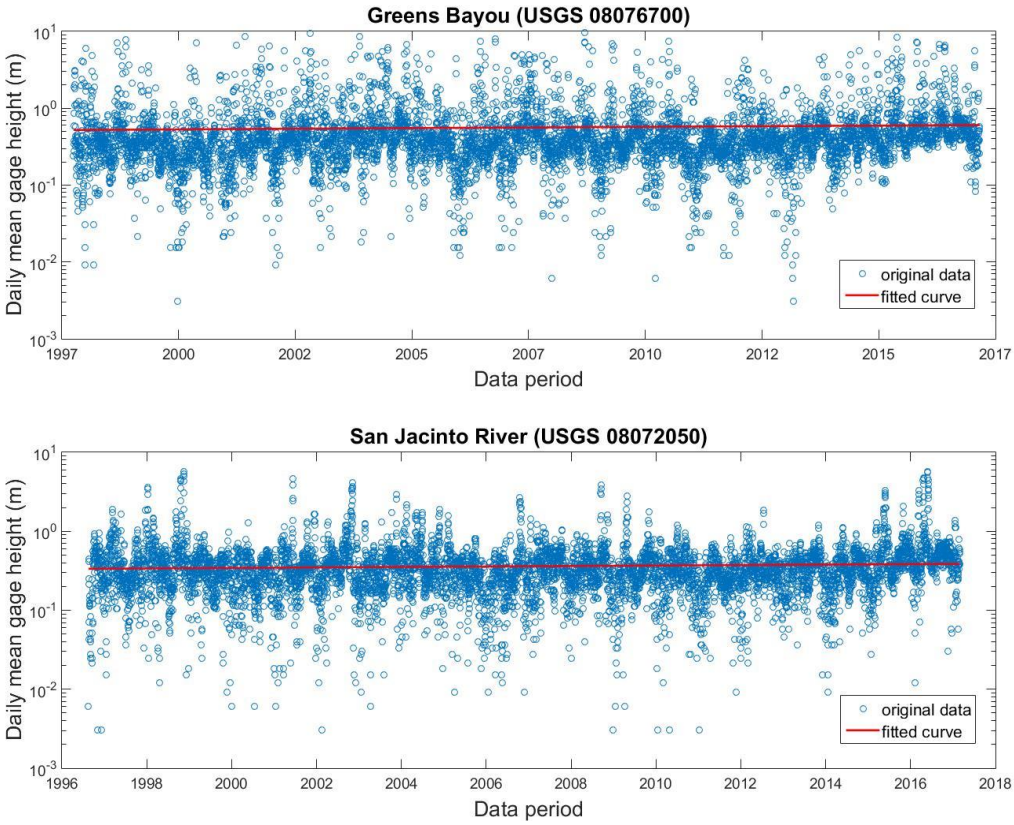
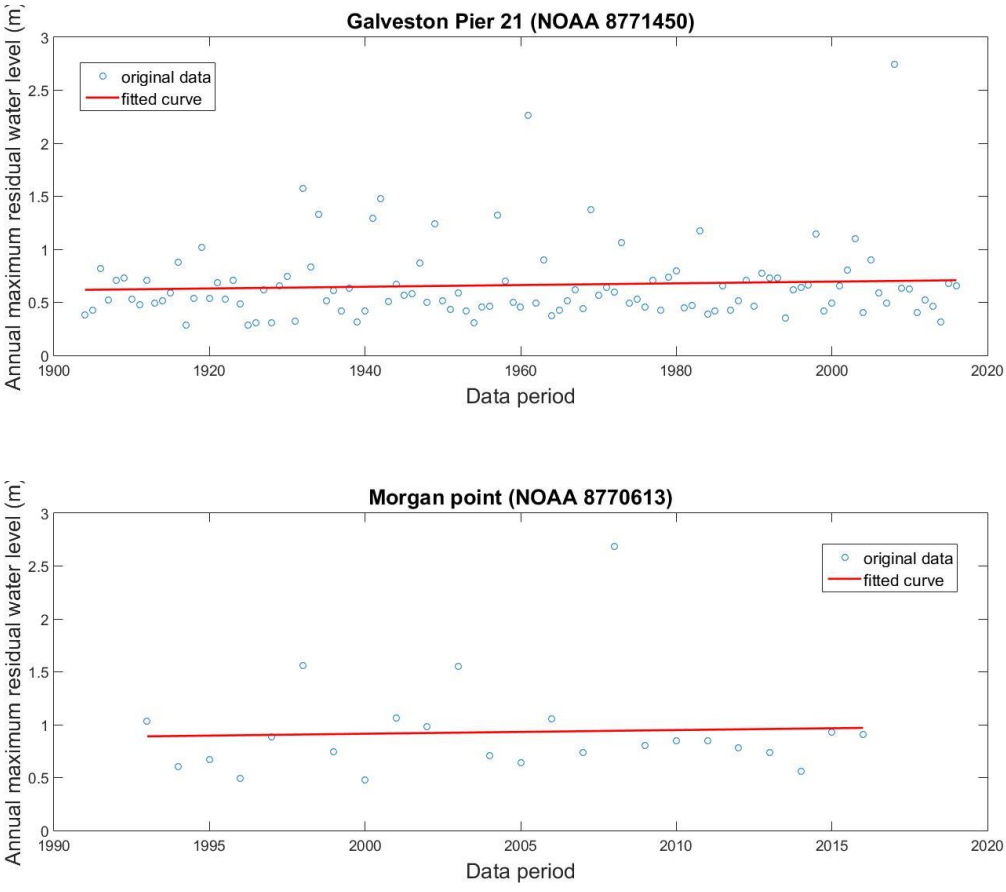
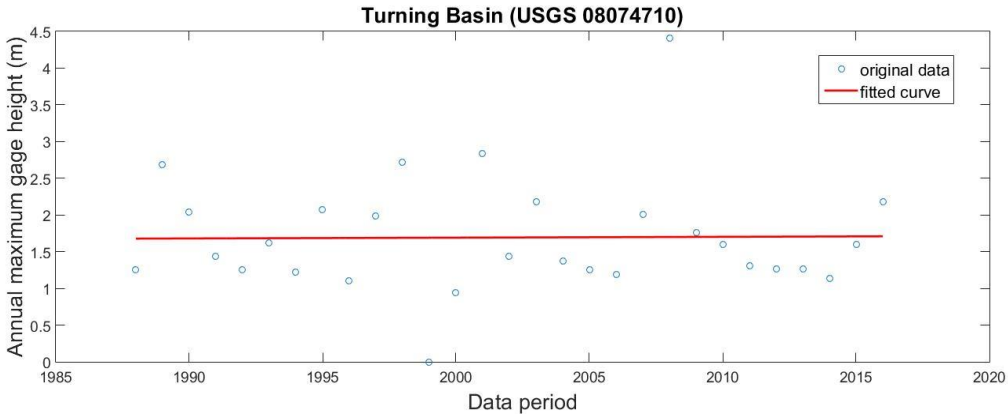
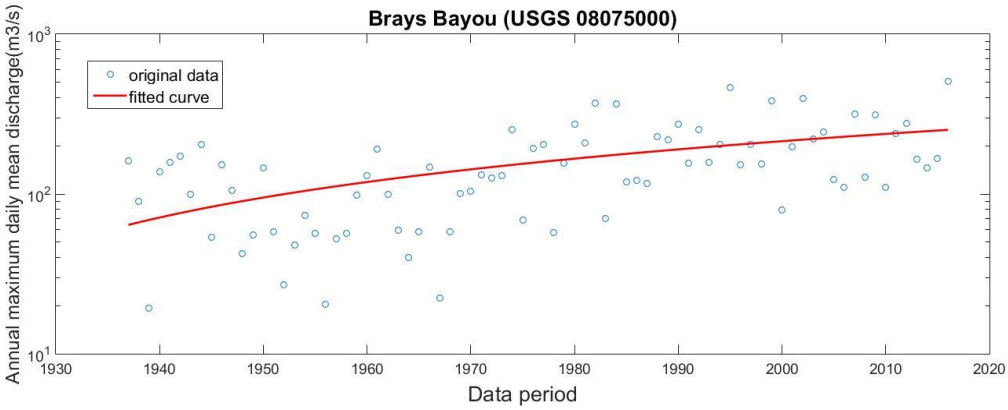


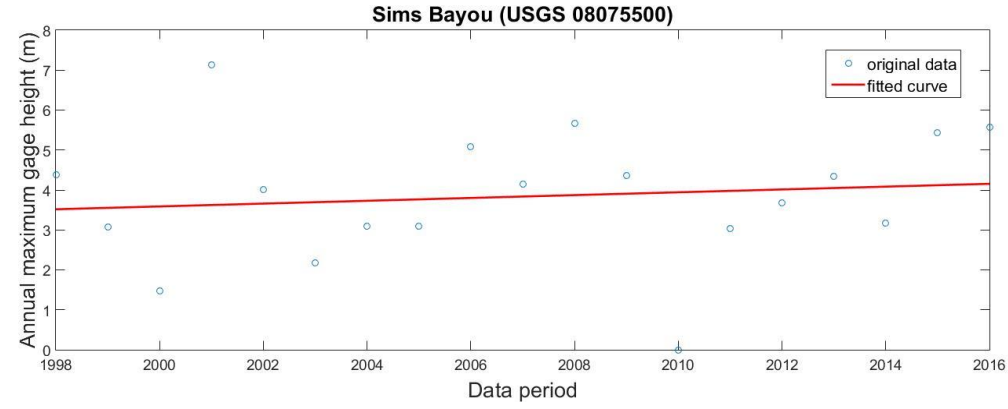

**Figure A.1: The daily data trend for all the original selected data. Source: NOAA, USGS.**

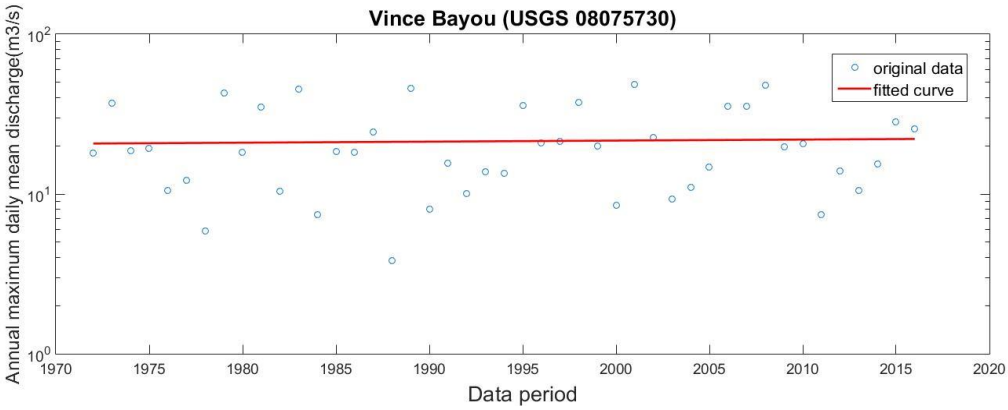
## A.1.2 Annual Maximum Data Trend




Turning Basin (USGS 08074710)
original data
fitted curve
Annual maximum gage height (m)
Data period


Brays Bayou (USGS 08075000)
original data
fitted curve
Annual maximum daily mean discharge(m3/s)
Data period


Sims Bayou (USGS 08075500)
original data
fitted curve
Annual maximum gage height (m)
Data period


Vince Bayou (USGS 08075730)
original data
fitted curve
Annual maximum daily mean discharge(m3/s)
Data period

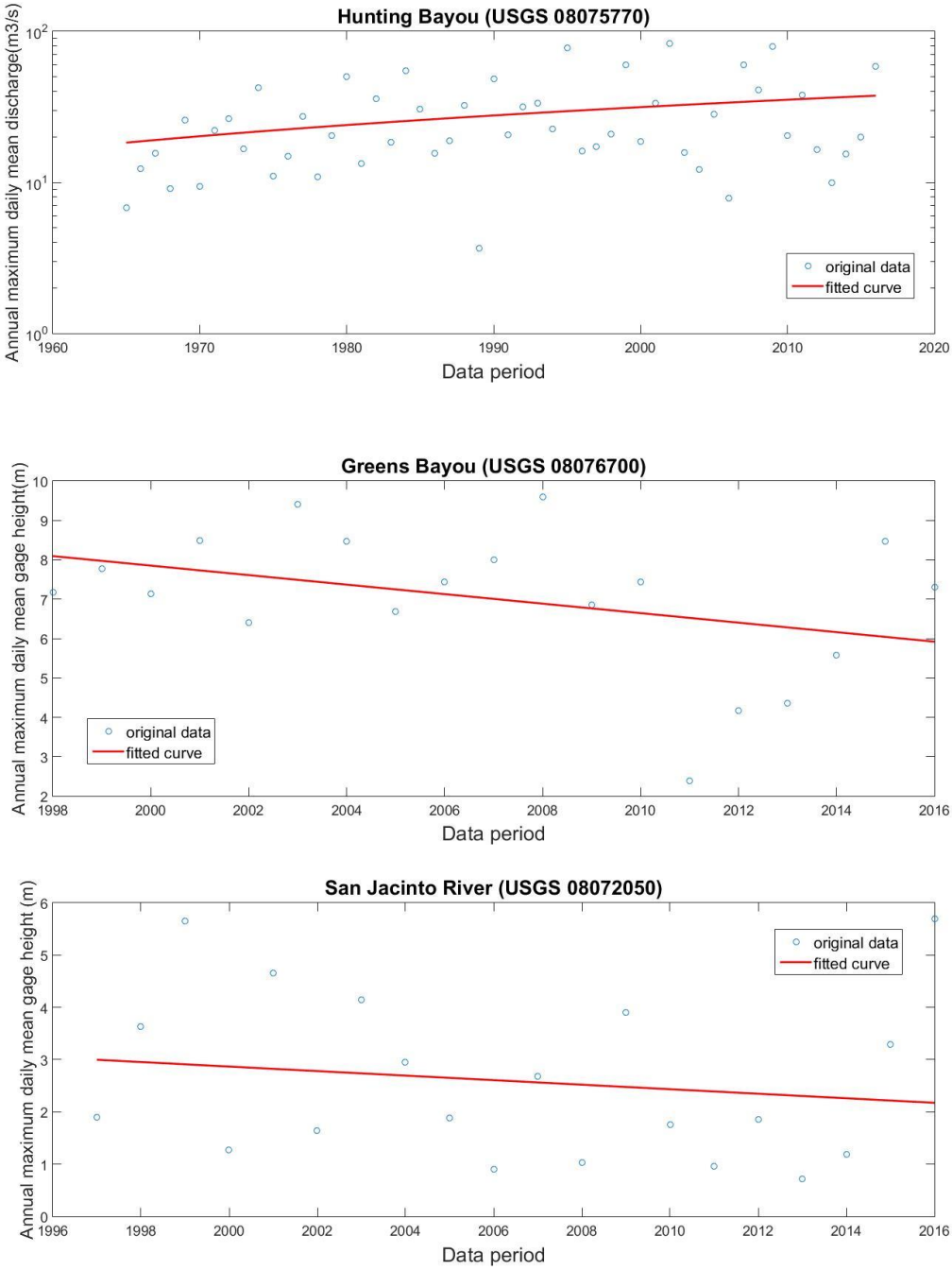



**Figure A.2: The annual maximum data trend for all the original selected data. Source: NOAA, USGS.**

## A.2. Rank Correlation

**Table A.1: Detailed information of rank correlations for all the variable pairs. Data source: NOAA, USGS.**

| Variable pairs | Common data | Number of all available data | Rank correlation | Common data without river trend | Number of available data without trend | Rank correlation |
|---|---|---|---|---|---|---|
| GP21-Morgan | 1993.3.17-2017.1.31 | 8722 | 0.8624 | No trend | | |
| GP21-TB | 1987.10.1-2017.1.31 | 10259 | 0.7092 | No trend | | |
| GP21-BB | 1936.5.25-2017.1.31 | 29472 | 0.2267 | 1985.1.1-2017.1.31 | 11719 | 0.3061 |
| GP21-SB | 1997.10.1-2017.1.31 | 6602 | 0.6867 | No trend | | |
| GP21-VB | 1971.10.1-2017.1.31 | 16555 | 0.2986 | 1982.1.1-2017.1.31 | 12813 | 0.2758 |
| GP21-HB | 1964.4.14-2017.1.31 | 18918 | 0.2925 | No trend | | |
| GP21-GB | 1997.10.2-2017.1.31 | 6929 | 0.6772 | No trend | | |
| GP21-SJR | 1996.8.17-2017.1.31 | 7426 | 0.6804 | No trend | | |
| Morgan-TB | 1993.13.17-2017.1.31 | 8265 | 0.7062 | No trend | | |
| Morgan-BB | 1993.3.17-2017.1.31 | 8722 | 0.2378 | No trend | | |
| Morgan-SB | 1997.10.1-2017.1.31 | 6602 | 0.6727 | No trend | | |
| Morgan-VB | 1993.3.17-2017.1.31 | 8721 | 0.2441 | No trend | | |
| Morgan-HB | 1993.3.17-2017.1.31 | 8354 | 0.2484 | No trend | | |
| Morgan-GB | 1997.10.2-2017.1.31 | 6929 | 0.6458 | No trend | | |
| Morgan-SJR | 1996.8.17-2017.1.31 | 7425 | 0.6716 | No trend | | |
| TB-BB | 1987.10.1-2017.3.1 | 10288 | 0.2413 | No trend | | |
| TB-SB | 1997.10.1-2017.3.1 | 6182 | 0.9059 | No trend | | |
| TB- VB | 1987.10.1-2017.3.1 | 10288 | 0.2172 | No trend | | |
| TB-HB | 1987.10.1-2017.3.1 | 9920 | 0.178 | No trend | | |
| TB-GB | 1997.10.2-2017.3.1 | 6512 | 0.7405 | No trend | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| TB-SJR | 1996.8.17-2017.3.1 | 7007 | 0.8139 | No trend | | |
| BB-SB | 1997.10.1-2017.3.1 | 6631 | 0.3170 | No trend | | |
| BB-VB | 1971.10.1-2017.3.1 | 16589 | 0.5167 | 1985.1.1-2017.3.1 | 11748 | 0.5805 |
| BB-HB | 1964.4.14-2017.3.1 | 18949 | 0.6719 | 1985.1.1-2017.3.1 | 11380 | 0.7171 |
| BB-GB | 1997.10.2-2017.3.1 | 6958 | 0.5025 | No trend | | |
| BB-SJR | 1996.8.17-2017.3.1 | 7455 | 0.2507 | No trend | | |
| SB-VB | 1997.10.1-2017.3.1 | 6630 | 0.2891 | No trend | | |
| SB-HB | 1997.10.1-2017.3.1 | 6278 | 0.2446 | No trend | | |
| SB-GB | 1997.10.2-2017.3.1 | 6506 | 0.7637 | No trend | | |
| SB - SJR | 1997.10.1-2017.3.1 | 6589 | 0.7863 | No trend | | |
| VB-HB | 1971.10.1-2017.3.1 | 16221 | 0.5693 | 1982.1.1-2017.3.1 | 12476 | 0.5739 |
| VB-GB | 1997.10.2-2017.3.1 | 6959 | 0.4418 | No trend | | |
| VB-SJR | 1996.8.17-2017.3.1 | 7455 | 0.2248 | No trend | | |
| HB-GB | 1997.10.2-2017.3.1 | 6593 | 0.4552 | No trend | | |
| HB-SJR | 1996.8.17-2017.3.1 | 7098 | 0.2217 | No trend | | |
| SJR-GB | 1997.10.2-2017.3.1 | 6913 | 0.7983 | No trend | | |

Note: In the above table, all the selected variables are expressed by their abbreviations. The meaning of each abbreviation can be referred to Table 3.4.

# Appendix B

## Pearson's Chi-Squared Test

### B.1. General Introduction

The Pearson's Chi-squared test is commonly used method for the goodness of fit. For a set of a random variable $X$ with a distribution function $F(x)$, it is assumed that $X_1,...,X_n$ is a sample extracted from the distribution and then the Pearson's Chi-squared test can be expressed as (Dahiya & Gurland, 1972):

$$D = \sum (n_i - N \cdot p_i)^2 / (N \cdot p_i) \qquad \text{[Eq. B.1]}$$

Where $p_i$ is the estimated probability that the random variable $X$ will be classified in the $i$ th class, $n_i$ is the number of observed sample in the $i$ th class and $N$ is the total number of the sample.

### B.2. Marginal Distributions of Selected Variables

In the following, the fitted continuous marginal distributions of daily maximum residual water level at NOAA 8771450 (GP21) and daily mean discharge at USGS 08075000 (BB) based on different empirical data are shown together with the corresponding results of Pearson's chi-squared test.

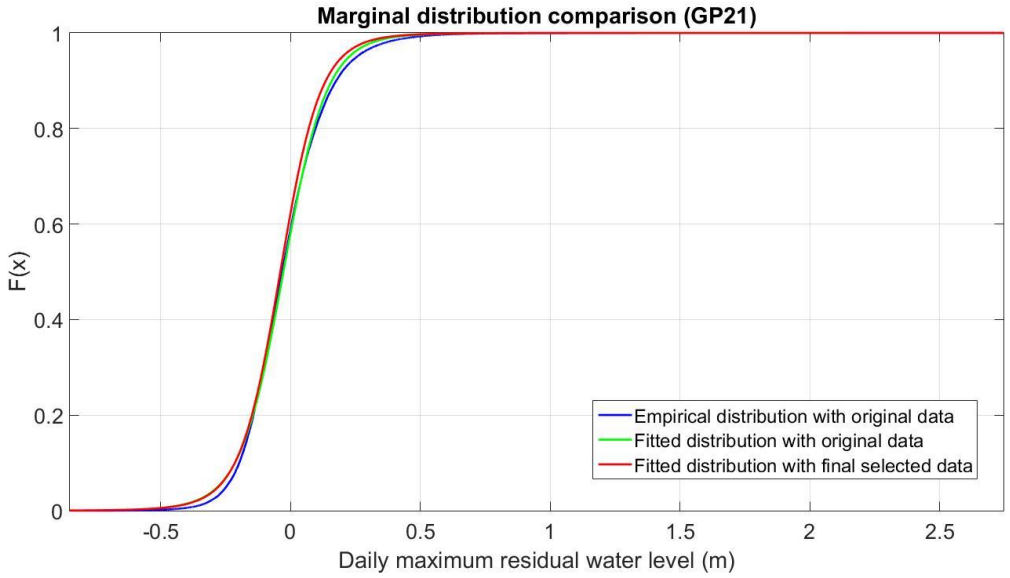



**Figure B.1: The comparison of empirical marginal distribution and fitted continuous marginal distributions of daily maximum residual water level at NOAA 8771450 (GP21).**

**Table B.1: The results of Pearson's chi-squared test for fitted continuous marginal distributions of daily maximum residual water level at NOAA 8771450 (GP21).**

**(a)**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| Daily maximum residual water level at NOAA 8771450 (GP21) from 01/01/1904 to 31/01/2017 | Normal | 0.053 |
| | Extreme value | 0.962 |
| | Generalized extreme value | 0.059 |
| | Logistic | 0.015 |
| | Rayleigh | 2.163 |
| | Tlocationscale | 0.014 |

**(b)**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| Daily maximum residual water level at NOAA 8771450 (GP21) from 02/10/1997 to 31/01/2017 | Normal | 0.065 |
| | Extreme value | 1.306 |
| | Generalized extreme value | 0.061 |
| | Logistic | 0.016 |
| | Rayleigh | 2.302 |
| | Tlocationscale | 0.015 |

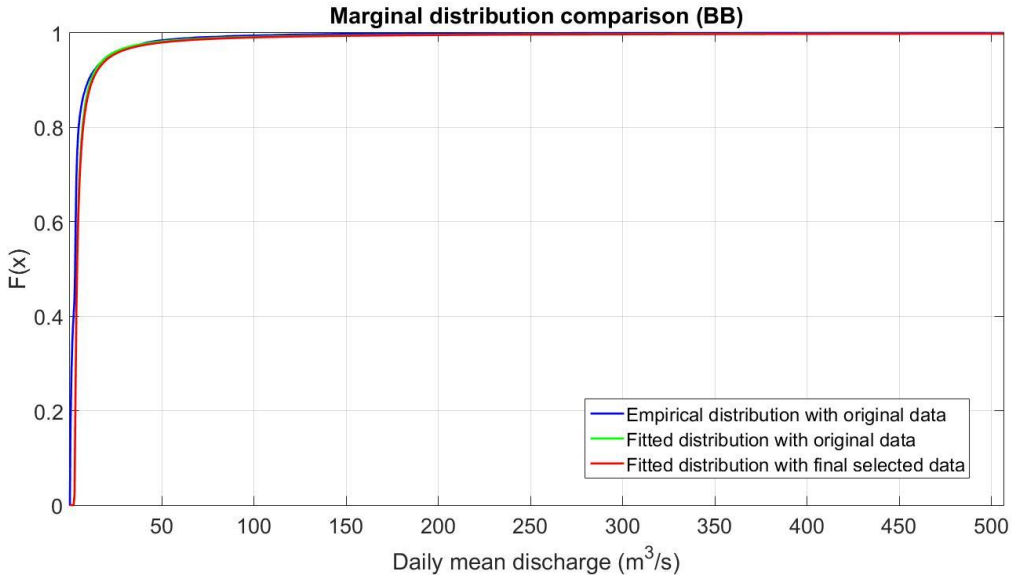



**Figure B.2: The comparison of empirical marginal distribution and fitted continuous marginal distributions of daily mean discharge at USGS 08075000 (BB).**

**Table B.2: The results of Pearson's chi-squared test for fitted continuous marginal distributions of daily mean discharge at USGS 08075000 (BB).**

**(a)**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| Discharge at USGS 08075000 (BB) from 01/01/1985 to 31/01/2017 | Normal | 0.4758 |
| | Birnbaumsaunders | 0.2913 |
| | Exponential | 0.3731 |
| | Extreme value | 0.6465 |
| | Gamma | 0.3648 |
| | Generalized extreme value | 0.0127 |
| | Inversegaussian | 0.2743 |
| | Logistic | 0.3768 |
| | Loglogistic | 0.1611 |
| | Lognormal | 0.2322 |
| | Rayleigh | 0.1307 |
| | Rician | 0.1307 |
| | Tlocationscale | 0.0311 |
| | Weibull | 0.4185 |

**(b)**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| Discharge at USGS 08075000 (BB) from 02/10/1997 to 31/01/2017 | Normal | 0.3954 |
| | Birnbaumsaunders | 0.2180 |
| | Exponential | 0.2859 |
| | Extreme value | 0.5381 |
| | Gamma | 0.2816 |
| | Generalized extreme value | 0.0144 |
| | Inversegaussian | 0.2008 |
| | Logistic | 0.3030 |
| | Loglogistic | 0.1131 |
| | Lognormal | 0.1683 |
| | Rayleigh | 0.1093 |
| | Rician | 0.1093 |
| | Tlocationscale | 0.0154 |
| | Weibull | 0.3199 |

# Appendix C

# Copula

## C.1. General Introduction

In nature, the hydrological phenomena often occur combining many different influential factors i.e. random variables. Therefore, to analyze the hydrological phenomena, it is vital to study the influential random variables and their dependencies with each other, which is usually solved by the joint modelling. The traditional way of exploring the dependences between the variable pairs is through bivariate distributions e.g. bivariate normal, gamma which however restrict the random variables with the same family of univariate distributions (Genest & Favre, 2007). To avoid this kind of limitation, the Copula models are developed based on the theory proposed by Sklar (1959). Through the copula models, the joint cumulative distribution function (CDF) $H(x, y)$ of variable pairs $(X, Y)$ can be expressed as:

$$H(x, y) = C\{F(x), G(x)\}, \quad x, y \in R \qquad \text{[Eq. C.1]}$$

Where $F(x), G(y)$ are marginal cumulative distributions and $C$ represent the copula (Genest & Favre, 2007).

In this project, Gaussian Copula, Gumbel Copula and Clayton Copula are selected to test the dependences of variable pairs opted which are commonly used in the hydrologic analysis and relatively simple because only one parameter needs to be determined (Favre et al., 2004; Sebastian, Dupuits, & Morales-Nápoles, 2017).

**Gaussian Copula**

The expression of Gaussian Copula is:

$$C_\rho(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \quad (u, v) \in (0,1) \qquad \text{[Eq. C.2]}$$

Where $\rho$ is the (conditional) product moment correlation, $\Phi_\rho$ is bivariate standard normal cumulative distribution and $\Phi^{-1}$ is the inverse of standard normal distribution.

**Gumbel Copula**

The expression of Gumbel Copula is:

$$C_\delta(u, v) = \exp\left[-(\left[-\log(u)\right]^\delta + \left[-\log(v)\right]^\delta)^{1/\delta}\right], \ \delta \in [1, \infty) \qquad \text{[Eq. C.3]}$$

Where $\delta$ is the parameter.

**Clayton Copula**

The expression of Clayton Copula is:

$$C_\alpha = \left[\max\left\{u^{-\alpha} + v^{-\alpha} - 1; 0\right\}\right]^{-1/\alpha},\ \alpha \in [-1, \infty) \setminus \{0\} \qquad \text{[Eq. C.4]}$$

Where $\alpha$ is the parameter.

One of the important characteristics of copulas is the ability to test the tail dependences of the variable pairs. For the upper tail dependence of variable pair $(X_i, X_j)$, it can be described by the upper tail coefficient $\lambda_U$ which is expressed as:

$$\lambda_U = \lim_{u \to 1} P(X_i > F_{X_i}^{-1}(u) \,|\, X_j > F_{X_j}^{-1}(u)) = \lim_{u \to 1} P(U > u \,|\, V > u) \qquad \text{[Eq. C.5]}$$

The similar definition is also suitable for the lower tail dependence. For the Gaussian copula, there is no tail dependence exhibited (i.e., $\lambda_U = 0$) while, for the Gumbel copula, upper tail dependence is displayed (i.e. $\lambda_U = 1 - 2^{1/\delta}$), and, for the Clayton copula, lower tail dependence is displayed (i.e. $\lambda_U = 2^{1/\alpha}$) (Sebastian et al., 2017)

# Appendix D

## Semi-Correlation Test

### D.1. Theoretical Background

Semi-correlation test is the method used to assess the fitness of a given copula on the tail dependence of a data set in which the marginal distributions of the variable pair are converted to the standard normal distributions and then the tail correlations of the standard normal variable pair are calculated (Joe, 2015). For bivariate variable pairs, the tail dependence can be divided into four parts: (1) the upper right tail correlation $\rho_{ne}$; (2) the lower left tail correlation $\rho_{sw}$; (3) the upper left tail correlation $\rho_{nw}$; (4) the lower right tail correlation $\rho_{se}$. In addition, $\rho_{ne}$ and $\rho_{sw}$ represent the correlations of positively correlated data while $\rho_{nw}$ and $\rho_{se}$ represent the correlations of negatively correlated data.

In more detail, for a variable pair $(X_i, Y_i)$, it is assumed that the marginal cumulative density functions (CDFs) are $F_X, F_Y$ and the standard normal CDF is $\Phi$, then the transformed standard normal variable pair is $(Z_X, Z_Y)$ where $Z_X = \Phi^{-1}(F_X)$ and $Z_Y = \Phi^{-1}(F_Y)$. Accordingly, the semi-correlations in different tails can be expressed as:

$$\rho_{ne} = \rho(Z_X, Z_Y \,|\, Z_X > 0, Z_Y > 0) \quad \text{[Eq. D.1]}$$

$$\rho_{sw} = \rho(Z_X, Z_Y \,|\, Z_X < 0, Z_Y < 0) \quad \text{[Eq. D.2]}$$

$$\rho_{nw} = \rho(Z_X, Z_Y \,|\, Z_X < 0, Z_Y > 0) \quad \text{[Eq. D.3]}$$

$$\rho_{se} = \rho(Z_X, Z_Y \,|\, Z_X > 0, Z_Y < 0) \quad \text{[Eq. D.4]}$$

### D.2. Results of Semi-Correlation for the Variable Pairs in the BN

**Table D.1: The detailed information of semi-correlation tests for all the variable pairs in the BN model.**

| GP21-Morgan | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.8880 | 0.2798 | 0.8189 | 0.7614 | 0.046 |
| Gaussian Copula | 0.8907 | 0.2544 | 0.7546 | 0.7563 | 0.1873 |
| Gumbel Copula | 0.8857 | 0.0062 | 0.8645 | 0.6561 | 0.0286 |
| Clayton Copula | 0.7858 | 0.088 | 0.2221 | 0.8624 | 0.0744 |

| GP21-SJR | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6902 | -0.2171 | 0.4798 | 0.5371 | -0.0256 |
| Gaussian Copula | 0.6667 | 0.1124 | 0.4285 | 0.4677 | 0.172 |
| Gumbel Copula | 0.6520 | 0.0621 | 0.6513 | 0.3134 | 0.0612 |
| Clayton Copula | 0.5640 | 0.0848 | 0.1011 | 0.6152 | 0.1198 |

| GP21-GB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6841 | -0.0916 | 0.3986 | 0.5597 | 0.0933 |
| Gaussian Copula | 0.6708 | 0.1775 | 0.4241 | 0.4623 | 0.0967 |
| Gumbel Copula | 0.6370 | 0.0969 | 0.6056 | 0.3077 | 0.1198 |
| Clayton Copula | 0.5826 | 0.0214 | 0.0753 | 0.6522 | 0.0817 |

| GP21-TB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7221 | 0.1987 | 0.6206 | 0.5226 | 0.0507 |
| Gaussian Copula | 0.7411 | 0.1328 | 0.524 | 0.5246 | 0.192 |
| Gumbel Copula | 0.7159 | 0.0617 | 0.6872 | 0.4215 | 0.0783 |
| Clayton Copula | 0.5792 | 0.0527 | 0.132 | 0.6651 | 0.0851 |

| GP21-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7030 | -0.1836 | 0.49 | 0.5523 | 0.1415 |
| Gaussian Copula | 0.6867 | 0.2047 | 0.4799 | 0.4897 | 0.1451 |
| Gumbel Copula | 0.6539 | 0.0734 | 0.6107 | 0.3037 | 0.0382 |
| Clayton Copula | 0.5876 | 0.0718 | 0.075 | 0.6744 | 0.0901 |

| Morgan-GB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6477 | 0.1146 | 0.4555 | 0.5585 | -0.1583 |
| Gaussian Copula | 0.6515 | 0.1918 | 0.3997 | 0.4276 | 0.1342 |
| Gumbel Copula | 0.6007 | 0.0565 | 0.5968 | 0.2618 | 0.0502 |
| Clayton Copula | 0.5423 | 0.1032 | 0.0933 | 0.6647 | 0.0588 |

| Morgan-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6820 | 0.2791 | 0.5083 | 0.5388 | -0.1758 |
| Gaussian Copula | 0.6675 | 0.2054 | 0.4079 | 0.4406 | 0.1992 |
| Gumbel Copula | 0.6369 | 0.0229 | 0.6439 | 0.3203 | 0.0224 |
| Clayton Copula | 0.5799 | 0.1219 | 0.1106 | 0.6472 | 0.1175 |

| Morgan-TB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7231 | 0.1798 | 0.6495 | 0.5192 | 0.2288 |
| Gaussian Copula | 0.7348 | 0.172 | 0.5231 | 0.5187 | 0.2092 |
| Gumbel Copula | 0.7036 | 0.0651 | 0.6734 | 0.3679 | 0.0287 |
| Clayton Copula | 0.5682 | 0.086 | 0.1068 | 0.6774 | 0.0664 |

| Morgan-SJR | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6899 | 0.0535 | 0.4879 | 0.5593 | -0.2153 |
| Gaussian Copula | 0.6648 | 0.1576 | 0.4486 | 0.47 | 0.1292 |
| Gumbel Copula | 0.6591 | 0.0162 | 0.655 | 0.3211 | 0.0805 |
| Clayton Copula | 0.5630 | 0.0894 | 0.0723 | 0.6865 | 0.1022 |

| SJR-GB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.8065 | -0.1685 | 0.6021 | 0.8614 | -0.2412 |
| Gaussian Copula | 0.7885 | 0.1534 | 0.6154 | 0.6144 | 0.1399 |
| Gumbel Copula | 0.7860 | 0.106 | 0.7629 | 0.5006 | 0.0904 |
| Clayton Copula | 0.7934 | 0.0393 | 0.2302 | 0.8597 | 0.0824 |

| SJR-TB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.8128 | -0.034 | 0.6187 | 0.7724 | -0.3182 |
| Gaussian Copula | 0.7846 | 0.1763 | 0.5979 | 0.617 | 0.1671 |
| Gumbel Copula | 0.7747 | 0.0118 | 0.7633 | 0.4793 | 0.0391 |
| Clayton Copula | 0.7412 | 0.1032 | 0.1947 | 0.8234 | 0.1083 |

| SJR-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7934 | -0.148 | 0.5371 | 0.7987 | -0.3637 |
| Gaussian Copula | 0.7581 | 0.206 | 0.5513 | 0.5598 | 0.2186 |
| Gumbel Copula | 0.7541 | 0.0904 | 0.7328 | 0.4581 | 0.1189 |
| Clayton Copula | 0.7433 | 0.1249 | 0.171 | 0.8226 | 0.0857 |

| GB-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7748 | -0.1901 | 0.599 | 0.7984 | -0.2679 |
| Gaussian Copula | 0.7805 | 0.1314 | 0.5552 | 0.5962 | 0.2054 |
| Gumbel Copula | 0.7657 | 0.1535 | 0.7475 | 0.4603 | 0.1197 |
| Clayton Copula | 0.7409 | 0.1033 | 0.1768 | 0.8101 | 0.0754 |

| GB-TB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7396 | -0.1603 | 0.4885 | 0.7652 | 0.0208 |
| Gaussian Copula | 0.7298 | 0.2305 | 0.4924 | 0.516 | 0.1941 |
| Gumbel Copula | 0.6890 | 0.1025 | 0.6778 | 0.3629 | 0.0926 |
| Clayton Copula | 0.6937 | 0.0773 | 0.1735 | 0.7761 | 0.0618 |

| GB-HB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.4492 | 0.2014 | 0.7481 | -0.1246 | -0.0277 |
| Gaussian Copula | 0.4611 | 0.1182 | 0.2232 | 0.2479 | 0.1692 |
| Gumbel Copula | -0.0005 | 0.0202 | 0.5282 | 0.2325 | 0.053 |
| Clayton Copula | -0.0082 | 0.0187 | 0.032 | 0.296 | 0.0641 |

| GB-BB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.4851 | 0.2214 | 0.7044 | -0.0486 | 0.0532 |
| Gaussian Copula | 0.4875 | 0.117 | 0.279 | 0.2692 | 0.1741 |
| Gumbel Copula | 0.0096 | 0.0716 | 0.522 | 0.2008 | 0.086 |
| Clayton Copula | -0.0162 | 0.0349 | 0.0123 | 0.2924 | 0.0773 |

| BB-HB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7232 | 0.0552 | 0.777 | 0.2829 | -0.1251 |
| Gaussian Copula | 1.0000 | 0.1717 | 0.5301 | 0.5129 | 0.1434 |
| Gumbel Copula | 1.0000 | 0.0729 | 0.7373 | 0.4465 | 0.0814 |
| Clayton Copula | 1.0000 | 0.0648 | 0.0571 | 0.5662 | 0.0383 |

| TB-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.9089 | -0.3583 | 0.7628 | 0.92 | -0.2387 |
| Gaussian Copula | 0.8881 | 0.1975 | 0.7612 | 0.7713 | 0.1362 |
| Gumbel Copula | 0.8990 | 0.1467 | 0.8702 | 0.7033 | 0.0588 |
| Clayton Copula | 0.9059 | 0.0372 | 0.435 | 0.9391 | 0.1032 |

Note: In the above tables, all the selected variables are expressed by their abbreviations. The meaning of each abbreviation can be referred to Table 3.4.

# Appendix E

## Rank Correlation Matrix in the BN

**Table E.1: Bayesian belief net rank correlation matrix.**

| | Morgan | SJR | GB | HB | TB | SB | BB | GP21 |
|---|---|---|---|---|---|---|---|---|
| Morgan | 1.0000 | 0.6735 | 0.6403 | 0.3012 | 0.7383 | 0.6789 | 0.3165 | 0.8932 |
| SJR | 0.6735 | 1.0000 | 0.7885 | 0.3681 | 0.7863 | 0.7609 | 0.3869 | 0.6653 |
| GB | 0.6403 | 0.7885 | 1.0000 | 0.4603 | 0.7241 | 0.7690 | 0.4840 | 0.6662 |
| HB | 0.3012 | 0.3681 | 0.4603 | 1.0000 | 0.3393 | 0.3594 | 0.7141 | 0.3130 |
| TB | 0.7383 | 0.7863 | 0.7241 | 0.3393 | 1.0000 | 0.8847 | 0.3565 | 0.7296 |
| SB | 0.6789 | 0.7609 | 0.7690 | 0.3594 | 0.8847 | 1.0000 | 0.3777 | 0.6909 |
| BB | 0.3165 | 0.3869 | 0.4840 | 0.7141 | 0.3565 | 0.3777 | 1.0000 | 0.3289 |
| GP21 | 0.8932 | 0.6653 | 0.6662 | 0.3130 | 0.7296 | 0.6909 | 0.3289 | 1.0000 |

**Table E.2: Empirical normal rank correlation matrix.**

| | Morgan | SJR | GB | HB | TB | SB | BB | GP21 |
|---|---|---|---|---|---|---|---|---|
| Morgan | 1.0000 | 0.6735 | 0.6403 | 0.2319 | 0.7383 | 0.6789 | 0.2353 | 0.8932 |
| SJR | 0.6735 | 1.0000 | 0.7885 | 0.2353 | 0.7863 | 0.7609 | 0.2634 | 0.6653 |
| GB | 0.6403 | 0.7885 | 1.0000 | 0.4603 | 0.7241 | 0.7690 | 0.4840 | 0.6662 |
| HB | 0.2319 | 0.2353 | 0.4603 | 1.0000 | 0.1737 | 0.2729 | 0.7141 | 0.2593 |
| TB | 0.7383 | 0.7863 | 0.7241 | 0.1737 | 1.0000 | 0.8847 | 0.2198 | 0.7296 |
| SB | 0.6789 | 0.7609 | 0.7690 | 0.2729 | 0.8847 | 1.0000 | 0.3376 | 0.6909 |
| BB | 0.2353 | 0.2634 | 0.4840 | 0.7141 | 0.2198 | 0.3376 | 1.0000 | 0.2809 |
| GP21 | 0.8932 | 0.6653 | 0.6662 | 0.2593 | 0.7296 | 0.6909 | 0.2809 | 1.0000 |

**Table E.3: Empirical rank correlation matrix.**

| | Morgan | SJR | GB | HB | TB | SB | BB | GP21 |
|---|---|---|---|---|---|---|---|---|
| Morgan | 1.0000 | 0.6899 | 0.6476 | 0.2229 | 0.7230 | 0.6816 | 0.2268 | 0.8879 |
| SJR | 0.6899 | 1.0000 | 0.8062 | 0.2215 | 0.8128 | 0.7931 | 0.2554 | 0.6900 |
| GB | 0.6476 | 0.8062 | 1.0000 | 0.4492 | 0.7397 | 0.7749 | 0.4850 | 0.6839 |
| HB | 0.2229 | 0.2215 | 0.4492 | 1.0000 | 0.1512 | 0.2407 | 0.7233 | 0.2598 |
| TB | 0.7230 | 0.8128 | 0.7397 | 0.1512 | 1.0000 | 0.9087 | 0.2068 | 0.7218 |
| SB | 0.6816 | 0.7931 | 0.7749 | 0.2407 | 0.9087 | 1.0000 | 0.3137 | 0.7023 |
| BB | 0.2268 | 0.2554 | 0.4850 | 0.7233 | 0.2068 | 0.3137 | 1.0000 | 0.2864 |
| GP21 | 0.8879 | 0.6900 | 0.6839 | 0.2598 | 0.7218 | 0.7023 | 0.2864 | 1.0000 |

# Appendix F

# Detailed Analysis of Boundary Conditions

## F.1. Marginal Distribution Analysis

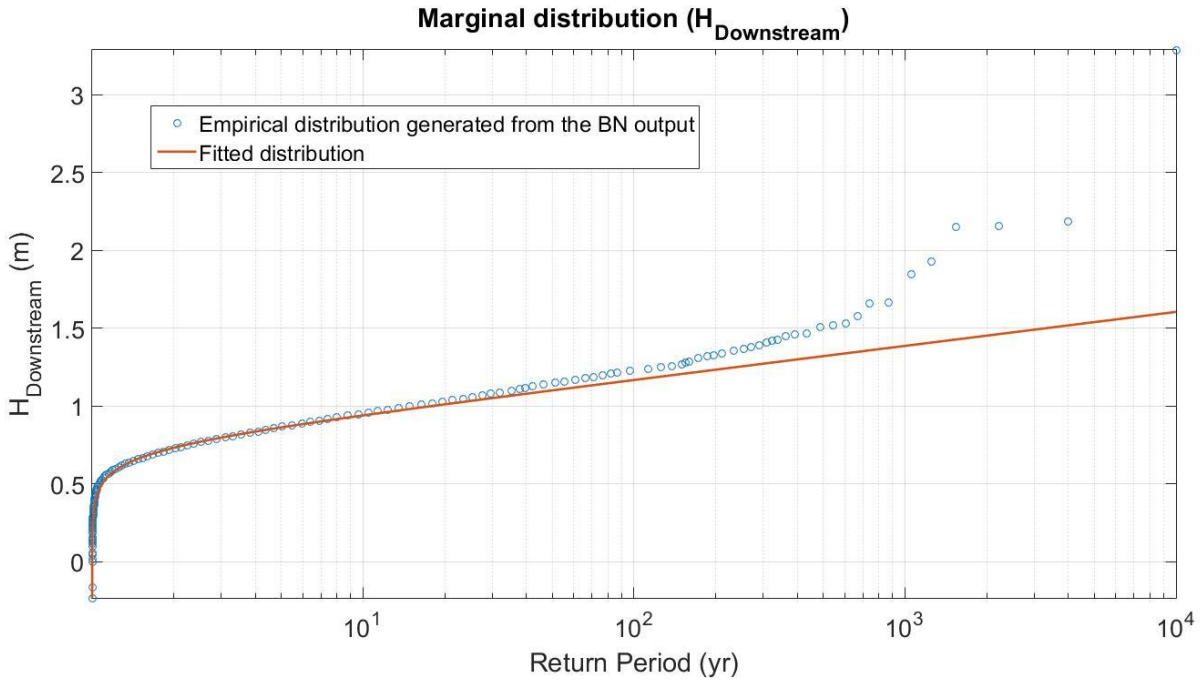



**Figure F.1: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the sea level at downstream boundary.**

**Table F.1: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the sea level at downstream boundary.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $H_{Downstream}$ | Normal | 0.0520 |
| | Extreme value | 0.8211 |
| | Generalized extreme value | 0.0982 |
| | Logistic | 0.0172 |
| | Rayleigh | 0.3878 |
| | Tlocationscale | 0.0182 |

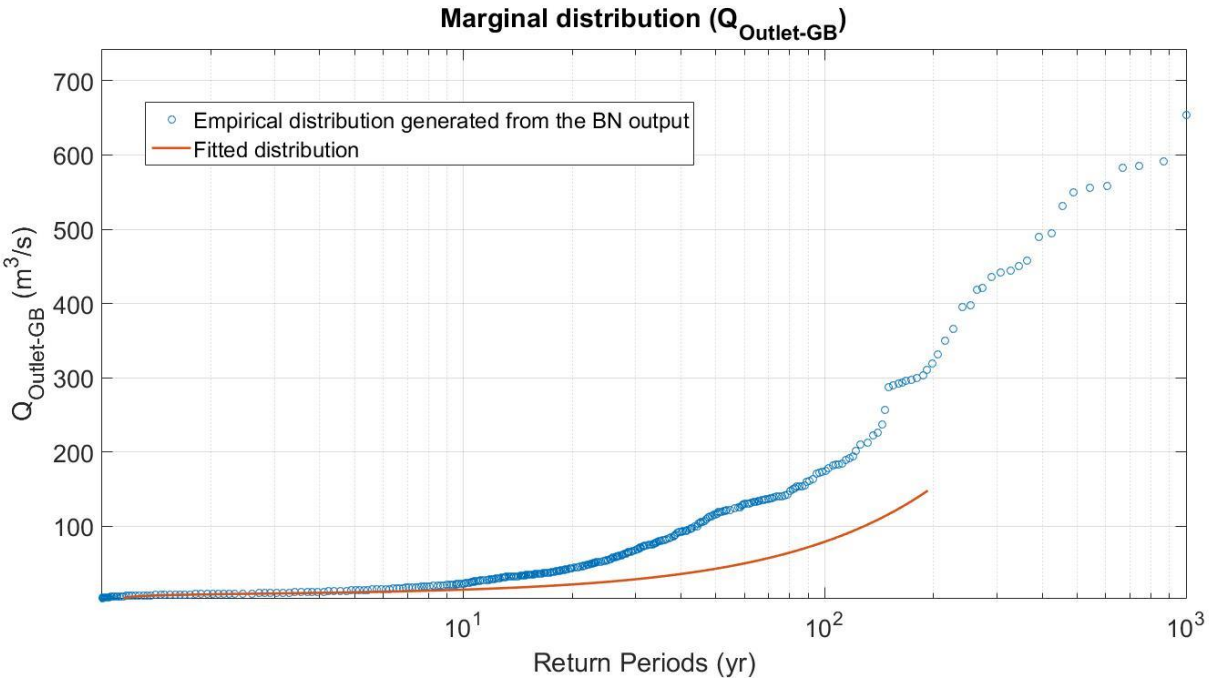



**Figure F.2: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at outlet of Greens Bayou.**

**Table F.2: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at the outlet of Greens Bayou.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Outlet\_GB}$ | Normal | 0.1367 |
| | Birnbaumsaunders | 0.0412 |
| | Exponential | 0.0648 |
| | Extreme value | 0.1818 |
| | Gamma | 0.0600 |
| | Generalized extreme value | 0.0084 |
| | Inversegaussian | 0.0365 |
| | Logistic | 0.0782 |
| | Loglogistic | 0.0109 |
| | Lognormal | 0.0260 |
| | Rayleigh | 0.0473 |
| | Rician | 0.0473 |
| | Tlocationscale | 0.0035 |
| | Weibull | 0.0688 |

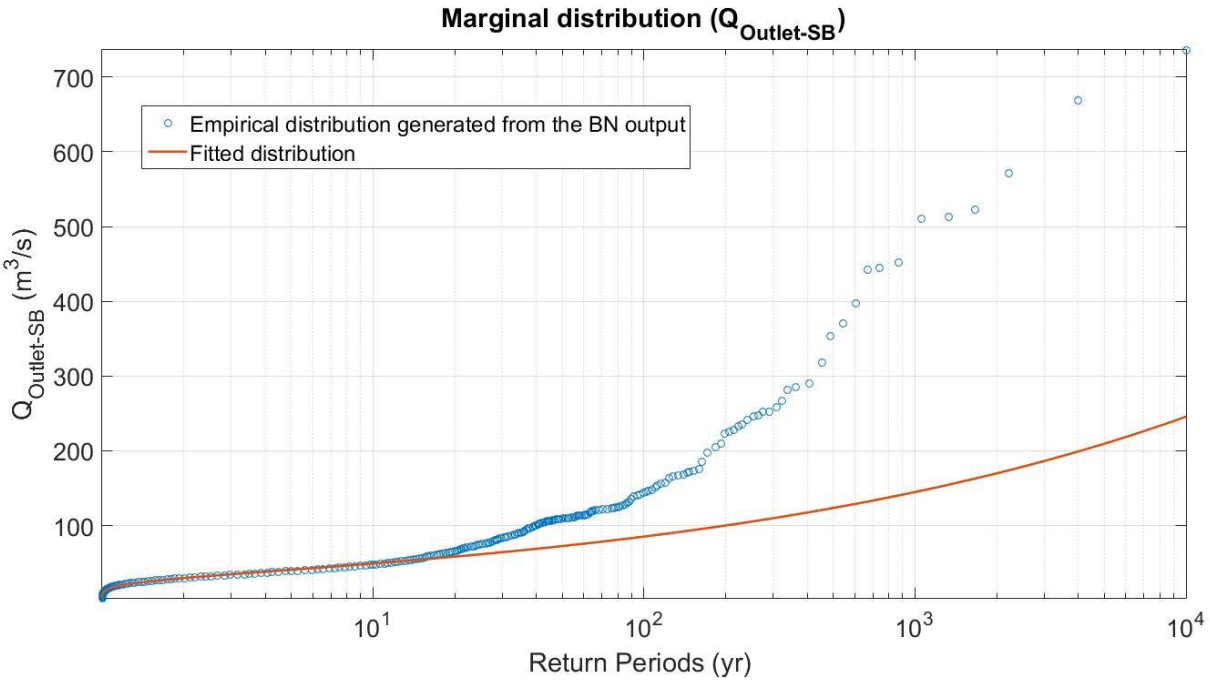



**Figure F.3: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at outlet of Sims Bayou.**

**Table F.3: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at the outlet of Sims Bayou.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Outlet\_SB}$ | Normal | 0.0182 |
| | Birnbaumsaunders | 0.0038 |
| | Exponential | 0.0203 |
| | Extreme value | 0.0488 |
| | Gamma | 0.0054 |
| | Generalized extreme value | 0.0016 |
| | Inversegaussian | 0.0037 |
| | Logistic | 0.0039 |
| | Loglogistic | 0.0004 |
| | Lognormal | 0.0026 |
| | Rayleigh | 0.0111 |
| | Rician | 0.0111 |
| | Tlocationscale | 0.0007 |
| | Weibull | 0.0121 |

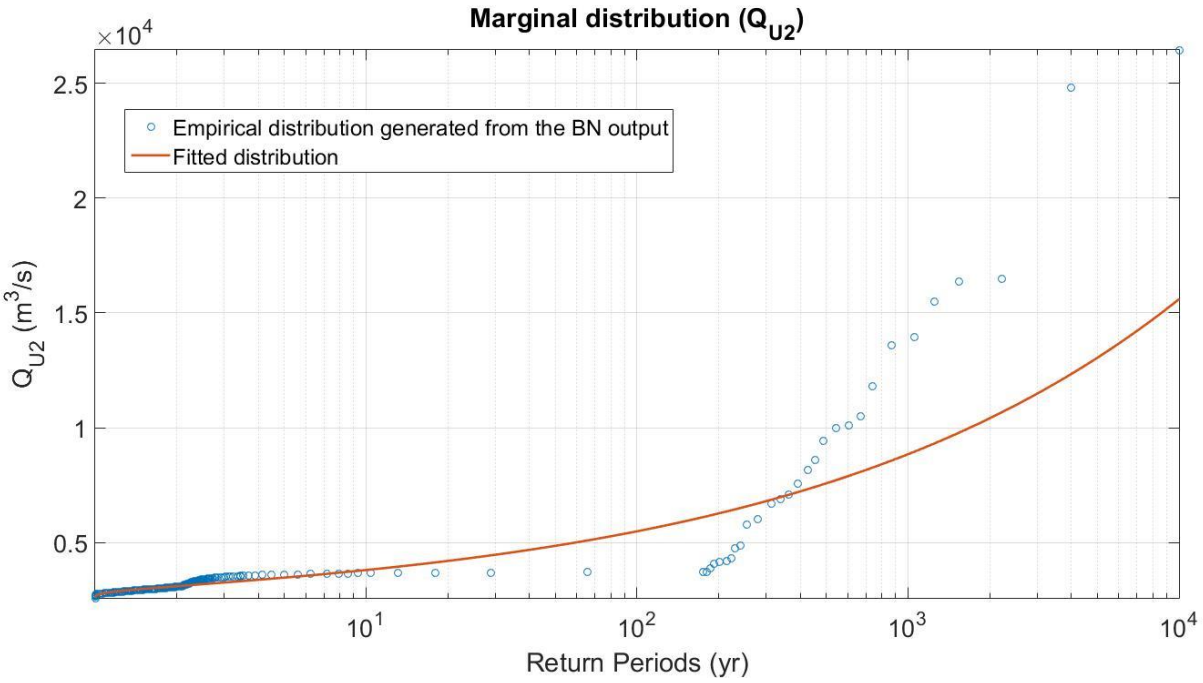



**Figure F.4: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at upstream boundary U2.**

**Table F.4: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at upstream boundary U2.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{U2}$ | Normal | 0.0006 |
| | Birnbaumsaunders | 0.0004 |
| | Exponential | 0.0033 |
| | Extreme value | 0.0022 |
| | Gamma | 0.0004 |
| | Generalized extreme value | 0.0002 |
| | Inversegaussian | 0.0004 |
| | Logistic | 0.0004 |
| | Loglogistic | 0.0004 |
| | Lognormal | 0.0004 |
| | Rayleigh | 0.0018 |
| | Rician | 0.0006 |
| | Tlocationscale | 0.0005 |
| | Weibull | 0.0012 |

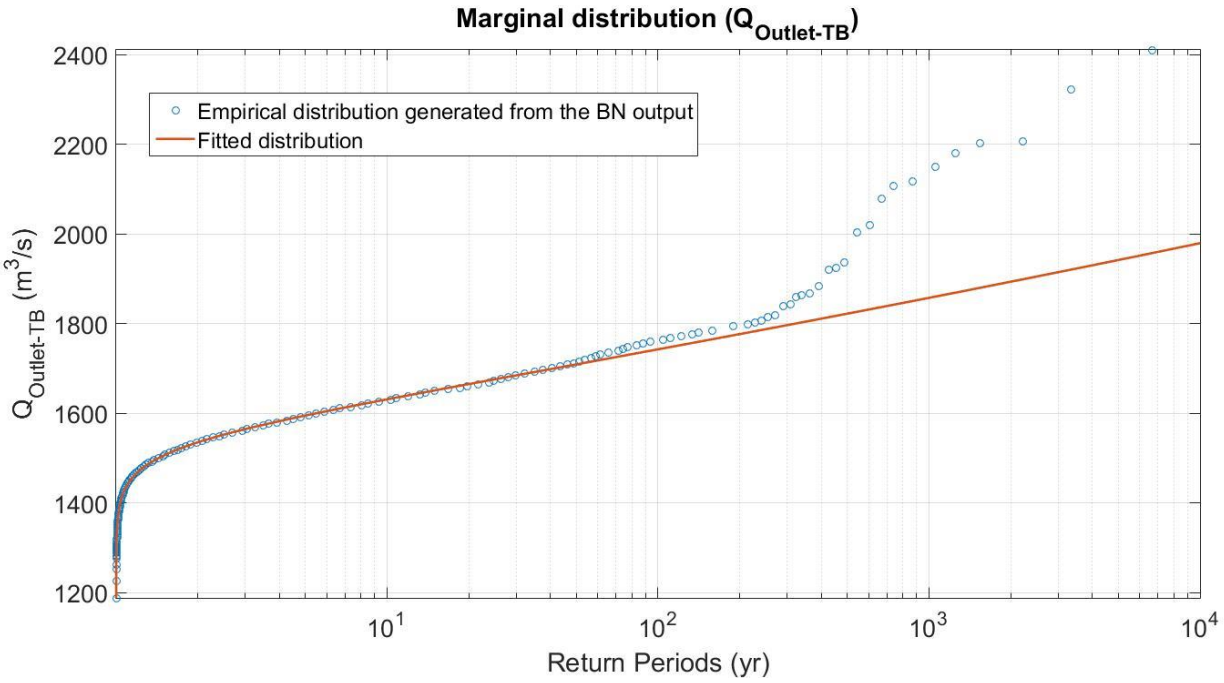



**Figure F.5: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at Turning Basin.**

**Table F.5: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at Turning Basin.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Outlet\_TB}$ | Normal | 0.000112 |
| | Birnbaumsaunders | 0.000089 |
| | Exponential | 0.015393 |
| | Extreme value | 0.001605 |
| | Gamma | 0.000093 |
| | Generalized extreme value | 0.000244 |
| | Inversegaussian | 0.000089 |
| | Logistic | 0.000008 |
| | Loglogistic | 0.000004 |
| | Lognormal | 0.000088 |
| | Rayleigh | 0.007895 |
| | Rician | 0.000111 |
| | Tlocationscale | 0.000011 |
| | Weibull | 0.001131 |

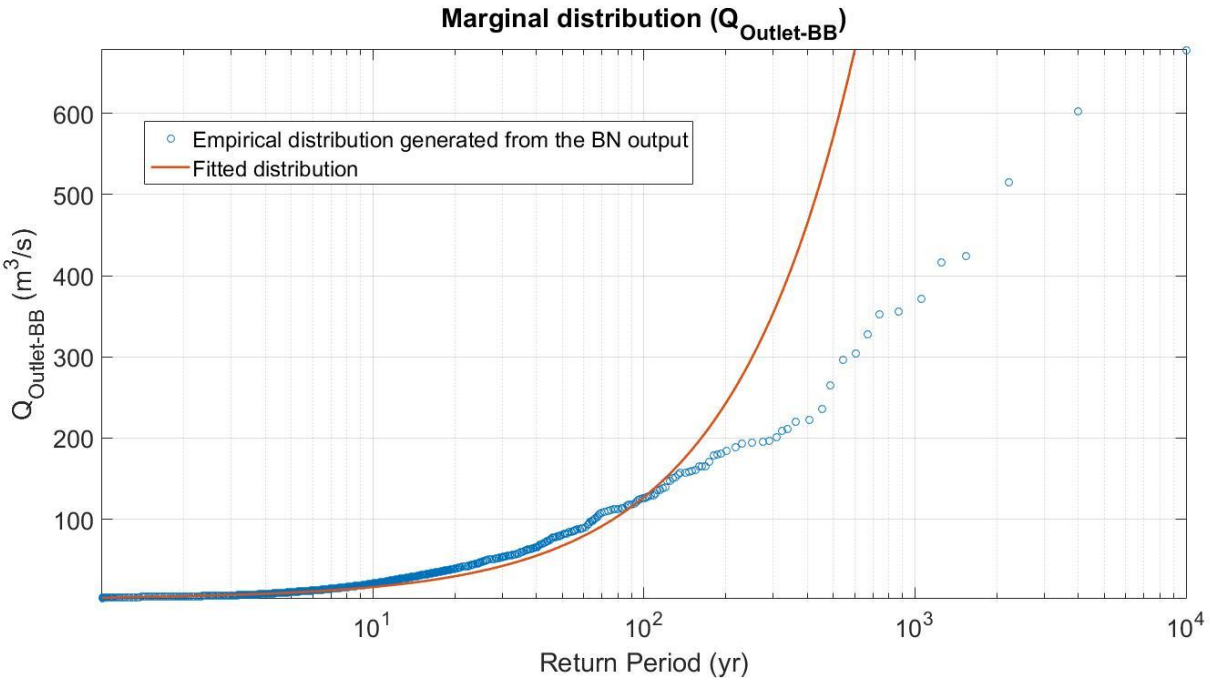



**Figure F.6: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at the outlet of Brays Bayou.**

**Table F.6: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at the outlet of Brays Bayou**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Outlet\_BB}$ | Normal | 0.2952 |
| | Birnbaumsaunders | 0.1600 |
| | Exponential | 0.2137 |
| | Extreme value | 0.4075 |
| | Gamma | 0.2075 |
| | Generalized extreme value | 0.0081 |
| | Inversegaussian | 0.1476 |
| | Logistic | 0.2242 |
| | Loglogistic | 0.0811 |
| | Lognormal | 0.1232 |
| | Rayleigh | 0.0848 |
| | Rician | 0.0848 |
| | Tlocationscale | 0.0142 |
| | Weibull | 0.2376 |

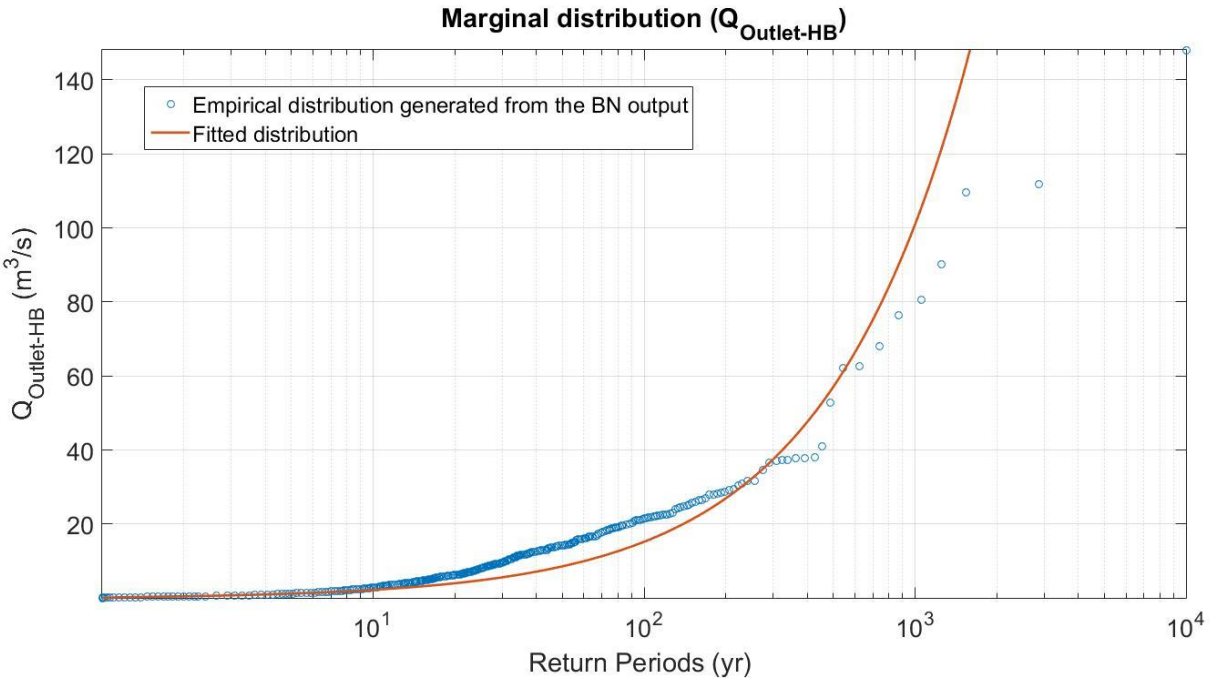



**Figure F.7: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at the outlet of Hunting Bayou.**

**Table F.7: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at the outlet of Hunting Bayou**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Outlet\_HB}$ | Normal | 1.6811 |
| | Birnbaumsaunders | 0.3968 |
| | Exponential | 0.6638 |
| | Extreme value | 2.3230 |
| | Gamma | 0.8603 |
| | Generalized extreme value | 0.0172 |
| | Inversegaussian | 0.1973 |
| | Logistic | 1.2029 |
| | Loglogistic | 0.1250 |
| | Lognormal | 0.2177 |
| | Rayleigh | 0.2636 |
| | Rician | 0.2636 |
| | Tlocationscale | 0.1120 |
| | Weibull | 0.9189 |

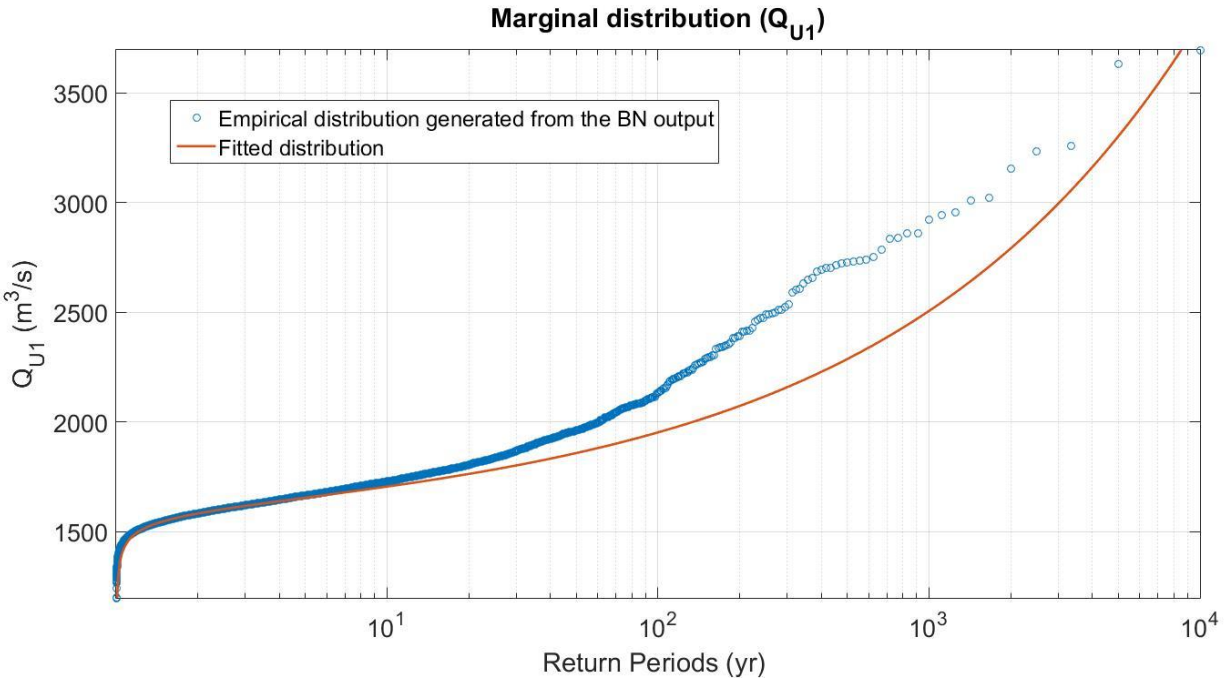



**Figure F.8: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the discharge at upstream boundary U1.**

**Table F.8: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the discharge at upstream boundary U1.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{U1}$ | Normal | 0.000570 |
| | Birnbaumsaunders | 0.000368 |
| | Exponential | 0.009242 |
| | Extreme value | 0.002533 |
| | Gamma | 0.000422 |
| | Generalized extreme value | 0.000215 |
| | Inversegaussian | 0.000368 |
| | Logistic | 0.000106 |
| | Loglogistic | 0.000065 |
| | Lognormal | 0.000361 |
| | Rayleigh | 0.005269 |
| | Rician | 0.000570 |
| | Tlocationscale | 0.000025 |
| | Weibull | 0.001777 |

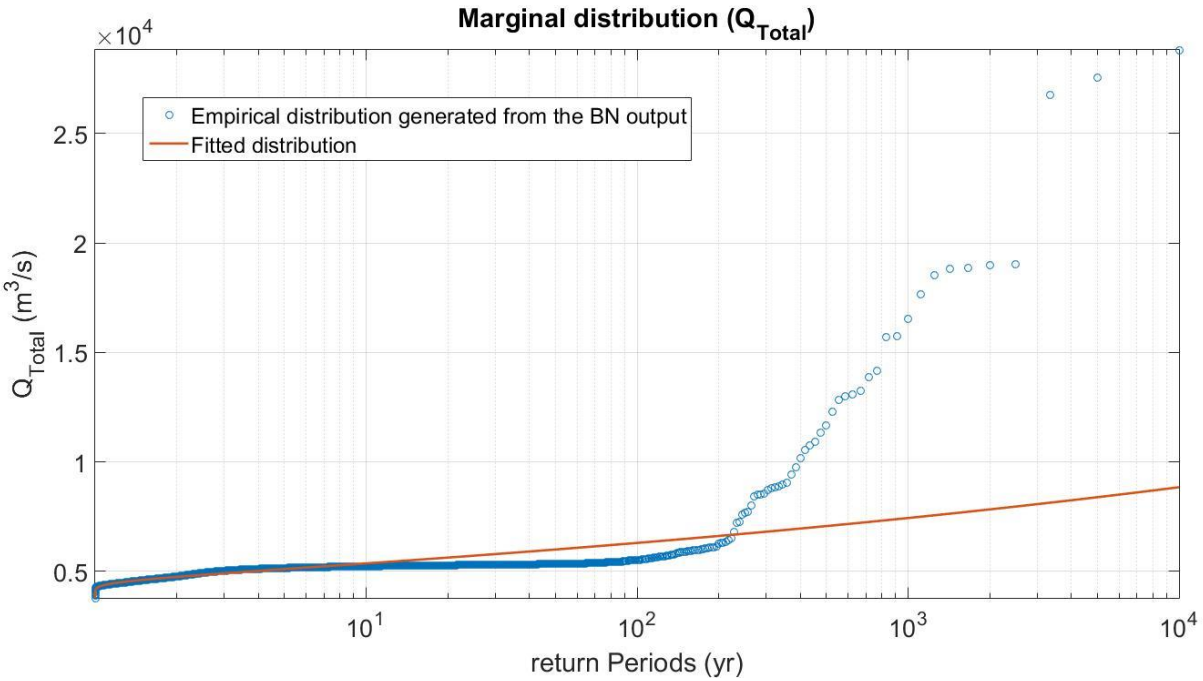



**Figure F.9: The comparison between the empirical marginal distribution and fitted continuous marginal distribution of the total discharge at upstream boundaries.**

**Table F.9: The results of Pearson's chi-squared test for fitted continuous marginal distributions of the total discharge at upstream boundaries.**

| Variable | Distribution type | Result of Chi-squared test |
|---|---|---|
| $Q_{Total}$ | Normal | 0.000309 |
| | Birnbaumsaunders | 0.000115 |
| | Exponential | 0.002651 |
| | Extreme value | 0.001611 |
| | Gamma | 0.000141 |
| | Generalized extreme value | 0.000092 |
| | Inversegaussian | 0.000115 |
| | Logistic | 0.000137 |
| | Loglogistic | 0.000144 |
| | Lognormal | 0.000109 |
| | Rayleigh | 0.001528 |
| | Rician | 0.000309 |
| | Tlocationscale | 0.000159 |
| | Weibull | 0.000901 |

## F.2. Dependence Analysis

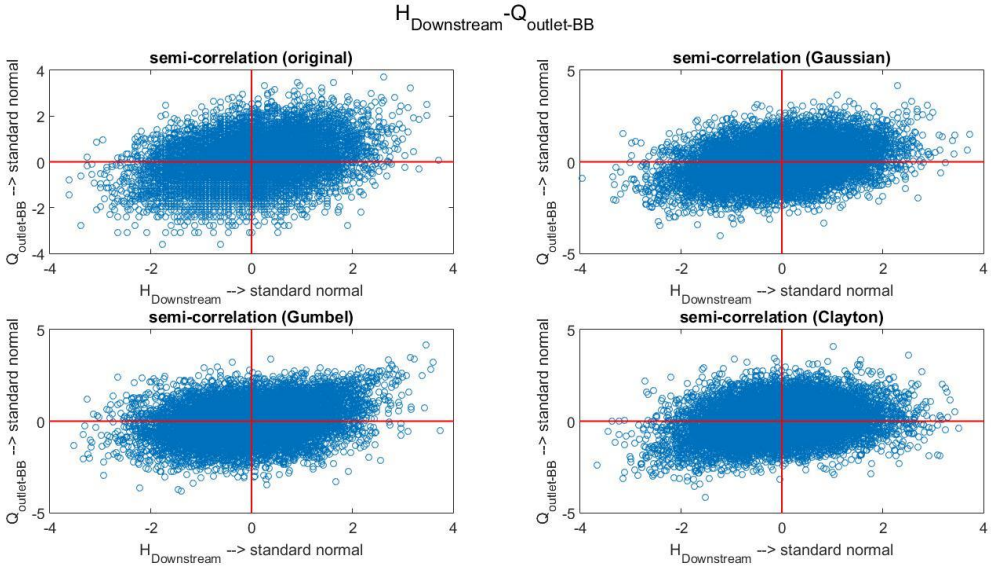



**Figure F.9: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at the outlet of Brays Bayou.**

**Table F.9: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at the outlet of Brays Bayou.**

| Morgan-BB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.3095 | 0.1211 | 0.1541 | 0.1419 | 0.1284 |
| Gaussian Copula | 0.3141 | 0.0761 | 0.1573 | 0.2083 | 0.0964 |
| Gumbel Copula | 0.2773 | 0.0078 | 0.3191 | 0.1102 | 0.0289 |
| Clayton Copula | 0.2269 | 0.0511 | 0.0345 | 0.2526 | 0.0167 |

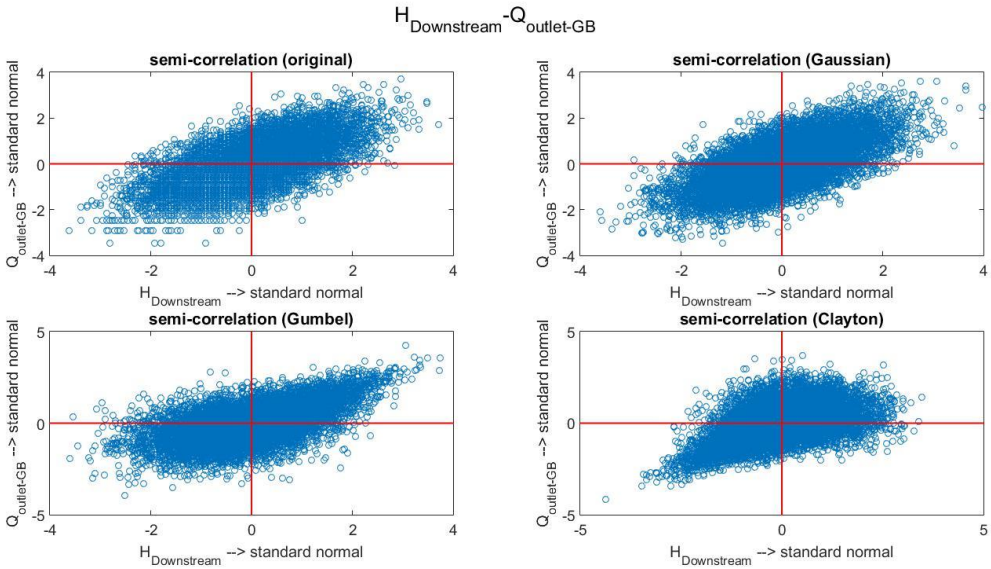



**Figure F.10: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at the outlet of Greens Bayou.**

**Table F.10: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at the outlet of Greens Bayou.**

| Morgan-GB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6392 | 0.1975 | 0.4264 | 0.3868 | 0.1754 |
| Gaussian Copula | 0.6458 | 0.2069 | 0.4285 | 0.3955 | 0.1777 |
| Gumbel Copula | 0.5863 | 0.0014 | 0.5889 | 0.2512 | 0.0526 |
| Clayton Copula | 0.4809 | 0.0606 | 0.0771 | 0.5800 | 0.0438 |

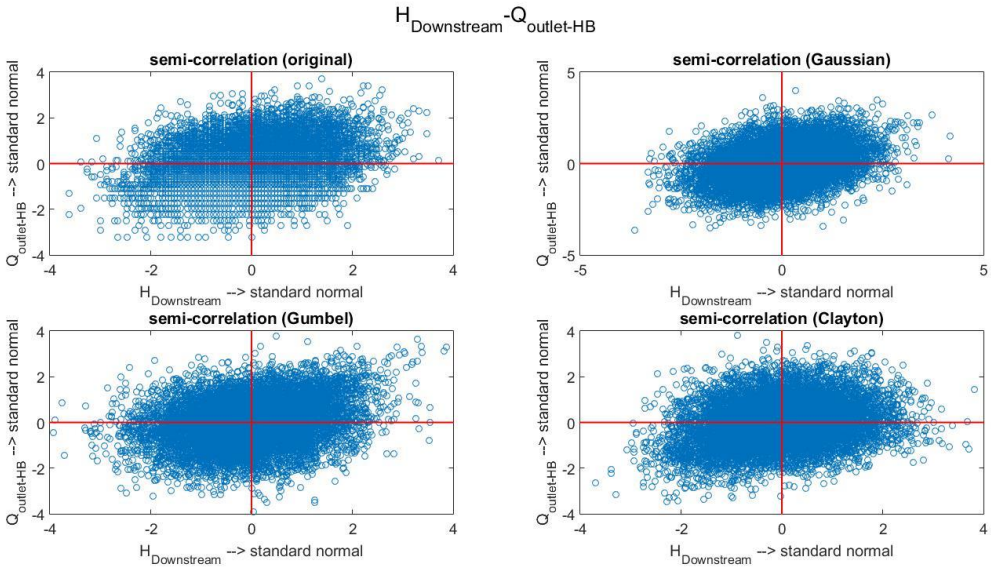



**Figure F.11: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at the outlet of Hunting Bayou.**

**Table F.11: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at the outlet of Hunting Bayou.**

| Morgan-HB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.2892 | 0.0537 | 0.1652 | 0.1457 | 0.1103 |
| Gaussian Copula | 0.3145 | 0.1507 | 0.1480 | 0.1383 | 0.1211 |
| Gumbel Copula | 0.2509 | 0.0362 | 0.2726 | 0.0628 | 0.0764 |
| Clayton Copula | 0.2149 | 0.0245 | 0.0277 | 0.2663 | 0.0747 |

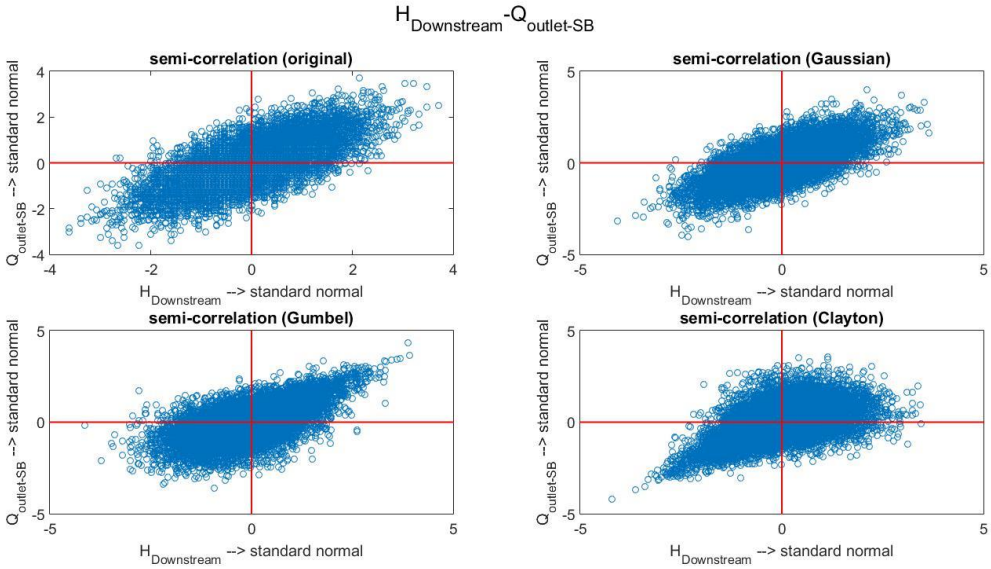



**Figure F.12: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at the outlet of Sims Bayou.**

**Table F.12: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at the outlet of Sims Bayou.**

| Morgan-SB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.6708 | 0.2050 | 0.4678 | 0.4473 | 0.2287 |
| Gaussian Copula | 0.6684 | 0.1681 | 0.4326 | 0.4743 | 0.1568 |
| Gumbel Copula | 0.6351 | 0.0621 | 0.6148 | 0.2800 | 0.0885 |
| Clayton Copula | 0.5226 | 0.0748 | 0.0640 | 0.6143 | 0.0675 |

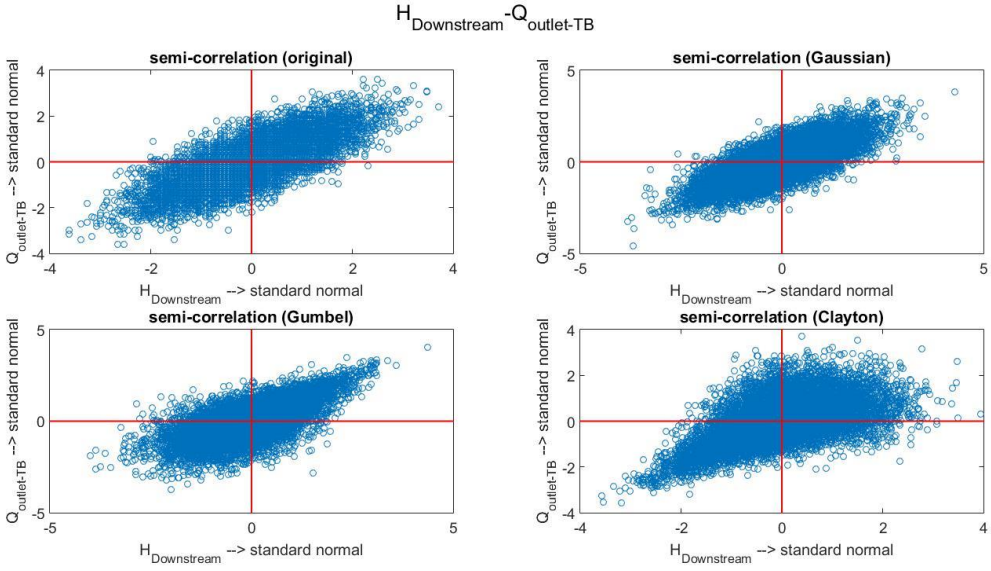



**Figure F.13: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at Turning Basin.**

**Table F.13: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at Turning Basin.**

| Morgan-TB | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7346 | 0.1999 | 0.5326 | 0.5175 | 0.2403 |
| Gaussian Copula | 0.7351 | 0.2076 | 0.5360 | 0.5367 | 0.1614 |
| Gumbel Copula | 0.6893 | 0.0964 | 0.6700 | 0.3564 | 0.0411 |
| Clayton Copula | 0.5841 | 0.1302 | 0.0851 | 0.6588 | 0.0608 |

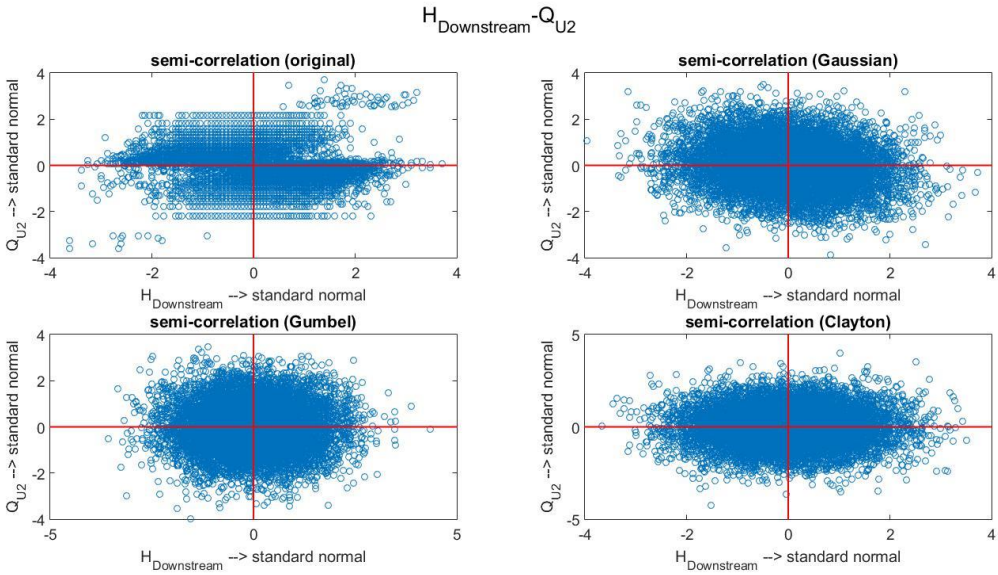



**Figure F.14: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at upstream boundary U2.**

**Table F.14: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at upstream boundary U2.**

| Morgan-U2 | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | -0.2606 | 0.2875 | -0.0696 | 0.0806 | 0.2505 |
| Gaussian Copula | -0.1520 | -0.0522 | -0.0375 | -0.0371 | -0.0977 |
| Gumbel Copula | 0.0018 | 0.0167 | 0.0025 | -0.0113 | -0.0152 |
| Clayton Copula | -0.0016 | 0.0415 | 0.0082 | -0.0336 | -0.0059 |

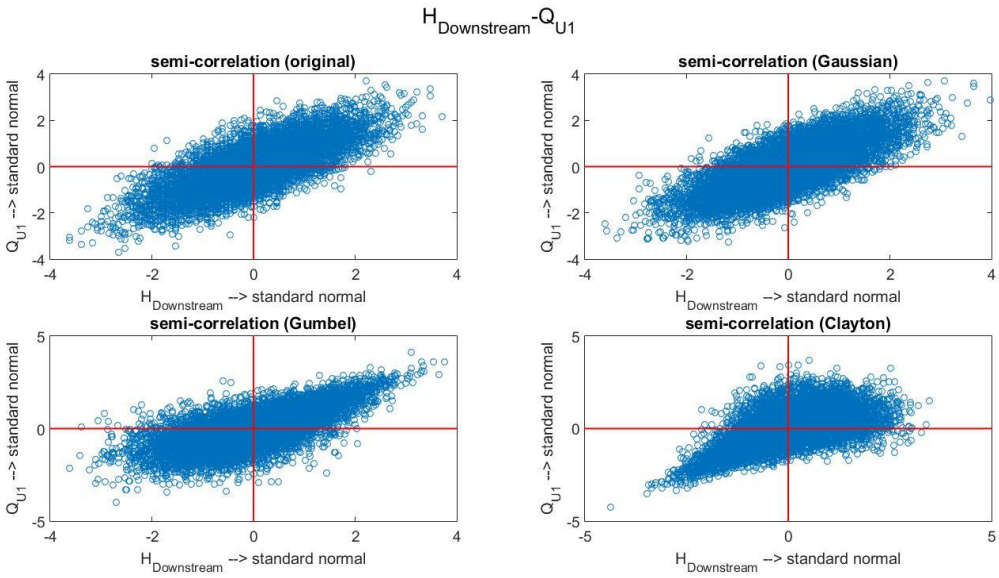



**Figure F.15: Graphical semi-correlation of the variable pair: downstream boundary sea level and the discharge at upstream boundary U1.**

**Table F.15: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the discharge at upstream boundary U1.**

| Morgan-U1 | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | 0.7359 | 0.2055 | 0.5251 | 0.5243 | 0.2717 |
| Gaussian Copula | 0.7404 | 0.1908 | 0.5402 | 0.5077 | 0.1960 |
| Gumbel Copula | 0.6833 | 0.0118 | 0.6803 | 0.3424 | 0.0929 |
| Clayton Copula | 0.5863 | 0.0664 | 0.1160 | 0.6813 | 0.0653 |

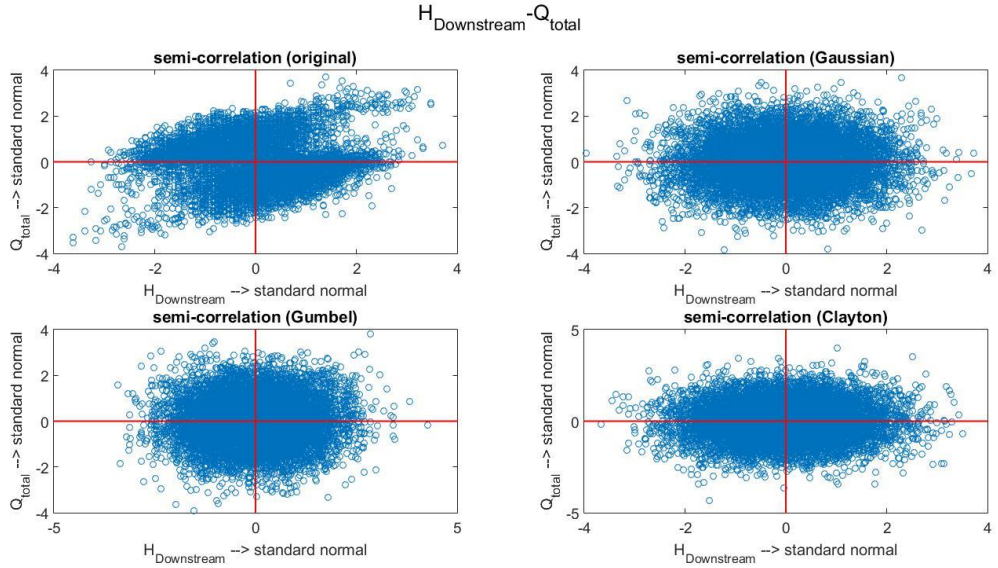



**Figure F.16: Graphical semi-correlation of the variable pair: downstream boundary sea level and the total discharge at upstream boundaries.**

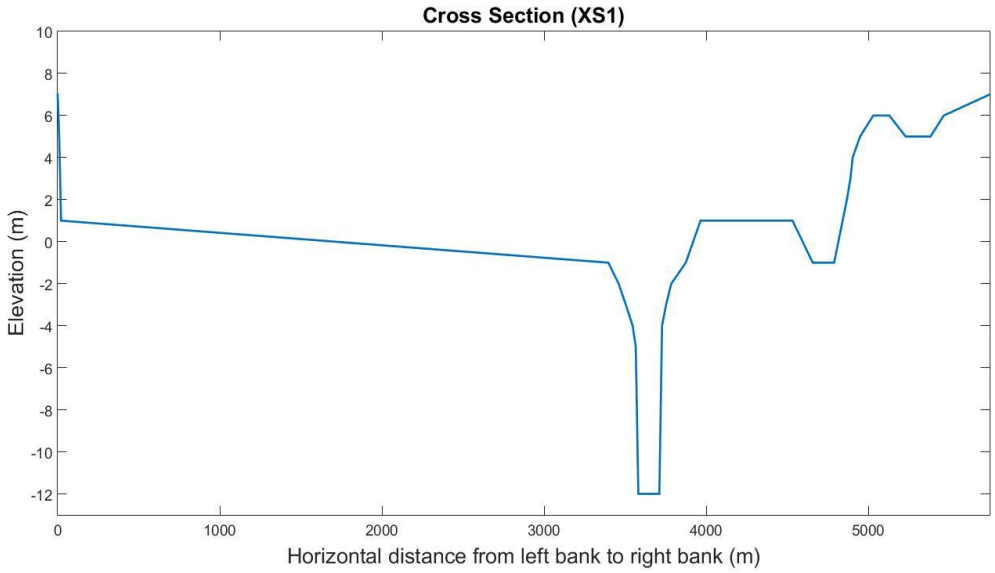
**Table F.16: The detailed information of semi-correlation tests for the variable pair: downstream boundary sea level and the total discharge at upstream boundaries.**

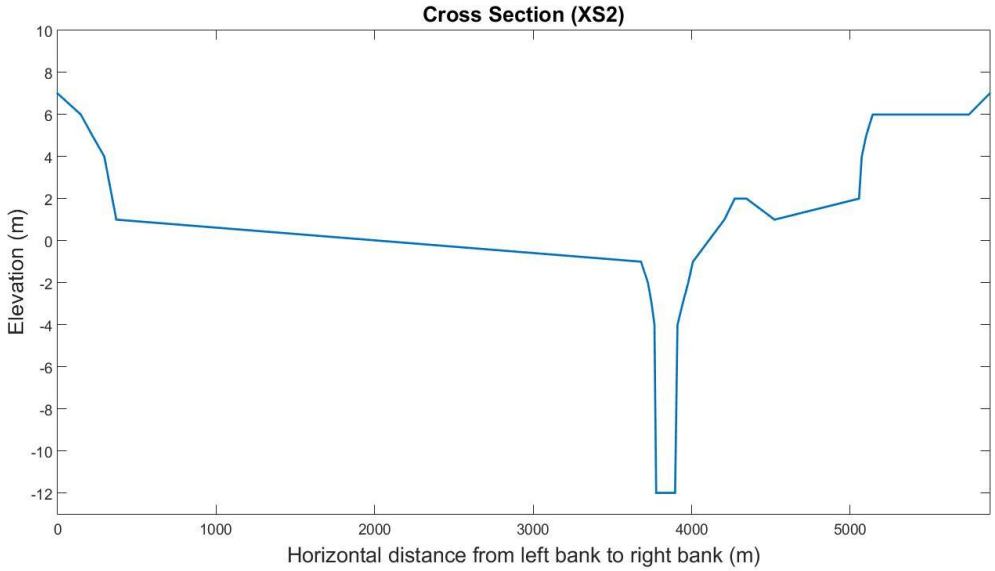
| Morgan-total | ρ | ρNW | ρNE | ρSW | ρSE |
|---|---|---|---|---|---|
| Original data | -0.1069 | 0.3204 | 0.0394 | 0.0876 | 0.3292 |
| Gaussian Copula | 0.0199 | -0.0079 | 0.0078 | 0.0328 | -0.0049 |
| Gumbel Copula | 0.0322 | -0.0472 | 0.0688 | 0.0245 | 0.0027 |
| Clayton Copula | 0.0185 | -0.0053 | 0.0290 | 0.0128 | -0.0098 |

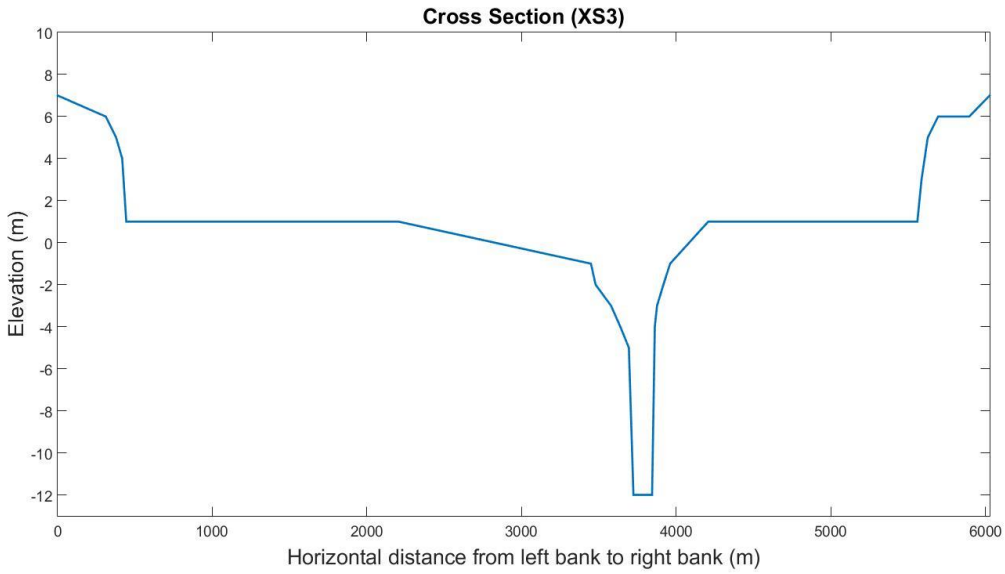
# Appendix G

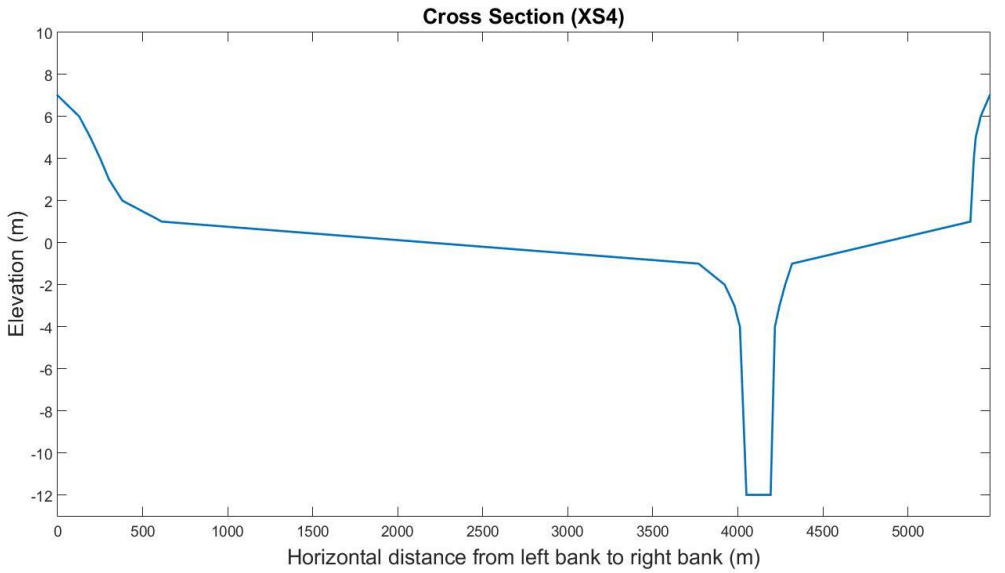
## Divided Cross-Section Profile

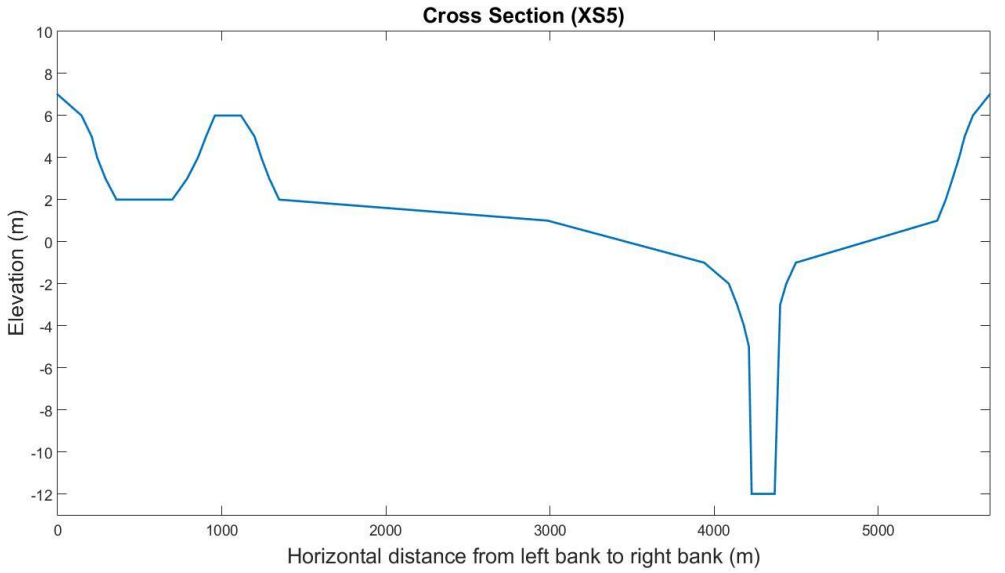
### G.1. The divided cross-section profiles in the study area

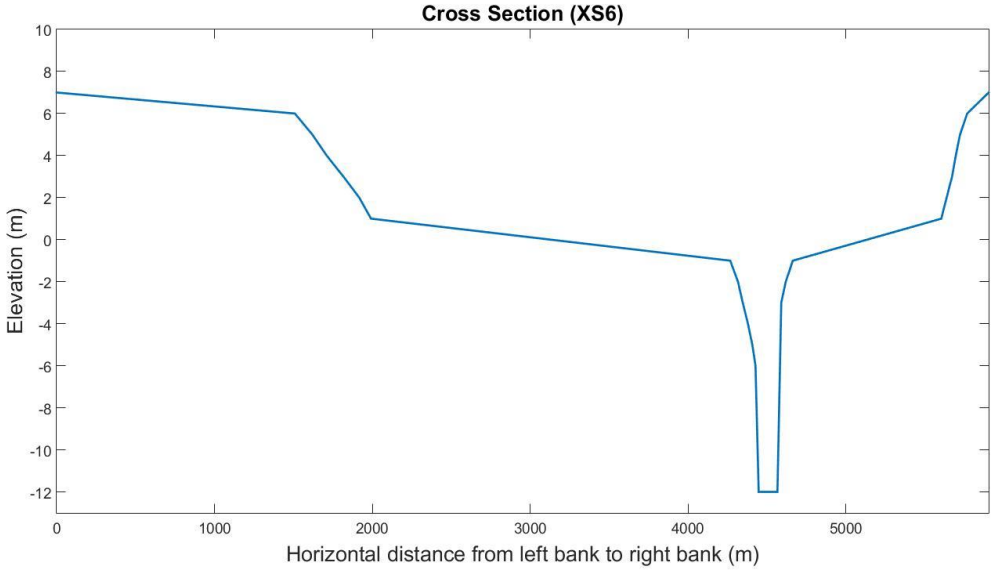


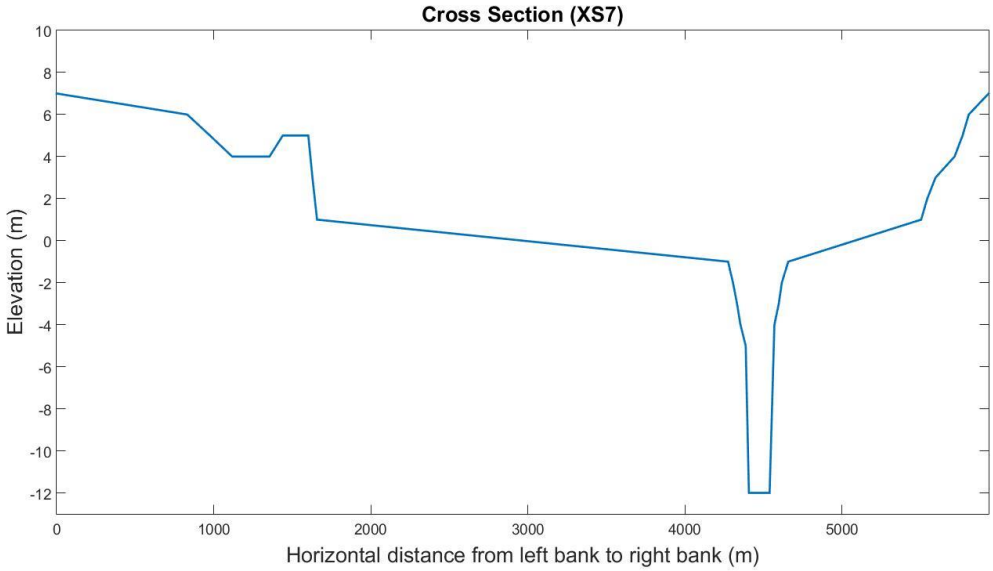



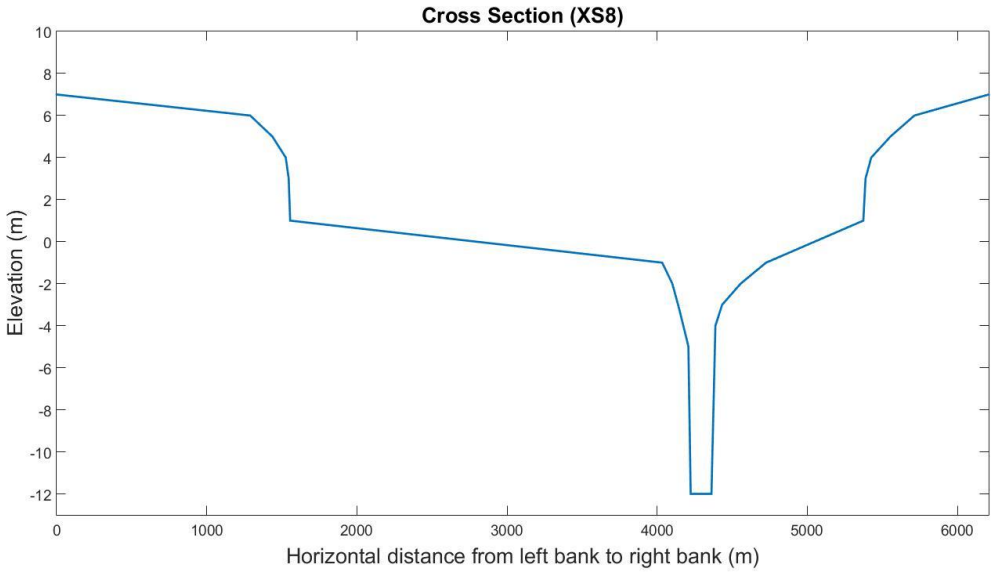

Cross Section (XS3)
Elevation (m)
Horizontal distance from left bank to right bank (m)

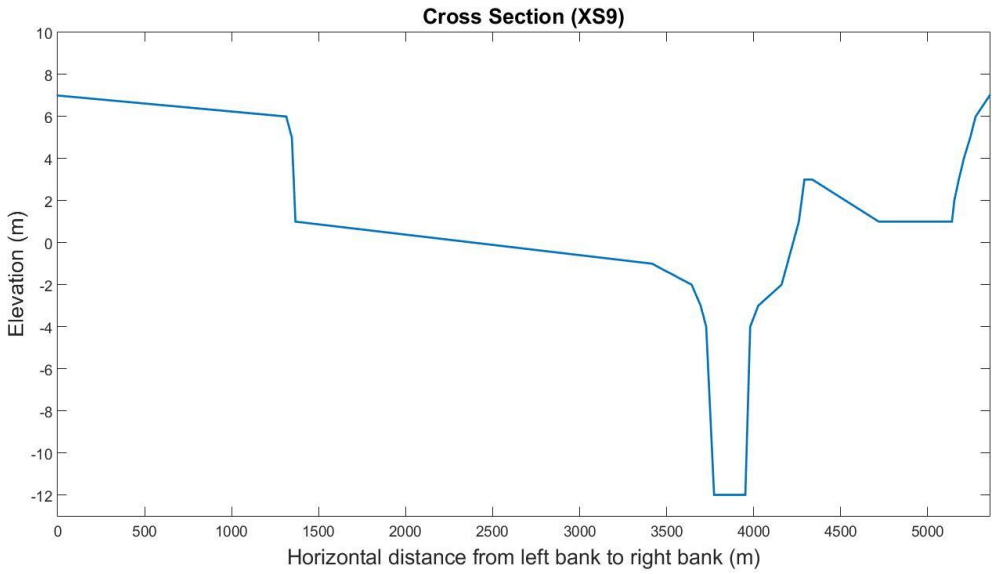

Cross Section (XS4)
Elevation (m)
Horizontal distance from left bank to right bank (m)

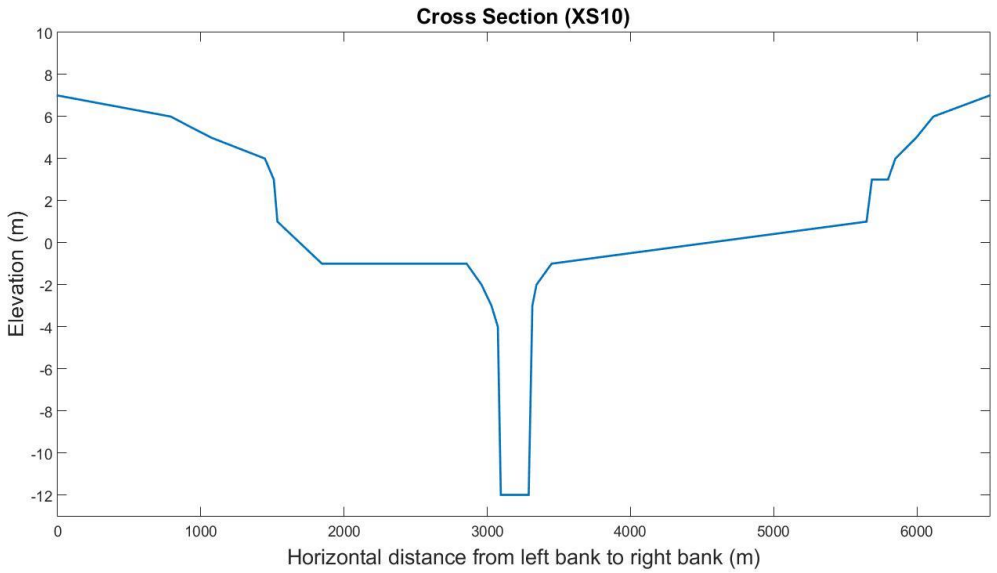

Cross Section (XS5)
Elevation (m)
Horizontal distance from left bank to right bank (m)

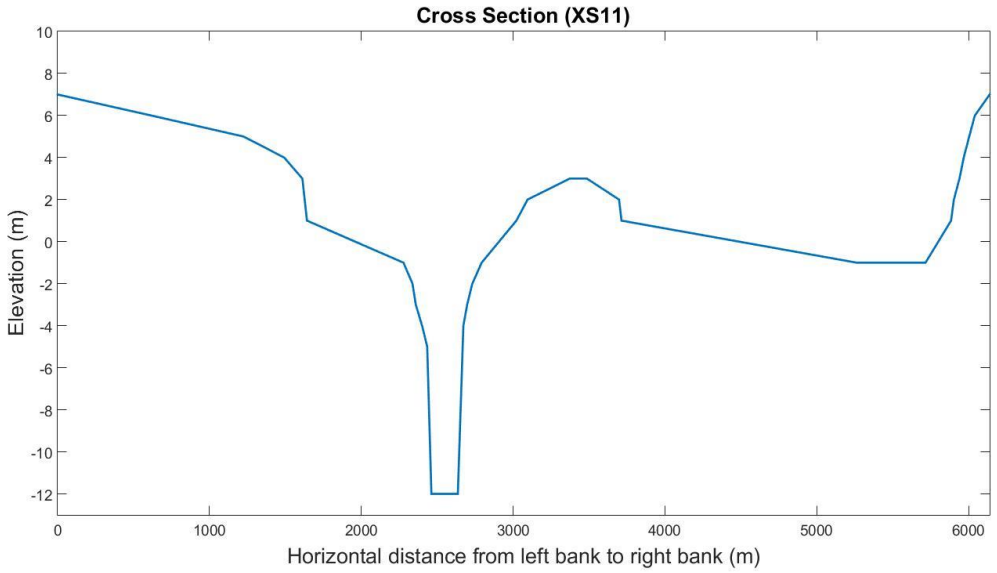

Cross Section (XS6)
Elevation (m)
Horizontal distance from left bank to right bank (m)

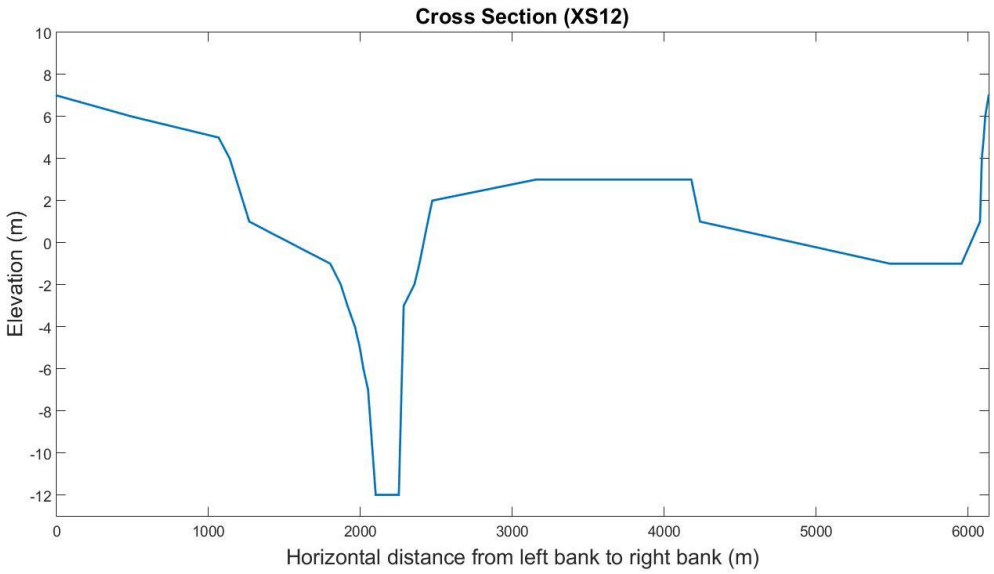

Cross Section (XS7)
Elevation (m)
Horizontal distance from left bank to right bank (m)

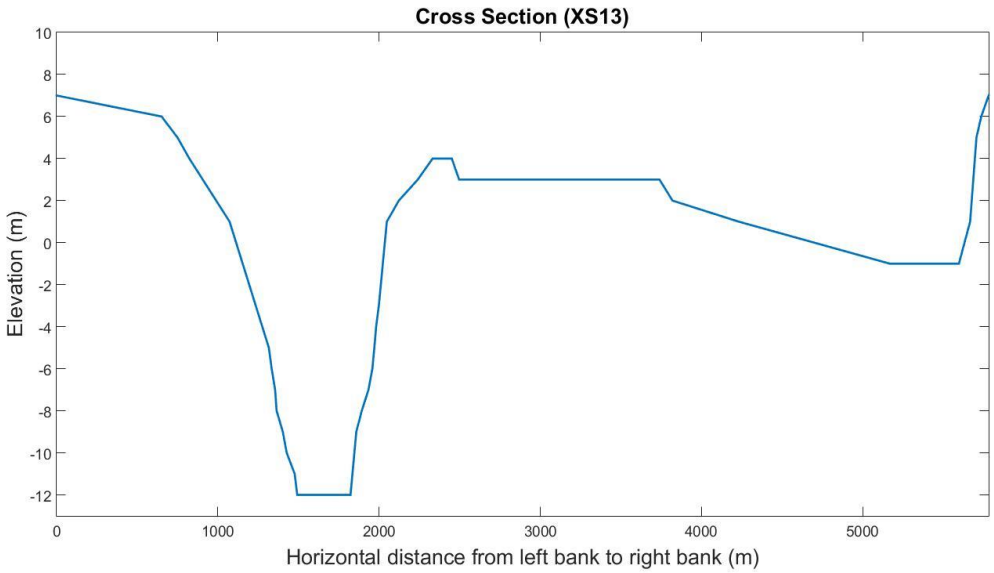

Cross Section (XS8)
Elevation (m)
Horizontal distance from left bank to right bank (m)

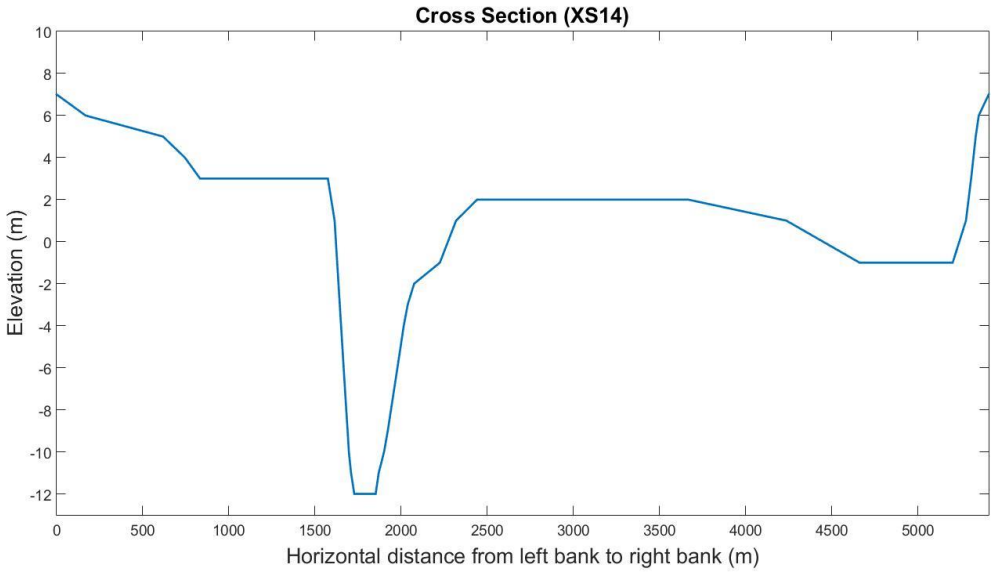

Cross Section (XS9)
Elevation (m)
Horizontal distance from left bank to right bank (m)

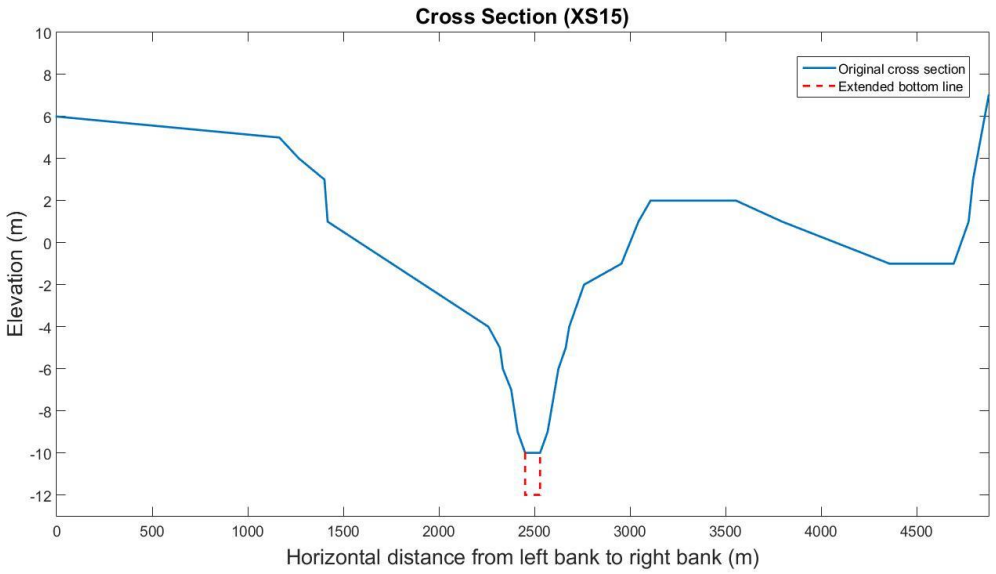

Cross Section (XS10)
Elevation (m)
Horizontal distance from left bank to right bank (m)

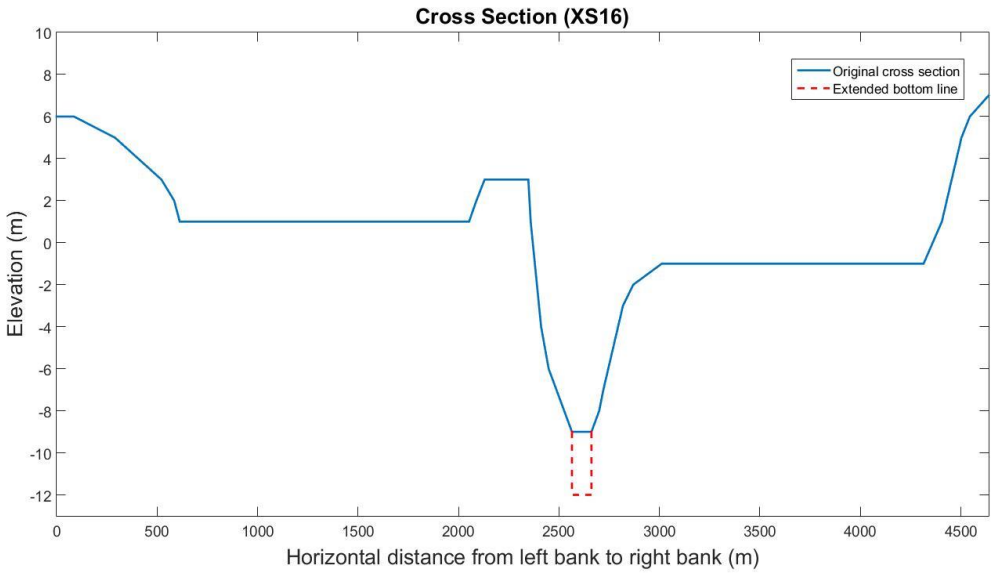

Cross Section (XS11)
Elevation (m)
Horizontal distance from left bank to right bank (m)

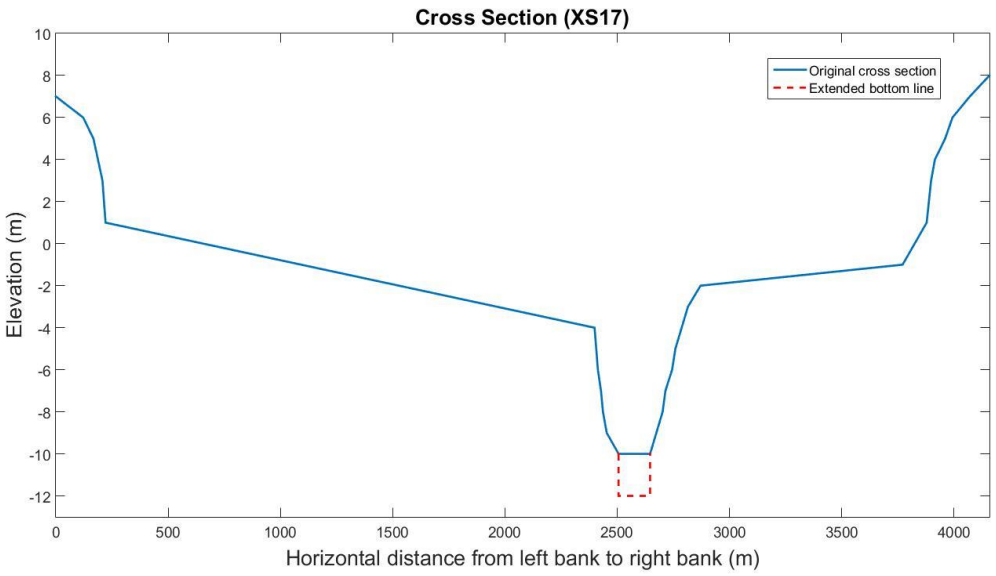

Cross Section (XS12)
Elevation (m)
Horizontal distance from left bank to right bank (m)


Cross Section (XS13)
Elevation (m)
Horizontal distance from left bank to right bank (m)


Cross Section (XS14)
Elevation (m)
Horizontal distance from left bank to right bank (m)


Cross Section (XS15)
Original cross section
Extended bottom line
Elevation (m)
Horizontal distance from left bank to right bank (m)


Cross Section (XS16)
Original cross section
Extended bottom line
Elevation (m)
Horizontal distance from left bank to right bank (m)


Cross Section (XS17)
Original cross section
Extended bottom line
Elevation (m)
Horizontal distance from left bank to right bank (m)

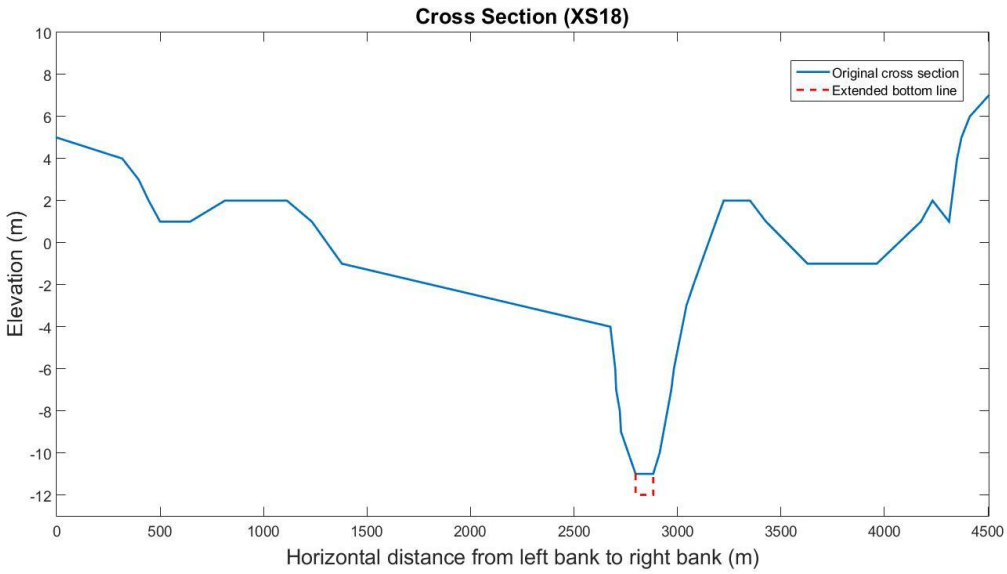



**Figure G.1: The divided cross-section profiles in the study area. Source: HCFCD.**