ABSTRACT

Title of Document: DEVELOPMENT OF A COMPREHENSIVE FLOOD HAZARD ASSESSMENT METHOD FOR MULTIPLE FLOOD SOURCES.

Stephanie Bray, Doctor of Philosophy, 2013

Directed By: Professor Richard McCuen, Civil and Environmental Engineering Department

For locations that could be influenced by multiple flood sources, considering the probability of flooding from each source individually does not provide a full understanding of the likelihood of a given flood depth occurring. Current methods of calculating the probabilities of flooding from multiple flood sources typically make the assumption that the flood sources are independent of each other. This research developed a method of calculating the probabilities of flooding for a location of interest based on multiple flood sources when a dependent relationship existed between those flood sources. Joint distributions were developed to describe the relationship between the flood sources using Archimedean copulas. Then, probabilities that correspond to total flood depths were calculated by taking integrals under the joint pdfs determined based on the copula equations. This process was carried out for both two flood sources and three flood sources, using both simulated

and observed data. Zero-flood years were found to significantly impact the fitting of marginal and joint distributions in this process.

A common assumption is that the flood sources are independent of each other. When a relationship between the flood sources is evident, the assumption of dependence or independence was observed to significantly impact the calculated exceedance probabilities. Flood risk estimates were made based on probabilities calculated while considering all flood sources, to determine the impact that this new flood hazard assessment would have on flood risk assessments. Finally, the expected error in the results of the flood hazard assessment was assessed through an uncertainty analysis. This analysis focused on the parameters used to calculate the flood depths from the observed gage measurements, which were primarily related to the physical characteristics of the location of interest, and on the distributions used to represent the total flood depths. A high level of uncertainty was found to exist in the results of the analyses, indicating the importance of using the most accurate information possible to determine the parameter values and of using care in selecting the best distribution to represent the total flood depths.

DEVELOPMENT OF A COMPREHENSIVE FLOOD HAZARD ASSESSMENT METHOD FOR MULTIPLE FLOOD SOURCES.

By

Stephanie Nicole Bray.

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Richard McCuen, Chair
Professor Kaye Brubaker
Professor Gerald Galloway
Professor Ed Link
Professor Adel Shirmohammadi

Acknowledgements

I would like to gratefully acknowledge the assistance provided by my advisor and all of my committee. Their suggestions and knowledge have greatly improved this document.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1. Risk and Risk Analysis

The term "risk" has a number of general definitions. For instance, it can be defined as the perceived extent of possible loss. A similar definition is the possibility of loss, disadvantage, or destruction. It could also be considered as the chance or likelihood that something with undesirable impacts will occur. Another way of saying this is that risk is the combination of the undesirable consequences of undesirable scenarios and the probability of these scenarios (Stamatelatos, 2002). This leads to a mathematical definition of the probability of an event multiplied by the cost or consequences of the event. The risk that arises from a particular hazard is determined by how often that hazard may occur and how much harm is likely to result from the hazard. It is possible to reduce risk in two ways: by making the undesirable hazard event less likely to occur or by making the associated consequences less serious (USNRC, 2007). To determine risk involves answering three questions: What can go wrong? How likely is it to go wrong? What are the associated consequences (Stamatelatos, 2002)?

Risk analysis is the process of systematically studying the risks that will be faced by an individual, organization, or community. The purpose of such an assessment is to identify the risks that may be faced, understand how and when they could occur, and estimate the impacts of adverse outcomes caused by the risks. A

risk analysis will typically focus on what could go wrong, but can also be useful in determining what could go right (Solver.com, 2010). The process of completing a risk analysis includes identifying the threats or hazards to be faced, estimating the associated risk, choosing methods to best manage the risk, and conducting regular reviews (MindTools, 2010).

Risk assessments have shifted from deterministic to probabilistic methods. Deterministic methods use point estimates of risk, which are frequently, though not always, representative of a worst-case scenario. Because deterministic risk assessment uses point estimates of input parameters, the use of average values could result in an underestimate of risk and the use of upper-bound estimates (representing worst case scenarios) could overestimate the risk. The distinction between deterministic and probabilistic risk assessments is that probabilistic methods allow for the use of distributions for the input parameters (Exponent, 2010). Probabilistic risk assessment can be used to systematically examine all of the components of a complex system and to determine how they work together. This can be used to specifically quantify risk and to identify the elements that could have the most serious effects on safety. The steps required to conduct a probabilistic risk assessment include specifying the hazard event, identifying initiating events that could lead to the consequence, estimating the frequency of each of the initiating events, and identifying all combinations of failures that could lead to the consequences. From this point the likelihood of each combination can be computed, and the probabilities of all failure combinations that lead to the same consequences can be added. In order to determine the potential frequency of a given outcome, the probabilities computed should be

multiplied by the frequency of the identified initiating events. The outcome of a probabilistic risk assessment will not be a single number, but will rather be a distribution or spectrum of possible outcomes (USNRC, 2007). Many government agencies, such as the NRC and EPA, have begun to use risk-based or risk-informed regulations, based on probabilistic risk assessment methods, to ensure safety without requiring unnecessarily high conservatism (Stamatelatos, 2000).

Monte Carlo simulation is frequently used in risk calculations because possible variations in each factor within the analysis can be considered, as can interactions between factors and imperfect knowledge (USNRC, 2007). Monte Carlo analyses can be used to examine the effect of uncertainty and natural variability on estimates of risk. The basis of a Monte Carlo analysis is a probability density function (pdf) for input parameters. Therefore, choosing the appropriate pdf is critical to ensure that meaningful results are found. After pdfs have been specified for the input parameters, a computerized routine is run repeatedly with input parameter values being selected according to the specified pdfs (Hayse, 2000).

1.2. Flood Risk

Flooding is one of the many natural hazards that communities could be at risk of experiencing. A flood can be defined as an event when the water level in a specific water body, such as a stream or river, lake, ocean, or a land-based point, rises above the normal limit (ASCE, 1996). Many sources of flooding are possible, including riverine or fluvial flooding, coastal or tidal flooding, pluvial or surface water flooding, and groundwater flooding (Pitt, 2008). Floods could also come from flash

floods, snowmelt events, ice jams, and mudflows (Simonovic, 2009). Sources of flooding suggested by FEMA (2010) include rainfall, river-flow, and tidal surge. Other factors that can increase the occurrence and severity of floods include the local topography, flood control measures, and development of the area. A failure to properly maintain infrastructure such as culverts and bridge openings can also lead to flooding. FEMA suggests that when conducting flood risk assessments, a study should evaluate statistical data for the river flow, storm tides, and rainfall, as applicable, hydrologic and hydraulic analyses, and an analysis of the local topography (FEMA, 2010). Floods are significant in the United States because floodplains consist of approximately 7% of US land area and serve as a valuable natural resource (NRC, 2000). Building within the floodplain is a significant factor.

1.2.1. Components and Use of Flood Risk Assessment

Bulletin 17B (IACWD, 1982) provided a definition for flood risk that suggested risk was the probability that one or more events will occur which exceed a given flood magnitude within a specified period of time. The definition of flood risk has shifted somewhat since then, such that flood risk is now typically described as a combination of the flood hazard, which is the probability of the event occurring, and the consequences of the event (Apel *et al.*, 2004). Usually the risk is determined by multiplying the probability of occurrence by the consequences of the event, if the consequences can be quantified. Another view of flood risk considers it as the combination of the threat of the event, which would be the probability, the vulnerability to the event, which would be the protection provided by flood risk reduction measures and the reliability of those measures, and the consequences of the

event (Carter, 2005). Again, the terms are combined by multiplication. A thorough flood risk analysis procedure must first define risk and then account for all of the relevant flood scenarios, the probabilities of occurrence associated with each scenario, the potential damages associated with each scenario, and the uncertainties that could impact the analysis in each scenario (Apel *et al.*, 2004). An interesting aspect of flood risk to consider is that lowering the probability of flooding, using flood risk reduction measures such as levees or floodwalls, may not necessarily reduce the flood risk. These measures may encourage further development in the area, raising the consequences and the risk from flooding. Therefore, it is important to develop resilience to floods rather than resistance (Dijkman and Heynert, 2003).

The US Army Corps of Engineers began to investigate the possibility of using risk-based analysis in their flood-control work in the early 1990s. Their risk-based method requires a probabilistic estimation of uncertainty relevant to important variables such as flood-frequency, stage-discharge, and stage-damage relationships. These estimates can be used in designing a levee or other projects (NRC, 2000). Since then, the US Army Corps of Engineers has been expected to use risk-based analysis in the formulation of their flood damage reduction projects, including dams and reservoirs, levees and floodwalls, diversions, channel modifications, bypass channels, and nonstructural measures. Historically, the Corps of Engineers used best estimates of flood hazard and damage potential, which could be reflected by stage-flow rating curves and stage-damage curves, in order to develop and evaluate project alternatives. In this method, performance was considered as a degree-of-protection concept and uncertainty could be dealt with by best professional judgment, sensitivity

studies, or adding a freeboard amount to levees and floodwalls. A factor of safety was also often incorporated into the design criteria for flood damage reduction projects. The new policies depart from the historical polices in that uncertainty can be explicitly quantified. The risk-based approach also allows performance to be stated in terms of the expected annual exceedance and reliability of achieving goals (USACE, 2000).

1.2.2. Probability Determination Through Flood Frequency Analysis

One of the primary components of risk is the probability of a hazard event occurring. In risk calculations, the probability or likelihood of an event occurring is multiplied by the consequences of the occurrence of the event. In some cases, the probability and consequences are also multiplied by the vulnerability, which is a measure of how well the area is prepared for and protected against the event. The vulnerability reflects the system that is in place to mitigate the event. The system may include a levee, a floodwall, a diversion channel, or any other measures taken to reduce the impact of a flood event. The ability of that system to withstand and reduce the consequences of the event is reflected in the vulnerability term. Statistical analyses are often used when calculating the probabilities of flood events by estimating the future frequency of occurrence based on information in the available hydrologic records. To determine the probability of flooding a frequency curve, which can relate the magnitude of a flood event to the exceedance probability with which that particular magnitude is exceeded, with the assumption that flood events occur at random, are used. Typically the frequency curve of annual maximum or annual minimum events is used (Beard, 1962).

Prior to the 1960's, a uniform Federal guidance for flood frequency analysis had not been enacted. The need for this was realized in the 1960's, and the Federal government took a first step toward providing that uniform guidance with Bulletin 15, a Uniform Technique for Determining Flood Flow Frequencies, in 1967. This was the first time that a uniform set of flood frequency techniques was utilized by all of the Federal agencies involved in flood-related work. In 1976 Bulletin 17, Guidelines for Determining Flood Flow Frequency, was published. This document was quickly revised to Bulletin 17A, which was published in 1977 with the purpose of resolving discrepancies in the method for incorporating historical data into the analysis. Further revisions were found to be necessary, and Bulletin 17B was published in 1981, attempting to resolve discrepancies in the treatment of low outliers and the weighting method for skew estimation. Bulletin 17B was again revised in 1982, and this is the document that contains the currently recommended flood frequency analysis methodologies used by the Federal agencies (Griffis and Stedinger, 2007a).

Floods are typically predicted based on frequency analysis, a statistical tool that can be used for any random variable. Precipitation events can also be analyzed using frequency analysis. The result of a frequency analysis is often a graph of the value of the hydrologic variable versus the frequency of occurrence. This graph represents the best estimate of the statistical population responsible for the sample values collected. A mathematical model could also be used in place of a graph. The equation typically used to model a flood is:

$$X = \bar{X} + KS \tag{1-1}$$

where X is the discharge, $\bar{X}$ is the mean flow, S is the standard deviation of the flows, and K is a frequency factor (McCuen, 2005).

The first step to a conducting a flood frequency analysis for a river with a flow gage is to obtain a series of annual maximum discharge values. The rank-order method is commonly used to order and plot this sample data. Using the rank-order method, the discharge data should be ordered from largest (with rank 1) to smallest (with rank n). A number of plotting position formulas could then be used, such as the commonly used Weibull equation. The Weibull plotting position formula is:

$$P_i = \frac{i}{n+1} \tag{1-2}$$

where i is the rank, n is the total number of records, and P_i is the exceedance probability for the event with rank i. This allows the collected annual maximum discharge events to be plotted. In order to estimate the population from which the sample data were drawn, the mean, standard deviation, and skew of the annual maximum flood record need to be computed. To derive a population model an underlying probability density function (pdf) must first be chosen. It is possible to use any pdf to represent the population curve; however, the lognormal and log-Pearson Type III pdfs are typically used in hydrological applications. If either of these distributions are used, the sample data must first be transformed to logarithms of the data, and then the sample statistics (mean, standard deviation, and skew) can be calculated. The sample moments calculated are equated with the parameters of the chosen pdf. From here it is possible to develop the curve that will represent the

underlying population of the floods, which can be used to estimate the magnitudes for floods of given return periods or exceedance probabilities (McCuen, 2005).

The log-Pearson Type III distribution has three parameters, with two interacting shape parameters. Though this distribution is commonly used in hydrology, some have expressed concerns about whether or not it is the appropriate distribution. Specifically, the existence of an upper bound for flood flows has caused concern, as an upper bound on natural flows does not make physical sense. Griffis and Stedinger (2007b) have studied this distribution extensively and provided some insight into scenarios in which the distribution is appropriate for use in flood analysis (Griffis and Stedinger, 2007b).

1.2.3. Probability of Flooding from Multiple Sources

Flooding may have a variety of sources, and in some cases may result from multiple sources simultaneously. Combined-population frequency analysis can be used to consider the probability of flooding caused by a variety of sources, which may or may not occur at the same time, while joint probability methods can determine the probability of multiple causes of a flood occurring simultaneously.

1.2.3.1. Combined-Population Frequency Analysis

In the case of flood risk, it is often necessary to consider flooding that is the result of multiple factors. Combined-population frequency analysis can be used to determine the probability of a flood event when multiple hydrologic factors could be causing the flooding.

Morris (1982) provided a thorough evaluation of combined-population frequency analysis. When different types of hydrologic phenomena could be responsible for flood events, special treatment may be needed in the process of developing a frequency curve. Examples of this are (1) locations where both cyclonic and convective rain storms can cause flooding and (2) locations where floods could be caused by rain storms or by melting snowpack. Morris (1982) defines a combined-population frequency curve as one that is derived from two or more separate frequency curves, each of which was developed from a separate population. The assumption is commonly made that the populations are independent. This method develops a frequency curve based on two or more sets of data, or populations, which result from different causal factors. The result is a frequency curve for the flood source of interest, for example riverine flooding, that is based on multiple causal factors. When events in a series do not occur every year, such as zero-flood years, special frequency analyses will be needed. A general equation has been developed for combining frequency curves from multiple annual flood series data. The equation used in these scenarios is:

$$P_c = 1 - \prod_{i=1}^{n}(1 - P_i) \qquad (1\text{-}3)$$

where P_c is the exceedance probability of the combined-population frequency curve for the selected discharge and P_i are the exceedance probabilities associated with the selected discharges from the frequency curve (Morris, 1982).

1.2.3.2. Joint Probability Analysis

Hawkes (2005) provides an explanation of joint probability. Joint probability is the chance that two or more conditions occur at the same time. In the case of flooding, this could mean riverine flooding occurring near a coast at the same time as a high tide occurs. Hawkes (2005) explains that the study of joint probability is important to flood evaluations because high or extreme loadings on a flood defense structure, such as a levee or flood wall, will often be caused by more than one variable. Therefore, the probability that a given load will occur will be a function of the combined probabilities of occurrence of all the variables involved. Two primary approaches have been used in joint probability analyses. The first is referred to as the desk-study approach. This approach requires the high and extreme values of two variables and some representation of the dependence of the two variables as input. The method provides a number of pre-calculated computations of the joint return period of the two variables for a number of alternative joint return periods and levels of dependence. The second approach is referred to as the analytical approach. This approach requires relatively long records of the simultaneous occurrence of the two variables and uses Monte Carlo simulation in order to develop thousands of years of simulated joint data. The simulation is based on project-specific statistical analyses. Example situations for which this may be necessary include a coastal site where it is necessary to consider both waves and sea level and in urban areas where it may be necessary to consider both river or coastal flooding and interior drainage (Hawkes, 2005). Hawkes (2008) clarifies that it is possible to consider more than two variables

in a joint probability assessment, but the higher the number of variables being considered, the more complex the calculations will be.

1.2.3.3. Use of Copulas to Determine the Joint Probability Distribution Function

The copula function is a function that joins (or couples) a multivariate distribution function to its associated marginal distribution function (Nelson, 1999; Balakrishnan and Lai, 2009). Copulas have been increasingly used in various applications, from financial to hydrologic, because they allow the marginal, univariate distributions of the variables of interest to be addressed separately from the dependence structure between these variables. Further, they allow joint probability distributions to be expressed as a function of the marginal probability distributions (Cherubini *et al.*, 2004). The marginal distributions for each random variable are also expressed as uniform over the range of [0, 1]. Essentially, copulas have been found to be a convenient method of obtaining the joint behavior of random variables in scenarios where the traditional assumptions made in multivariate analyses do not hold. These assumptions include the assumption of independence between the random variables and the assumption that the marginal distributions of each variable must be modeled using the same distribution.

Previously, univariate analyses methods have been favored over multivariate analyses because of the difficulty in meeting these assumptions. However, to fully characterize hydrologic events such as floods and storms, a number of characteristics must be considered. Thus, copulas enable multivariate analyses to be conducted more readily, which utilize more information and produces more informed results. Further,

the characteristics that describe hydrologic events are rarely independent and do not necessarily follow the same marginal distribution.

Sklar's theorem, developed in Sklar (1959), provides an understanding of the ability of copulas to relate a bivariate distribution to its univariate marginal distributions (Balakrishnan and Lai, 2009). This theorem is central to the theory of copulas. Cherubini *et al.* (2004) interpret the theorem to mean that a joint distribution can be determined in terms of a copula function with marginal distributions as arguments, and conversely, that a copula taking univariate marginal distributions as arguments can yield a joint distribution. This theorem underscores that copulas can be combined with marginal distributions in order to develop a joint distribution for several random variables that could not otherwise, or at least could not easily be developed.

1.2.3.3.1. Families of Copulas

There are many classes and families of copulas from which to choose. The selection of the appropriate class or family depends on the scenario being modeled. For instance, certain copula families are limited in the levels of correlation between the random variables for which they can be used. The three most common families used in hydrologic applications are the elliptical copulas, the extreme value copulas, and the Archimedean copulas (Klein et al., 2010; Favre *et al.*, 2004).

Elliptical copulas are related to elliptical distributions, such as the normal distribution and the Student's t distribution. The elliptical copula parameter has a one-to-one relationship with Pearson's correlation coefficient. A strength of elliptical

copulas is that they allow the specification of different levels of correlation between the marginal distributions. However, they do not have easy closed-form expressions (Balakrishnan and Lai, 2009; Cherubini *et al.*, 2004). Extreme value copulas are best used for modeling rare events. The extreme value copulas may be used with extreme value marginal distributions. For a series of values (X_i, Y_i) through (X_n, Y_n) with a common copula there will be a copula of componentwise maxima $X_{(n)}$=max X_i and $Y_{(n)}$=max Y_i. The extreme value copula for u and v between 0 and 1 will be calculated as:

$$C_{(n)}(u,v) = C^n(u^{\frac{1}{n}}, v^{\frac{1}{n}}) \tag{1-4}$$

Archimedean copulas are particularly easy to construct and a number of copula equations belong to this family, making them a very popular choice. Archimedean copulas are based on a generator function, φ, where φ(1) is equal to zero (Klein *et al*., 2010). The copula function is calculated as:

$$C_\theta(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \tag{1-5}$$

where C_θ is the copula and u and v are values of the random variable. The generator function of an Archimedean copula must be both decreasing and convex between 0 and 1. There is also a one-to-one relationship between Kendall's τ, a measure of rank correlation, and the parameter of Archimedean copulas (Favre *et al.*, 2004).

1.2.3.3.2. Modeling of Dependence Structure Between Variables

A number of statistical measures of dependence or association can be used with copulas to express the statistical dependence structure between the random

variables. Nelsen (1999), among many other authors, discuss the concept of concordance. Concordance indicates that "large" values of one variable tend to correspond with "large" values of the other variable, or that "small" values of one correspond with "small" values of the other, mathematically defined as $(x_i-x_j)(y_i-y_j)$ greater than zero. Discordance indicates that "large" values of one variable tend to correspond with "small" values of the other, mathematically defined as $(x_i-x_j)(y_i-y_j)$ less than zero. Kendall's τ is one measure of association that can be defined using the concept of concordance. The sample value for τ can be computed based on the number of concordant and discordant data pairs in the sample. Depending on the information that is known, τ can be calculated based on information that is known about the sample or population, and then used in calculating the copula, or if the copula function is known, the τ can be calculated from the copula to determine the dependence structure. Spearman's ρ is also based on the concepts of concordance and discordance, and can be used in a similar manner to τ. Though these two measures of association are based on the same concepts, they result in different values when calculated for the same problem. The relationship between the two measures varies depending on the family of copulas being considered (Nelson, 1999).

Though Kendall's τ and Spearman's ρ are the most commonly used measures of association based on concordance and discordance, there are others. Balakrishnan and Lai (2009) also describe Gini's coefficient, γ. This coefficient is interpreted as the expected distance between the point (U, V) and the diagonal of [0, 1] x [0, 1]. These authors also introduce tail dependence coefficients, which do not measure association based on concordance and discordance, but on the values of the variables

located in either the upper-right quadrant or the lower quadrant of [0, 1] x [0, 1]. More complete descriptions of measures of association and dependence are given by Nelson (1999) and Balakrishnan and Lai (2009).

It is important to note that there exist relationships between the measures of association and the parameters of the copula families. Frees and Valdez (1998) illustrate the one-to-one relationship between τ and α, with α being the parameter of copulas in the Archimedean class. Cherubini et al. (2004) also provide explanations and examples of the relationship between copula parameters and the measures of association and dependence that have been discussed.

1.2.3.3.3. Use of Copulas

The first step to using a copula to calculate the joint probability of two or more random variables is to develop the marginals. If using copulas to determine the joint probability of multiple sources of flooding, the marginals would be the flood frequency curves corresponding to each individual source of flooding. Thus, flood frequency curves should first be calculated using standard methods, including choosing the most appropriate underlying probability distribution to represent each source of flooding. The cumulative distribution values corresponding to values of the flood frequency curves should then be used as the copula marginal input values. The next step is to calculate the measure of dependence between the random variables, which may be expressed by Kendall's τ or another of several measures.

The next step is to choose the appropriate copula family to represent the joint probability distribution of the random variables. One method of choosing the most

appropriate copula family involves calculation of Akaike's Information Criteria (AIC). The AIC is a commonly used method of comparing copula families. The AIC can be used to determine which family best models the dependence between the two correlated random variables. The AIC can be calculated as:

$$AIC = -2 * \log likelihood + 2 * k \tag{1-6}$$

where k is the number of parameters in the copula equation being fitted. The most appropriate copula family can be identified as the family with the lowest calculated AIC value (Klein *et al.*, 2010). The log likelihood is calculated as:

$$\log likelihood = \sum_{i=1}^{n} \log(c_\alpha \left(\frac{R_i}{n+1}, \frac{S_i}{n+1}\right)) \tag{1-7}$$

where R_i and S_i were the ranks of the riverine and tidal data points within the series, respectively, and n is the sample size. The general (not family-specific) copula density function, c_α can be calculated as:

$$c_\alpha = \frac{\partial^2 C(u,v)}{\partial u \partial v} \tag{1-8}$$

where u and v are the riverine and tidal marginal distributions, respectively, and C is the copula cumulative density function calculated based on u and v. The smallest calculated AIC value indicates the most appropriate copula family.

Another method available to determine the appropriate copula family involves creating Quantile-Quantile (Q-Q) plots for each family and then comparing these plots. Genest and Rivest (1993) outline the procedure for calculating Q-Q plots. For

the several copula families being evaluated, for each value of the two random variables, U is calculated as:

$$U = \frac{\emptyset(X)}{\emptyset(X)+\emptyset(Y)} \tag{1-9}$$

where X and Y are values of the random variables, and ø is the generating function of each copula being evaluated. Then, V is calculated as the copula value corresponding to the values of X and Y. A parametric estimate of K(v) is calculated as:

$$K(v) = v - \frac{\emptyset(v)}{\emptyset'(v)} \tag{1-10}$$

The next step is to compute a nonparametric estimate by the following procedure:

$$V_i = \frac{\#\{(X_j,Y_j): X_j<X_i, Y_j<Y_i\}}{n-1} \tag{1-11}$$

where # is the cardinality of the set. The nonparametric estimate of K(v) is then calculated as:

$$K_n(v) = \sum_{i=1}^{n} \frac{\delta(v-V_i)}{n} \tag{1-12}$$

The parametric and nonparametric estimates of K(v) would then be plotted against each other in the Q-Q plots. If the copula family provides an acceptable fit to the data, the plotted points should fall along a straight line. The Q-Q plots developed for each copula family under consideration are then compared to determine which of the families are most appropriate for use.

Once a copula family has been chosen, the copula parameter(s) must be calculated. In many cases, the copula parameter should have some relationship to the

measure of dependence calculated between the marginals. Then, for each corresponding set of values from the random variables, the copula can be calculated. The copula calculation provides the joint probability of occurrence of the corresponding values of the random variables.

1.2.4. Consequences of Flood Events

While floods are naturally occurring events that provide vital ecosystem services such as improving biodiversity and sustainability of ecosystems, they are also the most taxing water-related disaster to humans, material assets, cultural, and ecological resources (Simonovic, 2009). According to a 2001 World disasters report (IFRCRCS, 2001), on average 211 million people are affected annually by natural disasters, and more than two-thirds of these people are affected by floods. Of the various possible natural disasters, floods appear to cause the most destruction, in terms of both magnitude and impact on humans (Purnell, 2002). In 2007, at least 200 major floods occurred across the globe, which killed more than 8,000 people, otherwise impacted over 180 million people, and caused over $23 billion in damages (Pitt, 2008). The increase in flood risk may be due to an increase in the flood hazard, or probability of occurring, to an increase in the amount of people and property located in the floodplain as urbanization increases, or to a combination of the two factors. This research will focus primarily on the hazard portion of the risk equation.

The 2008 floods in Iowa created a wide range of consequences. The floodwaters destroyed a number of homes and businesses, shut down city services, disrupted travel, and damaged farms and crop land. The agricultural impacts from flooding can be related to severe erosion of sediment or the deposition of sediment,

which may or may not be suitable for agricultural purposes, onto the crop lands (Mutel, 2010). Flood events can also disrupt emergency services to residents and lead to the growth of mold in homes and buildings, which may result in human health impacts. Floods often bring contaminated waters into the region, which cannot be used as potable water and which can cause a variety of human health impacts.

Damages caused by floods have been increasing in the last several decades. Between the 1950's and the 1990's annual economic damages caused by extreme weather events, including floods, increased by a factor of ten in inflation-adjusted dollars. Global climate change will affect flood patterns and damages over the coming decades. The primary cause of changes in flooding will be changes in the temporal and spatial patterns of precipitation caused by changes in atmospheric conditions. In areas where precipitation is expected to increase, the occurrence of floods is also expected to increase, while in areas where precipitation is expected to decrease, the occurrence of floods is expected to decrease. In areas where floods are typically caused by snowmelt events, as temperatures warm globally and snow decreases, floods caused by snowmelt will also decrease. While statistically significant long-term trends have not been observed globally, regional trends do seem to be present. These changes in flood occurrence will have significant impacts on the levels of damage caused by flood events in the future. Overall, it is apparent that the changes in flood frequency caused by climate change will be complex, and will be dependent on the source of flooding (Kundzewicz and Hirabayashi, 2010).

The damages caused by floods will be a function of the inundation depth and duration, the velocity of the floodwaters, and the water quality (Middelmann-

Fernandes, 2010). One method of quantifying the economic consequences of a flood event is a stage-damage curve. The economic damages corresponding to different depths of inundation can be estimated based on knowledge of land use and property type. The replacement value of the structures or property that are lost or damaged can be related to the depth of flooding through a stage-damage curve (Apel *et al.*, 2004). Stage-damage curves can also be used to estimate flood damages to crop lands. Stage-damage curves can be based on actual data from past flood events or they can be based on synthetic data (Smith, 1994).

1.3. Potential Sources of Flooding

There are a number of potential sources of flood waters, including riverine flow, coastal water, surface water flow, and groundwater. Both riverine and surface water floods are influenced by rainfall characteristics, while coastal floods are influenced by winds and barometric pressure. The impact that a flood event has will depend on the depth of water, the velocity of flow, the rate of rise of the water, and the duration of the flood event. When considering flooding from any of the sources mentioned, it is also important to consider the uncertainty associated with the event. The uncertainty could be expressed in location, timing, or intensity of the flood (Golding, 2009).

1.3.1. Riverine Flooding

The flow of water in rivers is a naturally variable quantity, with variations occurring over hourly, daily, seasonal, and yearly time scales (Poff *et al.*, 1997). Because of this variation in flows, riverine floods have always, and will always,

occur. This sort of flooding actually happens quite regularly, though it usually only draws attention when the flood event is a large disaster. Riverine flooding will occur when there is an excess of water draining from the land, from sources such as precipitation or snowmelt, than the river can contain within its banks. The antecedent moisture condition, which controls the volume and rate of surface runoff that can be infiltrated into the soil, will also influence the magnitude of a riverine flood (Mutel, 2010).

A number of factors influence riverine floods, such as watershed characteristics including size, shape, slope, and topography of the watershed, land use characteristics, precipitation event characteristics such as magnitude, duration, and intensity, snowpack, ice, particularly ice jams, the processes of erosion and sedimentation, and even the failure of a dam (ASCE, 1996). In addition to these characteristics, Poff *et al.* (1997) suggest that vegetative cover on the land, the terrain and soil texture of the watershed, and evapotranspiration rates will influence riverine flood events.

It is important to note that floods can occur on a seasonal basis, as local climate is one factor that influences the occurrence and timing of riverine floods. The floods that occurred in Iowa in 1993 and 2008 serve to illustrate this point. The winter and spring seasons prior to the occurrence of these floods were particularly wet seasons for the region. This created conditions of high antecedent soil moisture. In both years, the wet winter and spring seasons were followed by very intense rainfall over the summer months. Therefore, during the summers of 1993 and 2008

surface runoff exceeded the bankfull capacity of the rivers, which resulted in very large riverine flood events (Mutel, 2010).

As previously explained, riverine flooding is a particularly complex phenomenon. The interactions between flooding and the terrestrial system have been discussed, but interactions also occur between flood events and the socioeconomic and climates systems. The terrestrial, socioeconomic, and climate systems are the factors that control the flood hazard event and the vulnerability to flood events. The socioeconomic system is particularly related to the vulnerability to flood events and impacts how well a community can respond to and recover from a flood event (Kundzewicz *et al.*, 2010). Socioeconomic factors are a significant factor in increasing risk due to their influence on the vulnerability to flood events.

The probability of riverine floods of certain magnitudes occurring can be estimated using a flood frequency analysis. For riverine flooding, as with many hydrological applications, the log Pearson-Type III distribution is commonly chosen to model the annual maximum flood series (McCuen, 2005). To conduct a riverine flood frequency analysis a streamflow record must first be identified, and then the annual maximum flow events must be extracted from that record for further analysis. Once the annual maximum flow series has been identified, the methods described in Section 1.2.2. can be followed. If flow gages are not available, several other methods could be used to estimate the frequency of large floods. Paleohydrologic studies, which examine the debris left behind by flood events to determine the frequency with which flooding occurs, could be made. It is also possible that the historical damage

done by floods to living trees in the area could be assessed to determine the frequency of flooding (Poff *et al.*, 1997).

Because riverine flow gages are relatively rare, there must be methods for transferring flow gage measurements downstream if no gage exists at a particular location of interest. McCuen and Levy (2000) reviewed two such transposition or transfer methods: (1) Sauer's weighting function method, and (2) the drainage area-ratio method. The drainage area-ratio method is the simpler of the two methods, consisting only of the equation:

$$q_u = q_g \left(\frac{A_u}{A_g}\right)^n \tag{1-13}$$

where q_u is the ungaged discharge, q_g is the discharge measured at the gage, A_u is the ungaged drainage area, A_g is the gaged drainage area, and n is an empirical constant. Sauer's weighting-function method is more complicated. The equation is as follows:

$$q_{uw} = R_w * q_{ur} \tag{1-14}$$

where q_{uw} is the weighted estimate of the discharge at the ungaged site, q_{ur} is the discharge at the ungaged site calculated using USGS regression equations, and R is a weight, which is calculated as:

$$R_w = R - \frac{2*(R-1)*|A_g - A_u|}{A_g} \tag{1-15}$$

where A_g is the gaged drainage area and A_u is the ungaged drainage area. The value of R is calculated using:

$$R = \frac{q_{gw}}{q_{gr}} \quad (1\text{-}16)$$

where q_{gr} is the discharge estimated at the gaged location using the USGS regression equations and q_{gw} is a weighted peak flow estimate using the flood frequency estimate from the flow gage record, q_g, and the USGS regression equation estimate for the gaged site, q_{gr}. The weighted peak flow estimate can be calculated using:

$$q_{gw} = \frac{q_g N_g + q_{gr} N_r}{N_g + N_r} \quad (1\text{-}17)$$

where N_g is the gage record length and N_r is the equivalent record length of the regression equation (McCuen and Levy, 2000).

1.3.2. Coastal Flooding

A number of potential hazards affect coastal regions, including erosion, siltation, movement of sediment, and floods and high winds generated by storms. Of these, floods and wind storms are the highest hazard events (Doornkamp, 1998). Threats to coastal regions are of significant concern to the United States because 81% of the population lives in the coastal states, and those states generate 83% of the gross domestic product (GDP) (NOAA Digital Coast, 2010). According to Hoozemans and Hulsbergen (1995), approximately 200 million people lived in coastal floodplains, defined as a storm surge elevation less than the1- in-1000-year event, as of 1990. Purvis *et al.* (2008) state that 1.2 billion people worldwide currently live within 100 km of the coast and 100 m or less above sea-level.

Along coasts, a number of possible factors cause water level changes. Water levels could be changed by astronomical tides, storm surges, wave setup, wave runup,

tsunamis, or long-term changes in sea-level. Wave setup is the super elevation of mean sea-level caused by waves. Wave runup is the peak vertical elevation of waves along the face of a structure. Both wind setup and storm surge, caused by strong, long-duration winds on the shore and reduced atmospheric pressure, act as the principal components of coastal flood events. Other critical factors contributing to coastal flooding include climatic factors and topographic features. It is possible to conduct a statistical analysis of wave data similar to the flood frequency analysis process (ASCE, 1996).

McInnes *et al.* (2002) provided a thorough explanation of the processes contributing to coastal flooding. The processes that can contribute are extreme rainfall events, storm surges, and high waves. It was observed that meteorological events that result in high rainfall also typically produce severe winds over the coastal ocean that can lead to higher storm surges and waves. The increase in sea-level during a severe storm can be attributed to the addition of storm surge and breaking waves to the normal astronomical tides. Storm surge, which is the abnormal rise of water above expected astronomical tide levels caused by storm events, is one of the primary concerns from coastal flooding because it tends to be the greatest threat to life and property (NHC, 2010).

Another key characteristic is wave set-up, an increase in the still-water level caused by changes in the ocean's momentum flux due to breaking waves along the surface. This phenomenon is a function of the height and direction of the waves as they break. According to McInnes *et al.* (2002) the sea-level is a result of the nonlinear interaction of astronomical tidal levels, storm surges, and wave set-up, with

an acknowledgement that the total sea-level will be less than the sum of the individual components, due to losses from friction along the ocean floor.

Important characteristics of waves typically include wave height and period, while wave measurements are typically of surface elevation or subsurface pressure. The Gumbel distribution can be used to analyze the annual maximum significant wave heights (ASCE, 1996). The generalized extreme value (GEV) distribution, with three parameters, is also commonly used to represent coastal phenomena (Reeve *et al.*, 2004). The magnitude of coastal flooding is typically measured in terms of the elevation and the inland extent of the water (Doornkamp, 1998). In many cases, the measurement of concern is inundation depth, which occurs when normally dry land is covered with water, and can occur from both riverine and coastal flooding.

The Federal Emergency Management Agency (FEMA) conducts coastal flood frequency analyses to determine expected flood levels for the purpose of setting flood insurance rates. To conduct these analyses, FEMA uses data about storm events, including long-term tide gage records, which measure the water level, and observed high water marks from past flood events (FEMA, 2010). From tidal gages, it is possible to obtain a frequency curve of still-water elevation (SWEL), which is the elevation of the water surface without out considering the influence of waves (FEMA, 2003). The National Research Council (1977) has developed methods for accounting for the influence of waves on the water surface elevation.

Sea-level rise is a contributing factor to coastal flooding that must be considered. Increased sea-level rise appears to be occurring because of changing

climate, which may be impacted by human actions. Sea-level rise will have significant impacts on coastal wetlands and marshes, which serve as defense mechanisms against coastal flooding. According to the USGS (1997), increased sea-level rise could result in the conversion of marsh areas to open water and the conversion of upslope forested area to marsh. Sea-level rise can also be caused by the subsidence of land, which is a natural process, but which is also being accelerated due to human actions such as pumping groundwater faster than it can be recharged (Pethick, 1993). The term eustatic sea-level rise refers to sea-level rise caused by melting of ice and thermal expansion of the oceans. Local sea-level rise refers to sea-level rise caused by subsidence, while relative sea-level rise refers to a combination of eustatic and local sea-level rise (Kirschen *et al.*, 2008).

1.3.3. Surface Water (Pluvial) Flooding

Another potential source of flooding is pluvial or surface water flooding. This flooding can be caused by rainfall-generated overland flow prior to the water entering a water course or sewer. The rainfall events responsible for this flooding typically have very high intensities, but other conditions, such as melting snow over frozen or saturated ground, can cause pluvial flooding (Waterworlds, 2008). Snowmelt events can also provide water that runs off over the land surface, causing a surface water flood.

In the past few years, numerous cases of significant pluvial flooding have occurred. According to Hussain (2008), initial estimates suggest that two-thirds of the damage from floods in the United Kingdom in the summer of 2007 resulted from pluvial flooding. This flood event caused 13 deaths and flooded almost 45,000

properties (Hussain, 2008). The GeoInformation Group (2008) suggest that pluvial flooding will increase in frequency and severity in the future due to factors such as increasing urbanization and climate change. Pitt (2008) provided a thorough review of the United Kingdom flooding of 2007, acknowledging the high impact of surface water flooding. A particular concern is the complexity associated with pluvial flooding, where the capacity of the sewer and drainage system, the saturation of the ground, and the river stage all affect the occurrence and severity of flood events. The critical factors identified as impacting pluvial flooding include the volume of the rainfall, where the rainfall occurs, and the intensity of the rainfall (Pitt, 2008). Pluvial flooding was also a significant factor in the summer flooding of Washington, DC, in 2006, which seriously damaged a number of government buildings. This flooding was primarily caused by extremely high rainfall, as 7.09 inches of rainfall fell over a 24-hour period though the Potomac River was not above flood stage (Setty and Associates, 2006).

In urban areas pluvial flooding can be a particularly significant issue when intense rainfall overwhelms the capacity of the storm sewer system. The impervious surfaces in urban areas prevent infiltration of the rainfall, forcing high levels of surface runoff (Falconer *et al.*, 2009). Maksimovic *et al.* (2009) identify a number of factors that can cause urban flooding. These include the limited capacity of the drainage system, which may discharge water to the surface under extremely wet conditions, and preferential flow pathways that can create a surface flow network (Maksimovic *et al.*, 2009). Collier (2009) suggests three similar sources of urban pluvial flooding, including the inability of the natural water courses to handle

excessive rainfall-induced fluvial flooding, the inability of the urban drainage system to manage excessive rainfall-sewer discharging, and runoff over land causing local flooding in areas not typically considered at risk from natural or man-made water courses.

Precipitation is the primary factor responsible for pluvial flooding. A frequency curve can be computed for station precipitation records in much the same way they are computed for streamflow data. Typically, the cumulative precipitation depths or intensities for a specified duration are used for the analysis. The National Weather Service has frequently used the Fisher-Tippet Type I frequency distribution with Gumbel's fitting procedure to represent the precipitation data (Frederick *et al.*, 1977), but it is also possible to use lognormal, Pearson Type III, and log-Pearson Type III distributions to represent the precipitation data (USACE, 1993).

In addition to intense precipitation events, surface runoff can be caused by snowmelt water. Melting of the snowpack is an important source of water in many areas of the United States, providing groundwater recharge and replenishing surface water storage, but they can also cause floods. Water from melting snow can infiltrate into the ground, depending on soil characteristics, soil moisture content, and whether or not the ground is frozen. Melting of snow will not cause runoff until the soil storage has reached capacity, but if the ground is frozen the soil will not be able to store as much water, which will increase runoff during a storm event. Knowledge of the snow water equivalency (SWE) is necessary to conduct a snow runoff analysis. The frequency analysis procedure used for flow data and precipitation data can also be used with SWE data (USACE, 1998).

1.3.4. Groundwater Flooding

The largest source of water in the United States is groundwater. Groundwater accounts for approximately 97% of the available freshwater in the United States, and 23% of freshwater usage (USACE, 1999). Groundwater is also a potential source of flooding that is just beginning to be understood. England and Wales have only recently recognized groundwater as a potentially significant source of flooding. Groundwater flooding was not a contributing factor to the massive flooding experienced in the United Kingdom in the summer of 2007, but it has been a factor in other floods (Cobby *et al.*, 2009). Groundwater flooding has also been experienced in areas of the United States, including the states of Washington and Illinois (USGS, 2000). Though it is known that groundwater can be a source of water in rivers, such as in the Snake River, there have not been thorough studies of the interactions between surface and ground water (Johnson, 1991), which has resulted in a lack of understanding of flooding caused by groundwater.

Flooding caused by groundwater is not currently well understood (Pitt, 2008). Groundwater flooding can be defined as flooding that is caused by the emergence of water that originated from permeable strata in the subsurface. The primary cause of this flooding is unusually high groundwater levels, and it can inundate low-lying areas for months at a time. A number of consequences of groundwater flooding have been observed, including flooding of basements and even ground-level floors, buried utilities and other assets, the inundation of farmland, roads, and commercial and residential areas, and finally the overflowing of sewers due to groundwater ingress

(Cobby *et al.*, 2009). Pitt (2008) has also suggested that groundwater flooding can damage foundations of buildings by reducing their load-bearing capacity.

Groundwater flooding has been a problem in Puget Sound, in Washington State. Flooding typically occurs in this region after several consecutive years of higher than average precipitation. With enough precipitation, the water table rises high enough to intersect the ground surface and inundate low-lying areas. Groundwater flooding can also occur when water collects in low-lying areas of low permeability. These flood events are further complicated by the fact that they can take several months to recede. For example, in 1997 a small groundwater-fed lake expanded to an area of approximately 25 acres and remained flooded for months. A great deal of data on groundwater flooding does not currently exist; the best source of information about the occurrence and frequency of groundwater flood events remains anecdotal reports from local residents. In addition to multiple years of high rainfall, cool, wet springs followed by mild summers provide conditions favorable to groundwater flooding. The precipitation provides recharge to the groundwater system, leading to high groundwater levels, while the cool temperatures reduce losses to evapotranspiration, assisting in maintaining the high groundwater levels until groundwater storage is exceeded and flooding occurs (USGS, 2000).

1.4. Potential Causal Factors of Flooding

There are a number of potential causal factors that could contribute to flooding from each of the above-mentioned sources of flooding. These include convective rainfall events, cyclonic storm events, ice jams, snowmelt, and rainfall on

a snowpack. Coastal flooding can also be caused by intense storm events, such as cyclonic events, or tidal cycles.

1.4.1. Convective Rainfall Events

Convective rainfall events are a common cause of riverine floods, due to the production of runoff (Poff, 1997). When rainfall reaches the land surface some of it will infiltrate into the ground, while the rest will become surface runoff. The surface runoff will travel over the land surface, based on the local topography, until it reaches a stream or river. If the amount of surface runoff reaching the rivers and streams overwhelms their capacity, riverine flooding will occur. The amount of surface runoff produced by a given storm will be determined primarily by the antecedent moisture conditions, which are primarily a function of recent rain or snow storms and recent temperatures. The amount of urbanization in the watershed will also influence the amount of surface runoff generated, as this is related to the amount of impervious surface. The intensity of the rainfall will also influence how much of it can be absorbed by the ground and how much will become surface runoff (USACE, 1993). Excessive amounts of surface runoff reaching the river can result in a riverine flood event. In addition to the antecedent moisture conditions, the size and direction of the rainfall event also influence the amount of surface runoff reaching the river or stream. Intense summer rainfall events following a particularly wet winter and spring, resulting in high antecedent moisture conditions, were one factor responsible for the 2008 Iowa floods (Mutel, 2010).

1.4.2. Cyclonic Storm Events

Cyclonic storm events, such as hurricanes and other storms associated with frontal systems, have several mechanisms for causing flooding; however, very intense rainfall is the primary cause of riverine flooding. Cyclonic storms have different mechanisms for causing flooding in coastal regions and in rivers. As cyclonic storms travel over land they result in very intense rainfall over large areas. This rainfall overwhelms the capacity of the ground to infiltrate the water, thus producing significant amounts of surface runoff. As with other rain storms, the surface runoff then travels over the land surface towards streams and rivers. The capacity of these streams and rivers are overwhelmed by the addition of the surface runoff, resulting in riverine flooding. Tropical cyclones can also cause flooding in coastal regions, though in these areas the flooding is typically caused by high storm surges pushing waves onto the shore. Intense storm events such as hurricanes produce storm surge that can cause flooding in coastal areas.

1.4.3. Ice Jam Flooding

In many regions of the world with colder climates, riverine flooding can also be caused by the presence of ice in the river. Flooding can be caused by ice jams, which are accumulations of ice fragments within the river banks that restrict flow. The presence of ice can block flow through a jam, or it can cause added resistance to flow, by approximately doubling the wetted perimeter of the stream and reducing the area available for flow. Such flooding is typically very site-specific (USACE, 1991). Floods caused by ice jams have occurred in at least 36 states, primarily northern states, though mountainous areas of southern states such as Arizona and New Mexico

have also been impacted, and they are estimated to be responsible for at least $100 million in annual damages (USACE, 2002). In regions with colder climates, where ice forms on rivers, milder weather as the spring season starts leads to the break-up of sheets of ice into smaller blocks. These blocks move downstream until they encounter stationary ice cover, or potentially some constriction in the river channel, where they pile up and cause a jam (Beltaos, 2003).

Because ice-influenced flood events are caused by jams or by increased resistance to flow, the floods typically result in much higher stages than would be produced by a comparable open-water discharge. In fact, ice-influenced flooding typically occurs at very low discharge values. Because of this, ice-influenced flooding must be analyzed based on stage frequency, rather than discharge frequency (USACE, 1991). However, developing a stage-frequency for rivers where flooding is influenced by ice is not typically a simple matter. Where gages do exist to monitor stage, the ice may damage or interfere with the gage. (White *et al.*, 2000). Also, the typical stage-discharge rating curves are not applicable when ice influences flooding (Tuthill *et al.*, 1996).

To conduct a frequency analysis of riverine flooding during the winter season a mixed population approach must be taken because the annual maximum flood may be caused by ice or it may be caused by open-water flooding in any given year. If the data are not separated by causal factors, a mixed population analysis should be followed. If the data are separated into subsets based on causal factors, a combined population approach should be followed. The assumption, when using the combined population approach, that the frequency curves used to develop the combined curve

are independent is a valid assumption when considering open-water or ice-influenced flood events. It is also possible that an ice-influenced flood event will not occur every year, such that zero-flood year analyses will be needed (USACE, 1991). The procedure for conducting a flood frequency analysis for a river where ice-influenced flooding may occur involves two steps. First, flood frequency analyses are conducted for winter season flooding, for both ice jam and no ice jam events. These two frequency curves are combined into one winter season flood frequency curve. Then the winter season frequency curve is combined with the open water season flood frequency curve, to provide an annual stage-frequency analysis for the river (Tuthill *et al.*, 1996).

Bulletin 17B recommends the use of the log-Pearson Type III distribution with a weighted skew coefficient to model riverine flood events. However, as pointed out by Morris (1982), these recommendations were not made for data sets separated by causal factors. Morris (1982) suggests instead using the log-normal distribution to model ice-influenced flood events. It is not inherently clear what distribution ice-influenced flooding should follow, so the log-normal distribution is suggested because it is simple and because it has been used previously (USACE, 1991). The log-normal distribution was also suggested by Tuthill *et al.* (1996).

1.4.4. Snowmelt Flooding

In northern regions, snowmelt runoff serves as a major source of floodwaters. In the western United States, most rivers are supplied by snowmelt from higher elevation watersheds (Gottfried, 2003). Kattelmann (1990) provides a thorough discussion on the various mechanisms that can generate floods in the Sierra Nevada

region in the United States. In this region, snow is observed to accumulate in a snowpack for approximately six to seven months of the year, and to melt and cause runoff over three to four months. Snowmelt floods occur annually during the spring season, with sustained high flows, long durations, and large volumes. According to Kattelmann (19991), these snowmelt floods can become hazardous not because of their peak discharges, but because of their durations and volumes. Snowmelt floods can be characterized by high water levels for a period of several weeks.

Though snowmelt floods can be quite significant events, they are not typically responsible for the highest instantaneous peaks observed. In fact, rainfall events occurring on a snowpack have been responsible for the highest flood flows in the Sierra Nevada region over the last century (Kattelmann, 1990). Rain-on-snow events tend to produce much greater peak flows in maritime snow climates than do either rainfall or snowmelt events alone. In an examination of streamflow records for California, Kattelmann *et al.* (1991) observed that the largest floods in the record were caused by rain-on-snow events in 17 of 18 rivers included in the study.

1.5. Goal and Objectives of Research

Currently, methods for calculating both the probability of flooding and the flood risk from riverine and coastal sources are available. Methods for calculating the probability and risk from flood sources that are less common and not as easily understood, such as pluvial or groundwater, are not as well established. Established methods for calculating the probability of flooding that occurs from two or more sources at the same time are also available. However, a fully tested method of

estimating the probability that a given location will flood, given that the flood event may be caused by multiple potential flood sources or some combination of those flood sources is not available. Thus, it is not possible to provide a comprehensive flood risk assessment for such locations.

Providing separate probability estimates of a location experiencing a riverine flood and a coastal flood event does not provide the true likelihood of flooding at the location, it only provides partial information, which can lead to inaccurate estimates of risk. To fully understand the flood risk at a location, the likelihood of flooding from all potential sources must be accounted for and understood.

The goal of this research is to develop a method of combining flood probability curves from multiple sources (riverine, coastal, pluvial, and/or groundwater) to produce a comprehensive flood frequency curve for a given location and show how the calculation of flood frequency or probability impacts flood risk assessment. This method will provide the true probability of a flood of a given magnitude, or a given inundation level, occurring at a specific location of interest, taking into account that the flood may be caused by one or more potential sources. This asks the question "What is the probability of experiencing a flood inundation of three feet at a critical location, regardless of the source of the flood water?", as opposed to the question asked by a more traditional flood frequency analysis, which is "What is the probability that a riverine flood will cause an inundation of three feet at a critical location?" This analysis will shift the focus of the flood frequency study from the source of the flooding, where the stage or discharge of the river might be measured, to the level of flooding at the location of interest, where the inundation

experienced because of the flood event will be measured. By providing the probability of a certain level of inundation at a given location, regardless of the source or cause of the flooding, a comprehensive flood risk assessment can be conducted, based on information about the consequences associated with floods of different magnitudes and probabilities of occurrence. This comprehensive assessment of flood probability and flood risk will allow for more informed decision-making. This approach will be contrasted with the assumption that flood risk studies based on individual flood source probability curves give acceptable estimates of risk.

The work that will be done to meet this goal will be exploratory in nature. A methodology will be suggested based on tests for one location of interest. General conclusions as to the success of the methodology can be drawn from this small test, but definitive conclusions about the usefulness should not be drawn without further study and testing under different conditions. The factors to consider, as well as the results of the procedure, will likely differ for each location of interest. The results of the study conducted for this research will only provide an understanding of the potential of the methodology, and should only be used to determine whether further study could be warranted.

1.5.1. Assessment of Multiple Populations

In the process of conducting the research necessary to fulfill this goal, several underlying objectives will also be considered. The first objective is to assess the current methods for accounting for multiple populations in flood frequency analyses, and to develop an updated method if necessary. Multiple populations are different causal factors associated with a single source of flooding. For example, ice-jams and

rainfall events can each individually cause riverine flooding but they are separate populations and the floods they cause may have different characteristics. The most common method of accounting for multiple populations was described by Morris (1982). This method separates the annual flood frequency record by causal factors and develops frequency curves for each individual causal factor. Principles of probability are then used to combine the individual frequency curves into a single frequency curve for the source of flooding. This method requires an assumption that each population under consideration be independent, which may not be a reasonable assumption. Whether or not independence may be assumed will be determined, and an adjusted method will be developed for scenarios in which the assumption of independence should not be made. The length of record available at the location of interest will determine the ease with which multiple populations can be considered, as it will require that the record be separated into different records corresponding to each population.

When considering multiple populations, the concept of zero-flood years may need to be considered, as some hydrometeorological causal factors may not lead to flooding every year. Examples of these include hurricanes or ice jam floods. A record of annual maximum floods caused by hurricanes will have zero-values for all years in which a hurricane does not occur in the area of interest. While there are accepted methods for accounting for zero-flood years in flood frequency analyses, these methods typically do not perform well when a large percentage of the flood record consists of zero flood-years. Therefore, an alternative method of accounting for zero flood-years may be needed.

1.5.2. Assessment of the Assumption of Independence or Dependence

A second objective to consider is an evaluation of the impact of the assumption of independence of events that is commonly used in flood frequency analysis. The flood frequency analysis procedure outlined in Bulletin 17B (IACWD, 1982) assumes that the annual maximum peak floods can be considered as a sample of random, independent, and identically distributed events. However, there has been some debate as to whether or not they can be considered to be independent events. Olsen *et al.* (1999) documented some variation in flood risk over time that resulted in statistically significant trends, which they claimed challenged the assumption that flood events were a series of independent and identically distributed events.

The assumption of independence is particularly important when multiple sources or causal factors could cause flooding. The method of combined-population analysis developed by Morris (1982) requires an assumption of independence. In northern areas, where ice jams or snowmelt floods combine with rainfall-generated floods, this assumption may be reasonable. In Tuthill *et al.* (1996) the assumption was considered to be reasonable when analyzing ice jam floods in a northern region. However, in more southern regions, such as Sierra Nevada, California, studied extensively by Kattelmann (1990; 1991), rainfall may occur while a snowpack still exists, leading to larger floods than would occur from either rainfall or snowmelt alone. These events cannot be considered as independent events. Payton and Brenedecke (1985) demonstrated this by examining the correlation between rainfall and snowmelt events and finding a significant negative correlation between the two, which suggested that they are not independent events.

When considering whether or not the assumption of independence is valid, storm and watershed characteristics may also be a factor. Olsen *et al.* (1999) also noted that in regions where rainfall and snowmelt events could cause floods, rainfall floods would frequently follow snowmelt floods because of increased soil moisture and baseflow from the snowmelt event. Interdependencies of rainfall and watershed characteristics during a storm event may be responsible for causing a riverine flood and a surface water flood at the location of interest. In this case, the flood events certainly would not be independent.

1.5.3. Assessment of Uncertainty

A third objective to consider is an attempt to quantify the various sources of uncertainty that could be associated with each of the potential sources of flooding. The various sources of uncertainty associated with flood frequency analyses have long been understood to be a problem. However, these uncertainties are difficult to quantify, and their effects on flood risk assessments are difficult to evaluate. When considering a flood frequency analysis based on many different flood sources, it will be important to understand the various sources of uncertainty inherent in the analysis, and how these uncertainties will impact the flood frequency analysis.

Many sources of uncertainty enter into flood risk assessments. For instance, the exact location of a flood event cannot be specified with certainty, nor can the time or intensity of the event (Golding, 2009). Uncertainty in data influence how the inundation of the surrounding land will occur. This can be predicted based on knowledge of the topography, but there will always be some amount of uncertainty here. In addition to the uncertainty associated with the physical processes causing

flooding, uncertainty exists in the result of hydrologic and hydraulic models used to predict flooding. Understanding the uncertainties in both the physical processes and the modeling of floods will prove valuable in improving the prediction of flood events. By quantifying the uncertainties that may be associated with a flood frequency analysis and determining which of these uncertainties are most significant, an understanding can be developed of where to focus future research efforts to reduce these uncertainties and improve the accuracy of flood predictions. As this research will focus on developing a frequency curve based on multiple sources of flooding, understanding the various sources of uncertainty associated with each will be especially important. For each flood source, the uncertainties will be identified and generally categorized. The uncertainty associated with each category will be quantified for each flood source, and illustrated using box plots or confidence intervals.

1.5.4. Flood Risk Assessment

The final objective is to use the comprehensive flood frequency analysis method to improve flood risk assessment by providing a more detailed understanding of the flood hazard. An improved understanding of the flood hazard will result in an improved understanding of the flood risk. The hazard in this case is the level of water inundating the location of interest, which could come from a number of potential sources.

The vulnerability of the system to flooding will be determined after the location of interest is identified. The system will include any measures to reduce flood inundation depth. Whatever the system may be that is identified as protecting

the location of interest, there is some probability that it will fail. This will be identified as part of the vulnerability term of the risk equation (probability multiplied by vulnerability multiplied by consequences).

To complete the flood risk calculations, consequences of the flood event must be quantified. There are many ways in which consequences of a flood event could be categorized and quantified. For instance, there may be economic consequences, such as damages to structures and infrastructure, environmental damages, or loss of life or injuries as a result of the flood event. For the purpose of providing an example of the method, economic consequences will be considered because they can be easily calculated and understood by using depth-damage curves. The probability of flooding will be multiplied by the vulnerability and expected consequences in order to estimate the flood risk.

1.5.5. Benefits of Research

A number of benefits will come from the successful completion of this research. The proposed research would provide a complete methodology for estimating the likelihood of flooding and the flood risk for a given location. Many reasons why a complete understanding of the likelihood of flooding at a location would be beneficial can be cited. This understanding could lead to better decisions about land-use and development, including locating critical facilities, emergency services, and infrastructure. Having accurate knowledge of flood risk in an area would ensure that these critical facilities, services, and infrastructure were placed in areas where they were truly at the least risk. New land developments projects could also be prevented in places where flood risk is significant. If the decision is made to

place new developments or critical facilities in areas that have high flood risk, at least appropriate precautions could be taken to reduce the consequences of expected flood events. Full information about flood risk could also assist citizens in making informed decisions about the need for and purchase of flood insurance, or use of other methods to reduce their flood risk. With a full understanding of the flood risk of their location, many citizens may realize that even though they are not at high risk for riverine or coastal flooding, they are still at risk from other sources, and as such, should consider purchasing flood insurance to protect their homes and possessions.

It is also possible that the location of interest could be an area being considered for a restoration project. If, for instance, a constructed wetland were being considered, it would be helpful to know that the location had a high probability of flooding, to assist in the creation of hydric soils and wetland conditions. In creating these projects, it may not matter so much where the water is coming from so long as the water is there. Surface water or groundwater flooding may be just as likely to lead to successful wetland creation as riverine flooding, so a comprehensive flood probability analysis may suggest new areas where projects could be successful. These are just a few examples of scenarios in which a complete understanding of the probability of flooding and flood risk in a given area would be beneficial to society.

1.5.6. Potential Challenges to be Addressed

In the process of developing the objectives of this research, several potential challenges presented themselves. It was apparent that these potential challenges may need to be considered in the course of conducting this research. The first of these is the possibility of zero-flood years, which are years in which an occurrence of the

event of interest, e.g., snow-generated flood, did not occur. Such lack of occurrence can cause problems in calculating the probability of flooding when certain probability distributions are used. The second is multiple populations, which are when flooding can occur from multiple causal factors, such as rainfall and snowmelt. The assumption of independence of events is the third issue that will need to be addressed. It is possible for the same storm event to cause flooding from more than one of the potential causes. Fourth, different pdfs are frequently used to model the different sources of flooding. There may be difficulties associated with combining frequency curves that are based on different distributions. Finally, the various sources of uncertainty in these analyses, and their impact on the final calculation of flood probability and risk, must be considered.

CHAPTER 2

LITERATURE REVIEW

2.1. Flood Risk

Flood risk is the primary concept behind this research project. The research goal is to modify the procedure for determining the probability of a flood event occurring, in order to better assess the risk of flooding at a given location. An understanding of the recent research related to concepts of flood risk was, therefore, required.

2.1.1. Risk-Based Design and Operation

Plate (2002) defines risk management as a process that involves the following three different sets of actions, depending on the operators involved: actions necessary to operate an existing system, planning for a new or revised system, especially when the old system ceases to be acceptable, and the process of obtaining the best design for and constructing a new project. Plate applies these principles to flood risk management systems. When considering risk management for existing flood protection systems, Plate considers the following four steps: risk analysis provides the basis for long-term operations and management actions, continuous maintenance and improvement of the system, preparation for the possibility that the system could fail, and providing disaster relief in the event that the system does fail. When considering options for modern flood risk management, according to Plate, the availability of technology, financial resources, and the perception of the urgency of the need for

protection must all be considered; one solution will not serve in all situations. Important goals to consider when planning a flood risk management system include objectives for safety and the preservation of natural systems. Overall, according to Plate, it is not unusual to find, in planning a flood risk management system that no one solution actually meets all of the objectives identified in the planning process.

Jonkman *et al.* (2009) applied the risk-based design principles that have been used by the Dutch for flood protection since the 1950's to New Orleans, Louisiana. They present a perspective for long-term flood risk reduction in the coastal Louisiana area that could also strengthen the natural ecosystem functions of the Mississippi Delta, primarily focusing on hurricane protection. Several strategies were identified by the study team, but the preferred strategy presented included a combination of strengthened levees around the New Orleans metropolitan area and wetland stabilization measures. The method of economic optimization, which balances incremental investments in higher levels of safety with the risk reduction potential, was utilized. This approach was used because it can account for local economic factors and damage levels. The results of this study indicate that the optimal level of safety for different protected areas in South-East Louisiana can be determined using the economic optimization method described.

Buchele *et al.* (2006) considered methods to improve flood risk assessment through improved methodologies and improved data for quantification and mapping of risks. The studies conducted included the estimation of extreme events that exceed the design floods used in designing and constructing flood protection measures, assessing the flood hazard and risk over the entire spectrum of potential damaging

flood events, and estimating damage by considering various building- and event-specific influences on damage levels. Specifically, Buchele *et al.* discussed a regionalization method for developing state-wide flood probabilities in Baden-Wurttemberg, Germany. Their results indicated that this regionalization approach was capable of reproducing the shape of the statistical distribution of the flood event. However, confidence intervals indicated that there is substantial uncertainty, especially in areas of extrapolation. They also discussed flood damage estimation using damage functions, which relate monetary damages to the depth of inundation caused by the flood or the type or use of the building. They suggested that flood damage will not only depend on the water depth and the type of building, but will also depend on other factors such as flow velocity, duration of inundation, sediment concentration, flood warning systems, and the quality of external response to the flood event, but as single comprehensive approach for considering these other factors in estimating losses is not possible. In order to improve the ability of hazard mapping to cover the entire spectrum of flood events, hydraulic simulation was coupled to a GIS tool for flood damage assessment, based on established stage-damage functions. A methodology to reduce uncertainty in flood damage estimates when it was not possible to conduct an individual site estimate was also developed that considered a number of damage influencing factors other than water depth.

2.1.2. Bulletin 17B

One of the most important guidance documents for flood calculations currently is "Guidelines for Determining Flood Flow Frequency: Bulletin #17B of the Hydrology Subcommittee", also known as Bulletin 17B, developed in 1982. This

document is an update of previous documents that describes all of the major elements in the process of defining the flood potential at a given location, in terms of both peak discharge and exceedance probability. Bulletin 17B was developed for the annual flood peak discharge calculations in particular. Bulletin 17B contains revised procedures for weighting a station skew value with results from a generalized skew study, the detection and treatment of outliers, making comparisons between two stations, computing confidence limits for a frequency curve. Bulletin 17B acknowledges that risk and uncertainty are inherent in flood frequency analyses (IACWD, 1982).

Several important techniques were developed in Bulletin 17B. First, the log-Pearson Type III distribution, a Pearson Type III distribution with a log transformation of the data, was suggested as the base method for analyzing annual series of data using the generalized skew coefficient. A procedure for fitting the log-Pearson Type III distribution to the annual peaks was also suggested, and enabled the logarithms of discharge values at selected exceedance probabilities to be computed. The problem of zero-flood years, acknowledged to be a problem in arid and semi-arid environments, was also addressed. The method of accounting for zero-flood years includes conditional probability adjustments for determining the frequency curves. The concept of mixed populations was also considered. The bulletin points out that where flooding could be caused by mixed populations, the flood record may not necessarily be homogenous, and therefore may require special treatment during the analysis process. It is suggested that if two or more distinct and independent causes

of floods can be identified, it may be best to segregate the flood record by cause, analyze each cause separately, and then combine the data sets again (IACWD, 1982).

Stedinger and Griffis (2008) undertook to evaluate Bulletin 17B to determine whether or not an update was needed. They state that flooding causes an average of 140 deaths per year in the United States and cost approximately $6 billion per year, excluding the cost of flooding associated with Hurricane Katrina. In the opinion of Stedinger and Griffis, Bulletin 17B is a good document, but there are long-standing problems that were identified that need to be addressed. In fact, there have been recent advances into many of these problems, suggesting that it is time to update the Bulletin to maintain the statistical credibility of the guidelines and to provide accurate risk and uncertainty assessments. One significant problem identified is that the skew map provided in the document was first published 30 years ago and has not been updated since. Bulletin 17B includes a list of issues within the document that are recommended for additional study. This list includes flood frequency distribution selection and fitting procedures, identification and treatment of mixed distributions, identification and treatment of outliers, treatment of historical information, confidence limits for the Pearson Type III distribution, use of precipitation measurements in estimates of flood potential, estimation of flood potential at ungaged sites or in watersheds with limited records, and estimation of flood potential for watersheds that have undergone urbanization or have reservoirs. According to Stedinger and Griffis there has been progress relating to at least half of the items on that list. They claim that further research is still needed to update low outliers and historical flood procedures, plotting positions, and confidence interval computations.

Also, the regional skew map should be updated to account for the additional 30 years of data that are now available. Stedinger and Griffis also suggested that the flood management community in the United States should adopt the Expected Moments Algorithm (EMA), which provides a direct fit of the log-Pearson Type III distribution using the entire data set.

Griffis and Stedinger (2007a) provided a thorough review of the history of Federal guidance into flood frequency analysis, focusing primarily on the evolution of Bulletins 17 through 17B. The paper explains the original political motivations behind Bulletin 17, the changes that occurred between Bulletin 17, Bulletin 17A, and Bulletin 17B, and provides a review of recent research that might be considered in revising Bulletin 17B. The conclusion of this review is that although Bulletin 17B is a wonderful document that has provided strong guidance for 25 years, it is now time to update the guidance provided within it. Specific areas that need updating include outlier and historical flood procedures, plotting positions, confidence intervals, and regional skew estimators, to provide methods that are more efficient and consistent.

2.1.3. Assessing Probabilities of Flood Events

According to Apipattanavis *et al.* (2010) the traditional parametric methods used in design flood computations assume that the annual maximum floods in the data set are independent of each other, identically distributed, and also drawn from some homogeneous population with a known probability distribution function (pdf). The pdfs most often used are log-Pearson Type III, log-normal, and extreme value Type I distributions. While statistical tests do exist to discriminate between these distributions, it is often difficult to distinguish between the candidate models for a

given data set. Also, the best-fit criteria used to evaluate the distributions tend to emphasize the fit of the distribution as a whole, rather than emphasizing the fit in the tails. Overall, there remains considerable uncertainty as to which model provides the best estimation of the upper flood quantiles. One possible explanation for the overall poor fit of single population distributions to flood data is that floods are often caused by two or more distributions, as opposed to the single distribution assumed in the traditional method. Non-parametric methods; however, do not assume any distributional form for the data. This study evaluates a higher-order non-parametric estimation scheme, called local polynomial regression (LPR), which has improved upon the existing kernel quantile estimator. Significant detail is given on the LPR based estimator and its ability to reduce boundary problems, or biases in the tails, seen in the kernel-based method. The LPR quantile estimator is tested on a number of synthetic heterogeneous data sets composed of mixtures of the conventional distribution populations. The estimator is then tested on four streamflow data sets that were identified as exhibiting mixed population characteristics. For all of the parent distributions, the LPR estimator was observed to provide good performance, though some bias on larger return periods was identified. None of the traditional parametric and homogeneous distributions were observed to perform so well on the parent distributions, possibly due to their inability to recognize the mixed populations in the data.

Griffis and Stedinger (2007b) conducted a very thorough examination of the log Pearson Type III distribution and its appropriateness for use in flood frequency analyses. The purpose of this study was to evaluate the method-of-moments

estimator in log space relative to a number of other quantile estimators for the distribution. A Monte Carlo analysis was used to compare the method-of-moments estimator in log space and real space with maximum likelihood estimators and a method-of-mixed-moments estimation. The most commonly used estimators with the log Pearson Type III distribution are the method-of-moments in log space, the maximum likelihood, and the mixed moment estimators, though the method-of-moments estimator in real space is also typically reasonable. The method of moments estimator recommended by Bulletin 17B was observed to perform as well as the other methods for the range of parameters of interest. Of the methods not using regional skew information, the method of mixed moments was observed to perform best and was found to be comparable to the method recommended by Bulletin 17B. Overall, the log space method-of-moments estimator was found to perform better than the maximum likelihood estimator and was concluded to be a robust method that performed well when regional skew information was available.

2.1.4. Joint Probability Assessment

Joint probability has been thoroughly investigated in the United Kingdom. Samuels (2002) describes work done to better assess the risk of flooding along the shore of Cardiff Bay, after the Cardiff Bay Barrage was installed. This analysis was not done thoroughly prior to the installation of the barrage, as the computing power at the time was not yet advanced enough. The objective of this research was to assess the actual flood risk to the inland bay with the barrage in operation. In conducting these analyses, a formal joint probability analysis of water levels in the inland bay caused by the interaction of tides, river flows, and the operation of sluices on the

barrage, was used to determine the level of protection provided by the barrage. Samuels first developed cumulative exceedance probability functions for both the tide and the river flows and volumes, using observations and previous analyses. Then the JOIN-SEA software was used to establish the degree of correlation between the flood flows and tidal levels. From this study Samuels was able to conclude that the barrage had substantially reduced flood risk in Cardiff Bay (Samuels, 2002).

Hawkes (2008) has also conducted a number of studies focusing on joint probability. In Hawkes (2008) he suggests that the purpose of a joint probability analysis is to estimate the likelihood that two or more relevant source variables will take high values at the same time, which could lead to flooding. To conduct a joint probability assessment of extreme values it is necessary to know the distribution of each variable, the extreme values of each variable, and the dependence between each variable pair. Hawkes discusses three methods that are commonly used to present the results of a joint probability exceedance assessment, the joint probability density, the structure function, and joint exceedance extremes. He considers all three methods to have some value, and based on his test cases, believes that all three methods can give reasonable results if they are applied and interpreted correctly (Hawkes, 2008).

Yue *et al.* (2001) suggested that assessment using single-variable hydrologic frequency analysis could provide only a limited assessment of complex hydrological events. They believe that complex hydrological events such as storms and floods appear to be multivariate events that can be characterized by some number of correlated random variables. To fully understand the multivariate hydrological events requires the study of the joint probabilistic behavior of the identified correlated

variables being used to characterize the event. Though the gamma distribution is generally used in practice, this study evaluated the bivariate gamma or bigamma model that could be constructed from specified gamma marginals. A number of different bivariate gamma distribution models, including the Izawa bigamma model, the Moran model, the Smith-Adelfang-Tubbs (SAT) model, and the Farlie-Gumbel-Morgenstern (FGM) model, were considered. It was determined that the Izawa, Moran, and SAT models could be used to represent the joint probability distribution of two correlated random variables with different gamma marginals, as required.

Yue *et al.* (1999) did very similar work with mixed population flood frequency data and the Gumbel distribution. This study examined using the Gumbel mixed model, the bivariate extreme value distribution model with Gumbel marginals to represent the joint probability of flood peaks and volumes and the joint probability distributions of flood volumes and durations. The model was verified using data from the Ashaupmushuan basin in Quebec, Canada. It was determined that using this model, if the marginal distributions of the two random variables were able to be approximated by the Gumbel distribution, the joint probability distributions, conditional distributions, and associated return periods, could be readily determined (Yue *et al.*, 1999).

2.1.4.1. Box-Cox Transformations to Normal Distributions to Determine Joint Probability Distributions

Yue (2000) applied the ideas of joint probability analysis to annual maximum storm peaks. Storm severity is a function of both the storm peak intensity and the total amount of rainfall. The focus of this research was to use the bivariate normal

distribution to represent the joint probability distribution of storm peaks and amounts. Yue's procedure involved transforming the sample storm peak data to near the normal distribution using various Box-Cox transformations and representing the joint distribution with the bivariate normal distribution. The model was verified using data from two meteorological stations in different climatic regions of Japan. Through this method, Yue claimed that the joint cumulative probability distribution function, the conditional cumulative distribution function, and the associated return periods could easily be determined. This method was found to be able to provide information which would not be obtained by traditional single variable storm frequency analysis (Yue, 2000).

Loganathan *et al.* (1987) developed a methodology utilizing the Box-Cox transformation to determine the joint probability distribution for stream flows and tides in a tidal estuary. The Box-Cox transformation was used to transform the original stream flow and tidal measurements to approximately normal distributions. From this point, the joint normal distribution could be calculated as the joint distribution. This allowed exceedance probabilities to be calculated for various combinations of stream flow and tidal values. In laying out their approach, the authors also did not assume that there would be statistical independence between the random various under consideration. The methods were applied to the Rappahannock River, which drains into the Chesapeake Bay. Results indicated that the Box-Cox transformation gave satisfactory results for the joint probability distribution and indicated that the assumption of statistical independence may not result in conservative design calculations.

2.1.4.2. Use of Copulas to Determine Joint Probability Density Function

The literature available of the use of copulas in hydrologic applications is rapidly increasing. Though copulas first raised interest in the financial world, their applicability to hydrologic applications has long been recognized, and has increasingly studied in recent years.

2.1.4.2.1. Explanation of Copulas

Genest and MacKay (1986) provided an early view of copulas for mathematical and statistical use. This paper provides an overall introduction to the subject and use of copulas, with a focus on symmetric copulas classes and copulas with singular components. A discussion on the relationship between copulas and Kendall's τ, one of several measures of dependence that could be used with copulas, was also provided.

Genest and Favre (2007) have provided a much more detailed discussion of copulas and their use in hydrologic applications. A thorough overview and explanation of copulas is presented, along with a simple example problem to further illustrate the concepts explored. Copulas techniques are applied to a set of annual maximum streamflow and corresponding volume in the Harricana watershed in Quebec.

Salvadori and De Michele (2007) illustrate the importance of being able to model the joint dependence between various random variables that describe hydrologic events. They examined several of the more recent uses of copulas in hydrological applications. In particular, they discussed the calculation of conditional

probabilities, level curves of joint distributions, return periods of bivariate events, the concept of the secondary return period, trivariate models for storm events, calculation of storm depth, and calculation of convolution variance. The purpose of this study was to review the recent advances and to provide a more thorough understanding of the applicability of copulas to these scenarios.

2.1.4.2.2. Example Uses of Copulas

There have been many illustrations of copulas applied to hydrologic problems in the literature. De Michele *et al.* (2005) address the design of dam spillways using copulas. Both flood peak and flood volume were identified as important characteristics to consider in designing dam spillways, and the dependence between the factors was recognized. Thus, a bivariate probability distribution was developed using 2-Copulas for use in spillway design. The adequacy of this method was demonstrated with a synthetic data series.

De Michele and Salvadori (2003) considered the use of copulas in modeling rainfall events. Characteristics used to model these rainfall events are typically storm duration and average intensity, which have been shown to be dependent. This study suggested the use of heavy tailed Pareto-like distributions, rather than the typical exponential distributions, for modeling of the marginal distributions, and the use of a 2-Copula to model the dependence between variables. The 2-Copula was found suitable to model the marginal variability of the storm characteristics as well as the joint variability. The 2-Copula approach proposed was illustrated with a theoretical case study.

Klein *et al.* (2010) also evaluated the use of copulas for hydrologic design, by conducting a risk analysis of a flood control structure. They defined a flood event based on multiple characteristics, including peak flow, volume, shape, and duration and suggested that a multivariate analysis based on those variables would be necessary to obtain the probability of a given flood event. Copulas were considered to avoid the many problems that have been associated with the use of multivariate distributions for dependence modeling. An example based on the Unstrut River in Germany was provided. The results of this study indicated that it was quite beneficial to consider flood volume in addition to flood peak, as is currently typically done, in risk-based planning and design applications.

Favre *et al.* (2004) also suggest the use of copulas to avoid the drawbacks commonly associated with multivariate distributions. Two hydrologic applications of copulas are presented. The first focused on combined risk in a frequency analysis, with four different copulas evaluated to model peak flows in a Quebec watershed. The second considered the joint modeling of peak flows and volumes on a river in Quebec.

Wang *et al.* (2009) applied copulas to develop joint probability distributions for coincident flows. Their interest was in the protection of critical infrastructure located near the confluence of streams. The bivariate distribution of design flows as the confluence location was developed using a proposed Copula-based Flood Frequency (COFF) method, modeling the marginal distributions based on Archimedean copulas. The results of the COFF method were compared to the commonly used National Flood Frequency Program method of regional flood

frequency analysis. A case study in the Des Moines River in Iowa and Monte Carlos simulations were used to demonstrate the success of the proposed method.

Kao and Govindaraju (2010) assessed the ability of copulas to model the characteristics of droughts. The proposed method, based on the standardized index algorithm, was designed to account for seasonality in the marginals of the streamflow and precipitation. The dependence structure was characterized with multiple time windows, ranging from one to twelve months. A joint deficit index was developed based on the copulas, which could give a description of the drought conditions based on probabilistic analyses. The index was successful in predicting droughts, and could be used with hydrologic variables as well.

A unique application of the copula method by Wang *et al.* (2012) used copulas to assess the probability of eutrophication in water bodies. This analysis focused on the relationship between the Chlorophyll a concentration and various environmental factors such as dissolved oxygen, total nitrogen, or total phosphorous. Archimedean copulas were used to assess the relationship between Chlorophyll a and various environmental factors in Wulihu Lake, a eutrophic lake in China. For several environmental factors, conditional joint return periods are calculated, which provide information about the likelihood of algal blooms based on the concentrations of the environmental factors such as nitrogen and phosphorous. The method presented in this study provides an effective tool for managing and treating eutrophic lakes based on an understanding of the interactions between the variables causing eutrophication.

Singh and Zhang (2007) used copulas to develop intensity-duration-frequency (IDF) curves, which can be considered as joint frequency distributions of rainfall intensity and duration. Typically, IDF curves are derived using univariate methods; however, this has been because of the past difficulties in bivariate analyses. The ability of the copula to derive IDF curves through a bivariate method was investigated in this study. The study made use of the Frank copula family with rainfall depth and duration as the marginals. Hourly rainfall data were obtained from six rain gages in Louisiana and cumulative rainfall depths and average rainfall intensities were calculated for each storm. A series of annual maximum rainfall depths and corresponding durations resulted. Marginal distributions were fitted to these annual maximum series, and the dependence was calculated using Kendall's τ. From τ, the copula parameter was calculated and the joint distribution was calculated using the Frank copula equation. The IDF curves calculated based on the copula method were compared to IDF curves calculated empirically and based on TP-40. Reasonable agreement was observed between the copula-based and empirically-based IDF curves, both of which were bivariate analyses, while in some cases quite large differences were observed between the copula method and the method outlined in TP-40.

2.1.4.2.3. Copula Modeling of Multiple Flood or Storm Characteristics

Many of the hydrologic applications of copulas have focused on modeling the joint distribution or dependence of some combination of flood peak, volume, and duration, or rainfall depth, duration, and intensity. Grimaldini and Serinaldi (2006b) considered three characteristics of rainfall events, the critical depth, the peak, and the total depth. The multivariate statistical analysis approach was used to assess data

collected for these three characteristics. A 3-Copula was used to determine the distribution of the characteristics. This allowed a return period to be related to critical depths, which allowed for determination of peak and total depth in a probabilistic manner without preconceived expectations of the design hyetograph pattern.

Kao and Govindaraju (2007a) also considered a bivariate analysis of extreme rainfall events. Copulas were used to analyze rainfall total depth, duration, and peak intensity in order to describe the dependence between the variables. Bivariate distributions were developed for these characteristics and compared to univariate distributions developed based on conditional distributions using traditional methods. For rainfall events of short duration, the traditional methods were observed to perform reasonably well, but copula methods were found to perform better for longer storm events.

Zhang and Singh (2007a) considered the development of a trivariate rainfall frequency analysis based on copulas. The Gumbel-Hougaard copula was tested in this application to get away from the common assumptions that the rainfall variables are independent and normally distributed. Using the copula, joint conditional return periods were developed and the trivariate distribution was evaluated using data from the Amite River Basin in Louisiana. The copula method was also compared to a Box-Cox transformed trivariate normal distribution, and found to perform better.

Kao and Govindaraju (2007b) considered the relationship between rainfall and runoff using copula techniques. Rainfall storms were defined based on their average intensity and duration. The bivariate distribution developed using this information

was used to develop the probabilistic structure for surface runoff, based on the Natural Resources Conservation Service (NRCS) rainfall runoff model. The dependence between rainfall characteristics was found to significantly impact the rainfall excess properties, which influences surface runoff. A case study was illustrated for Indiana. It was observed that copulas had could offer simplifications to rainfall and runoff scenarios over more complicated multivariate stochastic models.

Zhang and Singh (2006) have also applied copulas to analyze multiple characteristics of flood events. The bivariate distribution of flood peak and volume was considered in this study. Joint conditional return periods, required for design applications, were obtained from copulas. One-parameter Archimedean copulas were chosen to determine the joint probability distribution of the flood characteristics, with the Gumbel-Hougaard copula observed to be the best fit. Both the Gumbel mixed distribution and the bivariate Box-Cox transformed normal distribution were compared to the copula-derived distribution, and the copula distribution was observed to better fit the data used in this study.

Zhang and Singh (2007b) conducted a similar study focusing on the trivariate distributions of flood peak, volume, and duration. Conditional return periods were again developed based on the copula distributions. For given recurrence intervals, the magnitude of peak discharge was determined using the trivariate Gumbel-Hougaard copula, the trivariate normal distribution, and a joint distribution that assumed the three variables were independent, with the copula method observed to best fit the data used in this study.

Grimaldi and Serinaldi (2006a) also attempted to model the trivariate joint distribution of flood peak, volume, and duration. To do so they used the asymmetric Archimedean class of copulas. After determining the joint distribution of the three flood characteristics, the authors defined the bivariate distribution of the flood volumes and durations conditioned on the flood peak discharge values. Simulations were used to illustrate differences between the use of symmetric and asymmetric Archimedean copulas.

2.1.4.2.4. Use of Different Copula Families

Many authors have investigated the use of different families of copulas in hydrologic applications. Renard and Lang (2007) considered the use of Gaussian copulas for extreme value analysis. The article provides a series of case studies related to hydrologic applications. The case studies demonstrating the use of Gaussian copulas highlight applications to field significance determination, regional risk analysis, discharge-duration-frequency models, and regional frequency analysis.

Genest and Rivest (1993) focused on the Archimedean class of copulas, which includes a number of well-known bivariate distributions. The focus of this study was the selection of the Archimedean copula best able to represent the existing dependence structure between the sample values of the random variables. The article offers a one-dimensional empirical distribution that is applicable regardless of whether the Archimedean copula is appropriate or not, and regardless of the marginal distributions of the random variables. This method provides a strategy for choosing the most appropriate Archimedean copula family for a given set of data.

Serinaldi and Grimaldi (2007) described an inference procedure for asymmetric Archimedean copulas which would allow for trivariate frequency analyses. A secondary purpose of this study was to further stress the importance of asymmetric Archimedean copulas in applications involving flood and sea wave data. Differences between asymmetric and symmetric Archimedean copulas are highlighted in addition to illustrating the inference procedure. Finally, a number of goodness-of-fit tests applicable to choosing the best-fitting model are described.

2.2. Potential Sources of Flooding

As discussed previously, there are many potential sources of flooding, any combination of which might potentially impact a given location. Therefore, a thorough understanding of the current research related to each of the identified flood sources was necessary to advance this project. The goal was to identify methods being used to model each source of flooding, as well as any recent suggestions for determining the expected probability of flooding from each source.

2.2.1. Riverine Flooding

Riverine flooding is one of the most commonly recognized sources of flooding throughout much of the world. Recent research focused on modeling of riverine flood events, and the calculation of the probability of riverine flood events was examined.

2.2.1.1. Modeling of Flood Events

Cullmann *et al.* (2009) presented a methodology using Monte Carlo simulation to model flood inundation and flood forecasting. A distributed rainfall-

runoff model with multiple parameter sets assigned using artificial neural networks was used. The purpose was to evaluate the possibilities of communicating flood warnings. The problem with the current method of modeling flood inundation based on the concept of return periods is that it suggests a level of security to people living in risky areas because the majority of the public are not able to understand and interpret the statistical meaning of return periods. Warnings based on probability can better reflect the uncertainty in flood forecasts. The ability of the model developed to produce fast probabilistic flood warning was tested on the Freiberger Mulde catchment in Eastern Germany. While the model was observed to under-predict some floods, it was overall determined that using the absolute frequency to produce flood warnings was more comprehensive and easier to understand than using the relative flood frequency. The results indicate that it is possible to predict future meteorological events in near-real time, and to combine rigorous physically-based catchment modeling and probabilistic evaluation of flood events. The results also suggest that the uncertainty of the initial hydrological conditions can be evaluated.

Bales and Wagner (2009) also considered the impact of uncertainty in flood predictions. They acknowledge a growing understanding of the need for more and better flood information. They were particularly interested in FEMA flood maps as a flood risk communication method. They point out that there is an amount of uncertainty associated with these flood maps that is not always recognized by users of the maps. This study was specifically interested in the uncertainty associated with flood inundation maps produced using LiDAR topographic data and one-dimensional hydraulic models. Identified sources of uncertainty in these maps come from

hydrologic data, topographic data, and from the hydraulic modeling. They concluded that the most important factors necessary for accurate flood inundation maps were high-quality topographic data and the appropriate application of hydraulic modeling. They also noted that the assumption of steady flow, made in developing FEMA's maps, can have major effects on the simulation of flood inundation, and in some cases unsteady-flow models may be required. Bales and Wagner (2009) offer the suggestion that it may be necessary to replace the typical flood inundation boundaries with zones of a certain probability of flooding, as the given boundaries are not likely accurate due to the uncertainties inherent in the mapping process.

Wheater (2000) reviewed the current state of flood modeling, with an emphasis on rainfall-runoff, climate, and flood routing models. He states that for the best results from modeling of inundation, a dynamic model would be needed to represent the effects of transient storage. The alternative would be a steady-state analysis based only on the peak flow of the flood. However, according to Wheater's review, most two- and three-dimensional modeling use steady-state computations over fairly short lengths of river reach with conventional numerical schemes. Wheater also assessed the state of uncertainty modeling, stating that uncertainty can now be represented in flood inundation simulation exercises, but this idea has not yet taken firm hold in decision-support systems. Wheater suggested that ultimately, some modeling framework that integrates models for rainfall-runoff, climate, and flood routing, will be required for adequate decision-support.

2.2.1.2. Probabilities of Occurrence of Riverine Floods

Karmakar and Simonovic (2008a) attempted to better select the marginal distribution functions for flood characteristics using both parametric and nonparametric estimation methods and to demonstrate how these marginal distributions could be used in determining joint distribution functions. In traditional methods of flood frequency analyses marginal distribution functions of key flood characteristics (peak flow, volume, and duration) are considered to follow some parametric family of distribution functions; however, it does not seem necessary that all three characteristics should follow the same distribution family, nor is there a universally accepted distribution associated with hydrologic variables. The limitation of parametric methods is that a distribution function must be assumed, while for nonparametric methods no such method is necessary. To evaluate the various distribution functions based on parametric and nonparametric methods, 70 years of flow data from the Red River near Grand Forks, North Dakota, were used. Two methods of estimating marginal distributions were tested: a nonparametric univariate kernel estimator with optimal bandwidth for each flood characteristic determined and a method based on orthonormal series. It was observed that the orthonormal distribution method was better able to model a distribution with a binomial shape than any of the parametric methods used for comparison, and that the orthonormal series method is more appropriate than the kernel method for determining marginal distributions. The primary advantage of the orthonormal method is that it can estimate the marginal probability distribution function over the entire range of possible values.

Kamakar and Simonovic (2008b) evaluated the ability of copulas to model the dependence structure between flood characteristics (peak flow, volume, and duration) independent of the marginal distributions of those characteristics. Use of copulas to model joint distribution allows marginals to be selected for flood characteristics from different marginal families. Bivariate copulas were used to determine bivariate distributions between the three flood characteristics. Three widely used Archimedean copulas were evaluated in this study: the Ali-Mikhail-Haq, the Cook-Johnson, and the Gumbel-Hougaard families. The Gumbel-Hougaard was found to be the best family for this particular application and was used to obtain joint cumulative frequency distributions for various combinations of the flood characteristics. Conditional probabilities and their corresponding return periods were also calculated using this methodology for different combinations of the flood characteristics.

2.2.2. Coastal Flooding

Coastal flooding is the other most commonly recognized source of flooding, as throughout the world, large populations live along the coasts. Recent research on the modeling of coastal flood events, the effects of anticipated sea-level rise due to climate change on coastal flooding, and joint probability concepts were examined.

2.2.2.1. Modeling of Flood Events

Zerger and Wealands (2004) developed a deterministic flood model as a mitigation measure to predict flood inundation. This model was intended to serve as a decision-support tool that could integrate the output of the flood model with a GIS system. This model is able to incorporate consequences into the risk model. The model was used to create a hazard database for the coastal community of Cairns,

Australia. A total of 24,669 commercial and residential buildings in Cairns were examined and integrated into a GIS model. Of particular concern to the model were building attributes such as floor heights and building materials. A two-dimensional hydrodynamic storm surge model was used to examine coastal inundation. The inundation model is believed to be the necessary starting point for addressing coastal management challenges. The next step is to integrate this model with other databases to examine other problems, such as the need to evacuate.

Poulter and Halping (2008) examined some of the processes used in modeling of coastal flood inundation. They assessed the impact of horizontal resolution of elevation data sets and connectivity rule applied to the elevation grid cells influenced the extent of modeled flood inundation. Elevation data were obtained from LiDAR in 6-m versus 15-m DEMs. Three connectivity approaches were also examined. The first method, referred to as the 'bathtub' approach, defined a grid cell as flooded if its elevation was less than the predicted sea-level elevation. The other two approaches were very similar to each other and allowed a grid cell to be flooded if the elevation was below the sea level and an adjacent cell was flooded. The difference between these two methods was how adjacent cells were defined, using the four-side or the eight-side rule. The study was conducted in the coastal plain of North Carolina and a number of inundation scenarios were computed and mapped. Large differences were observed in the inundation extent depending on the horizontal resolution of the DEM and the connection scenario used. For instance, the 15-m DEM was observed to predict greater inundation extents than the 6-m DEM. In terms of connectivity, the 'bathtub' approach always produced the largest inundation extent, followed by the

eight-side rule, with the four-side rule producing the smallest inundation extent. This study should be taken as a note of caution, suggesting that data-processing and modeling choices can significantly impact the model outcome.

Golding (2009) evaluated the uncertainty in the parameters associated with flooding from various sources. A modeling exercise was conducted focusing on a coastal flood event with an element of surface water flooding near the end of the coastal event. A combination coastal and surface water flooding event was chosen because fluvial flooding has already been thoroughly studied. The specific event simulated was an extreme astronomical spring tide with a meteorologically generated storm surge combined with a severe thunderstorm. This simulation was based on real events so that the uncertainty in the forecast could be simulated. This modeling exercise illustrated that forecasts of exceptional storm surge could be forecast with accuracy several days in advance and illustrated the importance of including an estimate of uncertainty in the modeling process.

2.2.2.2. Incorporation of Sea-Level Rise into Coastal Flooding

Kirshen *et al.* (2008) examined various potential sea-level rise scenarios with respect to the possible changes in recurrence intervals of storm surges in the northeast United States. Data were obtained for five cities, Boston, MA, Woods Hole, MA, New London, CN, New York, NY, and Atlantic City, NJ. The first step in this analysis involved removing any evidence of trends from the data sets so that sea-level anomalies could be examined. They used the Generalized Extreme Values (GEV) distribution to conduct a frequency analysis on the annual maximum anomaly values. The return periods of the maximum anomalies during several well-known coastal

events in 2005 were determined with this methodology. The results of this study indicate that the storm event currently classified as the 100-year event will become the storm event expected to occur at least once every 30 years, or possibly more frequently, depending on the different sea-level rise scenarios.

Purvis *et al.* (2008) developed a methodology that could estimate the probability of coastal flooding in the future, considering the uncertainty related to sea-level rise. Because coastal flood defense structures typically have design lives between 50 and 100 years, it is important to estimate future sea-level rise as projects are planned. Previously developed methods have typically been based on deterministic modeling of inundation, which masks uncertainty and does not allow for expert knowledge to be considered. Various sea-level rise scenarios provided by the IPCC Third assessment were used to suggest a plausible distribution for the scenarios. Monte Carlo simulation was used to sample from this distribution and produce sea-level rise values that could then be added to water levels estimated for given recurrence intervals. This method allows predicted flood extents to be weighted by the probability of the sea-level magnitude used to make the prediction. Accumulating these weighted values, the probability of flooding for particular recurrence intervals can be estimated while considering the uncertainty associated with future sea-level rise. The methodology was applied to a 24 square km area in a coastal estuary of the River Severn in Somerset, England. A 1-in-200 year coastal flood event was estimated using this methodology and the traditional methodology for the sake of comparison. This illustrated that using a single value of sea-level rise can significantly underestimate the monetary impacts of a flood event, as this method

does not account for events with low probability and high consequence. Purvis *et al.* conclude that the developed methodology is better able to predict future coastal flood risk considering uncertainty, and they state that the methodology could be applied to other sources of uncertainty in the analysis, as well.

Murdersbach and Jensen (2009) developed an approach for calculating exceedance probabilities for coastal flooding based on annual maximum sea level data based on using an extreme value analysis. The analysis accounted for the possibility of nonstationarity in the data. The first step was tests on the mean and the variability of the data to assess the data for nonstationarity. To account for nonstationarity the parameters of the GEV were replaced with time-dependent parameters, which allowed the results of the analysis to vary with time. Annual maximum water-level data collected from 1849 to 2007 were used, with annual mean tidal high water levels being used to normalize the data. The analysis was compared to an analysis performed using the assumption of stationarity. The results indicated that a parametric approach was needed to extrapolate time-dependent parameters in an extreme value analysis. Murdersbach and Jensen recommend this nonstationary GEV approach because it is able to extrapolate future design water levels using a physical basis and it can be used to incorporate climate change scenarios.

Pirazzoli *et al.* (2006) examined recent trends in sea-level, storm surges, and related meteorological parameters. Their analyses focused on the eastern coast of the English Channel. They also considered possible changes in local flood risk. The goal of this research was to provide information that could assist in understanding erosion trends in the English Channel coastal area and in validating tide-surge-wave models.

A total of 114 equivalent full years of hourly tidal records from six British gage stations and over 170 years of data from six French gage stations were used. Prediction software was used to calculate astronomical tidal values for each hourly data point, and the surge was defined as the difference between the observed sea-level value and the predicted astronomical tidal value. This procedure can provide information about the occurrence of surge events from various storms. Sea-levels were observed to be rising comparatively rapidly at each gage station used, which was believed to be the most significant factor in increasing future flood risk.

2.2.2.3. Joint Probability Considerations

Hawkes *et al.* (2002) state that determining the probability of failure of a coastal defense mechanism is a key step in protecting against extreme coastal conditions. In estimating the probability of failure, waves and water level both play a role, so a joint probability analysis is necessary. Because waves and water level are not typically independent, the joint probability analysis is somewhat complex. They consider three potential joint probability approaches: a structure variable approach, a joint exceedance approach, and a joint density approach. The joint density approach outlined requires the preparation of input data (records of wave height, wave period, and water level), fitting statistical distributions to each variable, simulation of large samples of each variable, and then an extremes analysis. This method was found to be consistent and reliable, even for high return periods. This approach appears to be more objective and flexible than alternate methods that have previously been used.

Coastal flooding is typically the result of the joint occurrence of high water levels and large waves, which are partially correlated. Garrity *et al.* (2007) identified

and tested three approaches that could be used to estimate coastal flood response levels and their probabilities, accounting for the joint probability of the water level and the wave events. The three approaches are forcing event selection, time series response, and Monte Carlo response. Each approach was used to estimate the 100-year flood event for a location in the Strait of Georgia region of the Puget Sound in Washington. The flood response parameter selected was the wave runup elevation. For forcing event selection three water level and wind and wave events were estimated to approximate the 100-year flood response. For time series response the wave runup elevation was calculated for a 24-year data series of measured water level and wind data, and then the 100-year flood response was estimated using extreme value analysis with a Gumbel (Extreme Value Type I) distribution. For Monte Carlo response a large statistical sample of the forcing parameters was simulated. This joint probability was calculated using a program known as JOIN-SEA, and a bivariate normal distribution was used. Advantages and disadvantages of each method were identified. In using forcing event selection, the level of uncertainty is highest and the accuracy cannot be known without actual flood data. However, forcing event selection is an efficient engineering tool and it is in line with local engineering practices based on experience in the field. The most conceptually straight-forward method is time-series response, which involves the least amount of judgment in application. However, the accuracy is largely dependent on the flood response calculation method used and the extreme probability distribution chosen. Finally, the Monte Carlo response method was found to be the most conceptually and computationally intensive, but it was also found to yield the most information.

2.2.3. Surface Water (Pluvial) Flooding

Flooding caused by surface water flooding, also sometimes referred to as pluvial flooding, is not currently a well-recognized or understood cause of flooding. The recent research identified focused on modeling of pluvial flood events, determining the probability of occurrence of these events, and providing warning and prediction services.

2.2.3.1. Modeling of Flood Events

Leitao *et al.* (2009) examined digital elevation models (DEMs) for use in modeling surface water flooding. In order to accurately model pluvial flooding and calculate the volume of water flowing over the land surface, a detailed representation of the land surface is required. To ensure model accuracy, ponds and flow paths must be accurately represented. DEMs can be developed at high resolution from LiDAR data, are available and can be used to conduct a detailed analysis of surface flow paths. Leitao *et al.* (2009) compared surface flow paths based on DEMs of three watersheds, two in the United Kingdom and one in Portugal. These watersheds had different slopes and types of terrain, which would have significant impacts on the surface flow patterns. Different techniques were used to generate DEMs for the watersheds. Identification of ponds is particularly important for delineation of surface flow paths, but the area covered by ponds was different for each of the DEMs. This is likely due to errors in the DEMs rather than actual ponds on the surface. Ponds with small depths and areas, which are more likely to be DEM errors, can be filtered out to get a better representation of the surface. Higher DEM resolution led to better pond representation and the delineation of the overland flow network.

2.2.3.2. Probability of Flood Occurrence

Ten Veldhuis and Clemens (2009) acknowledge that decisions must be made, even with a lack of data, in order to establish the necessary protection levels against pluvial flood risks. There is considerable uncertainty surrounding the frequency and consequences of this flooding. The purpose of this research was to investigate the nature and magnitude of these uncertainties with a case study in the Prinsenbeck watershed in the Netherlands. Rainfall data were collected from a weather station and design storms with return periods of 1- and 2-years were calculated and used as input to the model. The model then determined the occurrence of urban pluvial flooding based on occurrence of water above street level. Flood events were identified based on citizen complaints. Theoretical calculations indicate that pluvial flooding is expected at least once a year, which agrees well with observed data.

2.2.3.3. Flood Prediction and Warning

Much of the literature relating to pluvial flooding has focused on prediction and warning systems for pluvial flood events. Collier (2009) suggested two important components to pluvial flood forecasting: forecasting of intense rainfall events and forecasting of locations where water will collect or pond on the surface. With the recent improvements in high resolution Numerical Weather Prediction models, rainfall can be predicted using radar data with reasonable accuracy. Predicting where water will collect and pond on the surface (other than natural water courses) is somewhat more complex. This will require a detailed understanding of the ground surface at the location and an understanding of urban morphology. Methods using airborne laser data are under development.

The European Commission (2009) suggested that a method for warning of pluvial flooding had been developed and could be implemented shortly. The method combined extreme rainfall forecasting technology with maps of regions at high risk for pluvial flooding. The areas at risk from pluvial flooding were identified using an initial high-level, rapid screening process to identify depressions in the land surface based on their depth. Site inspections followed up the screening to verify the results and to identify local hazards that could contribute to flooding. Flood hazard was then assessed considering other contributing factors such as land use and consequences. The benefit of this method is that resources can be directed to the regions with the highest risk, and it provides earlier warning of flood events than currently available.

Falconer (2007) contended that a warning system is technically feasible for pluvial flooding. He proposed a warning system that involves identifying vulnerable locations in urban areas using rapid screening processes and obtaining targeted rainfall forecasts for high intensity events. The proposed process includes five steps. First, a rapid screening process is used to identify locations vulnerable to pluvial flooding. Second, inspection of the sites identified verifies the screening results. Third, potential mitigation methods can be considered. The fourth step uses contingency planning to raise awareness of the potential risk and consider potential contingency measures. Finally, the results are reviewed and refined with more detailed hydraulic and hydrologic studies. It has been observed that screening of depressions in combination with identification of flow paths is a powerful combination. It was also suggested that use of ifSAR or LiDAR data in urban areas could be a cost-effective means of obtaining rapid-screening data over the catchment.

Simoes *et al.* (2010) consider real-time flood prediction to be the next step in urban flood modeling. They used a model based on the dual drainage concept, which consists of a flow network of open channels and ponds connected to a sewer or drainage system. The goal was to predict flood extent for a period of three hours within 15 minutes of receiving rainfall data. The watershed used in this study was Vale das Flores, in Coimbra, Portugal, which is a highly urbanized 220 ha watershed. A number of surface water flow paths were identified in the model. Simplifications in the sewer network and surface flow paths were necessary to reduce computation time. Simoes *et al.* (2010) compared two methods of simplifying the networks. They simplified the sewer network and the overland flow paths separately and in combination. Because the two networks were connected, it was observed that using the simplification techniques on each network individually resulted in greater reduction of elements in the system than using the techniques on the combined sewer and overland flow paths networks.

2.2.4. Groundwater Flooding

Groundwater may currently be the least-recognized and understood potential source of flooding. Very little data are available, and few studies have focused on groundwater flooding, but some recent research relating to modeling of these flood events and assessing the risk from groundwater flood events were identified.

2.2.4.1. Modeling of Flood Events

Finch *et al.* (2004) considered the spatial distribution of groundwater flooding in England, where groundwater flooding occurs primarily in chalk aquifers. They identify two forms of groundwater flooding. In the first, groundwater enters the

floodplain from alluvial deposits that have become so saturated that the floodplain floods before the river has left its banks. In the second, high groundwater levels in major "solid rock" aquifers cause flooding in the upper reaches of streams and rivers, such that the area in which the water table intersects the ground surface expands up the valley. This type of flooding was observed to occur over the winter between 2000 and 2001 in the watershed of the River Pang near West Berkshire Downs. The flooding resulted in prolonged inundation of farmland, homes, and roads. Several methods were used to estimate where flooding was being caused by groundwater as opposed to other sources. One of these methods was to measure the water temperature, as a difference in temperature had been observed between water that had been exposed to the surface and water that had been underground. An increase in temperature was believed to indicate an area where there was groundwater effluent into the surface water. From the results of this study, they concluded that by studying the interactions between the regional water table, the topography, and the variations in the permeability of the aquifer in determining the locations of surface water bodies the locations most at risk of groundwater flooding can be better identified.

Visocky (1995) studied the potential for groundwater flooding in Illinois after severe floods occurred over the course of several months in the fall of 1993. During this time, the area experienced record-setting precipitation in combination with high stages in the Illinois River, causing very high groundwater levels. The Federal Emergency Management Agency and the Illinois Emergency Management Agency requested that the Illinois State Water Service provide information regarding the frequencies of groundwater flooding and the locations most prone to groundwater

flooding. Maps of areas subject to a 1% frequency of groundwater flooding were also requested. A predictive model, based on combined precipitation, the Illinois River stage, and groundwater data, was used for flood frequency analyses and generating information about the 100-year groundwater flood stage. This study involved five primary stages. First, the well-head elevations for approximately 75 observation wells were determined and potentiometric surface maps in the study area were revised based on this information. Second, statistical analyses were performed to identify the return period for the flooding that occurred in 1993 and to determine the 100-year flood event. The log-Pearson Type III distribution was used in conducting the flood frequency analysis. Third, detailed, high-quality orthomaps and digitized aerial photos were developed. Fourth, maps of the areas at risk from the 100-year flood were prepared. Finally, map products from various GIS coverages were generated.

2.2.4.2. Assessment of Flood Risk

Cobby *et al.* (2009) reviewed and summarized previous studies focusing on groundwater flood risk. The information presented included methods for data collection and accumulation, mapping of the hazard, and provision of flood warning services. They also examined various approaches for managing the risk from groundwater flooding and suggested ways these methods might benefit from advances in data management, hazard mapping, and warning services. They suggested that there is a significant lack of data relevant to groundwater flood risk and the technology for warning services lags behind that for other flood events. A number of methods for assessing the frequency of occurrence of groundwater flooding were assessed. These included methods of assessing the frequency of

observed flood events directly, methods of assessing the frequency from the drivers of flood risk, such as antecedent moisture condition, or using surrogate measures to assess the frequency of flooding. Also suggested were using a combination of drivers and surrogate measures to assess the frequency of occurrence, and using mathematical modeling to predict the frequency of groundwater flood occurrence.

Macdonald *et al.* (2007) discussed collaborative efforts, which were undertaken between the Environment Agency and the British Geological Society that focused on groundwater flooding in connection with the Oxford Flood Risk Management Strategy. The purpose of the strategy was to identify options to reduce the flood risk in Oxford, which had more than 3600 properties located within the boundaries of the 1% annual chance flood from the Rivers Thames and Cherwell. In this region, sediments underlie the floodplain and limit the ability of groundwater to flow through the aquifer, causing high levels of groundwater. The groundwater levels are no more than one to two meters below the ground surface. While this study is still ongoing, there has been significant collection of data to identify the problem and to understand the dynamics of the groundwater and surface water system.

2.3. Potential Causal Factors

There are a variety of factors that could cause flooding from the different sources. These factors include rainfall events, hurricanes, snowmelt, ice jams, and rainfall on snowpacks. Research into some of these potential causal factors is described below.

2.3.1. Snowmelt

Brooks and Boll (2005) describe and test a GIS-based snow accumulation and melt (SAM) model that can be integrated into distributed hydrologic models for watershed management. In this paper they compare the simulated snow water equivalent to point-based measurements and spatial patterns made on a 10x15 m grid over a small test watershed. The model used both mass and energy balance equations to predict the accumulation and melting of snow with time. Data input included hourly solar radiation, relative humidity wind speed, air temperature, and precipitation, all of which are fairly easy to obtain. The model predictions were found to agree quite well with observed data over a three year period, indicating that the model was capable of modeling the accumulation and melt of a snowpack. The advantage of this model over others is that it is simple to use, incorporates GIS, and does not require excessive data.

Carlson and Fox (1976) acknowledge that there are several factors making computation of flood frequency estimates in northern regions difficult. These consider the arctic and sub-arctic regions of Alaska and Canada, where floods can occur from rainfall or snowmelt events, but the two do not typically occur together. A flood frequency model previously developed based on rainfall data was modified to predict the frequency of snowmelt floods and then applied to the Chena River in Alaska. The model used the kinematic wave method to estimate peak streamflow. After refining the model parameters to apply to snowfall events, the model was tested. Several important inconsistencies were identified, but overall the model was

found to be able to provide flood frequency estimates for northern watersheds experiencing snowmelt floods.

Troendle and Porth (2000) re-evaluated the USGS regional flood frequency estimation equations for Colorado and Wyoming, which have high-elevation, snow zone watersheds. The objective of this study was to improve the ability to estimate the 1.5-, 5-, 10-, and 25-year maximum instantaneous flow and the maximum mean daily flow from these watersheds. They tested both the Weibull distribution and the log-Pearson type III distribution, finding very similar estimates of flood magnitudes for both distributions. They evaluated the applicability of a number of potential variables for their regression equations, but typically found only precipitation and area to be significant variables. These equations were compared to the USGS equations and were found to have a much smaller range of errors.

Payton and Brendecke (1985) also evaluated the frequency of rainfall and snowmelt events in the Boulder watershed. From local climate stations, data on precipitation, relative humidity, temperature, solar radiation, barometric pressure, snow depth and snow water equivalent were obtained. It was determined that an asymptotic probability distribution would be best in this situation, so the Gumbel, lognormal, gamma, and exponential distributions were tested. Between May and September rainfall records for storms with durations form 30 minutes to 24 hours were used. Because a historic record was not available for snowmelt events, a snowmelt model was used to generate records. The model was used to predict potential 24-hour melt rates between April and July, with the exponential distribution observed to provide the best fit. Comparing the rainfall and snowmelt frequency

estimates, it was observed that for event durations of 4-6 hours, the snowmelt frequency curves typically provided much lower intensities for the given duration and return period. Similarly, for short duration events, the snowmelt intensities were much lower than those for rainfall events. The correlation between rainfall and snowmelt events was also examined, finding a significant negative correlation between the two types of events, which would imply that rainfall and snowmelt events cannot be assumed to be independent.

Zuzal and Greenwalk (1985) attempt to address the lack of quantitative data on magnitude and frequency on rain-on-snow events. Probabilities were developed for the frequencies of events in Moro, Oregon. Daily weather data, including minimum and maximum air temperature, precipitation, snowfall, and amount of snow on the ground were required. They developed several conditions which must be met in order for a rain-on-snow event to have occurred. Precipitation must have occurred, but it must not have been snow. The maximum air temperature must have been greater than 3^{o} Celsius. Finally, there must have been snow on the ground on the previous day. If all these conditions are met, then a rain-on-snow event can be assumed to have occurred. The events were correlated with the crest gage record to develop probabilities, and the lognormal distribution was determined to adequately represent rain-on-snow events. A threshold of approximately five mm of rain per day was observed to be necessary in order to create runoff from a rain-on-snow event.

Marks *et al.* (1998) examined the climate conditions that could influence the magnitude of runoff produced by rain-on-snow events, acknowledging that these events were quite common in the snow transition zone of the Pacific Northwest.

They found that detailed data on the climate conditions, snow properties, and snowmelt processes occurring during rain-on-snow events was unfortunately limited. The climate conditions considered included wind speed, temperature, humidity, and precipitation. Previous studies had indicated that runoff generation from rain-on-snow events was actually most sensitive to wind speed, and that higher temperatures and humidity only important factors if they occurred with high wind speeds. The purpose of this study was to compare runoff generated from rain-on-snow events in forested watersheds to clear-cut watersheds. Collected data were used to drive a model based on the snow cover energy balance, to simulate the development and ablation of the snow cover prior to, during, and after an event. In comparing the two types of watersheds, the forested watersheds were, as expected, found to produce less runoff than a clear-cut watershed.

2.3.2. Ice Jams

Tuthill *et al.* (1996) assessed the ability to model ice-jam flood events. Because sufficient stage data do not exist in many locations, it may be necessary to synthesize stage values from the existing discharge record. Tuthill *et al.* compared the method they developed to the method developed by Gerard and Calkins (1984). Gerard and Calkins developed a single compromise ice-related rating curve that could be used to convert the probability distribution for ice-related discharge to a probability distribution for ice-related stage. Tuthill *et al.* developed a discrimination test that could sort the winter peak discharge record into mutually exclusive populations of ice-influenced or not ice-influenced events. The criteria used to discriminate between populations were based on an analysis of the factors that were

conducive to formation of ice. The historical record of ice-jam floods was examined to identify discriminating factors, which could be used to assist in finding other, unrecorded ice-jam flood events. The factors identified include the time period in which ice-jam floods typically occur, the lower discharge bound for the historic population of ice-jam floods, the antecedent conditions that are most conducive to the formation of a strong, thick ice cover on the river, and the condition of the ice cover when it breaks up. Once the winter floods have been separated into ice-influenced and not ice-influenced events, separate probability distributions can be developed for the peak stage records in each population. These distributions can be combined to form a single stage frequency curve for the winter season. The log-normal distribution was used to fit the separate populations because this distribution was observed to fit the data well and because it had been used previously. Tuthill *et al.* illustrated the method with a case study by modeling the flooding that occurred in Montpelier, Vermont on March 11, 1992 on the Winooski River.

Another example of developing ice-influenced frequency curves comes from the US Army Corps of Engineers, St. Paul District. On the Red River of the North a gage has been recording peak discharges since 1942. Of the annual instantaneous peak discharge record, half of the data points were known to be associated with ice, while the other half were known not to be associated with ice. The mixed population approach was used, with the data separated by season. The ice-affected stage frequency curve was developed based on 56 years' worth of data from a USGS gage. However, because ice is known to be unstable at high discharges, the upper end of the frequency curve, from 0.2 to 0.3% exceedance frequency elevations, were determined

based on the annual all-season peak discharge-frequency curve. Stage-frequency curves were computed for a number of channel cross-sections along the reach, for the 10, 50, 100, and 500 year events. This analysis was used to identify the 100 year and 500 year floodplain boundaries and the 100 year floodway (White *et al.*, 2000).

2.4. Potential Challenges to be Addressed

Several challenges have been identified that will need to be addressed in conducting this research. These topics have been previously addressed to some degree in the literature, as discussed below.

2.4.1. Zero-Flood Years

The methodology recommended for accounting for zero flood years in Bulletin 17B was developed by Jennings and Benson (1969). They noted the problem associated with the log-Pearson Type III distribution, in that the logarithm of zero is negative infinity, but they also noted that none of the generally used theoretical probability distributions had the ability to fit a set of data in which part of the data could be represented by a curve and part by a straight line of constant value, as would happen when zero values or below-threshold values occurred. Their method uses a base flood value, Q_b, which can be zero or some other value above which flood discharges are measured. They set the probability P(x) that in a given year a flow event exceeding Q_b to be n/N, where n is the number of events above Q_b and N is the total number of events. The result of this method is a frequency curve with two segments, one of which is a continuous frequency curve and one of which is a discrete or mass probability segment. If so desired, a theoretical distribution, such as

the log-Pearson Type III distribution, can be applied to the flood series by fitting the distribution to values above Q_b and then multiplying the given probabilities by n/N. Advantages of this method are that it is based on probability theory and it does not ignore any of the data, as did the method suggested by Beard (1962), which was intended to correct for the effects of "abnormal dry years". Jennings and Benson compared their method to the method recommended by the Water Resources Council Subcommittee on Hydrology (1966), which suggested adding a small amount of discharge, such as 0.1 cfs, to all events, then fitting the log-Pearson Type III distribution to the adjusted data set, and found that their method generally had better fits to the data.

McCuen and Beighley (2003) considered a similar problem in seasonal flow frequency analysis. They point out that many design problems are based on analysis of annual maximum flood series; however, this may not be appropriate in all scenarios. Two specific topics were addressed by this research. First, a method for analyzing incomplete flow records needed to be developed in order to develop seasonal flow frequency estimates for sites with flow gages. Second, because seasonal flow frequency discharges are frequently needed at sites without flow gages, a method for adjusting annual maximum discharge estimates into seasonal estimates was needed. McCuen and Beighley developed three options that could be used for estimating moments for incomplete seasonal flood records. They then used Monte Carlo simulation to evaluate the bias and accuracy of the estimators, with bias expressed as a fraction of the mean and standard error expressed as a fraction of the standard deviation. The approach that seemed to provide the best results involved

replacing unmeasured values below some threshold value with the threshold value X_o. They compared these options to the method for incomplete record adjustment presented in Bulletin 17B. The results suggested that the maximum likelihood approximation method could be more widely applied than the method presented in Bulletin 17B. Finally, they applied their method to four test watersheds with varying sizes and annual maximum record lengths. The results showed that the seasonal 100-year discharge could vary significantly from the annual 100-year discharge, further supporting the use of seasonal flood frequency analysis. For cases where gaged data was not available, or was not adequate, McCuen and Beighley suggested modifications to regionalized equations or the use of the index flood approach.

2.4.2. Multiple Populations

Morris (1982) describes the most common method of addressing flood records based on multiple populations. Waylen and Woo (1982) had a somewhat different approach to improving prediction when multiple populations were involved. They suggested that the peak discharge data in these cases should be considered to be drawn from subpopulations with completely different statistical characteristics. The method proposed in this study recognizes and incorporates the distinctly different characteristics of these different populations. The Gumbel distribution, which is typically adequate for representation of floods generated by one process, was used in this study. The method involves using the method of moments to derive parameters for the Gumbel distribution for each of the flood-producing processes. To determine the overall annual flood distribution, considering both flood-producing processes, a Gumbel distribution should be calculated by multiplying the two Gumbel

distributions, with parameters previously identified for each flood-producing process, together. The advantages of this method are that it does not reduce the sample size and it does not involve the estimation of a weighting factor. This methodology was tested using data from the Coquihalla River, where floods could be caused by rainfall or by snowmelt events. To determine whether data points from the flood record were caused by rainfall or snowmelt they analyzed the antecedent rainfall conditions and found a break-point that could be used to classify the data points as rainfall-generated or snowmelt-generated. From this point, the Gumbel distribution procedure developed was used. Using this method was observed to improve the fit between the flood prediction model and the data; thus, advancing the ability to predict the probability of a flood event at a given location.

2.4.3. Assumption of Independence of Events

While evaluating the suitability of the Box-Cox transformation method for evaluating the combined probability of riverine and tidal flooding in an estuary, Loganathan *et al.* (1987) also evaluated the assumption of independence between the riverine flow data and the tidal data. Their findings suggest that high riverine flows and high tides are often correlated, thus making the assumption of independence between the random variables invalid. In fact, they suggested that the assumption of independence could lead to under-prediction of flood magnitudes, and thus, under-design of flood control structures.

The assumption of independence has also been found to be invalid when considering floods caused by multiple populations, such as riverine floods fed by rainfall and snowmelt. Payton and Brendecke (1985) studied the Boulder watershed

in Colorado, which is influenced by both rainfall and snowmelt events. They wished to determine whether the probable maximum flood event used to determine dam spillway capacity should consider both rainfall and snowmelt events. Because the study region was mountainous, and snowpack remained until well into the spring season, the joint probability of rainfall and snowmelt was of interest. Calculation of the joint probability depended on whether or not the two data sets were independent, so Payton and Brendecke (1985) calculated the correlation coefficient between the two rainfall and snowmelt data sets and determined that they were not independent. Thus, the joint probability was calculated using conditional probabilities. They also determined that calculation of a dam spillway capacity based on the PMF, calculated from rainfall data, was not necessarily appropriate in mountainous regions.

2.4.4. Uncertainty in Analyses

A number of researchers have considered the various sources of uncertainty and effects of uncertainty on hydrologic modeling. Merwade *et al.* (2008) focused on the effects of uncertainties associated with the key hydrologic variables in flood inundation mapping efforts. A number of sources of uncertainty were identified. In hydrologic models, uncertainty could be introduced through model parameters, structure, and assumptions, watershed characteristics and conditions, elevation data, and stream flow data. In hydraulic models, uncertainty could be introduced through the geometric description of the cross-section, model parameters, representation of hydraulic structures, and even model type. In summary, uncertainties in flood inundation maps come from uncertainties in design flows, terrain elevations, water surface elevations, and the accuracy of the modeling techniques. Therefore, Merwade

et al. (2008) suggest use of probabilistic inundation maps that could better account for and illustrate the effects of uncertainty.

Bales and Wagner (2009) also considered the sources of uncertainty affecting flood inundation maps. They focused on the uncertainty associated with flood inundation maps produced using LiDAR topographic data and 1-dimensional hydraulic models. The relevant sources of uncertainty identified include hydrologic data, topographic data, hydraulic modeling techniques and input data, and assumptions such as that of steady flow. They point out that it is not typical for the data to be available to perform rigorous and thorough analysis of uncertainty. Thus, it is vital that those who create and use flood inundation maps be aware of the sources of uncertainties that could influence those maps. This is particularly important in considering the boundaries of flood areas shown on inundation maps, which represent the expected boundaries of inundation. Bales and Wagner (2009) concluded that uncertainty in inundation mapping increases with distance away from the river. That these boundaries are uncertain is a vital fact to understand and communicate.

Golding (2009) focused on uncertainties in the flood events themselves and on the propagation of uncertainties in flood simulation. Uncertainties associated with specific flood events were identified as location, timing, and intensity of the event. There were also identified parameter uncertainties, including depth and velocity of water and the rate of rise. However, these uncertainties, and their effects, will depend on the source of the flood waters. Uncertainty characterization has focused primarily on the parameters controlling runoff generation in hydrologic models and on the propagation of the flood wave; however, meteorological input is another significant

source of uncertainty that should not be neglected. Finally, the response of flood risk reduction structures, where relevant, is increasingly uncertain as water levels rise. Golding (2009) evaluated a number of methods of propagating uncertainty throughout a flood event simulation. They concluded that it was possible to model the propagation of error through real-time flood models.

Ten Veldhuis and Clemens (2009) considered uncertainty associated with urban flooding. Specifically, they were interested in uncertainties associated with assessing the frequency and the consequences of urban flooding, as these are the two components that comprise urban flood risk. This study focused on identifying the nature and size of uncertainties in urban flood frequency estimation. The uncertainty sources identified related to the capacity and condition of the urban drainage system, flow processes in the urban environment, rainfall input, and flood consequences. While some of these sources of uncertainty can be reduced by obtaining more detailed information, others cannot be so easily managed. Ten Veldhuis and Clemens (2009) successfully modeled urban flood frequency, but further research and data collection will be needed to verify the assumptions made.

CHAPTER 3

EVALUATION OF THE COPULA METHOD

3.1. Introduction

The first step to conducting this research was to identify and verify a method for calculating probabilities of dependent events. The copula method was chosen and verified as the most appropriate method of doing so. Once the method was selected, the focus was on combining two sources of flooding; using both simulated and observed data. A hypothetical scenario was developed to use with simulated riverine and tidal data. The methodology was then applied to observed riverine flow gage and tidal gage measurements after an appropriate location of interest had been selected. Based on the results of this method for combining two flood sources, the methods could then be expanded to account for three or more flood sources, as needed.

3.2. Verification of Copula Methods

Hydrologic variables generally have distributions that are different from the normal distribution. For example, annual maximum discharges are considered to be log-Pearson Type III distributed, while rainfall depths are often represented using an extreme value distribution. In many hydrologic analyses, the design variable depends on the values of two causal variables, each with its own probability density function (pdf). For example, flooding is a function of both the peak discharge and the volume of runoff. The pdfs may be known from theory or identified via sampling of measurements. One task of this study was to identify the joint distribution for

hydrologic variables. An analytical expression exists for the joint normal distribution. Some distributions have analytical approximations, but their applicability depends on adherence to certain constraints; for example, an approximation is available for the joint exponential distribution. However, the joint distribution is rarely, if ever, known; therefore, it is important to verify that a method used to approximate joint probabilities provides accurate estimates.

The copula method of calculating a joint probability distribution for two or more random variables was of interest; however, some verification of the accuracy of the method was considered necessary. Because the copula method could be used to combine different distributions, it was necessary to verify that the values of the joint random variable would be accurately distributed. This was done by comparing joint distributions calculated using the copula method to joint distributions calculated using theoretical bivariate equations for two normally distributed random variables and two exponentially distributed random variables as inputs. In both cases, the Gumbel-Hougaard copula family was utilized for simplicity of calculations and because it seemed to be particularly widely used in hydrologic applications (Zhang and Singh, 2007; Genest and Favre, 2007).

3.2.1. Uncorrelated Bivariate Normal Analyses

The first test made to verify the accuracy of probabilities of a joint distribution that are calculated by the copula method used two normally distributed samples as input. The Monte Carlo generated values were uncorrelated and assumed to be from standard normal populations. The bivariate normal distribution was chosen for the first test because an analytical expression exists. It is based on the two parameters of

each marginal distribution and the correlation coefficient between the two random variables.

3.2.1.1. Description of Methodology

Two normally distributed random samples, each with a sample size of 1000, were generated. Both samples were drawn from a population with a mean value of 0 and a standard deviation of 1. The bivariate normal distribution that corresponds to these two marginals can be calculated using the following equations:

$$f_{x_1,x_2}(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{z}{2(1-\rho^2)}\right) \tag{3-1a}$$

$$z=\frac{(x_1-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2} \tag{3-1b}$$

in which x_1 and x_2 are the values of the random variables, μ_1 and μ_2 are the means of the random variables, σ_1 and σ_2 are the standard deviations of the two random samples, ρ is the coefficient of correlation between the two random samples, and z is the standard normal variate that corresponds to the values of x_1 and x_2. Equations 3-1a and 3-1b provide the ordinates of the joint density function. Probabilities can be obtained by integrating Equations 3-1a and 3-1b over ranges of both variables.

Since the copula provides probabilities of the cumulative joint distribution, it was necessary to integrate Equations 3-1a and 3-1b to compute the theoretical estimates for comparison. A three-dimensional numerical integration of the trapezoidal rule was used. For specific values of x_1 and x_2, say x_{1u} and x_{2u}, the intervals $(-\infty\leq x_1\leq x_{1u})$ and $(-\infty\leq x_2\leq x_{2u})$ were divided into 200 equal sections, which resulted in the ordinates in $f(x_1, x_2)$ for each of the 40,000 cells. These were used to

compute the volume of each cell, the sum of which approximates the joint cumulative probability:

$$F(x_1,x_2) = \frac{\Delta x_1 * \Delta x_2}{4} \sum\sum [f(x_1,x_2) + f(x_1+\Delta x_1, x_2) + f(x_1, x_2+\Delta x_2) + f(x_1 + \Delta x_1, x_2 + \Delta x_2)] \quad (3\text{-}2)$$

The lower limits of integration for each of the two marginal distributions were -3.5, as very little of the pdfs exist below those values. Thus, the volume for a cell for negative values of x_1 and x_2 will likely be more accurate than when x_1 and x_2 are positive. Testing indicated that increasing the number of intervals beyond 200 did not result in a significant improvement in accuracy.

Probabilities for the joint distribution were also calculated using the Gumbel-Hougaard copula family. The cumulative distribution probabilities that correspond to the two normally distributed sample values from the marginals were first calculated for use in the copula equation. The equation to calculate the joint distribution, $H(x_1, x_2)$ is as follows:

$$H(x_1,x_2) = \exp\left\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\right\} \quad (3\text{-}3a)$$

$$\alpha = \frac{1}{1-\tau} \quad (3\text{-}3b)$$

in which x_1 and x_2 and the two normally-distributed random variables, u is the cumulative distribution value corresponding to x_1, v is the cumulative distribution value corresponding to x_2, and α is a copula parameter, which is related to Kendall's τ, a measure of rank correlation, by equation 3-3b. The copula provided the

cumulative joint distribution probabilities that corresponded to the values of x_1 and x_2 from two normally-distributed marginal distributions. The two joint distributions, one based on Equations 3a and 3b and one based on the copula of Equations 3-3a and 3-3b, were then compared to determine whether or not the joint distribution calculated using the copula agreed with the distribution calculated from the analytical expression of the bivariate normal distribution. For Equation 3-3a, u and v were calculated as the normal cdf corresponding to the x and y samples. The comparison was based on the 1000 pairs of x_1 and x_2. Each pair was used to compute the joint probability and the 1000 differences between the two estimates were used to compute the bias, relative bias, standard error, and relative standard error. These reflect the systematic and total error variations.

3.2.1.2. Description of Results

Two methods were used to compare the joint probability distributions calculated based on Equations 3-1a and 3-1b and based on Equations 3-3a and 3-3b. First, the differences were evaluated. The minimum difference, mean difference, and maximum difference computed between the two cumulative distributions were observed to be 0.0, 0.0038, and 0.0077, respectively. The goodness-of-fit statistics were also calculated. For this scenario, the bias, relative bias, standard error, and relative standard error between the two joint distributions were calculated to be -0.0038, -0.0142, 0.0044, and 0.0100, respectively. These indicate high accuracy.

Values of the two cumulative joint distributions were also plotted against each other (see Figure 1). If the joint distribution calculated using the bivariate normal equation (Equations 3-1a and 3-1b) agreed well with the joint distribution calculated

using the copula (Equations 3-3a and 3-3b), then the scatter points should fall approximately along a straight line passing through the origin at a 45^{o} angle. This would indicate that the joint distributions calculated by the two different methods were virtually identical. In examining Figure 3-1, it is evident that the points do fall very closely along a straight line. A straight line was plotted against the scatterplot. Very good agreement is observed between the scatterplot and the straight line, indicating good agreement between the joint distributions calculated by copula and by bivariate normal equation.

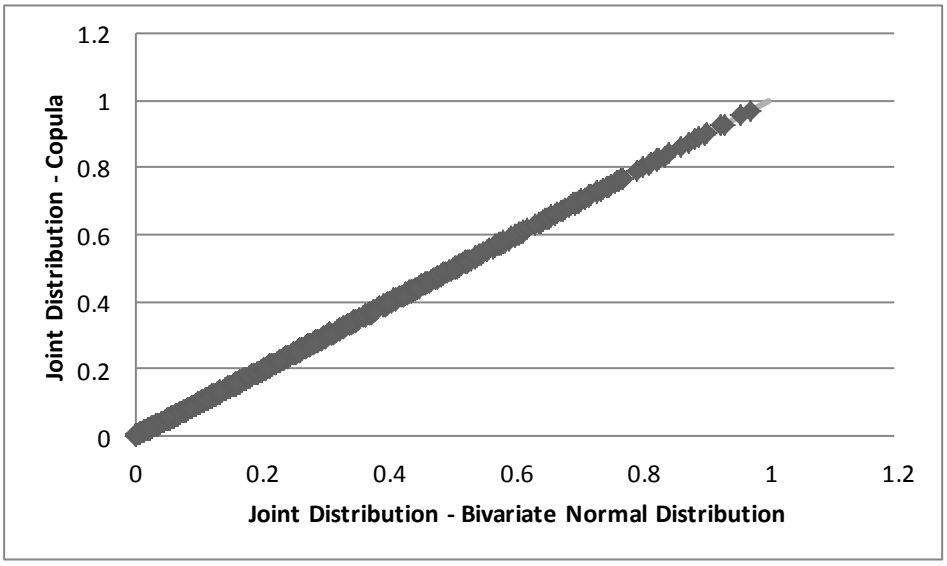

Figure 3-1: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Normal Equation and a Copula Equation for Uncorrelated Random Variables

3.2.2. Correlated Normal Bivariate Analyses

The next step in the verification procedure was to generate two correlated random variables from normal distributions and to calculate their joint cumulative distribution using both the copula method and the bivariate normal equation. This will verify that the copula is adequately accurate in developing joint distributions for correlated variables, which are common and expected when evaluating hydrologic

phenomena. Evaluations were conducted with low correlation between the samples, specifically a value of 0.2, and with high correlation, specifically a value of 0.8.

3.2.2.1. Description of Methodology

Generating correlated random, normally-distributed variables is fairly straightforward. The first normally-distributed sample, with 1000 random values, is generated using typical Monte Carlo techniques. The second, correlated normally-distributed sample is calculated using the following equation:

$$x_2 = \overline{x_2} + \rho \frac{S_{x_2}}{S_{x_2}} (x_1 - \overline{x_1}) + z S_{x_2} \sqrt{1 - \rho^2} \qquad (3\text{-}4)$$

where $x_1$and x_2 are individual values of the two random variables, $\overline{x_1}$ and $\overline{x_2}$ are the means of the two random variables x_1 and x_2, respectively, S_{X1} and Sx_2 are the standard deviations of the two random variables, and ρ is the correlation coefficient between the two variables. Random samples were generated with correlation coefficients of 0.2 and 0.8.

Once the random samples were generated, the joint cumulative distribution was generated using the Gumbel-Hougaard copula, as described previously, using Equations 3-3a and 3-3b. The joint probability distribution was calculated, using equations 3a and 3b, and then the cumulative distribution was approximated using Equation 3-2. Comparisons were made between the joint distributions calculated by copula and by bivariate equation, for both levels of correlation between the samples.

3.2.2.2. Description of Results

For each pair of random x_1 and x_2 values, the corresponding joint cumulative distribution value was calculated by the copula and by the bivariate normal equation. Thus, the two joint cumulative distribution values could be compared. The differences between the joint cumulative distribution values were calculated for each of the 1000 pairs. For a correlation of 0.2, the minimum, mean, and maximum differences were 0.0, 0.0179, and 0.0370, respectively. For this scenario, the bias, relative bias, standard error, and relative standard error were -0.0179, -0.0627, 0.0209, and 0.0473, respectively. For a correlation of 0.8, the minimum difference was 0.0001, the mean difference was 0.0813, and the maximum difference was 0.1505. For this scenario, the bias, relative bias, standard error, and relative standard error were -0.0813, -0.2067, 0.0924, and 0.1994, respectively. The differences in joint distribution values and goodness-of-fit values indicate a high degree of accuracy between the two joint distributions calculated. It is; however, obvious that there is a better fit between the joint distribution calculated by copula and the joint distribution calculated by bivariate normal equation when the correlation between the two samples is near zero.

Figure 3-2 shows a scatterplot of the joint distribution values calculated by the bivariate normal equation versus the joint distribution values calculated by the copula for the random variables with a correlation of 0.2. If the agreement between the two joint distributions is good, the scatter points should fall along a straight line passing through the origin at a 45^o angle. In examining Figure 3-2, it is evident that there is strong agreement between the two joint distributions. When a straight line is graphed

along with the scatterplot values for low correlation values, it is evident that strong agreement exists between the joint distributions calculated by copula and by bivariate normal equation.

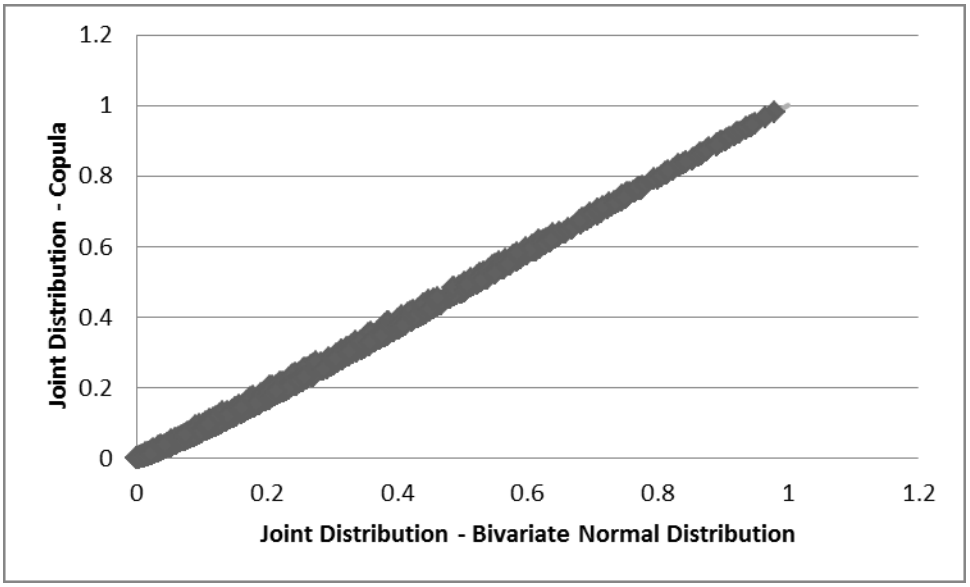

Figure 3-2: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Normal Equation and a Copula Equation for Random Variables with a Correlation of 0.2

Figure 3-3 shows a scatterplot of the joint distribution values calculated by the bivariate normal equation versus the copula for the random variables with a correlation of 0.8. Again, it would be desirable to have the points fall along a straight line to indicate that the two methods, bivariate normal equation and copula, give similar joint distribution values. In examining Figure 3-3, it is obvious that the trend deviates from the linear relation. A slight curvature is evident for the joint distribution values below approximately 0.5, and the scatterplot points do not agree so well with the desired straight line. The scatterplot is characterized by local biases. When comparing Figure 3-3 to Figures 3-1 and3- 2, this verification test indicates that the copula will lose a bit of accuracy when calculating joint distributions for highly

correlated random variables. Nevertheless, these results do not suggest that the copula method should not be applied.

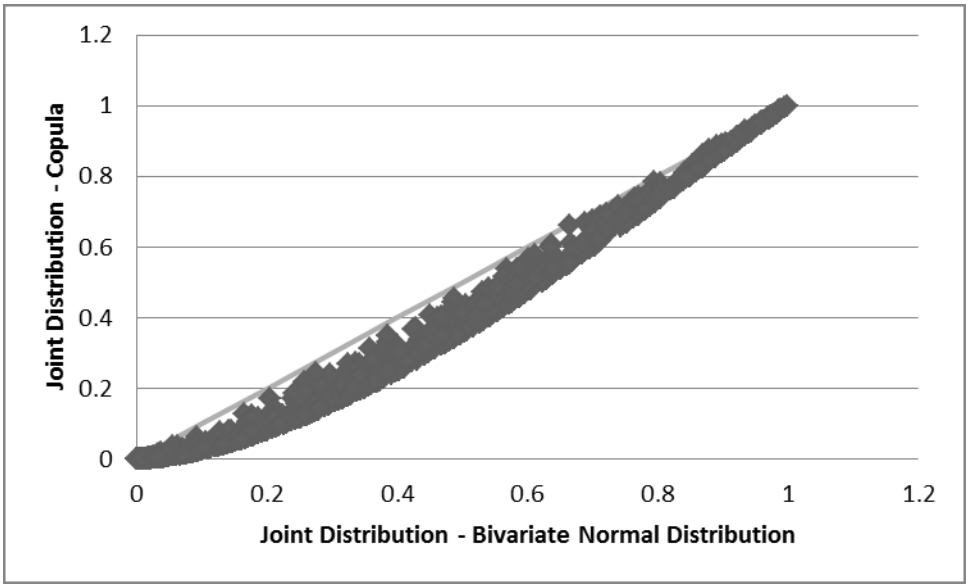

Figure 3-3: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Normal Equation and a Copula Equation for Random Variables with a Correlation of 0.8

3.2.3. Uncorrelated Bivariate Exponential Analyses

The next verification of the copula method used two random samples of exponential marginals. The bivariate exponential distribution is not as well-known as the bivariate normal distribution, but several approximations do exist for calculation of the joint distribution. After evaluation of several possibilities, the equation presented by Nagao and Kadoya (1971) was chosen for use. This equation requires a value from the modified Bessel function of the first kind of zero order; however, when the correlation coefficient is zero, the value of the Bessel function equals 1.0.

3.2.3.1. Description of Methods

Two exponentially-distributed random samples of 1000 values each were generated. Exponential variates X_i are easily generated using the mean value of b:

$$X_i = -b * lnU_i \tag{3-5}$$

where U_i is a uniform variate (0,1). The first random sample was drawn from an exponential population with a mean of 1/2, while the second random sample was drawn from an exponential population with a mean of 2/3. The joint exponential distribution equation used is as follows (Singh *et al.*, 2007):

$$f_{x_1,x_2}(x_1,x_2) = \frac{\alpha\beta}{(1-\rho)}\exp\left(-\frac{\alpha x_1}{(1-\rho)} - \frac{\beta x_2}{(1-\rho)}\right) I_0\left\{2\frac{\sqrt{\rho}}{(1-\rho)}\sqrt{x_1 x_2 \alpha\beta}\right\} \tag{3-6}$$

in which x_1 and x_2 are the values of the exponentially-distributed random samples, α is the parameter for the first exponential population, β is the parameter for the second exponential population, ρ is the coefficient of correlation between the two random samples, and the term inside of the bracket after I_0 is the argument for the Bessel function. Because the two samples were generated independently, the correlation coefficient of the population was equal to zero; therefore, the value of the Bessel function equaled 1.0. The three-dimensional version of the trapezoidal rule was used to obtain estimates of the joint probabilities.

The next step was to calculate the joint distribution, $H(x_1, x_2)$ using the Gumbel-Hougaard copula family. Again, the cumulative distribution probabilities of the random exponentially-distributed samples were calculated. Then, equations 3-3a and 3-3b were used to calculate the joint cumulative distribution as previously specified. Then the joint distribution calculated using the copula, Equations 3-3a and 3-3b, was compared to the joint distribution using a known bivariate exponential distribution equation, Equation 3-5, to verify the accuracy of the copula method.

3.2.3.2. Description of Results

The two exponential joint distributions were compared in two ways. First, the differences between the joint distribution cumulative probabilities calculated by the known Equation 3-6 and calculated by the copula for the same exponential variates x_1 and x_2 were calculated. The minimum difference, the mean difference, and the maximum difference between the two joint distribution probabilities for any set of (x_1, x_2) were found to be 0.0, 0.0002, and 0.0022, respectively. The bias, relative bias, standard error, and relative standard error were found to be 0.0001, 0.0002, 0.0003, and 0.0006, respectively. These indicate a high degree of accuracy.

The two cumulative joint distribution probabilities were plotted against each other (see Figure 3-4). If the two joint distributions were in good agreement, it would be expected that on an (x,y) scatterplot, the points would approach a straight line at a 45^o angle through the origin. In evaluating Figure 3-4, it is observed that the data points do very closely approximate a straight line passing through the origin at 45^o. The scatterplot is in near perfect agreement with the straight line. The small differences between the calculated joint distribution values and the graphical representation shown in Figure 3-4 provide support for the conclusion that the copula method can accurately calculate the joint distribution for two random samples drawn from exponentially-distributed populations.

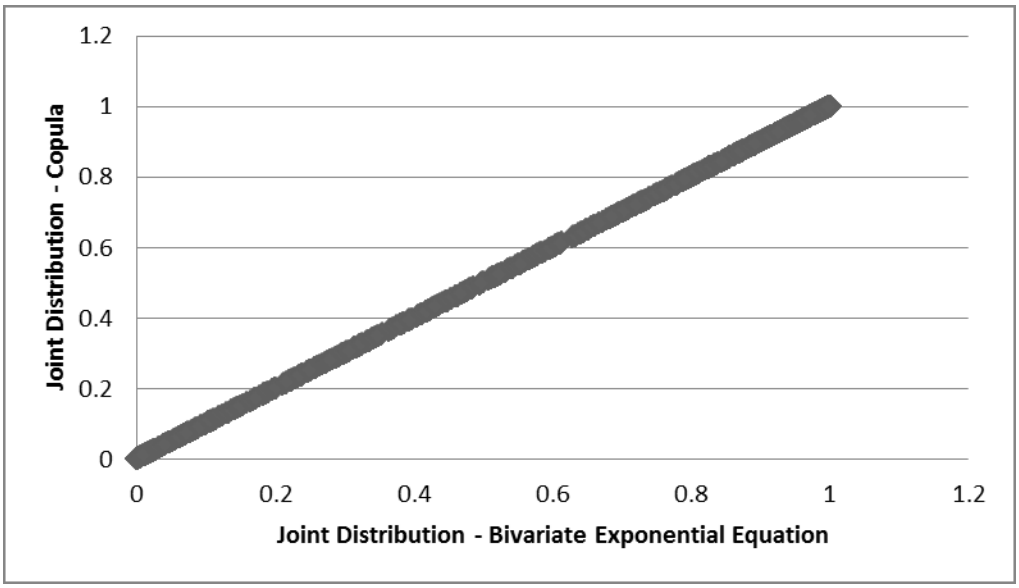

Figure 3-4: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Exponential Equation and a Copula Equation

3.2.4. Correlated Bivariate Exponential Analyses

The final verification experiment involved two correlated exponentially-distributed random samples. The joint distribution was calculated using the Gumbel-Hougaard copula and the bivariate exponential distribution equation. Because the two samples were generated to be correlated, the Bessel function had to be calculated as:

$$I_o(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{k!^2} \tag{3-7}$$

where z is the argument of the Bessel function. The two joint cumulative distributions were calculated for three levels of correlation, with correlation coefficients of 0.2, 0.6, and 0.8, and then compared.

3.2.4.1. Description of Methods

First, two samples of correlated, exponentially-generated random variables were generated. This was a two-step process. First, two random, exponentially-distributed variables were generated, and then the second value was adjusted based on the desired correlation coefficient in order to obtain correlated random samples. The following equation was used to generate the second sample, which was correlated to the first sample:

$$x_2 = \rho x_1 + (1-\rho)\dot{x}_2 \tag{3-8}$$

where x_1 is a randomly generated variate drawn from an exponential distribution with a mean of 1/2 as in Equation 3-5, $\dot{x}_2$ is a randomly generated variate drawn from an exponential distribution with a mean of 2/3 as in Equation 3-5, ρ is the desired correlation coefficient between the two exponentially-distributed random samples, and x_2 is a variate drawn from an exponential distribution that is correlated to x_1. Using this equation, two sets of 1000 correlated, random, exponentially-distributed values were generated.

Once the random samples were generated, the joint cumulative distribution was calculated using the Gumbel-Hougaard copula, as provided in Equations 3-3a and 3-3b. In Equation 3-3a, u and v were calculated by fitting an exponential cdf to the values of x and y. The joint cumulative distribution was also calculated using the bivariate exponential equation, which is provided in Equation 3-6. Joint distributions were created from samples with correlation of 0.2, 0.6, and 0.8. Comparisons were

made between the joint distributions calculated by copula and by bivariate equation for the three levels of correlation.

3.4.2.2. Description of Results

For each pair of random samples, a joint cumulative distribution value was calculated using both the copula and the bivariate exponential equation. The difference between the two joint distribution values was calculated, to determine how well the copula method agreed with the approximate bivariate equation. For a correlation of 0.2 between the two random samples, the minimum, mean, and maximum differences calculated were 0.0, 0.0207, and 0.0376, respectively. The bias, relative bias, standard error, and relative standard error were also calculated to be -0.0207, -0.0655, 0.0232, and 0.0510, respectively. These values suggest reasonably good agreement between the joint distributions calculated by copula and by bivariate equation. For a correlation of 0.6 between the two random samples, the minimum, mean, and maximum differences calculated were 0.0001, 0.0635, and 0.1034, respectively. The bias, relative bias, standard error, and relative standard error were calculated as -0.0628, -0.1739, 0.0703, and 0.1536, respectively. The calculated differences and goodness-of-fit values indicates less accuracy between the joint distributions calculated by copula and bivariate exponential equation than those for the lower correlation, but they still indicate reasonably good agreement between the two joint distributions. For a correlation of 0.8 between the two random samples, the minimum, mean, and maximum differences calculated were 0.0002, 0.1022, and 0.3827, respectively. The bias, relative bias, standard error, and relative standard error were calculated as -0.0455, -0.1298, 0.1201, and 0.2608, respectively. These

values suggest that at higher levels of correlation, the joint distribution calculated by copula does not agree as well with the joint distribution calculated by the bivariate exponential equation.

Figure 3-5 provides a scatterplot of the joint distribution values calculated by the bivariate exponential equation versus the joint distribution values calculated by the copula for the random variables with a correlation of 0.2. The scatter points should fall along a straight line passing through the origin at a 45^{o} angle if the agreement between the two joint distributions is strong. In examining Figure 3-5, very strong agreement between the two joint distributions is evident through the close fit to the straight line. Though there is some disagreement between the copula equation and the approximate bivariate exponential distribution at higher values of the joint cdf, this confirms that for low correlation values the agreement between the joint distributions calculated by copula and by bivariate normal equation is strong.

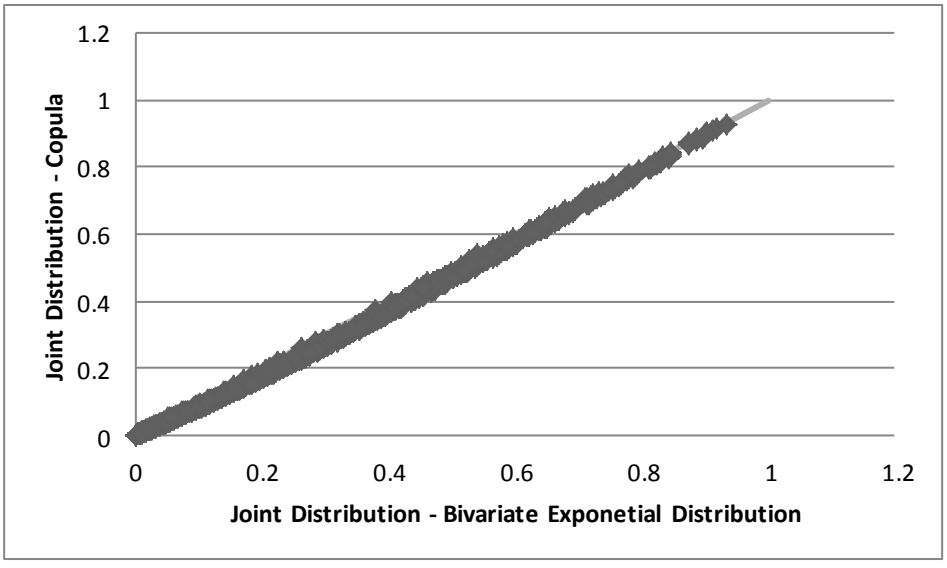

Figure 3-5: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Exponential Equation and a Copula Equation for Random Variables with a Correlation of 0.2

Figure 3-6 provides a scatterplot of the joint distribution values calculated by the bivariate exponential equation versus the joint distribution values calculated by the copula for the random variables with a correlation of 0.6. The scatterplot is again compared to a straight line passing through the origin at a 45^{o} angle to assess agreement between the two joint distributions. As was seen with the highly correlated normal distribution analysis, a distinct curvature is visible in Figure 3-6. As compared to the bivariate exponential equation, the copula tends to underestimate the lower probabilities and overestimate the higher probabilities, though because the bivariate exponential equation is only an approximation, it cannot be stated with complete certainty that the error is in the results of the copula.

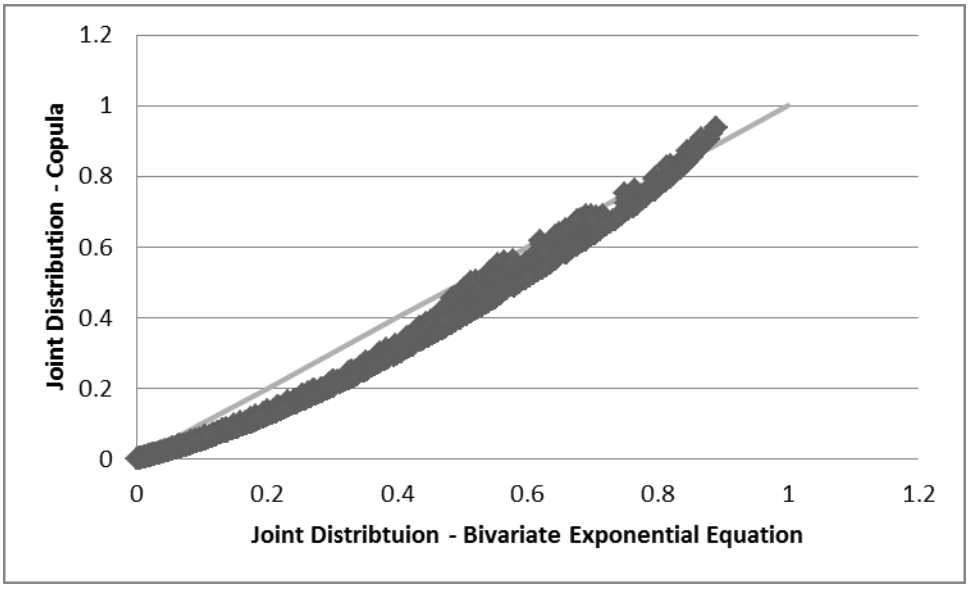

Figure 3-6: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Exponential Equation and a Copula Equation for Random Variables with a Correlation of 0.6

Figure 3-7 provides a scatterplot of the joint distribution values calculated by the bivariate exponential equation versus the copula for the random variables with a correlation of 0.8. Again, the points should fall along a straight line passing through the origin at a 45^{o} angle if the two methods, bivariate exponential equation and

copula, give similar joint distribution values. In examining Figure 3-7, it is clear that there is significant deviation from the linear trend. The copula appears to underestimate low values of the joint cumulative distribution while overestimating higher values of the joint cumulative distribution. These results suggest that the copula will lose some accuracy when calculating joint distributions for highly correlated random variables. However, it must be noted that, unlike the bivariate normal equation, the bivariate exponential distribution equation used is an approximation, thus it is possible that the errors may result from the bivariate equation rather than the copula. While these results do suggest some reason for concern about the results of the copula calculations, they should not prevent the use of the copula methodology.

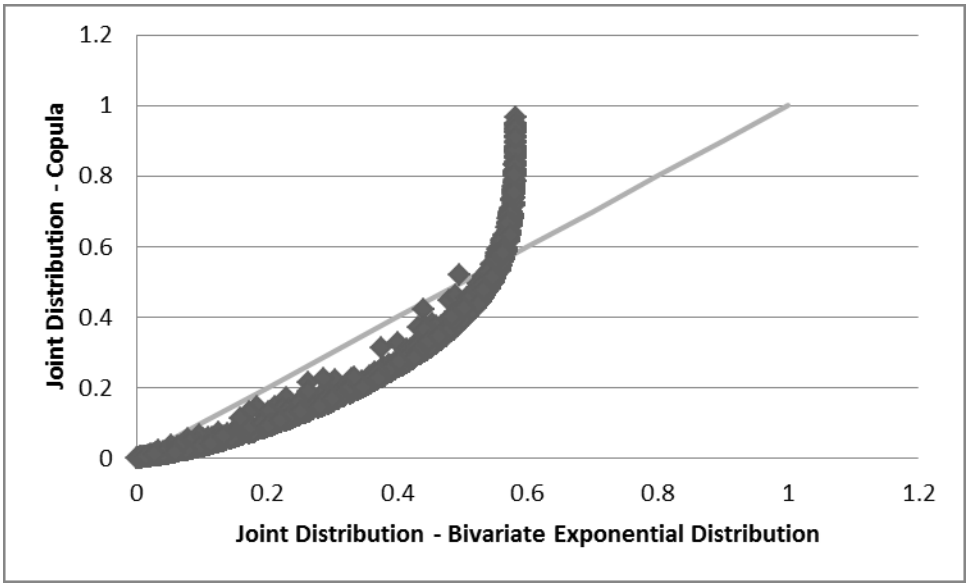

Figure 3-7: Comparisons of Joint Distributions Calculated Using a Theoretical Bivariate Normal Equation and a Copula Equation for Random Variables with a Correlation of 0.8

3.2.5. Conclusions

The results of these investigative studies suggest excellent agreement between the joint distributions calculated by copula equations and by known bivariate

distribution equations for uncorrelated and moderately correlated random variables. Some accuracy in the copula calculations is lost when the random samples are highly correlated. This suggests that, at least when the two random input variables are drawn from the same distribution, the copula method is able to calculate the joint distribution with reasonable accuracy. Further, though the copula method is independent of the distributions of the input variables, because the copula output agrees well with the theoretical output, it can be concluded that the copula is able to provide the appropriate distribution for the joint distribution. Thus, it is possible to feel confident that the joint distribution calculated by a copula when the input variables are not drawn from the same distribution will be accurate and will have the appropriate distribution.

CHAPTER 4

TWO SIMULATED FLOOD SOURCES

4.1. Introduction

The preliminary study presented in Chapter 3 compared the joint distributions calculated by copula to several known bivariate distributions and determined that reasonable agreement existed between the copula method and the known bivariate distributions. With an understanding of how copulas compared to known distributions, it was possible to have confidence in the ability of the copula method to calculate joint distributions in cases for which a known distribution did not exist. Thus, copulas were chosen for further analyses as part of the method to develop a multifactor flood frequency curve for a given location.

Analyses based on data obtained from multiple gages will have a number of complicating factors. These include potential inaccuracies in one or more of the gage records, temporal, and spatial variations in hydrometeorological conditions, use of different reference datum by each gage, and differences in the frequency of gage measurements, among others. Therefore, it is beneficial to use a hypothetical scenario and simulated data as a preliminary step to verify the procedures developed herein and to better understand the expected results.

4.2. Assessment of the Flood Hazard

The first step of a risk assessment is to assess the hazard. In this case, the hazard was assessed by calculating the probability of certain flood inundation depths occurring at the location of interest. The following sections will describe the methods used to fit marginal distributions to each flood source, the use of the copula procedure to determine joint probabilities of given flood events, and the determination of the probabilities that correspond to the total flood depths, as well as presenting the results of each of these steps.

4.2.1. Generation of Correlated Random Samples

Typically, when simulated correlated random variables are needed, a copula is used to generate those variables. However, because it was necessary to evaluate the applicability of the copula methodology for the analysis of the joint probability of flood events, the use of a copula to generate random variables for the simulation study would yield simulated values that were not independent of the copula analysis. Therefore, an alternate method of generating correlated random variables was developed.

The first step of this new approach was to generate the cumulative distribution function (cdf) of the first random variable. Then, a set of random values was generated from that cdf. These values reflect the flood characteristics of the first flood source. Each value of the second variable was generated with the first random variable as the center point of a range within which the generated value of the second variable would lie. The half-width of this bound was specified as input to the

generating program and was one of the factors that controlled the degree of correlation between the two random variables. When the value of the first random variable generated was close to either 0 or 1 in the cdf, the half-width was compressed so the half-width remained symmetrical around the value of the first random variable, but so that the second random variable could not equal 0 or 1 within the cdf. If the compression of the half-width was not made, then the values of the second variable would be biased, which would then bias the degree of correlation between the two variables.

The overall objective of this data generation procedure was to ensure that the generated series had the correlation that would reflect the correlation typical of the two flood sources. While this half-width approach did influence the level of correlation between the two random variables, it still tended to result in higher levels of correlation than were desired. A second technique was then employed to further reduce the correlation between the two variables. If the value generated from the cdf of the second variable was above or below specified limits, specified as input to the program, that value was exchanged with the next generated value within the series. This resulted in lower, more realistic correlations between the two random variables. The final step was to use the second cdf to calculate values of the second random variable. The result of this program was a set of two random, correlated variables that could represent riverine and tidal flood depths at the location of interest.

4.2.2. Overview of Simulated Scenarios

The simulation approach was used to evaluate two concerns. First, the effect of sample size was assessed, as this influences the level of sampling variation. It is

fairly rare for a long gage record to exist, so it was important to understand the limitations of the analysis based on short record lengths. While holding both the correlation between the riverine and tidal flood depths and the marginal distribution parameters roughly constant, sample sizes of 10, 25, 50, and 100 annual maximum events were generated and analyzed. The minimum of 10 years was chosen because that is the smallest sample size recommended for use in Bulletin 17B, while the maximum of 100 years was chosen because very few longer observed records exist. The correlation between the samples was held roughly constant while samples of varying record length were generated, in order to understand the effect of sample size independently.

Second, the effect of correlation between the riverine and tidal flood depths was assessed. It was not possible to determine the effect of correlation between the riverine and tidal data using observed data, as this would require evaluating multiple locations, which would introduce uncontrollable variables; additionally, it is rare to find riverine and tidal gaged locations that are near to each other. Four different sets of correlated variables were generated, spread over a wide range of correlations. The four correlation coefficients evaluated were 0.27, 0.31, 0.37, and 0.60. These values were not on a constant interval of separation because the generation program cannot allow explicit control of the correlation between the two variables; however, these values should provide a thorough understanding of the effect of correlation because the range of values varies from low to high dependency.

4.2.3. Fitting Marginal Distributions to Annual Maximum Event Samples

The first step to the hazard assessment was to determine the marginals. These were determined by fitting probability distributions to the generated riverine and tidal flood inundation depth samples. The following section will describe the methods used to fit marginal distributions and then present the results.

4.2.3.1. Description of Methods

The data generated in the simulation approach would represent riverine and tidal flood depth. In order to determine the appropriate populations from which to generate these samples it was necessary to determine the inundation depths for an observed location of interest based on measured gage data. Once samples had been generated, probability distributions were fitted using Method of Moments or Maximum Likelihood Estimation and the most appropriate distribution to represent each flood source was identified. The following sections describe the methods used to fit the marginal distributions.

4.2.3.1.1. Calculation of Riverine-Caused Flood Inundation Depths

One variable of interest in this research was the flood inundation depth caused by a given flood source; thus, one of the simulated random variables represented riverine flood inundation depths at the location of interest. Though the simulated data represented flood inundation depth at the location of interest, in order to determine the appropriate populations from which to simulate these flood inundation depths, a set of observed data from a location in Florida was examined and assessed. The

analysis of these data provided information as to the appropriate riverine flood depth population at the location in Florida. Discharge data from flow gage 02249007 on the Eau Gallie River at Heather Glen Circle at Melbourne were obtained. Information from this location, including elevations and channel characteristics, was also used to develop the simulation scenario. The population from which riverine flood depths were simulated was based on observed riverine flow gage discharge measurements translated to flood inundation depths at the location of interest. The process of transforming the discharges to inundation depths at the location of interest is explained here. The process involves two steps: (1) transposing the discharge from the gaged site to a site on the river adjacent to the point of interest and (2) transforming the river discharge to an inland flood depth.

The procedure used to calculate flood inundation depths at the location of interest in Florida, based on riverine discharge measurements taken by a local flow gage, is explained in detail in Chapter 5, which discusses the analyses conducted for two flood sources based on observed data. For the simulation study, based on information obtained from analysis of the location in Florida, the data generated represented flood inundation depths at the location of interest; thus, it was not necessary to carry out the procedure to transpose and transform simulated discharge measurements to flood inundation depth measurements. Once a sample of annual maximum flood inundation depths was generated, an appropriate distribution was fitted to the simulated data. Fitting a marginal distribution to the annual maximum events will be discussed in more detail in Section 4.2.3.1.3.

4.2.3.1.2. Calculation of Tidally-Caused Flood Inundation Depths

For the simulation studies, tidal flood inundation depths at the location of interest were simulated. In order to determine the appropriate distribution parameters to use as the basis for the simulation studies, the following procedure was applied to tidal gage measurements for a location in Florida. The tidal gage from which measurements were obtained was 8721604, located near Trident Pier. This tidal gage is approximately 20 miles from the location of interest, which is not ideal, but a closer location was not available.

The procedure used to calculate flood inundation depths at the location of interest in Florida, based on tide height measurements taken by a local tidal gage, is explained in detail in Chapter 5, which discusses the analyses conducted using observed data from two flood sources. For the simulation study, based on information obtained from analysis of the location in Florida, data generated represented flood inundation depths at the location of interest; thus, it was not necessary to carry out the procedure to calculate simulated tidal flood inundation depths from tidal gage measurements. Once a sample of annual maximum flood inundation depths was generated, an appropriate population was fitted to the simulated data. Fitting a marginal distribution to the annual maximum events will be discussed in more detail in Section 4.2.3.1.3.

4.2.3.1.3. Calculation of Flood Frequency Curves for Each Source Individually

To develop a flood frequency curve for a single flood source requires identification of the annual maximum flood series. The annual maximum events were ranked from largest to smallest, with the largest magnitude assigned to a rank of

1 and the smallest magnitude a rank of n. After developing a ranked list of annual maximum flood inundation depths at the location of interest caused by an individual flood source, the Weibull probability equation was used to determine the exceedance probability of each event. The Weibull plotting position formula is:

$$P_i = \frac{i}{n+1} \tag{4-1}$$

where P_i is the exceedance probability, i is the rank of the given flood depth, and n is the number of sample values in the record. The flood depths were plotted against the non-exceedance probability, calculated as 1.0 minus the exceedance probability, which provides the probability of the flood depth not being equaled or exceeded in a given year. Based on the pairs of data (non-exceedance probability, flood depth), an equation was fitted that would provide non-exceedance probabilities for any flood inundation depth value.

Though Bulletin 17B recommends the Log-Pearson Type III distribution be used to model riverine flood flows and the GEV distribution is frequently used to model tidal flood heights, clear evidence that either one of these distributions provides greater accuracy than that provided by the other distribution does not exist. Further, the variable of interest in this research is flood depth at a location of interest, rather than discharge or tidal height; the typical distributions may not be the most appropriate distributions. Thus, several distribution functions were fitted to the data to determine the most appropriate distribution for each data set. All tested distributions were plotted against the simulated data for visual comparison and probability plot correlation coefficients were calculated to assess the goodness-of-fit

of each distribution. For simulated riverine data, the distributions tested included the log-Pearson type III, lognormal, gamma, and Weibull distributions. For simulated tidal data, the distributions tested included the extreme value, generalized extreme value, and Rayleigh distributions. This procedure was first conducted with the observed data from Florida, in order to determine appropriate populations from which to generate samples for the simulation study. It was also used in the simulation studies to calculate the distributions best fitting the generated samples.

4.2.3.1.4. Calculation of Flood Frequency Curves Considering Two Flood Sources

To develop a joint flood frequency curve requires calculation of the marginal distributions, or individual flood frequency curves that correspond to the two individual flood sources. The typical process for conducting a flood frequency analysis, outlined in section 4.2.3.1.3., was followed to determine the marginals, with one exception. When calculating the flood frequency curve for an individual flood source, the annual maximum flood event caused by the individual source was identified. However, when two potentially interacting flood sources must be considered, it is possible that the maximum flood event experienced at the location of interest in a given year may not be caused by one individual source, but instead by a combination of the two sources. Thus, to identify the annual maximum flood depth at the location of interest, the simulated riverine and tidally-caused flood inundation depths at the location of interest for corresponding time periods were summed. This provided the total flood depth that occurred at the location of interest for each time period. The assumption that the contributing flood depths from each source could be summed was a convenient assumption to make for the purpose of conducting this

research; however, this methodology was not determined based on a technical analysis of the optimum approach for combining flood sources. This was completed first using the observed data from the location of interest in Florida, in order to identify the appropriate riverine and tidal flood frequency distributions. The same process was also used with the generated samples of riverine and tidally-caused flood inundation depths at the location of interest, as part of the simulation studies.

The simulated data represented corresponding annual maximum flood depths; however, it was necessary to use the observed data from the location in Florida in order to determine the appropriate populations from which to generate the samples. Using the observed data, it was necessary to identify the annual maximum events while considering both potential flood sources. As each of the gages from which observed data were obtained took hourly measurements each day, these measurements were summed and a single daily maximum flood depth at the location of interest was identified. Next, the annual maximum flood depth was identified based on the daily maximum events. This does not suggest that the annual maximum event could only occur through a combination of the two sources; this procedure could still identify an annual maximum event that was caused by only one source individually. The riverine contributions to the annual maximum flood inundation depth and the tidal contributions were separated into two flood series and flood frequency curves were developed for each. These flood frequency curves, based on the observed data from Florida, served to identify the appropriate populations from which samples of flood inundation depths should be generated for the simulation

study. When the same process was used with generated samples, the flood frequency curves served as the marginals, which were inputs to the copula procedure.

4.2.3.2. Description of Results

The results of fitting marginal distributions to samples of varying size and correlation are presented below. Any trends identified in the marginal distribution parameters are also discussed.

4.2.3.2.1. Effects of Varying Sample Size

Marginal distributions were fitted to generated riverine and tidal samples to assess the impact of sample size on the marginal distributions. Tables 4-1 through 4-4 provide the fitted parameters and the probability plot correlation coefficients for the riverine and tidal samples for each sample size investigated. Figures 4-1, 4-3, 4-5, and 4-7 illustrate the four cumulative distributions plotted against the generated riverine flood inundation depths for each sample size, while Figures 4-2, 4-4, 4-6, and 4-8 provide the same information for the generated tidal flood inundation depths for each sample size. Results and implications will be discussed for each sample size after the tables and figures have been presented.

Table 4-1: Parameter Values for Distributions Fitted to a Sample Size of 10 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	9.1654	-0.0017	0.5875	0.6916	N/A	0.2355	N/A
Scale	-0.5648	1.7099	4.8801	2.1642	2.0086	0.7648	2.2262
Location	5.1748	N/A	N/A	N/A	3.6407	2.0978	N/A
Probability Plot Correlation Coefficient	0.9460	0.9763	0.9156	0.9328	0.7553	0.8986	0.8374

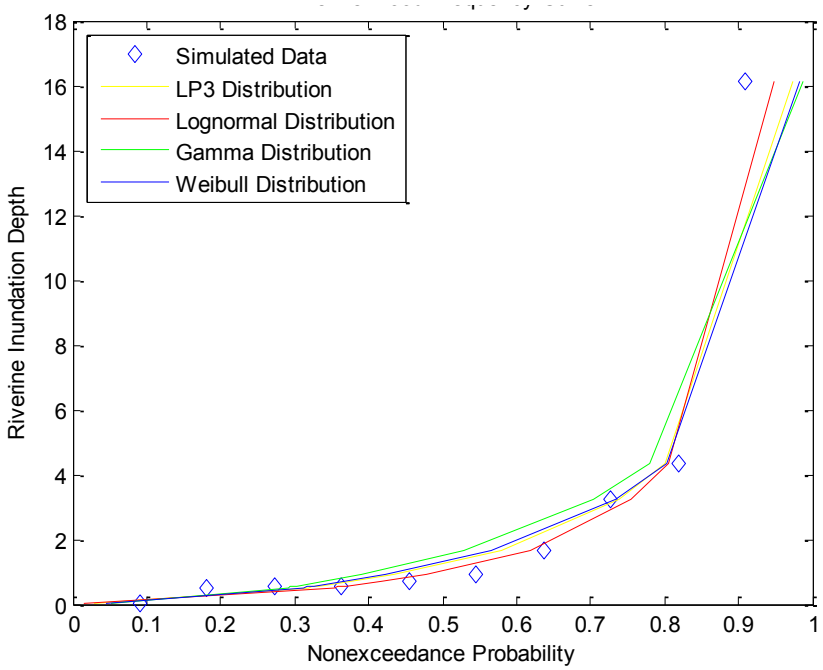

Figure 4-1: Comparison of Marginal Distributions Fitted to Riverine Sample Size of 10, where Riverine Inundation Depth Units are in Feet

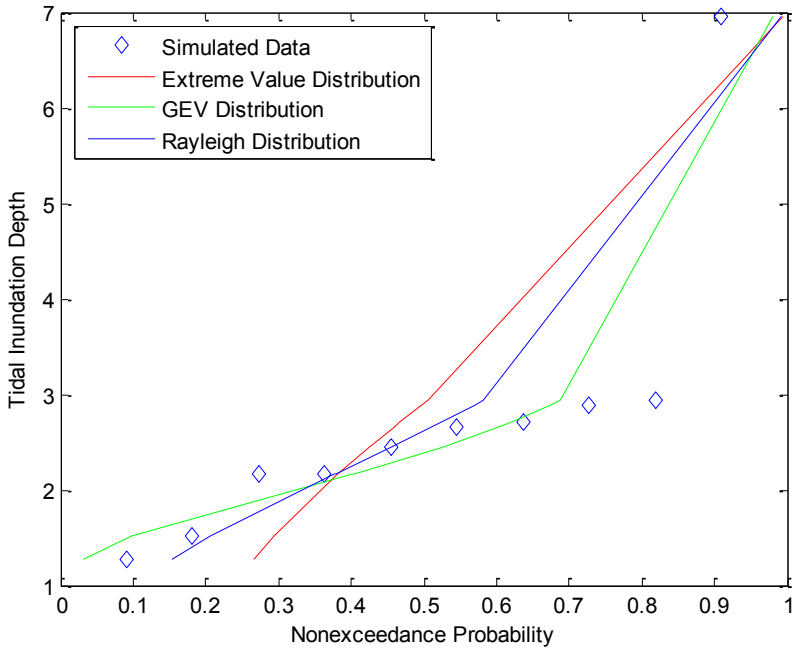

Figure 4-2: Comparison of Marginal Distributions Fitted to Tidal Sample Size of 10, where Tidal Inundation Depth Units are in Feet

Table 4-2: Parameter Values for Distributions Fitted to a Sample Size of 25 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	17.0895	0.3160	0.6155	0.7134	N/A	-0.0476	N/A
Scale	-0.3925	1.6227	6.0585	2.9447	1.6908	1.0514	2.1206
Location	7.0243	N/A	N/A	N/A	3.4265	2.1591	N/A
Probability Plot Correlation Coefficient	0.9262	0.8943	0.9362	0.9356	0.8745	0.9570	0.9421

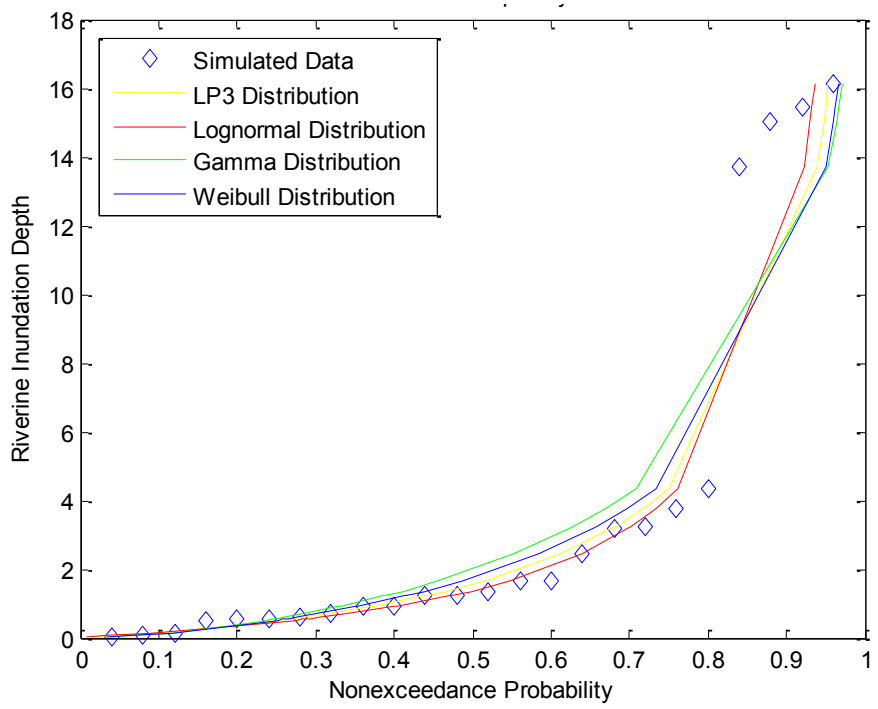

Figure 4-3: Comparison of Marginal Distributions Fitted to Riverine Sample Size of 25, where Riverine Inundation Depth Units are in Feet

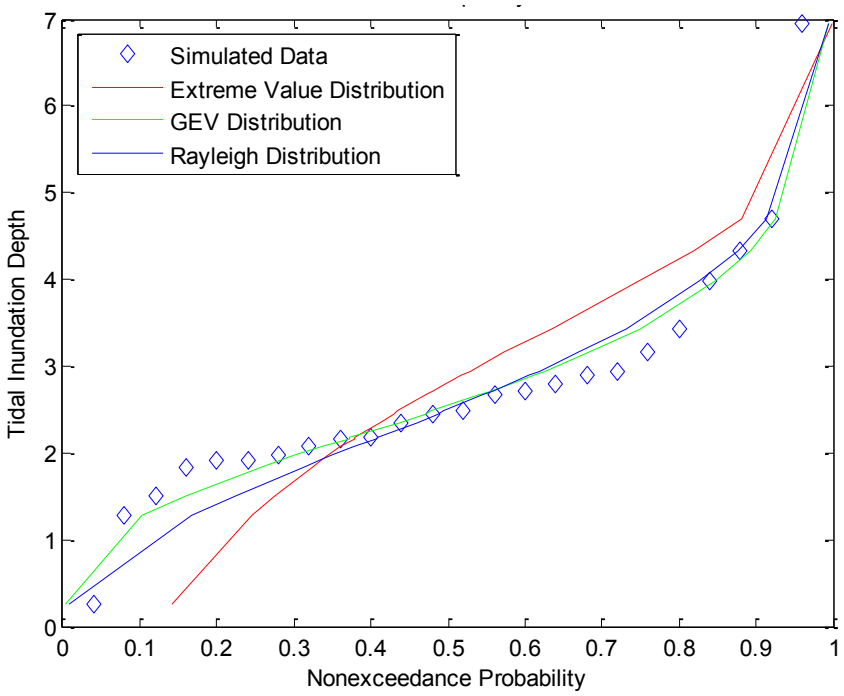

Figure 4-4: Comparison of Marginal Distributions Fitted to Tidal Sample Size of 25, where Tidal Inundation Depth Units are in Feet

Table 4-3: Parameter Values for Distributions Fitted to a Sample Size of 50 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	35.7042	0.0879	0.7043	0.7725	N/A	-0.0484	N/A
Scale	-0.2379	1.4212	3.6542	2.1634	1.4939	0.8731	1.9524
Location	8.5803	N/A	N/A	N/A	3.1461	2.0969	N/A
Probability Plot Correlation Coefficient	0.9463	0.9401	0.9245	0.9352	0.8759	0.9640	0.9485

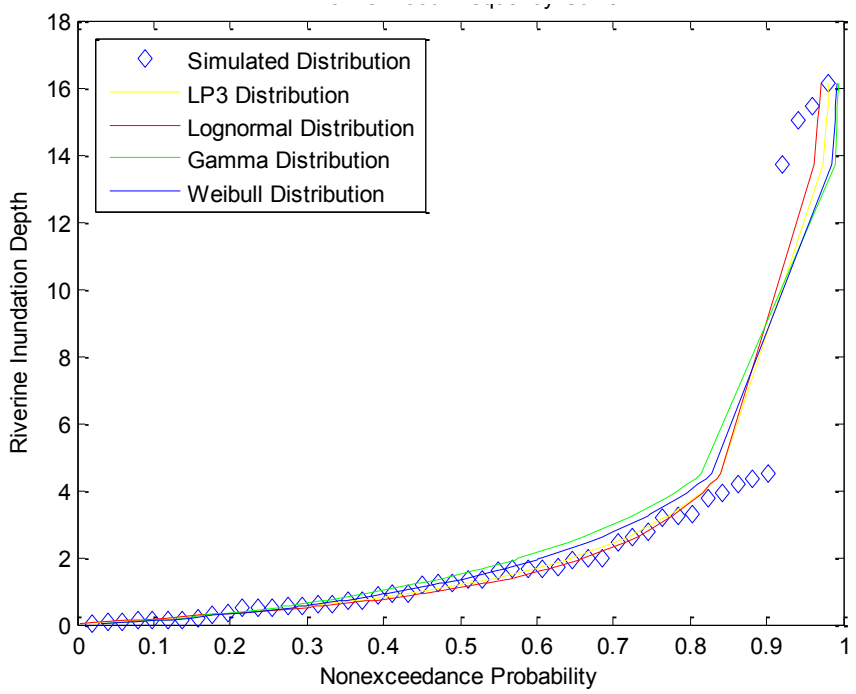

Figure 4-5: Comparison of Marginal Distributions Fitted to Riverine Sample Size of 50, where Riverine Inundation Depth Units are in Feet

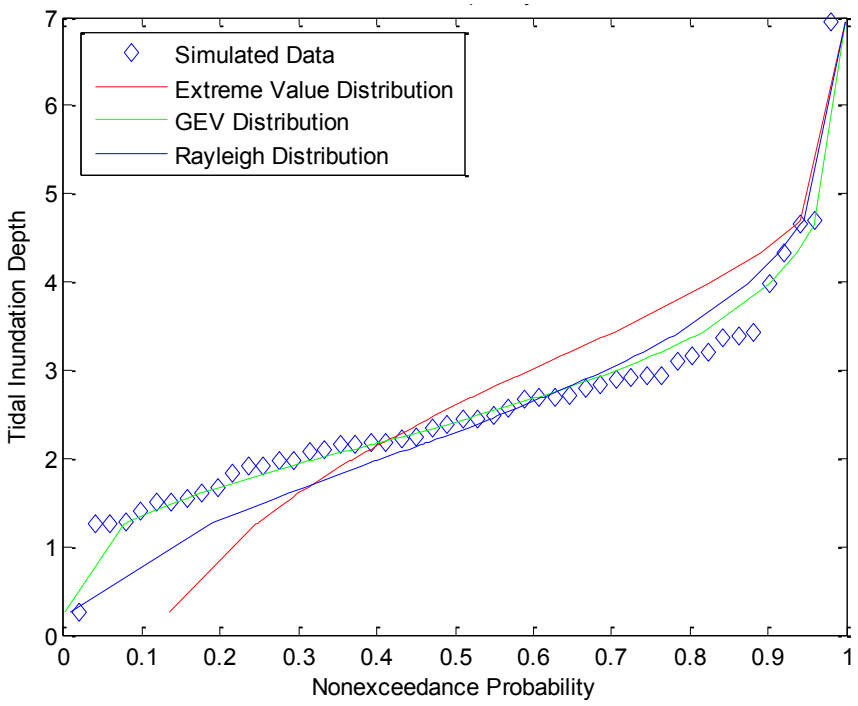

Figure 4-6: Comparison of Marginal Distributions Fitted to Tidal Sample Size of 50, where Tidal Inundation Depth Units are in Feet

Table 4-4: Parameter Values for Distributions Fitted to a Sample Size of 100 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	24.0706	-0.1268	0.5376	0.6544	N/A	0.0062	N/A
Scale	-0.3597	1.7647	5.2704	2.0462	1.5096	0.9328	1.9614
Location	8.5311	N/A	N/A	N/A	3.1572	1.9514	N/A
Probability Plot Correlation Coefficient	0.8983	0.8272	0.9375	0.932	0.8934	0.9968	0.9801

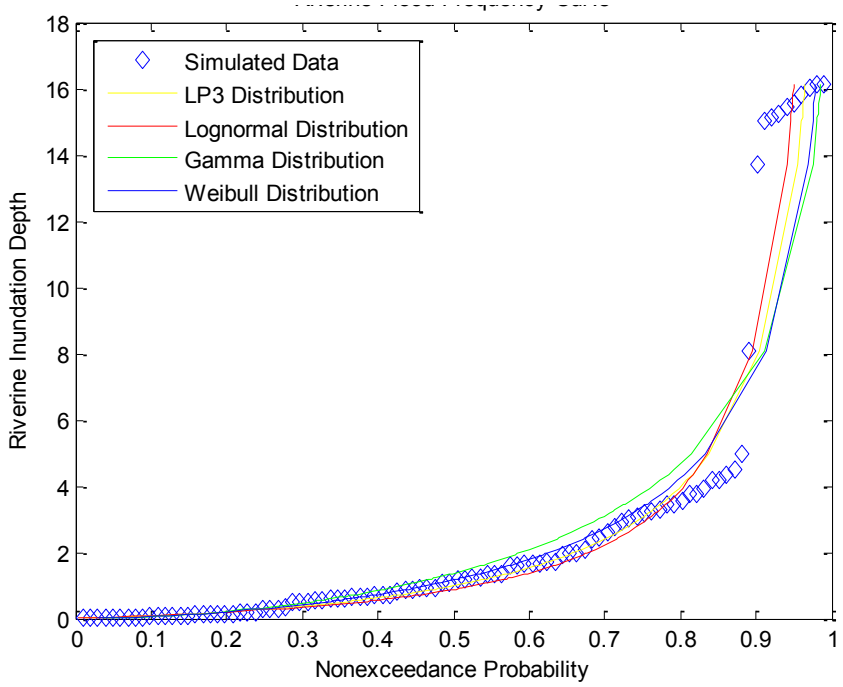

Figure 4-7: Comparison of Marginal Distributions Fitted to Riverine Sample Size of 100, where Riverine Inundation Depth Units are in Feet

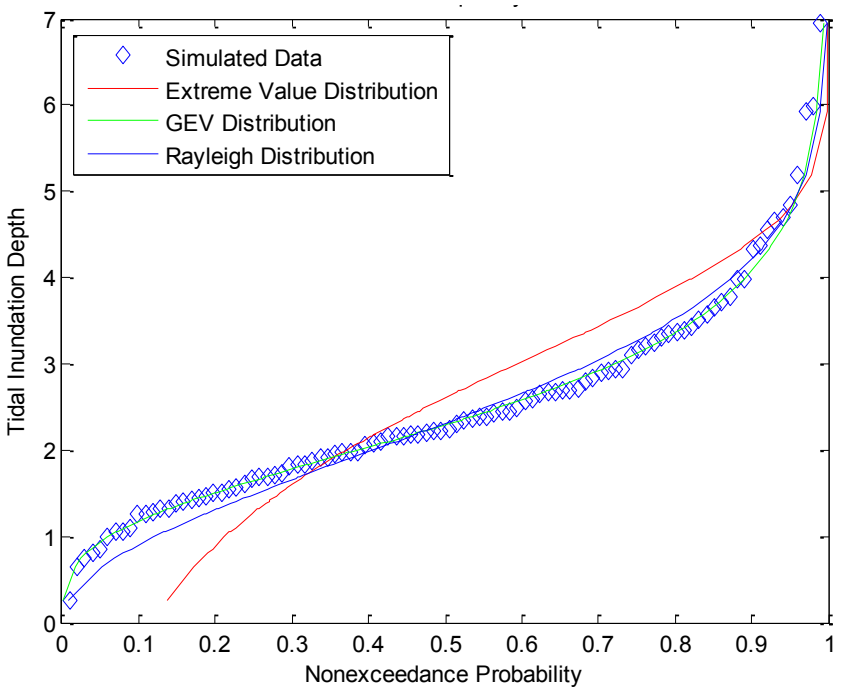

Figure 4-8: Comparison of Marginal Distributions Fitted to Tidal Sample Size of 100, where Tidal Inundation Depth Units are in Feet

The results of fitting probability distributions to the marginal riverine and tidal samples are not significantly impacted by the sample size. For the riverine samples, the probability plot correlation coefficients were typically above 0.90 for all four of the distributions, indicating that any of the distributions would fit the generated data reasonably well. In comparing the fitted distributions against generated data in Figures 4-1, 4-3, 4-5, and 4-7, very little difference in performance between the distributions considered was evident. The four distributions vary only minimally in representing the lower sample points and vary slightly more in representing the larger generated flood depths. Though any of the distributions could adequately represent the riverine data, the gamma distribution was previously determined (to be presented in Chapter 5) to be the most appropriate distribution to represent the observed data. To maintain consistency with the analysis of the observed data set, and because the difference in performance of the distributions was so minimal, the gamma distribution

was selected for further use in representing the generated riverine flood depth samples for all sample sizes.

More variation was typically observed in the performance of the three distributions considered to represent the generated tidal samples. The probability plot correlations calculated for these samples indicated that the GEV and Rayleigh distributions typically would be best able to represent the samples. The probability plot correlation coefficients calculated for the GEV distribution was always above 0.89, and was above 0.95 for the larger sample sizes. For the Rayleigh distribution, the probability plot correlation coefficient was low for a sample size of 10, only 0.84, but above 0.94 for the larger samples. The difference in performance between these two distributions was typically very small, but the probability plot correlation coefficients and Figures 4-2, 4-4, 4-6, and 4-8 all indicate that the GEV distribution was the best choice of distribution. For the smaller sample sizes, the extreme value distribution was unable to fit the data; however, as the sample size increased the extreme value distribution did become better able to fit the samples. However, the probability plot correlation coefficient reached a value of only 0.89 for the largest sample, and was as low as 0.76 for the smallest sample. For all four sample sizes, the GEV distribution was selected for further use in representing the generated samples.

For both the riverine and the tidal samples, it is clear that the small sample size limits the ability of any of the distributions to provide an excellent fit to the sample. The figures presented for each sample size, not unexpectedly, become much smoother as the sample size increases, indicating that it was easier to fit distributions to the larger sample size. Further, the marginal distributions fitted to the larger

sample sizes can be expected to be more accurate than those fitted based on small sample sizes, as the larger samples better represent the population from which the sample was generated.

Overall, for each sample size, the marginal distributions that were fitted provided good fits to the generated riverine and tidal samples. The parameters of the marginal distributions were evaluated for any trends that could be attributed to the variation in sample size; but such trends were not found to be evident. Though the sample size certainly impacted the fitted distributions, and the accuracy of the fitted distributions, specific trends could not be attributed to the marginal distributions as sample size varied. This suggests that the sample size did not have a systematic impact on the marginal distributions that were fitted to the generated riverine and tidal flood depth samples.

4.2.3.2.2. Effects of Varying Sample Correlation

The effect of correlation between the generated riverine and tidal samples on the marginal distributions was also investigated. Tables 4-5, 4-7, 4-9, and 4-11 provide the calculated parameters for each distribution fitted to the riverine and tidal flood depth samples generated with a correlation coefficient of 0.27, 0.31, 0.37, and 0.60, as well as the probability plot correlation coefficient calculated for each. Figure 4-9 provides a plot of each distribution fitted to the riverine flood depth sample for a correlation of 0.27. Because the method of generating correlated samples only changed the generated tidal flood depths as the correlation coefficient varied, Figure 4-9 also corresponds to each of the three other scenarios discussed. The riverine plot will not presented again as it does not vary with varying levels of correlation. Figures

4-10, 4-11, 4-12, and 4-13 provide plots of each distribution fitted to the tidal flood depth samples for each level of correlation. All tables and figures will be presented and then the results and implications will be discussed.

Table 4-5: Parameter Values for Distributions Fitted to a Sample with a Correlation Coefficient of 0.27 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	35.7042	0.0879	0.7043	0.7725	N/A	-0.0615	N/A
Scale	-0.2379	1.4212	3.6542	2.1634	1.4846	0.9213	1.9846
Location	8.5803	N/A	N/A	N/A	3.1896	2.1171	N/A
Probability Plot Correlation Coefficient	0.9436	0.9401	0.9245	0.9352	0.8973	0.9703	0.9595

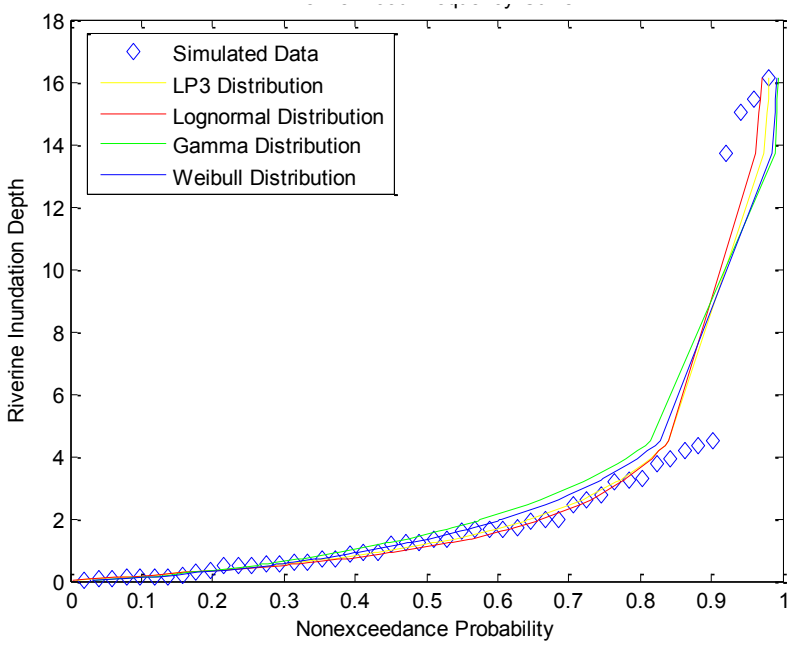

Figure 4-9: Comparison of Marginal Distributions Fitted to Riverine Sample for all Correlation Coefficients Studied, where Riverine Inundation Depth Units are in Feet

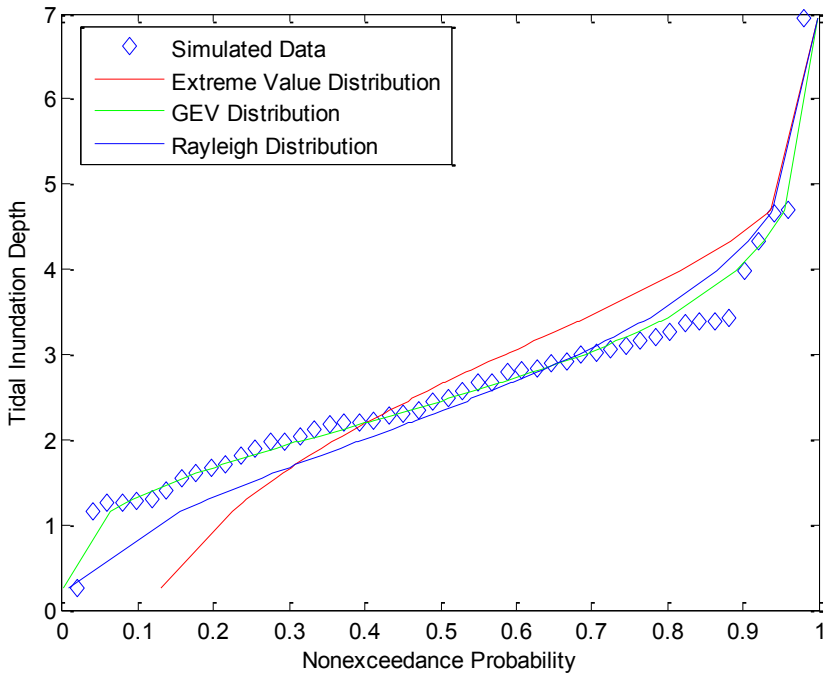

Figure 4-10: Comparison of Marginal Distributions Fitted to Tidal Sample with Correlation Coefficient of 0.27, where Tidal Inundation Depth Units are in Feet

Table 4-6: Parameter Values for Distributions Fitted to a Sample with a Correlation Coefficient of 0.31 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	35.7042	0.0879	0.7043	0.7725	N/A	-0.0408	N/A
Scale	-0.2379	1.4212	3.6542	2.1634	1.5005	0.8392	1.9252
Location	8.5803	N/A	N/A	N/A	3.1084	2.0782	N/A
Probability Plot Correlation Coefficient	0.9436	0.9401	0.9245	0.9352	0.8575	0.9568	0.9375

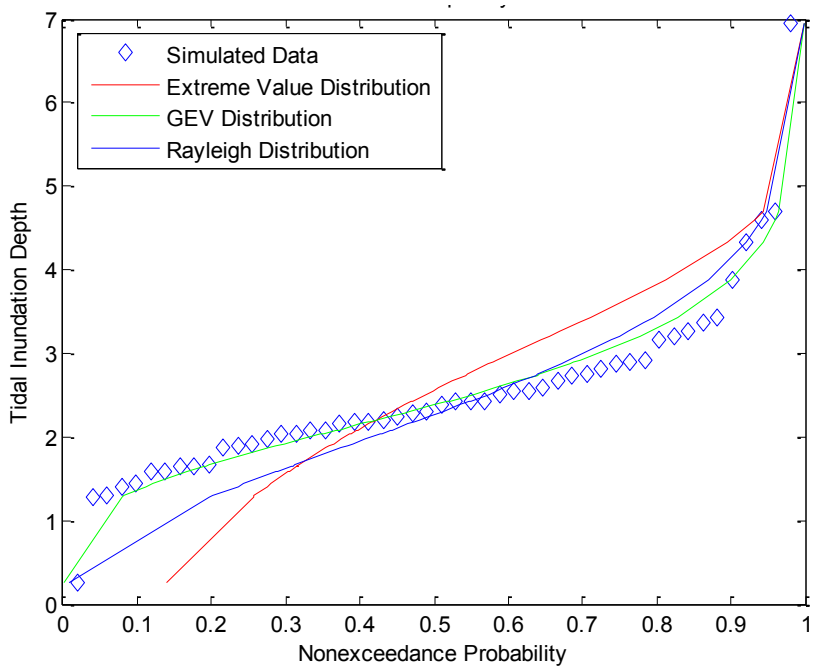

Figure 4-11: Comparison of Marginal Distributions Fitted to Tidal Sample with Correlation Coefficient of 0.31, where Tidal Inundation Depth Units are in Feet

Table 4-7: Parameter Values for Distributions Fitted to a Sample with a Correlation Coefficient of 0.37 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	35.7042	0.0879	0.7043	0.7725	N/A	-0.0099	N/A
Scale	-0.2379	1.4212	3.6542	2.1634	1.0050	0.7534	1.8860
Location	8.5803	N/A	N/A	N/A	3.0503	2.0599	N/A
Probability Plot Correlation Coefficient	0.9463	0.9401	0.9245	0.9352	0.8427	0.9529	0.9273

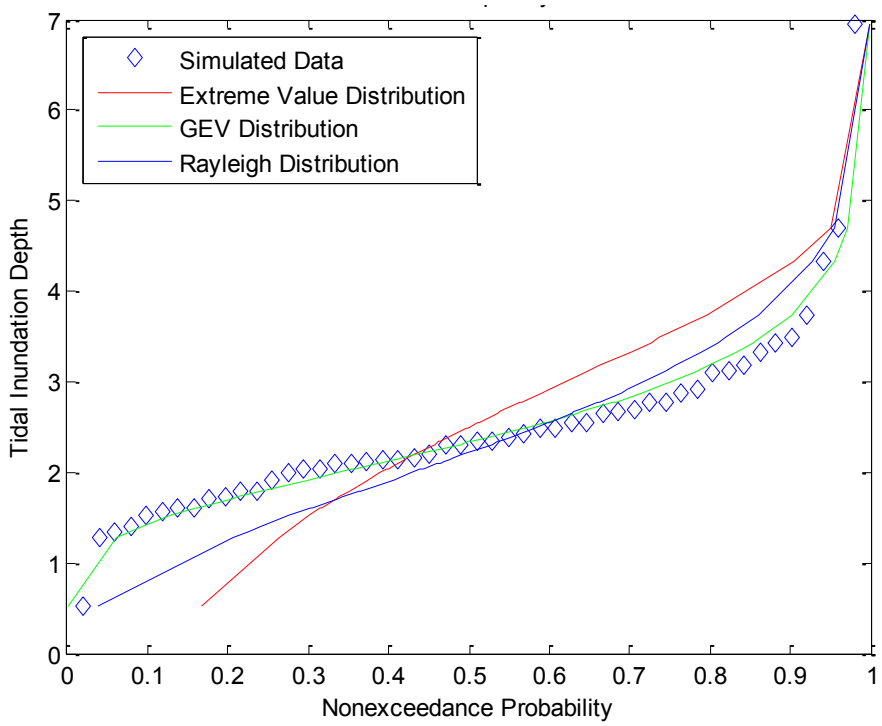

Figure 4-12: Comparison of Marginal Distributions Fitted to Tidal Sample With Correlation Coefficient of 0.37, where Tidal Inundation Depth Units are in Feet

Table 4-8: Parameter Values for Distributions Fitted to a Sample with a Correlation Coefficient of 0.60 and Corresponding Probability Plot Correlation Coefficients

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	35.7042	0.0879	0.7043	0.7725	N/A	0.0324	N/A
Scale	-0.2379	1.4212	3.6542	2.1634	1.0484	0.6002	1.7898
Location	8.5803	N/A	N/A	N/A	2.8484	2.0331	N/A
Probability Plot Correlation Coefficient	0.9463	0.9401	0.9245	0.9352	0.8788	0.9823	0.9647

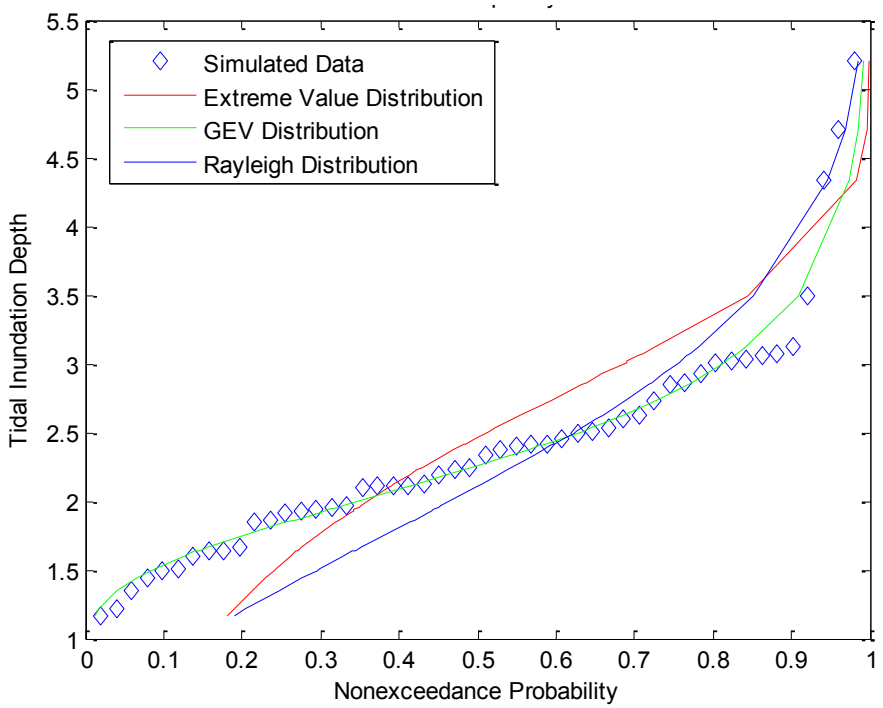

Figure 4-13: Comparison of Marginal Distributions Fitted to Tidal Sample with Correlation Coefficient of 0.60, where Tidal Inundation Depth Units are in Feet

The information presented in the tables and figures enable the goodness of fit of the potential marginal distributions to be investigated for each level of correlation. Good fits to the generated samples were observed in all cases. The distributions fitted to the riverine flood inundation depth sample all perform almost equally well, as was observed when the effects of varying sample size were investigated. Though the gamma distribution does not provide the highest probability plot correlation coefficient of the distributions representing the riverine sample, the differences in performance between the distributions were so minimal that it can safely be used to represent the riverine data. In order to remain consistent with the observed data set, which was observed to be best fit by the gamma distribution, the gamma distribution was selected for use in this scenario.

In the process of generating correlated samples, the riverine sample was held constant and the tidal sample was varied in such a way as to vary the level of correlation between the two samples. For all four levels of correlation investigated, both the probability plot correlation coefficients and Figure 4-10 through 4-13 suggest that both the GEV and Rayleigh distributions would be able to represent the samples well. The probability plot correlation coefficients for the GEV distribution were above 0.95 in all cases and above 0.93 for the Rayleigh distribution in all cases. However, from the figures, the GEV distribution is clearly superior to the Rayleigh distribution. The extreme value distribution was not observed to result in a good fit to the generated samples for any level of correlation. The probability plot correlation coefficients calculated for this distribution did not exceed 0.89 for any of the four levels of correlation, and the plots indicated that it was unable to fit the generated data. Based on these results, the GEV distribution was selected for further use in representing the generated tidal sample in all cases.

It is also interesting to compare the results of fitting distributions to the tidal data sets for the varying levels of correlations. In comparing Tables 4-5 and 4-6 and Figures 4-10 and 4-11, it appears that the level of correlation has little impact on the fitted distributions, though this is probably due to the fact that correlations of 0.27 and 0.31 are not significantly different. As the level of correlation between the samples increases further, some trends in the marginal distributions fitted to the tidal flood depth samples became evident, which are likely attributable to the change in correlation coefficient. The GEV shape parameter increased, and the scale and location parameters decreased as the level of correlation increased. Further, the

probability plot correlation coefficient slightly decreased, suggesting that as the correlation between the samples increases, it become more difficult to fit a distribution to the tidal flood depth sample. These changes are likely due to changes in the generated tidal flood inundation depth samples, in order to maintain higher degrees of correlation with the riverine flood inundation depth sample, though it is possible that sampling variation could be the cause.

4.2.4. Use of Copula Equations to Develop Joint Distributions

The next step in the process was to develop joint distributions for flood inundation depths from the riverine and tidal sources. The joint distributions provided probabilities of specific combinations of riverine and tidal flood depths. The following sections will discuss the methods used to develop the joint distributions and present the results.

4.2.4.1. Description of Methods

Once the marginals had been developed based on the generated riverine and tidally-caused flood inundation depths at the location of interest, a copula could be used to determine the joint probability of depths for the two flood sources. The first step was to calculate the appropriate dependence structure between the two flood sources. Typically, Kendall's τ is used, though other measures of association or dependence may also be used. For bivariate scenarios, using Kendall's τ is particularly convenient because a direct relationship exists between τ and the copula parameter for Archimedean copulas. Before Kendall's τ could be selected for use, it first had to be determined that it could serve as a measure of hydrologic dependence. Kendall's τ is a measure of statistical dependence, but hydrologic dependency was the

factor of interest to this research. Hydrologic dependence was defined as floods from multiple sources that occur simultaneously or being attributed to a common underlying cause. Because Kendall's τ is a measure of how often both variables under consideration experience high or low values at the same time and how often one variable experiences a high value while the other variable experiences a low value, it was determined that Kendall's τ could be used as a measure of the hydrologic dependence between the flood sources. Thus, Kendall's τ was calculated between the generated samples of riverine and tidally-caused flood inundation depths.

Once the measure of dependence was calculated, the appropriate family of copulas to fit to the data had to be determined. Numerous families of copulas are available from which to choose. Several families were fitted and compared using Akaike's Information Criteria (AIC), as described in Chapter 1. When AIC values are calculated, the smallest AIC value indicates the most appropriate copula family (Klein *et al.*, 2010).

The Gumbel, Clayton, and Frank copula families were chosen for comparison, as they are three of the more popular Archimedean families. For each family, the copula parameter was calculated as a function of Kendall's τ, a measure of association between the two data sets. The equations relating tau and the copula parameters for the Gumbel-Hougaard, Clayton, and Frank families, respectively, are as follows:

$$\tau = 1 - \alpha_{GH}^{-1} \tag{4-2a}$$

$$\tau = \frac{\alpha_C}{\alpha_C + 2} \tag{4-2b}$$

$$\tau = 1 - \frac{4}{\alpha_F}\left(D_1(-\alpha_F) - 1\right) \tag{4-2c}$$

where α is the copula parameter, τ is Kendall's τ, and D_1 is the first-order Debye function. The first-order Debye function is calculated as:

$$D_1(-\alpha) = \frac{1}{-\alpha^1}\int_0^{-\alpha} \frac{t^1\, dt}{e^t - 1} \tag{4-3}$$

Next, the copula cumulative distribution functions for the Gumbel-Hougaard, Clayton, and Frank families, respectively, were calculated as:

$$C(u, v) = \exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\} \tag{4-4a}$$

$$C(u, v) = (u^{-\alpha} + v^{-\alpha} - 1) \tag{4-4b}$$

$$C(u, v) = -\frac{1}{\alpha}\ln\left[1 + \frac{(\exp(-\alpha u) - 1)(\exp(-\alpha v) - 1)}{\exp(-\alpha) - 1}\right] \tag{4-4c}$$

where α was the copula parameter and u and v were the marginal distributions, which were calculated from the cumulative distribution functions of the original random variables, riverine and tidal flood depths. The calculated copulas provided the cumulative joint distribution of combinations of generated riverine and tidally-caused flood inundation depths for the location of interest. Joint pdfs that correspond to the variables u and v were calculated by taking the second derivative of the joint cdf with

respect to both u and v. To obtain joint pdf values that correspond to the riverine and tidal variables required that the joint pdfs calculated for the variables u and v then be multiplied by the marginal distribution pdf values (Wang *et al.*, 2009). The equations for the joint pdfs, expressed in terms of the riverine and tidal variables, are as follows:

$$f(x,y) = \frac{\partial^2 C(u,v)}{\partial u \partial v}\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} \quad (4\text{-}5)$$

where u and v represent the marginal distributions fitted to the riverine and tidal flood depths and x and y represent the riverine and tidal flood depths, respectively.

4.2.4.2. Description of Results

The results of calculating joint distributions for each of the eight simulation studies are presented in the following sections. Based on the joint distributions, trends were identified and discussed based on sample size and sample correlation.

4.2.4.2.1. Effects of Varying Sample Size

The following sections will discuss the results of the use of copulas to develop joint distributions for each of the four sample sizes investigated. The results of fitting the copula equations and identifying the most appropriate copula family for each scenario will first be provided. Next, plots of the joint distributions are presented and discussed.

4.2.4.2.1.1. Copula Fitting to Develop Joint Distributions

To develop joint distributions, three Archimedean copula families were fitted to the marginals previously identified for each sample size. The best copula family for each sample size was chosen based on calculation of Akaike's Information

Criterion. Table 4-9 provides Kendall's tau for each sample size, the calculated copula family parameters, and Akaike's Information Criterion calculated for each copula family.

Table 4-9: Calculated Copula Parameters (Alpha) and Akaike's Information Criteria for Each Family and Sample Size

N	Tau	Gumbel-Hougaard Alpha	Gumbel-Hougaard AIC	Clayton Alpha	Clayton AIC	Frank Alpha	Frank AIC
10	-0.1111	N/A	N/A	N/A	N/A	-1.0101	1.6862
25	-0.0362	N/A	N/A	N/A	N/A	-0.3264	2.2656
50	-0.0612	N/A	N/A	N/A	N/A	-0.5527	2.3538
100	0.2132	1.2709	-13.8725	0.5418	-8.464	1.9925	-10.1514

The calculation of the copula parameter is dependent on the calculated value of Kendall's τ. In cases where Kendall's τ has a negative value, only the Frank copula family can be used to develop the joint distribution. For sample sizes of 10, 25, and 50, negative values of Kendall's τ were calculated, so the Frank family was the only family that could be used in these scenarios. Table 4-9 demonstrates that larger values (in magnitude) of Kendall's τ correspond to larger magnitudes of copula parameter as well. An investigation of the sampling distribution of the copula parameters revealed that the copula parameters had a fairly wide standard error, such that the differences in copula parameters reported in Table 4-9 are likely not significant. The sample of 10 riverine and tidal flood depth values was calculated using a copula parameter value of 2.0, which was approximately the copula parameter calculated for the sample of 100. The difference in joint cdf values between the original analysis and this analysis with a modified copula parameter ranged from 1% to 7%, which indicates minimal impact overall. The difference in joint pdf values

between the original and modified copula parameter was typically only 1% to 3%; however, the difference in the peak was a more significant 15%. These results indicate that the sample size did not have a significant impact on the calculated copula parameters. The copula parameters calculated for the smallest and largest samples resulted in the calculation of nearly identical joint distributions for the sample. Thus, at least for this scenario, it would appear that a reasonable joint distribution can be calculated based on a small sample size. However, strong conclusions on the impact of sample size on the development of joint distributions cannot be made based on a single scenario.

For a sample size of 100 a positive value of Kendall's τ was calculated. For this scenario, all three of the copula families could then be evaluated. The magnitude of Kendall's τ was larger than for the smaller sample sizes, which resulted in a larger magnitude of copula parameter than calculated for the smaller sample sizes. It can also be noted that the sign of the copula parameter corresponds to the sign of Kendall's τ. The most appropriate copula family can be determined based on the calculated AIC values for this scenario. The Gumbel-Hougaard copula results in the lowest AIC value, suggesting that this is the most appropriate copula family for use in developing joint distributions for this scenario. The sign of the calculated value of Kendall's τ was not believed to be a function of sample size, as Kendall's τ is calculated based on the ranks of the data sets.

4.2.4.2.1.2. Plotting Joint Distributions Through Copulas

To examine the joint distributions, both pdfs and cdfs of the joint distributions were plotted using all three copula families, if possible, for each sample size. This

enabled a visual assessment of the differences that would exist in the joint distributions depending on the choice of copula family. Plots of both pdf and cdf are presented for sample sizes of 10, 50, and 100, to illustrate several points about the joint distributions. Little difference between the plots for sample sizes of 10, 25 and 50 were evident so the plots for a sample size of 25 were not presented.

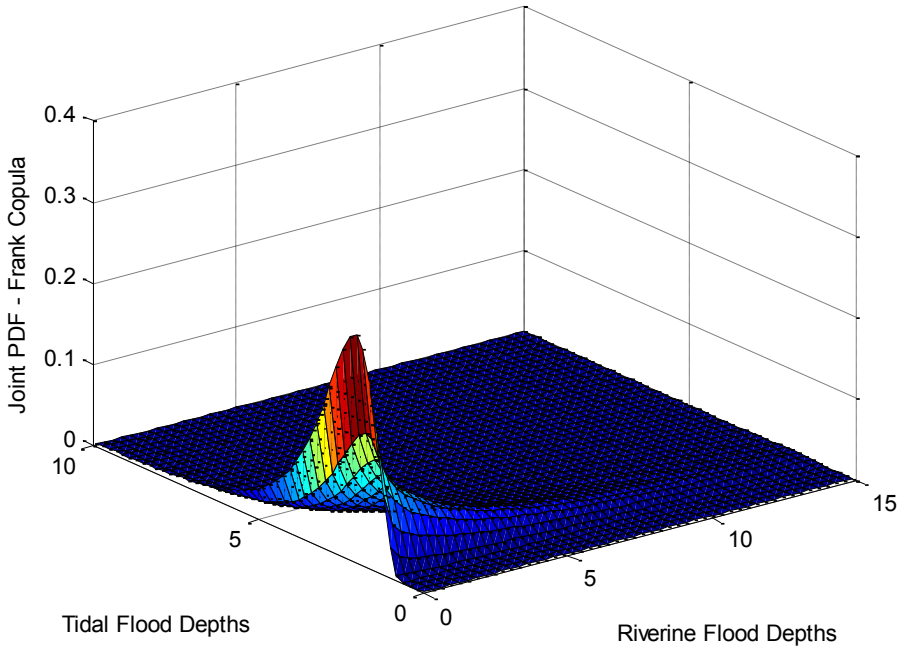

Figure 4-14: Joint PDF Calculated by the Frank Copula Family for a Sample Size of 10, where Flood Depths are in Feet

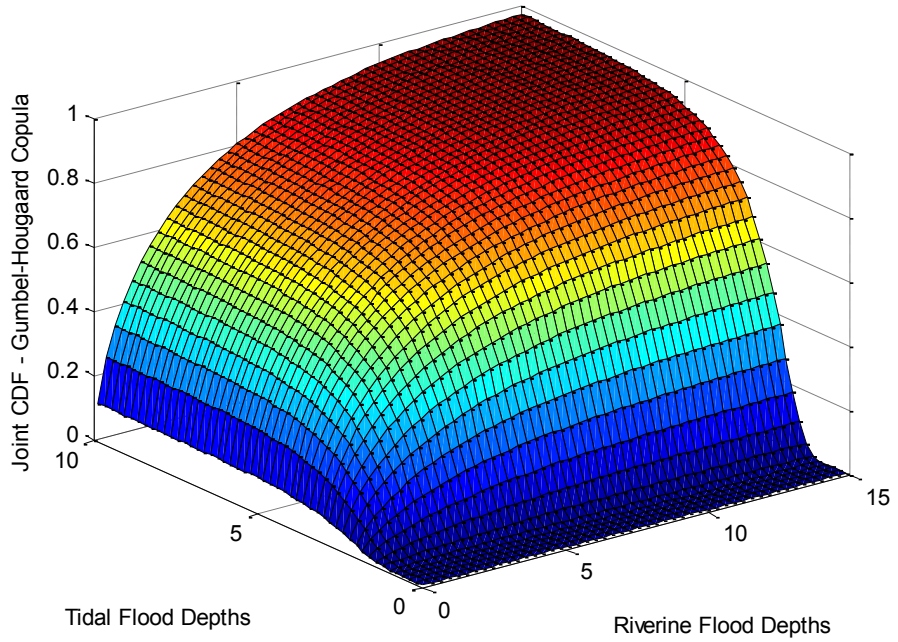

Figure 4-15: Joint CDF Calculated by the Frank Copula Family for a Sample Size of 10, where Flood Depths are in Feet

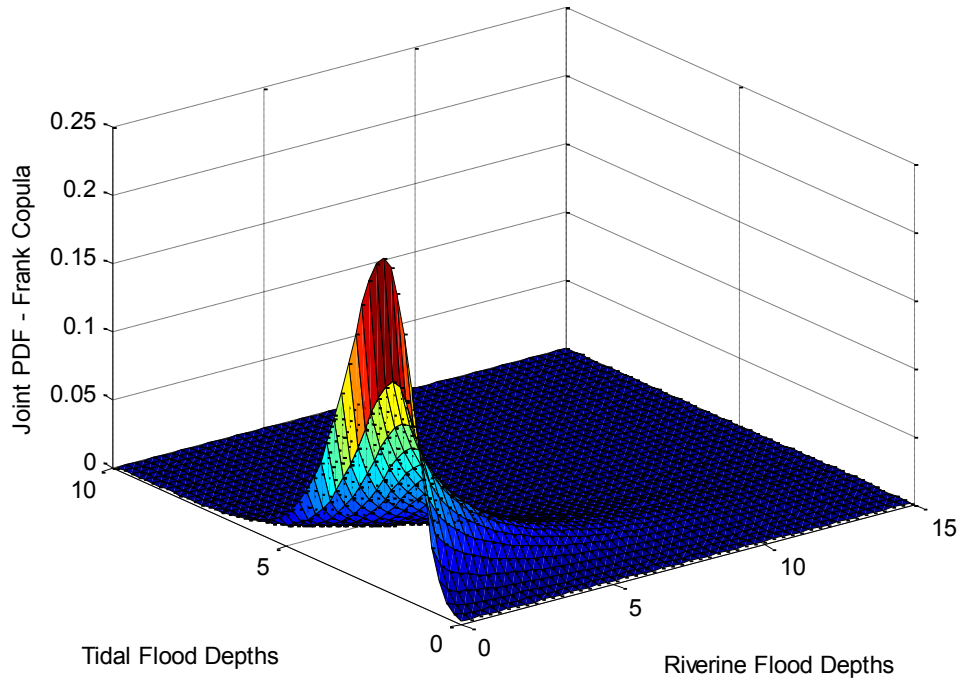

Figure 4-16: Joint PDF Calculated by the Frank Family for a Sample Size of 50, where Flood Depths are in Feet

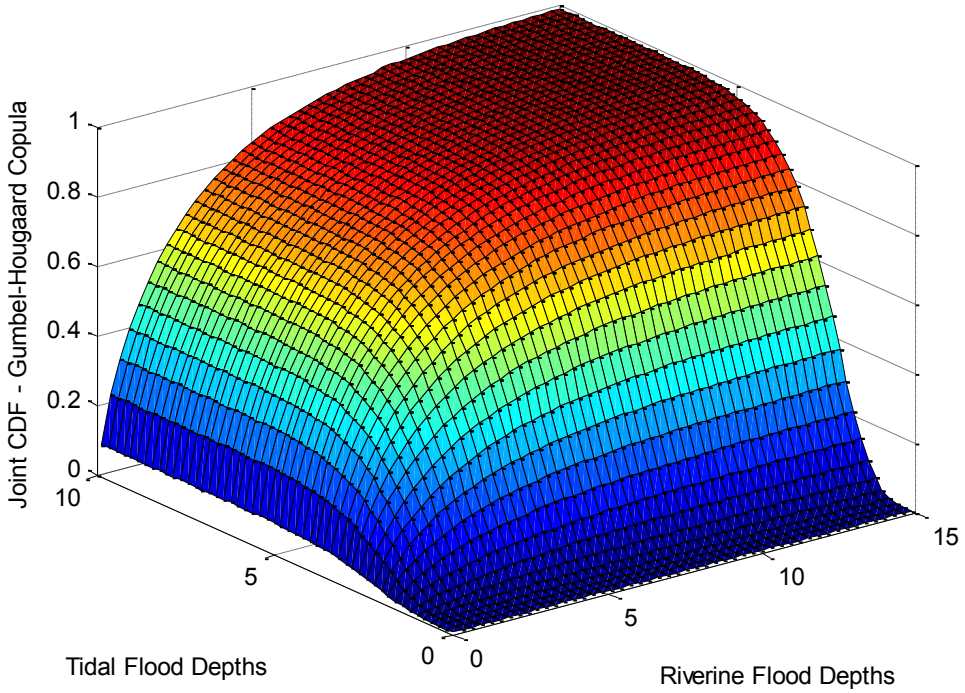

Figure 4-17: Joint CDF Calculated by the Frank Family for a Sample Size of 50, where Flood Depths are in Feet

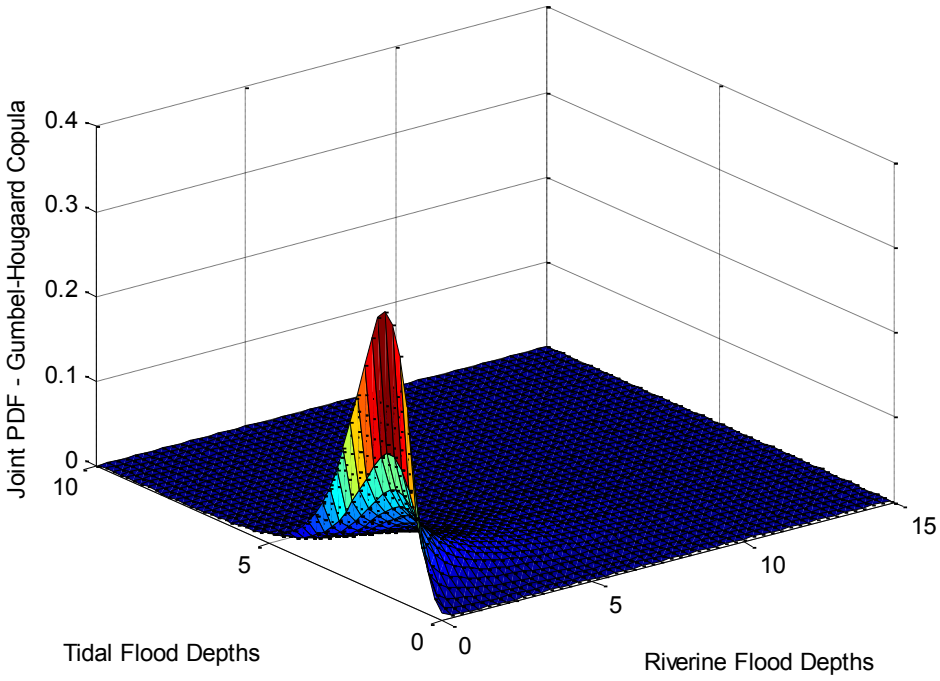

Figure 4-18: Joint PDF Calculated by the Gumbel-Hougaard Copula Family for a Sample Size of 100, where Flood Depths are in Feet

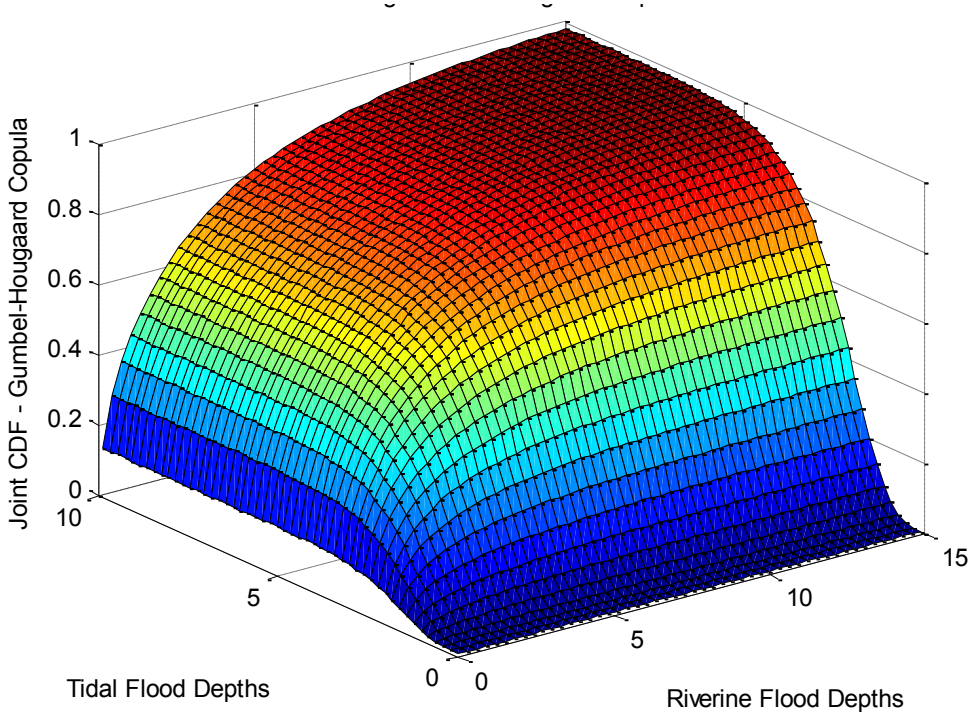

Figure 4-19: Joint CDF Calculated by Gumbel-Hougaard Copula Family for a Sample Size of 100 , where Flood Depths are in Feet

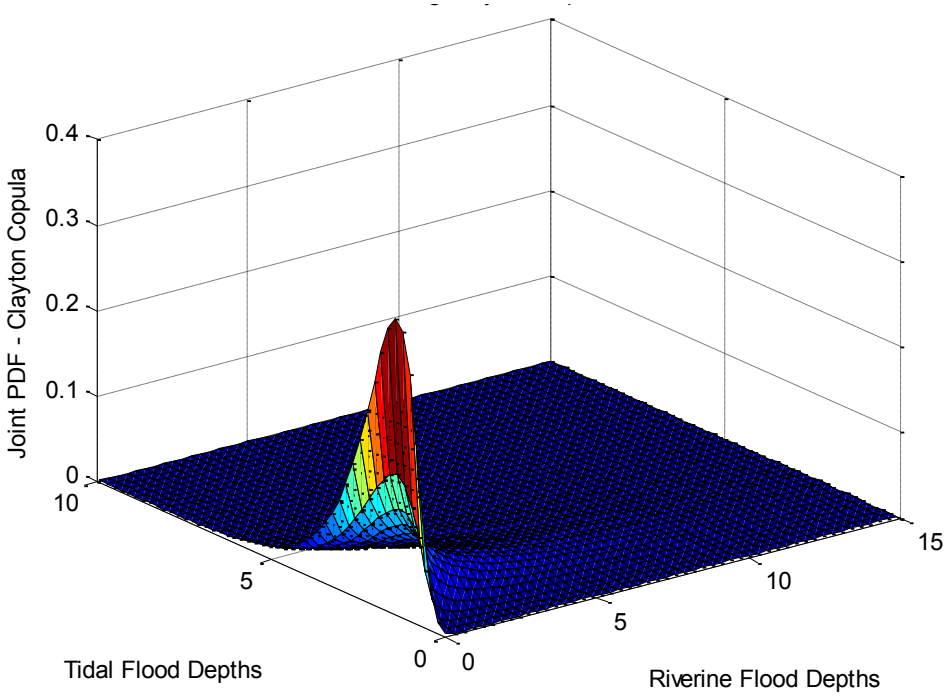

Figure 4-20: Joint PDF Calculated by the Clayton Copula Family for a Sample Size of 100, where Flood Depths are in Feet

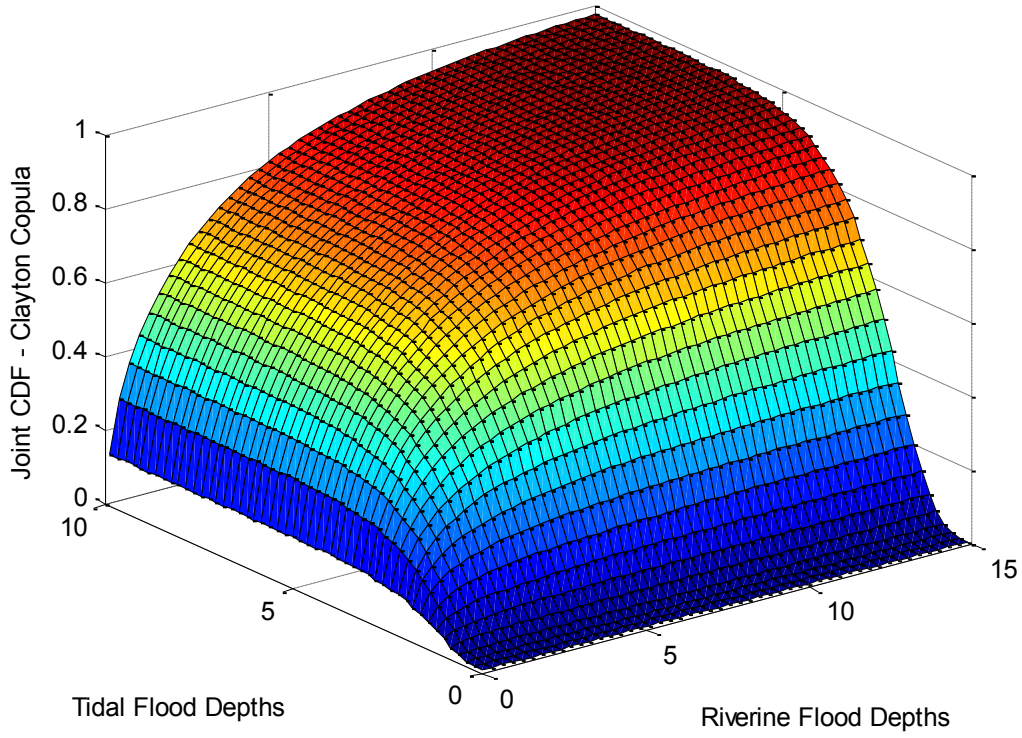

Figure 4-21: Joint CDF Calculated by Clayton Copula Family for a Sample Size of 100, where Flood Depths are in Feet

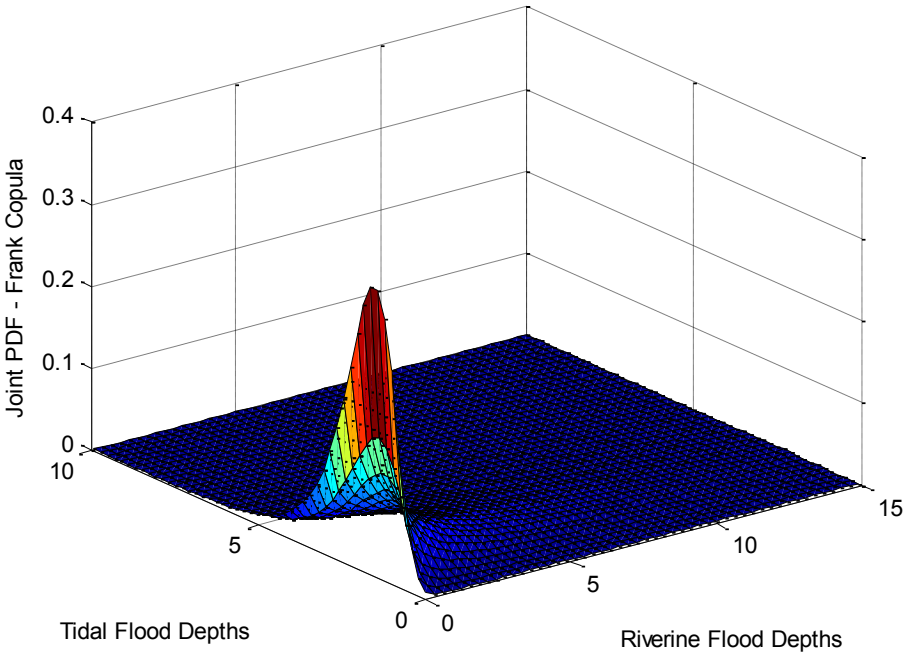

Figure 4-22: Joint PDF Calculated by the Frank Copula Family for a Sample Size of 100, where Flood Depths are in Feet

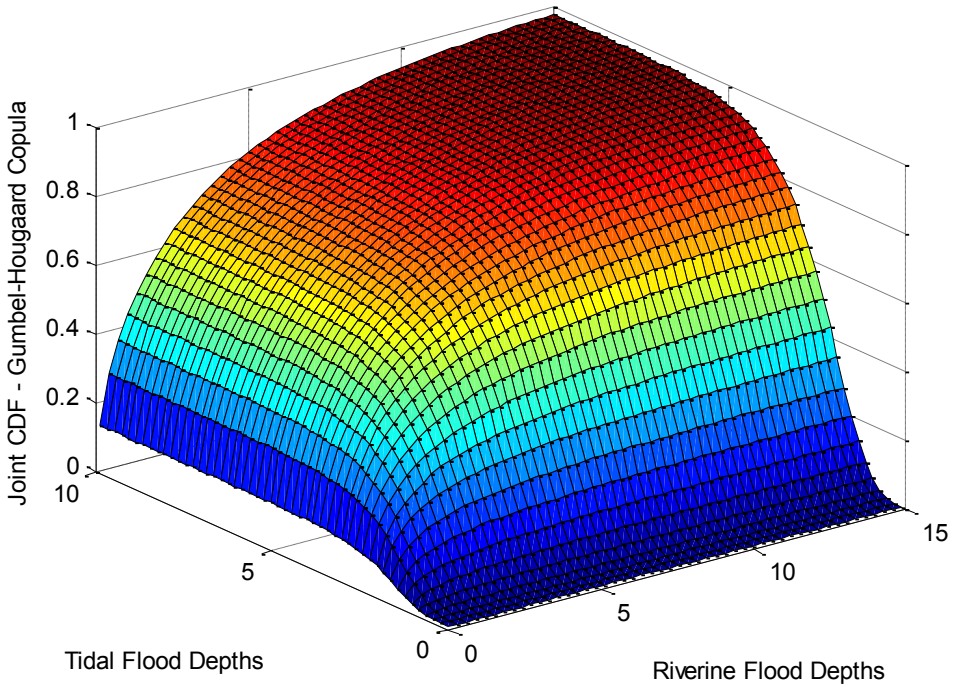

Figure 4-23: Joint CDF Calculated by Frank Copula Family for a Sample Size of 100, where Flood Depths are in Feet

The joint distributions created for each simulation scenario serve as an intermediate step to developing a comprehensive flood frequency analysis for two flood sources. In Figures 4-14, 4-16, and 4-18 the joint probability distributions calculated for sample sizes of 10, 50, and 100, are provided while Figure 4-15, 4-17, and 4-19 provide the corresponding joint cumulative distributions. These figures illustrate several key points about the joint distributions. The peaks of the joint pdfs for all sample sizes occur at a moderate tidal flood depth value and a very low riverine flood value. Given the values of the riverine and tidal generated samples, this location of the pdf peak seems reasonable. The results of these studies are based on only one location. For different locations, different relationships would exist between the flood sources, which may lead to differences in the shapes and locations of the joint peaks. Though the riverine samples typically included a few large flood

depths, the samples were overall very low. Differences in the shapes of the joint pdfs are not evident as the sample size increases. This can be explained by the similarity of the copula parameters calculated for each of the sample sizes. Though the individual riverine and tidal samples generated for each sample size differed, they were all generated from the same underlying populations, and they all had reasonably similar levels of Kendall's τ. Thus, the joint distributions calculated for each of the four sample sizes were very similar, as can be observed in the figures. This suggests that reasonable joint distributions could be developed with only a small amount of data available, at least for this scenario. However, this may not be the case for other scenarios. For most hydrologic applications, small samples are all that are available, but this limited amount of available data remains a significant challenge to hydrologic analyses. Due to the high variability typically evident in hydrologic data, analyses based on small samples are generally not considered to be highly accurate or reliable.

Though specific differences are not clearly evident in the joint pdf figures, a comparison of the joint pdf values calculated for the first ten samples of generated riverine and tidal flood depths can be made, as these ten samples were the same for all four of the sample sizes generated. Differences were evident in the joint pdf values calculated for each combination of riverine and tidal flood depths, with the difference appearing to be more significant in the peak of the distribution. Though these small differences can be attributed to differences in the copula parameter for each sample size, specific trends that relate to the varying sample size were not evident. However, it can be expected that the joint distributions developed for larger

sample sizes would be more accurate than the joint distributions developed for smaller sample sizes, as larger sample sizes can better define the population.

The joint cdfs (Figures 4-15, 4-17, and 4-19) also had very similar appearances despite the varying sample sizes. The shape and steepness of the joint cdf curve can be explained by the riverine and tidal flood depth samples. The riverine samples tended to have primarily very low depth values with just a few higher values. This results in a very steep rise parallel to the riverine axis over the lower flood depths followed by a more gently sloping rise toward the highest flood depths. The tidal flood depth samples did not typically have near zero values; thus, the slope parallel to the tidal axis does not begin until the flood depths reach values of 1 or 2 feet. The tidal samples also typically did not have depths greater than approximately 6 feet; therefore, the slope parallel to the tidal axis rises fairly steeply through that range of flood depths, and then becomes more gently sloping in the range of flood depths that were unlikely to occur. In examining the joint cdf values calculated for the first ten sets of riverine and tidal samples for each of the sample sizes, differences in the cdf values ranging from 1% to 10% are evident, with large differences observed in the moderate probability range rather than the particularly low or particularly high probability values. This suggests that some variation in the shape of the joint cdf occurs as the sample size increases; however, these differences are small enough not to be visible when the joint cdfs are plotted. It would be expected that as the sample size increases the joint population would be better defined by the samples; thus, the joint cdfs calculated based on the larger sample size should be expected to be more accurate representations.

For a sample size of 100 the calculated value of Kendall's τ was positive, which meant that all three copula families could be used to determine joint distributions. Though the AIC value indicated that the Gumbel-Hougaard copula family would be the most appropriate to represent the sample, the Clayton and Frank families were also used to calculate joint distributions for this sample in order to determine the significance of the choice of copula family. Figures 4-18, 4-20, and 4-22 provide the joint pdfs calculated for a sample size of 100 using the Gumbel-Hougaard, Clayton, and Frank families, respectively, while Figures 4-19, 4-21, and 4-23 provide the joint cdfs calculated for a sample size of 100 for those same copula families. No significant difference was observed in either the joint pdfs or cdfs based on the choice of copula family. The general shapes and location of the peaks for the joint pdfs and the general shapes and slopes of the joint cdf surfaces all appear to be very similar between the three copula families. Further, an evaluation of the calculated joint pdf and cdf values for each combination of riverine and tidal flood depths in the sample revealed insignificant differences based on the choice of copula family. Given that the primary difference in the copula families is their representation of the dependence between the riverine and tidal flood depth samples, this result was not surprising, given that the sample had a fairly low level of dependence between the flood sources.

4.2.4.2.2. Effects of Varying Sample Correlation

The following sections will discuss the results of the use of copulas to develop joint distributions for each of the four levels of correlation between the marginals that was investigated. The results of fitting the copula equations and identifying the most

appropriate copula family for each scenario will first be provided. Next, plots of the joint distributions are presented and discussed.

4.2.4.2.2.1. Copula Fitting to Develop Joint Distributions

In order to develop joint distributions, three Archimedean copula families were fitted to the marginals previously identified for each degree of correlation between samples. The most appropriate copula family for each scenario was chosen based on Akaike's Information Criterion. Table 4-10 provides Kendall's τ calculated for each scenario, the parameter calculated for each copula family for each scenario, and the Akaike's Information Criterion calculated for each scenario.

Table 4-10: Calculated Copula Parameters (Alpha) and Akaike's Information Criteria for Each Family and Correlation Coefficient

Sample Correlation	Tau	Gumbel-Hougaard Alpha	Gumbel-Hougaard AIC	Clayton Alpha	Clayton AIC	Frank Alpha	Frank AIC
0.27	-0.1216	N/A	N/A	N/A	N/A	-1.1080	2.0115
0.31	0.0351	1.0364	0.2929	0.0728	2.6367	0.3162	1.4532
0.37	0.3192	1.4688	-4.9883	0.9376	4.8640	3.1382	-8.3930
0.60	0.8253	5.7243	-20.1704	9.4486	-70.2054	21.1133	-93.9415

For a sample correlation of 0.27, a negative value of Kendall's τ was calculated due to the ranks of the generated samples; thus, only Frank's family could be used to represent the joint distributions for this scenario. It should be noted that Kendall's τ appears to increase with increasing correlation between the samples, which is not surprising given that Kendall's τ is a measure of correlation. However, because Kendall's τ is calculated based on ranks of the samples, if a different set of random samples were generated with these same levels of correlation between them, it is likely that different values of Kendall's τ could result. As was noted in

examining the scenarios of varying sample size, the value of the calculated copula parameter has the same sign as Kendall's τ, and as τ increases the copula parameters also increase. Thus, as sample correlation increases, the copula parameters also increase.

As was discussed in the evaluation of the results for varying sample sizes, the copula parameter tends to have a fairly large standard error. For the lower levels of correlation differences in the copula parameters are not significant; however, for the higher levels of correlation the difference in copula parameters is significant. The joint distributions were calculated for the samples with a correlation of 0.60 while using the copula parameters originally calculated for a correlation of 0.31. The joint cdfs varied significantly based on the copula parameters used. The lowest and highest joint cumulative probabilities remained similar but otherwise, a difference of 10-20% in joint cumulative probabilities was evident for the same combinations of riverine and tidal samples. The lower copula parameters calculated for a correlation of 0.31 resulted in consistently smaller cumulative probabilities than the copula parameters originally calculated for a correlation of 0.60. The copula parameters were also observed to have a significant impact on the joint pdfs developed. The peak of the joint pdf was most impacted by the change in parameters, with the peak decreasing when the lower parameter values were used to calculate the joint pdfs. Thus, for higher levels of correlation between the riverine and tidal samples, the copula parameter likely has more impact on the joint distributions developed. As the purpose of the copula is to model dependencies between variables, it was expected

that there would be differences such as these in the joint distributions developed as the level of dependence between the variables increased.

The calculated AIC values were used to determine the most appropriate copula family to use to represent the joint distributions for the scenarios with correlations of 0.31, 0.37, and 0.60 between the generated riverine and tidal samples. For a sample correlation of 0.31 the Gumbel-Hougaard copula family produces the lowest AIC value, which suggests that the Gumbel-Hougaard copula should be used for this scenario. For sample correlations of 0.37 and 0.60 the Frank family produced the lowest AIC value, which suggests that the Frank family should be used to develop the joint distributions. The magnitude of the AIC values increases with increasing sample correlation, as does the difference in calculated AIC values for the three copula families, which suggests that the differences in the three copulas may become more significant as the correlation between the riverine and tidal samples increases. This would not be surprising as each of the copula families typically performs best for different levels of Kendall's τ. One factor that frequently influences the decision as to which copula family should be used to represent a given set of data is the level of dependence between that data set, as each family typically has a certain range of dependence over which it performs better than the others.

4.2.4.2.2.2. Plotting Joint Distributions Through Copulas

Once the appropriate copula family had been chosen for each scenario, the joint distributions could be plotted for visualization. Both pdfs and cdfs of the joint distributions were plotted using the copula families previously identified as most appropriate for each degree of correlation between the marginal distributions. Due to

similarities in results, the copula pdfs and cdfs will be presented for correlations of 27, 0.37, and 0.60 only. The plots for a correlation of 0.31 were very similar to those for a correlation of 0.27 so they will not be presented. The joint pdfs and cdfs calculated using the copula procedure were plotted against the generated riverine and tidal flood inundation depths. These figures provide an understanding of the likelihood of occurrence of any combination of riverine and tidal flood inundation depths. Figure 4-24 provides the joint pdf for a correlation of 0.27 using the Frank family and Figure 4-25 provides the joint cdf for this scenario. Figures 4-26, 4-28, and 4-30 provide the joint pdfs calculated for all three copula families for a correlation of 0.37 while Figures 4-27, 4-29, and 4-31 provide the joint cdfs. Figures 4-32, 4-34, and 4-36 provide joint pdfs calculated for all the copula families for a correlation of 0.60 and Figures 4-33, 4-35, and 4-37 provide joint pdfs.

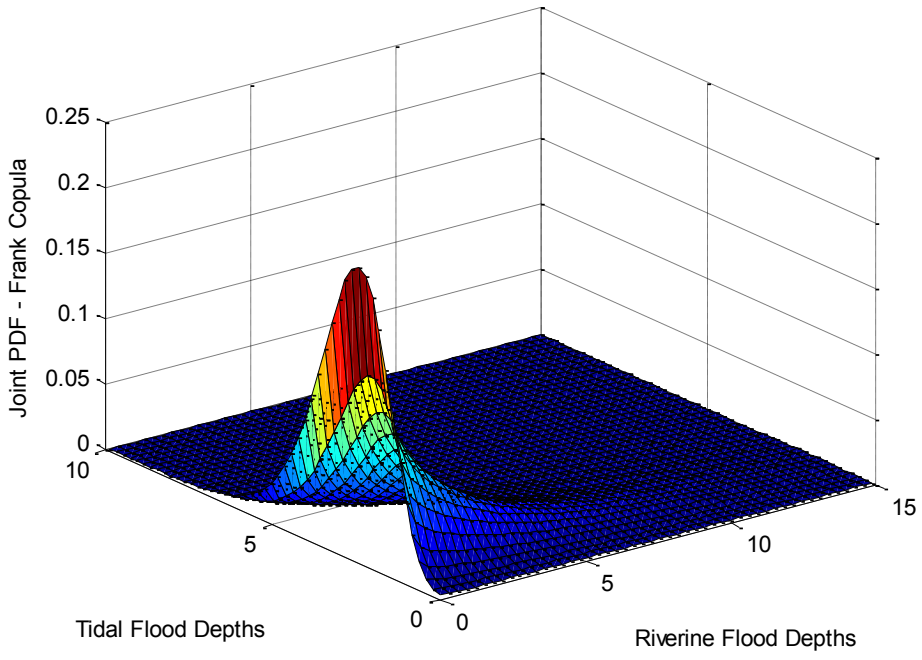

Figure 4-24: Joint PDF Calculated by Frank Family for Correlation Coefficient of 0.27, where the Flood Depths are in Feet

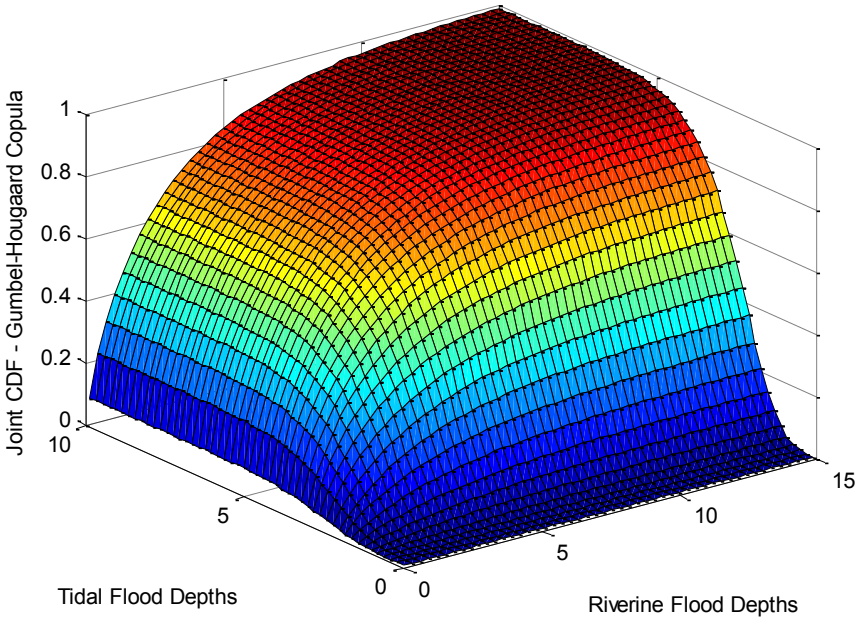

Figure 4-25: Joint CDF Calculated by Frank Family for Correlation Coefficient of 0.27, where the Flood Depths are in Feet

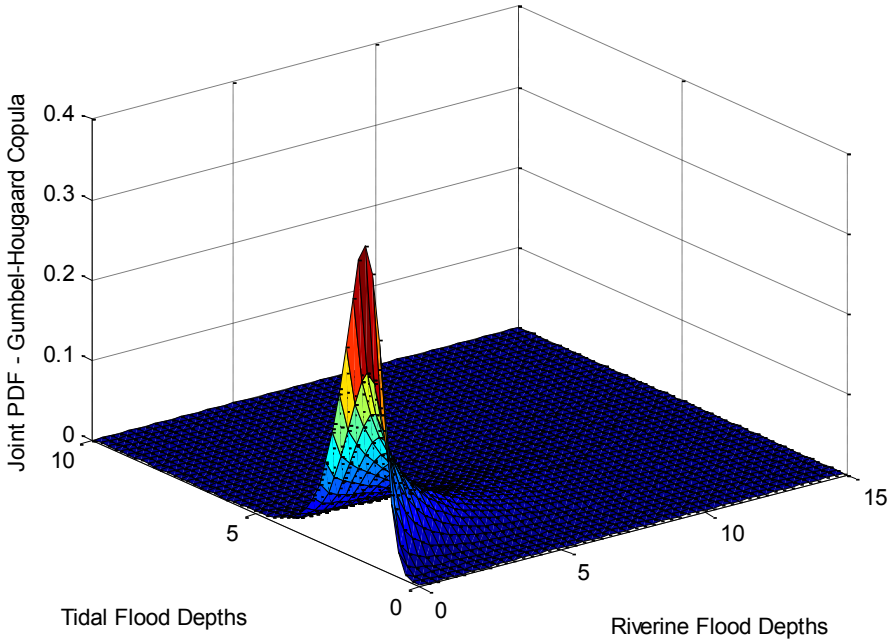

Figure 4-26: Joint PDF Calculated by Gumbel-Hougaard Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

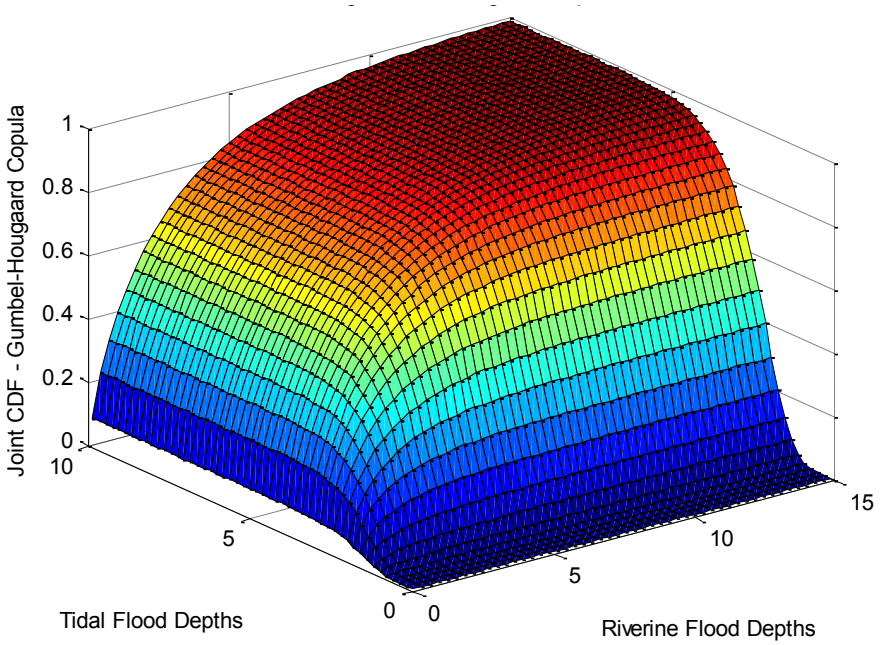

Figure 4-27: Joint CDF Calculated by Gumbel-Hougaard Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

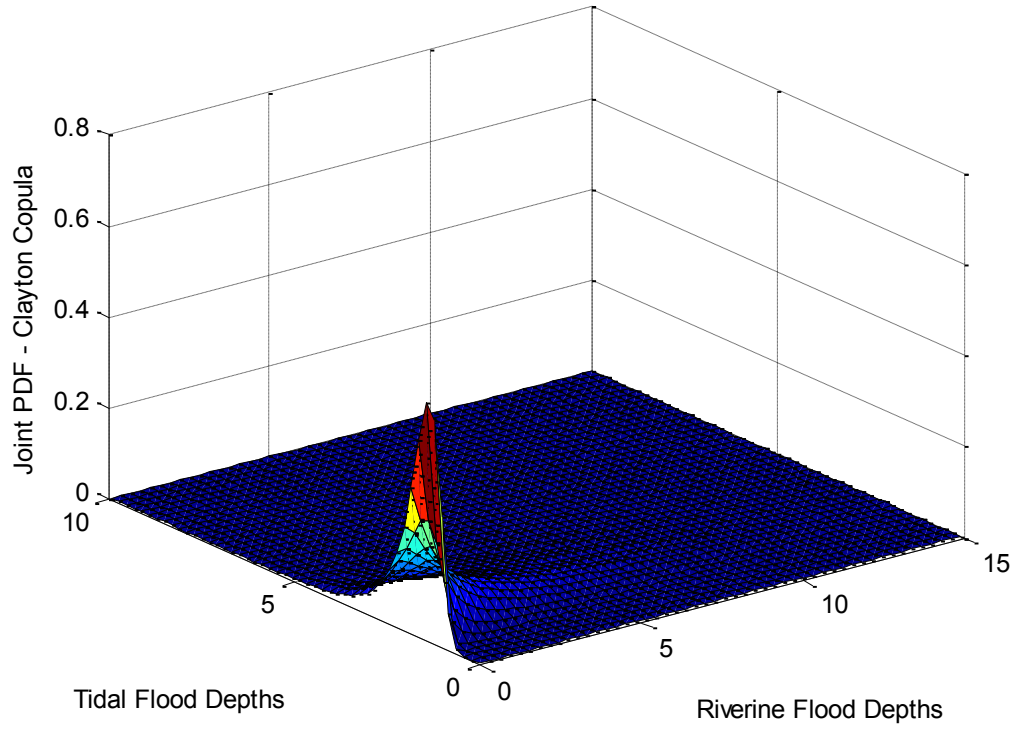

Figure 4-28: Joint PDF Calculated by Clayton Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

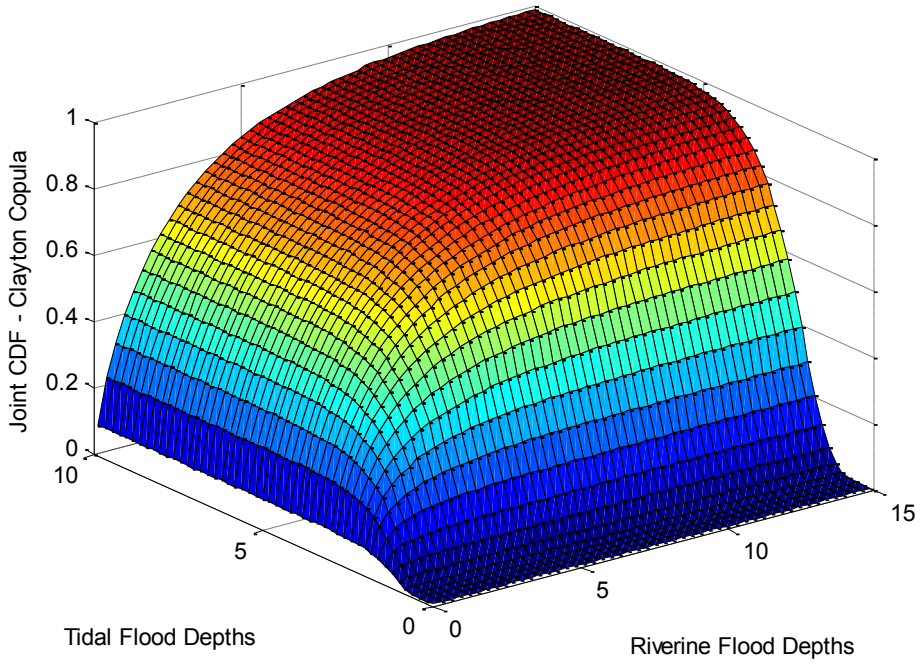

Figure 4-29: Joint CDF Calculated by Clayton Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

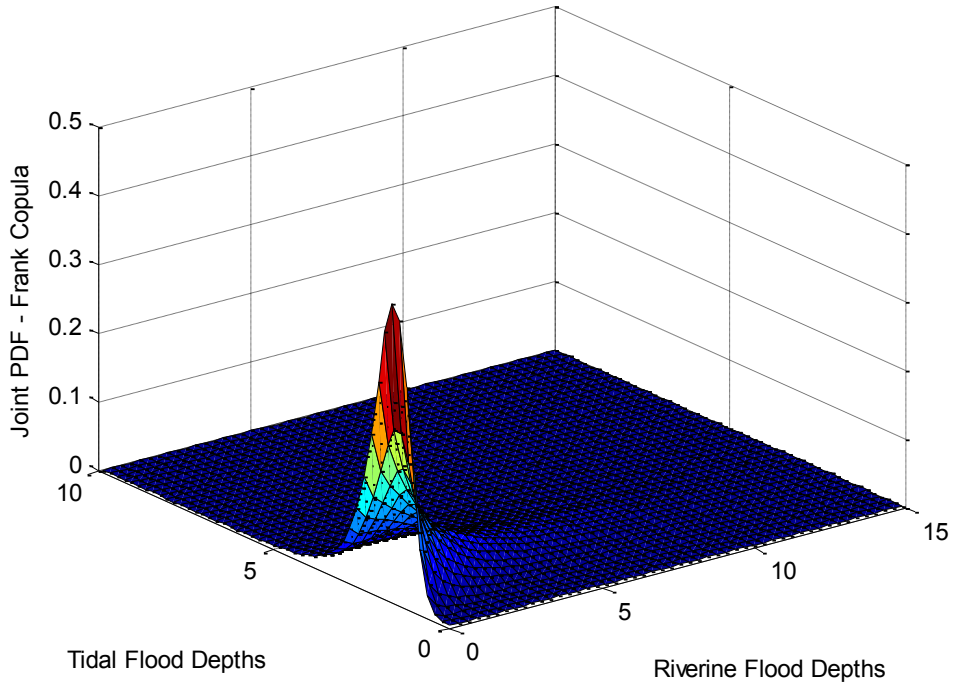

Figure 4-30: Joint PDF Calculated by Frank Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

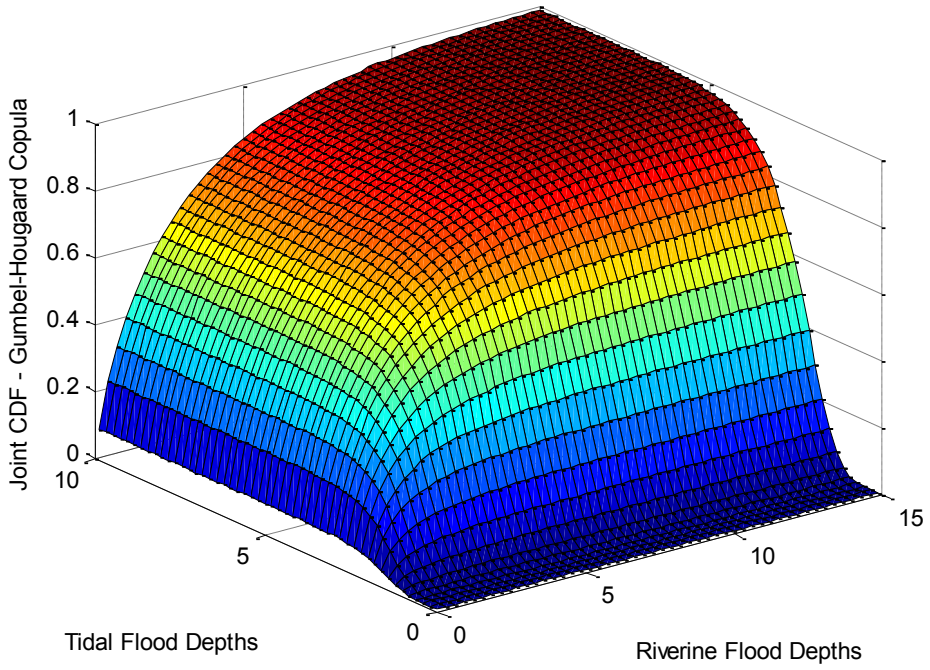

Figure 4-31: Joint CDF Calculated by Frank Copula for Correlation Coefficient of 0.37, where the Flood Depths are in Feet

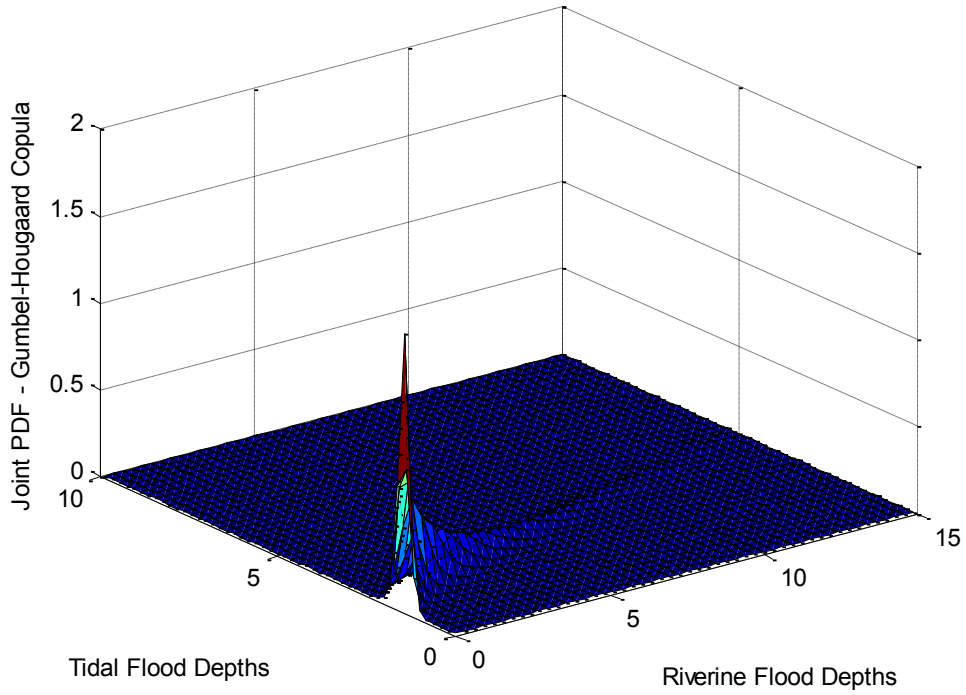

Figure 4-32: Joint PDF Calculated by Gumbel-Hougaard Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

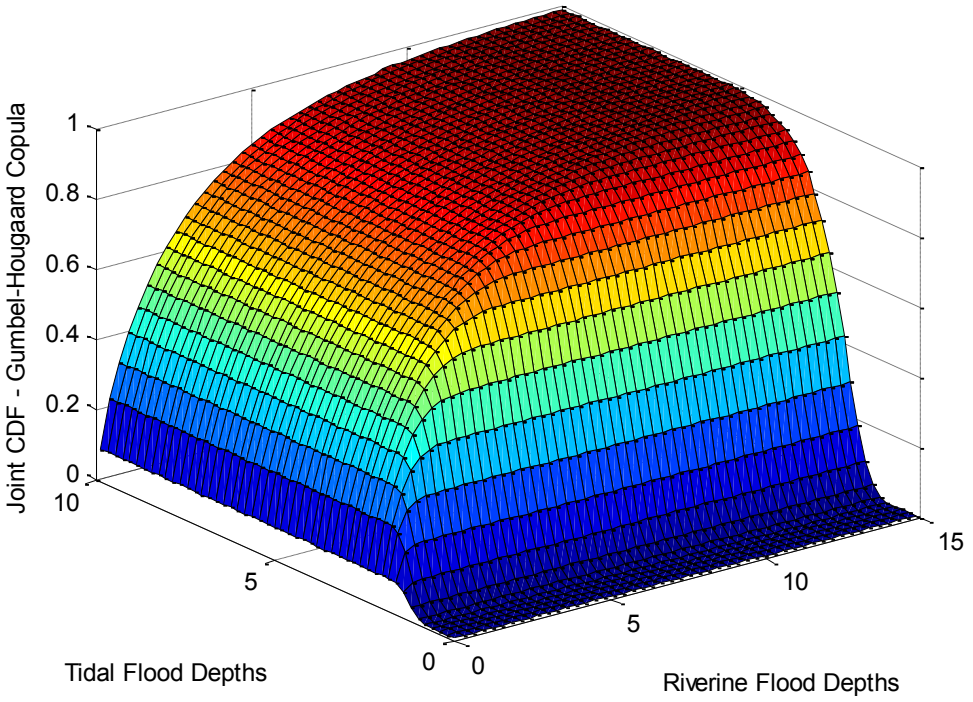

Figure 4-33: Joint CDF Calculated by Gumbel-Hougaard Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

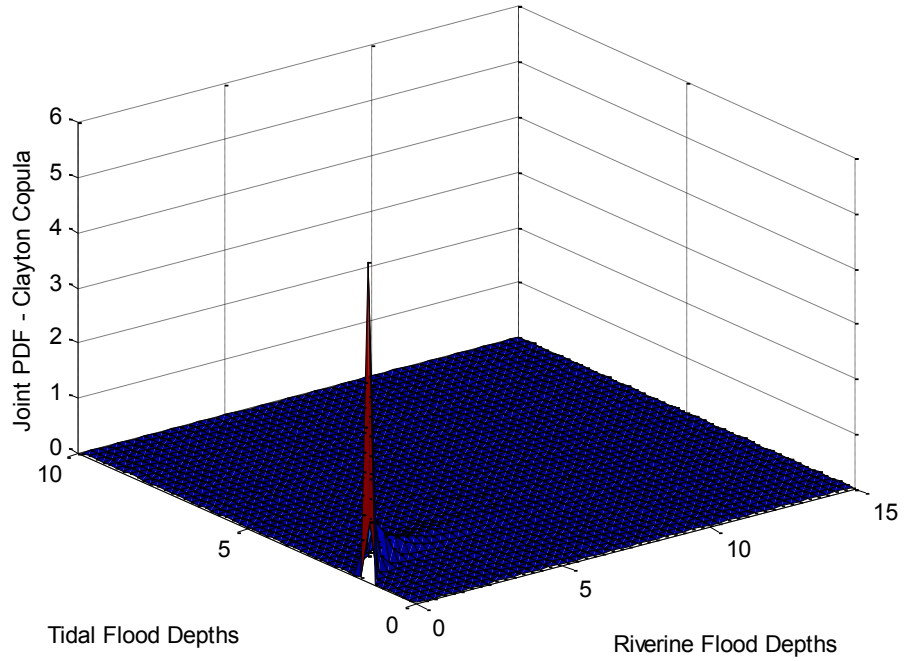

Figure 4-34: Joint PDF Calculated by Clayton Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

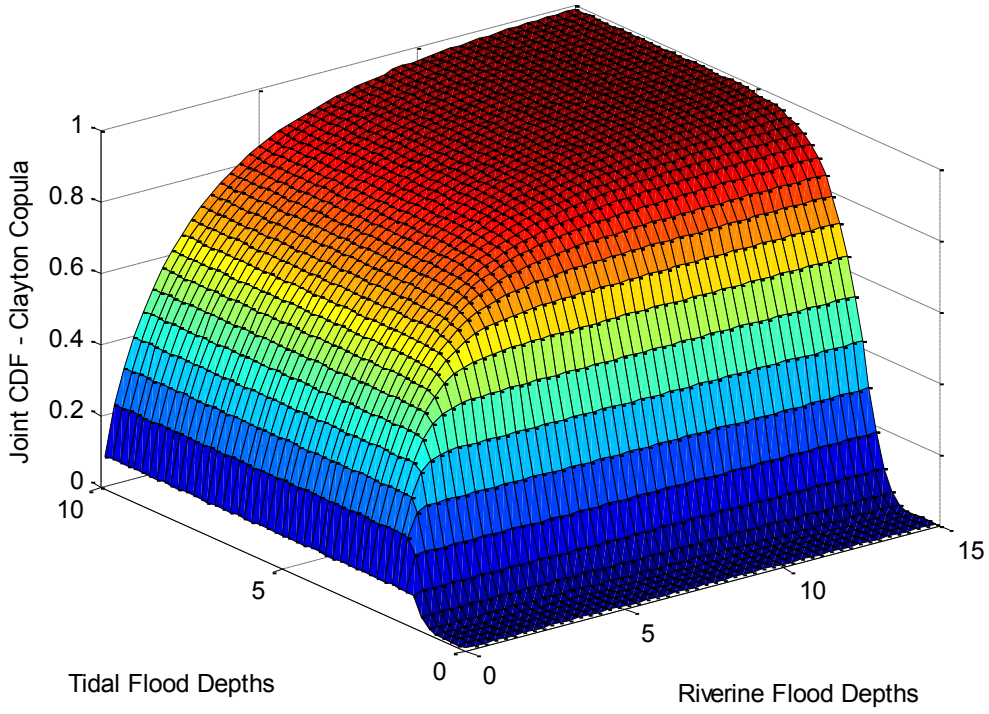

Figure 4-35: Joint CDF Calculated by Clayton Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

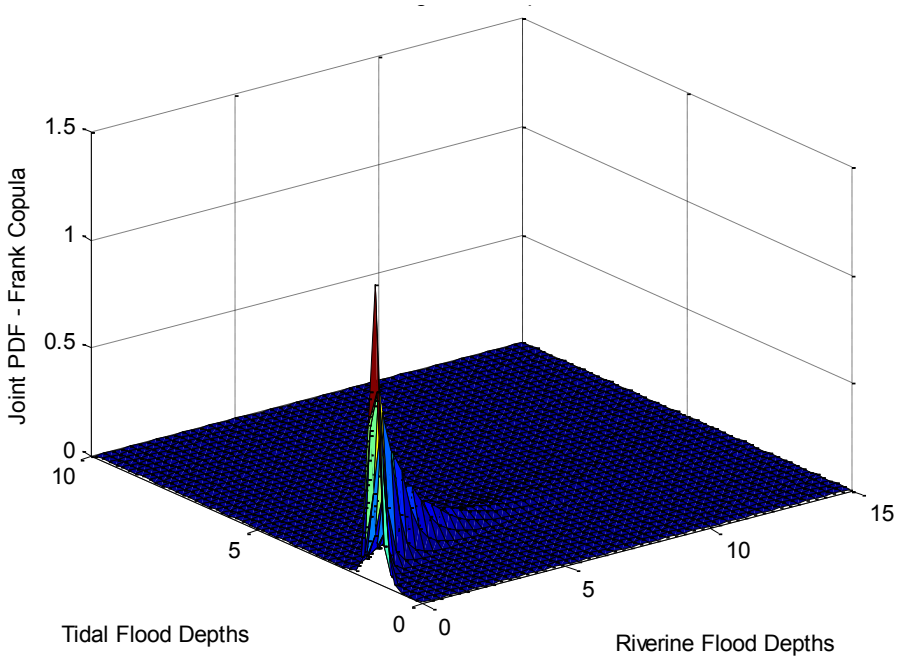

Figure 4-36: Joint PDF Calculated by Frank Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

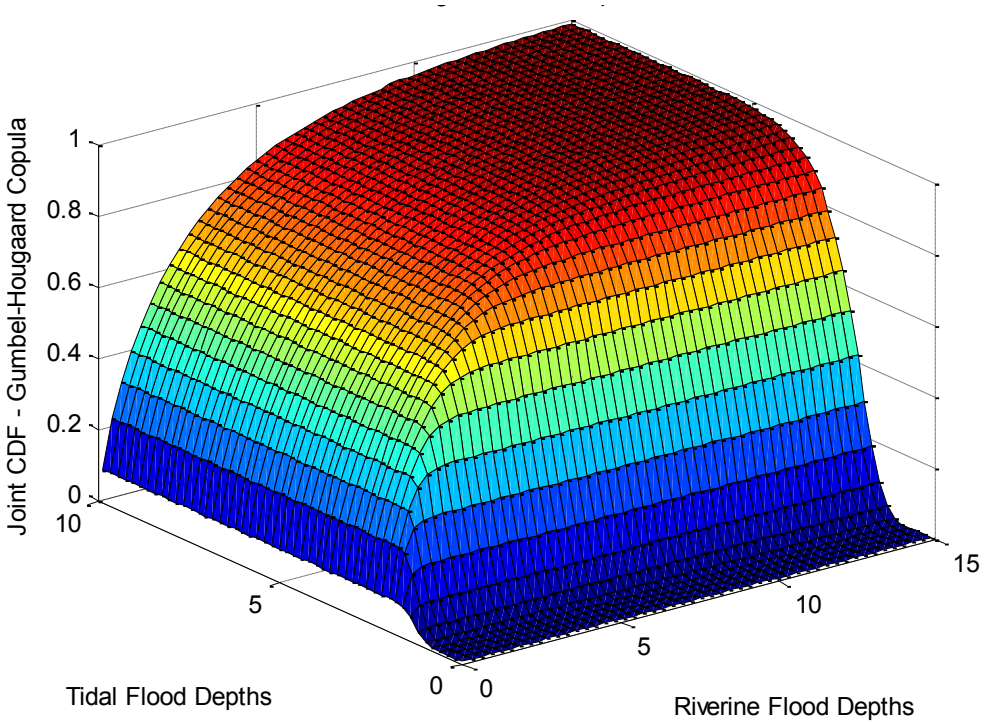

Figure 4-37: Joint CDF Calculated by Frank Copula for Correlation Coefficient of 0.60, where the Flood Depths are in Feet

For this low correlation scenario of 0.27 a negative value was calculated for Kendall's τ, such that only the Frank family could be used to develop joint distributions. As was observed for the joint distributions calculated based on varying sample sizes, the location of the peak, at a moderate tidal flood depth and very low riverine flood depth, seemed reasonable given the generated riverine and tidal flood depths. It was also noted that for this low level of correlation the peak was fairly broad and rounded. The steepness and shape of the joint cdf can also be traced to the generated riverine and tidal samples, as discussed for the scenarios with varying sample size. Much of the riverine sample consisted of very low depth values, with a few larger depth values, which explains the very steep slope parallel to the riverine flood depth axis at low flood depths and the more gradual slope as the riverine flood depths increased. The tidal sample contained depths primarily below 4 feet, which

explains the steeper slope parallel to the tidal depth axis over these low flood depths and the much more gradual slope over the larger flood depths. The trends in shape, size, and location of the joint pdf peaks would likely differ from those observed here for studies based on different locations.

For the higher correlation scenarios a positive value was calculated for Kendall's τ, so all three copula families could be used to develop the joint distributions. Based on the calculation of the AIC values, the Frank family would best represent these samples, but all three sets of joint distributions will be presented in order to assess the impact of choice of copula. As the correlation coefficient increased from 0.27 to 0.60 the peak of the joint pdf was observed to become narrower and steeper. The narrower peak indicates that similar values of riverine and tidal flood depths are more likely to occur jointly, which would be expected when higher levels of correlation exist between the samples. Similarly, the tail of the joint pdfs, which typically shows in the figures as a raised bump behind the peak of the joint pdf, was also observed to change as the level of correlation between the riverine and tidal samples increased. As the correlation increased, the tail became more diagonal, which further suggests the increased likelihood of flood depths of similar magnitudes occurring jointly from both flood sources. The tail indicates there is a higher likelihood of occurrence of a combination of a high riverine flood depth and a high tidal flood depth than there is of a combination of a high riverine flood depth and a low tidal flood depth.

Some variation was also evident in the joint cdfs as the level of correlation between the samples increased. Very small differences were evident for the lower

levels of correlation. However, some difference was observed in joint cdf values calculated for similar combinations of riverine and tidal flood depth combinations. The calculated joint cdf value calculated for similar combinations of riverine and tidal flood depths was typically, though not in all cases, observed to increase as the correlation between the riverine and tidal values increased. As the higher levels of correlation were approached the shape and slope of the joint cdf was observed to change. The joint cdfs became steeper parallel to the tidal flood depth axis as the correlation increased. This suggests an increased likelihood of larger flood depths occurring from the tidal flood source. These results suggest that the level of correlation between the samples has a much more significant impact on the corresponding joint distributions than did the sample size. Because the correlation between the samples influences the calculation of Kendall's τ and the copula parameters, it was expected that the level of correlation between the samples should influence the joint distributions.

In comparing the performance of the three copula families, fairly small differences in joint pdfs and cdfs were observed for the lower levels of correlation. For the lower levels of correlation, differences in the joint pdf values calculated by each family were typically small, on the range of 1-2%, though in some instances, larger differences in joint pdf values on the order of 10% were observed in the peaks. The calculated joint cdf values typically varied by approximately 1-3% between copula families for given combinations of riverine and tidal flood depths, which would typically not be considered significant. As the level of correlation increased, differences became more evident in both the joint pdfs and the joint cdfs based on the

copula family used. For instance, the peak of the joint pdf calculated using the Clayton family appears to be narrower and steeper than the peaks of the joint pdfs calculated by the Gumbel-Hougaard and Frank families. While the joint cdf values calculated for each family are similar over much of the distribution, differences are clearly evident, especially about the peaks. That the choice of copula family used to develop the joint distributions has a higher degree of impact at high levels of correlation between the samples was expected because the primary difference between the copula families is their modeling of dependence between the variables. Higher levels of correlation between the riverine and tidal flood depth samples suggest a higher level of dependence between the samples. Thus, differences between the copula families should be more readily apparent when a high level of dependence exists between the samples, as the copulas are known to perform differently for various levels of dependence between the samples.

4.2.5. Calculation of Combined Flood Frequency Curve

The final step in the procedure required using a double integral to calculate the area under the joint distributions corresponding to total flood depths. This provided the non-exceedance probabilities that correspond to total flood depths that were influenced by both flood sources. The following sections will describe the methods used and present the results of this procedure.

4.2.5.1. Description of Methods

The results of the copula procedure described in section 4.2.4 were joint pdfs and cdfs, which can provide the probability that corresponds to a specific joint event. However, the primary interest is the probability of a certain flood depth occurring,

when that flood depth might be due to either of the sources individually, or a combination of the two, such that many possible combinations could result in the desired flood depth. Because the riverine and tidal contributions to flood inundation at the location of interest were summed in the process of identifying annual maximum events, flood depths were also summed to identify probabilities corresponding to total flood depths. Within the joint pdf, the area corresponding to a specific total flood depth was identified as bounded by a triangle of the riverine and tidal inundation depths axes and a line connecting the points (0, tidal depth of interest) and (riverine depth of interest, 0). In other words, the points where the flood inundation depth of interest was caused only by one of the sources, either riverine or tidal, were identified and connected by a line. All of the points on this line summed to a total depth equal to the depth of interest and all points inside the bounding triangle summed to total depths less than the depth of interest. This is illustrated in Figure 4-38.

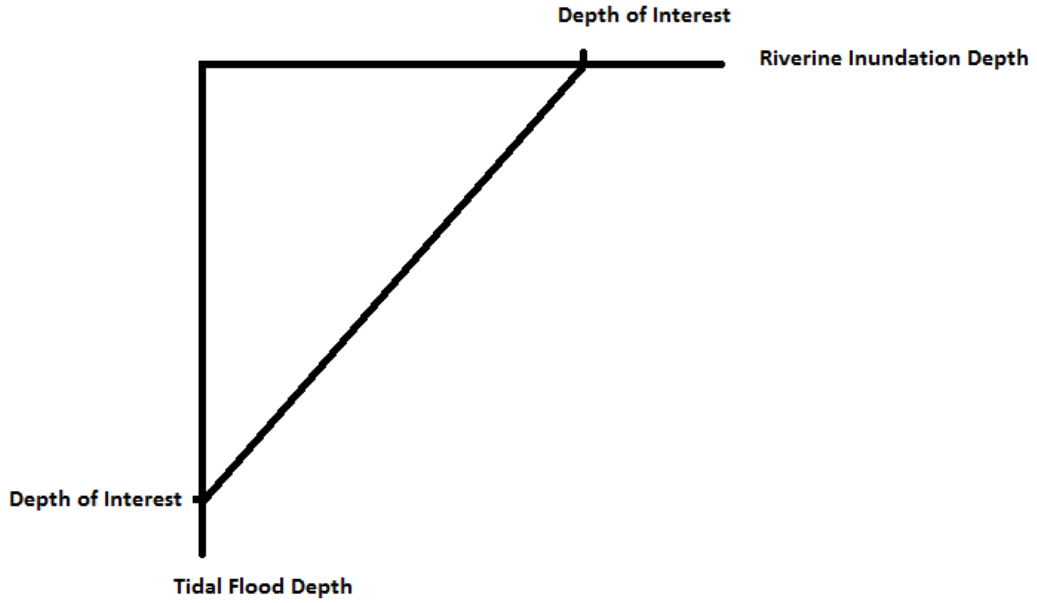

Figure 4-38: Graphical Illustration of Area Corresponding to Certain Flood Inundation Depth

The volume under the joint probability distribution within this bounding region represents the non-exceedance probability for a given flood depth. This was calculated using a double integral. The lower and upper bounds of the outer integral were 0 and total depth of interest, respectively. The lower and upper bounds of the inner integral were 0 and:

$$Upper\ Bound = Total\ Depth\ of\ Interest - Riverine\ Contribution\ to\ Total\ Depth \quad (4\text{-}6)$$

As an example, if the riverine flood depth was 3.5 feet, and the corresponding tidal flood depth was 3 feet, the total flood depth of interest would be 6.5 feet. In this scenario, the outer integral lower and upper bounds would be 0 feet and 6.5 feet, respectively, and the inner integral bounds would be 0 and (6.5 feet – 3.5 feet), or 3 feet. The double integral was calculated in Matlab using the adaptive Simpson's method, a numerical evaluation method, which uses a recursive algorithm to approximate the integral based on the error in estimates calculated using Simpson's rule. The non-exceedance probability of each value of total flood depth was calculated using this double integral procedure, and then the probabilities were plotted against the total depth values.

The final task to determining the probability of total flood depths was to identify the appropriate distribution to represent the probabilities that correspond to the total flood depths. The following distributions were fitted to the total flood depths: the LP3, the GEV, the gamma, and the normal. The GEV and gamma

distributions were chosen for consideration because they were the two marginal distributions used to represent the simulated riverine and tidal data. The LP3 distribution was selected for consideration because it is commonly used to represent flood data. The normal distribution was selected for consideration because the purpose of the copula is to develop a joint distribution when the marginals are represented by different distributions; thus, it was believed that the results of the copula might not follow either of the marginal distributions.

These four distributions were fitted to the total flood depth samples using either Maximum Likelihood Estimation or Method of Moments, and they were compared to the probabilities calculated for each copula family based on the double integral procedure. The fitted distributions were plotted against the total flood depths and compared to a plot of the non-exceedance probabilities calculated by double integral versus the total flood depths. These plots, as well as probability plot correlation coefficients, were used to determine which of the four distributions could best represent the total flood depth populations for each simulated scenario. The result of this process was an inferred population from which non-exceedance probabilities corresponding to total flood depths for the location of interest could be identified. From this distribution, the true nature of the flood hazard for the location of interest can be understood, which was the ultimate goal of this research.

4.2.5.2. Description of Results

The effects of both sample size and correlation between marginal samples on the probabilities corresponding to total flood depths were investigated. The results of these analyses are presented in the following sections.

4.2.5.2.1. Effects of Varying Sample Size

While the joint cumulative distribution calculated using the copula provides the non-exceedance probabilities corresponding to specific joint events, it is the non-exceedance probability corresponding to the total flood inundation depth that is needed, to which both sources may contribute. The parameters calculated in fitting the LP3, GEV, gamma, and normal distributions to the total flood depths for each sample size are presented in Tables 4-11, 4-12, 4-13, and 4-14. Figures 4-39, 4-40, 4-41, and 4-42 compare the fitted distributions to the calculated non-exceedance probabilities for the total flood depths for each sample size. Table 4-15 provides flood depths calculated using each distribution for several common exceedance periods for all sample sizes for another method of comparing the distributions. The results will be discussed after all tables and figures have been presented.

Table 4-11: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Size of 10

	LP3	GEV	Gamma	Normal
Shape	2.6304	0.6599	2.3729	N/A
Scale	0.3973	1.3948	2.3775	4.9626
Location	0.4599	3.2766	N/A	5.6417
Probability Plot Correlation Coefficient	0.9859	0.9749	0.9882	0.9233

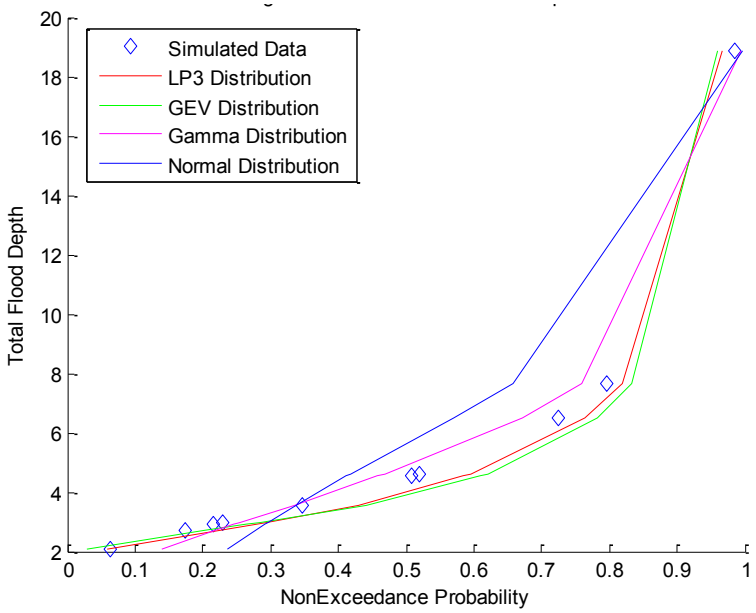

Figure 4-39: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample of 10, where Total Flood Depth Units are in Feet

Table 4-12: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Size of 25

	LP3	GEV	Gamma	Normal
Shape	2.5850	0.6787	1.9991	N/A
Scale	0.4223	1.4645	3.2182	5.8394
Location	0.4993	3.5182	N/A	6.4335
Probability Plot Correlation Coefficient	0.9948	0.9905	0.9958	0.9639

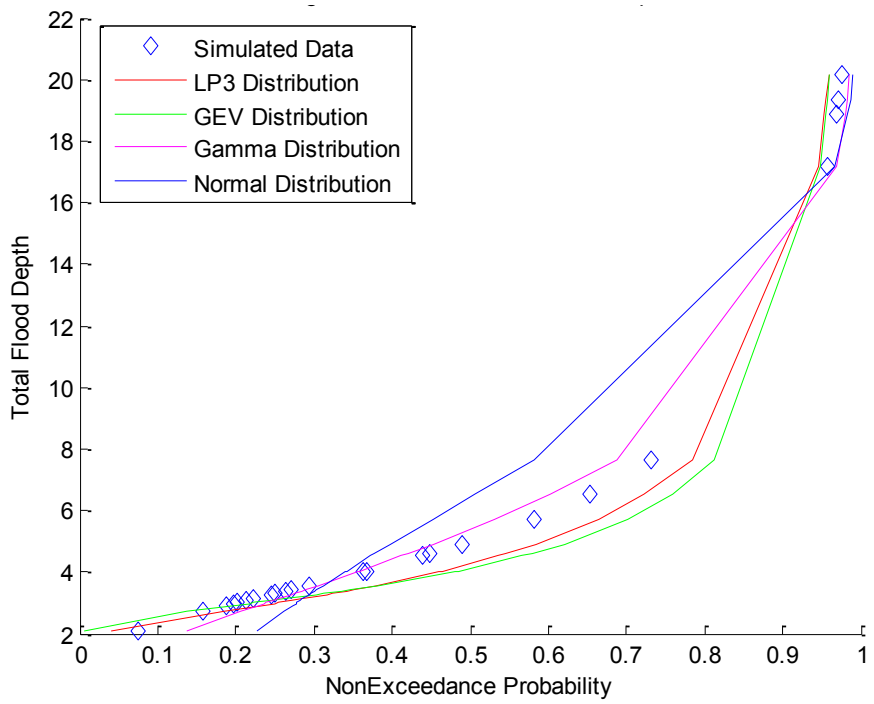

Figure 4-40: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample of 25, where Total Flood Depth Units are in Feet

Table 4-13: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Size of 50

	LP3	GEV	Gamma	Normal
Shape	1.8656	0.4841	2.6750	N/A
Scale	0.4099	1.2495	1.9136	4.3523
Location	0.6697	3.2552	N/A	5.1187
Probability Plot Correlation Coefficient	0.9766	0.9766	0.9913	0.9337

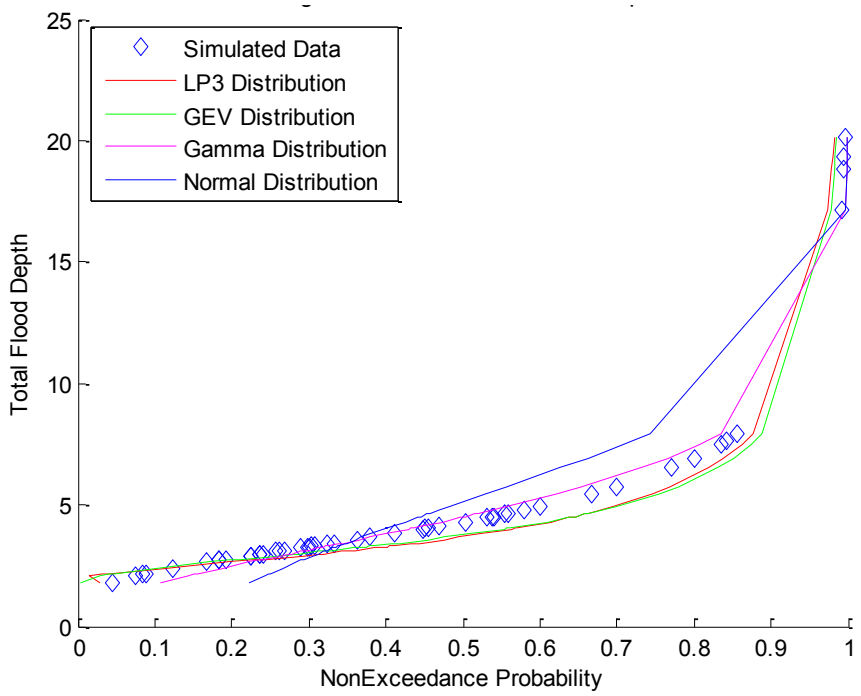

Figure 4-41: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample of 50, where Total Flood Depth Units are in Feet

Table 4-14: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Size of 100

	LP3	GEV	Gamma	Normal
Shape	18.8680	0.4327	1.7842	N/A
Scale	0.1721	1.8991	2.9872	5.0315
Location	-1.8797	2.9538	N/A	5.3297
Probability Plot Correlation Coefficient	0.9961	0.9929	0.9890	0.8885

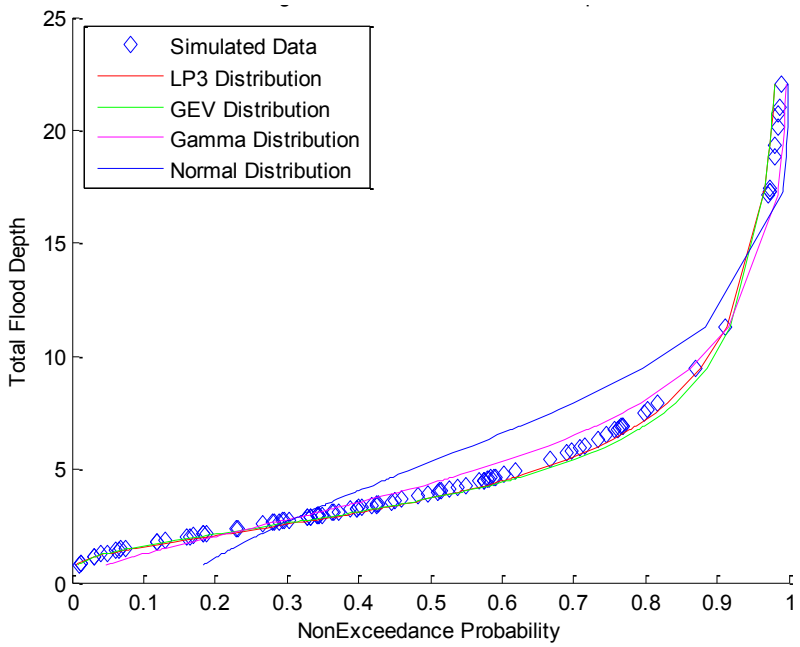

Figure 4-42: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample of 100, with Total Flood Depth Units in Feet

Table 4-15: Flood Depths in Feet Calculated for Exceedance Probabilities of 10%, 2%, and 1% for Each Distribution

Exceedance Probability	LP3	GEV	Gamma	Normal
	Sample Size 10			
0.1	10.681	10.4948	10.5457	12.0015
0.02	24.6788	28.9149	15.4051	15.8336
0.01	34.7293	45.1571	17.3978	17.1864
	Sample Size 25			
0.1	12.1915	11.2989	12.5136	13.917
0.02	29.555	31.8485	18.7696	18.4262
0.01	42.4278	50.3306	21.3582	20.018
	Sample Size 50			
0.1	8.8606	8.3464	9.3135	10.6964
0.02	19.3524	17.7409	13.3678	14.0572
0.01	26.7693	24.6043	15.0203	15.2437
	Sample Size 100			
0.1	10.5357	10.186	10.6502	11.7778
0.02	21.7673	22.3119	16.2717	15.6631
0.01	28.6309	30.6879	18.612	17.0347

For all four sample sizes, the probability plot correlation coefficients indicate that the LP3, GEV, and Rayleigh distributions would all provide acceptable fits to the generated data. The probability plot correlation coefficients also suggest that the normal distribution would not provide as adequate a fit to the total flood depths for all four sample sizes. The differences in probability plot correlation coefficients calculated for the LP3, GEV, and gamma distributions for all sample sizes are minimal, indicating that the three distributions would perform similarly in representing the samples.

The graphical comparisons of the fitted distributions to each sample size provide similar results to the probability plot correlation coefficients. For all four sample sizes, the graphical comparisons indicate that the LP3, GEV, and gamma distributions would adequately fit the total flood depths while the normal distribution would not adequately fit the data. The difference in the ability of the LP3, GEV, and gamma distributions to fit the generated flood depths appears to be minimal, as suggested by the probability plot correlation coefficients. The graphical comparisons also suggest that the LP3 and GEV distributions would perform almost identically and would be an excellent fit to the data. Though the gamma distribution appears to fit the generated data fairly well for all sample sizes, it typically is not observed to match the curvature of the generated data as well as the LP3 and GEV distributions. Based on these conclusions, the LP3 distribution was selected to compute the probabilities corresponding to total flood depths for all four sample sizes.

The impact of the choice of distribution on flood depths calculated for various exceedance probabilities was examined in Table 4-15. While the LP3, GEV, and

gamma distributions were all observed to perform very similarly over the range of the data, the results in Table 4-15 indicate that there would be large differences in the predicted flood depths corresponding to the larger flood events, such as the 1% annual chance event. The LP3 and GEV distribution continue to perform approximately equally for the larger flood events; however, the gamma distribution predicts much lower flood depths than the other two distributions for the larger events. Differences in the tails of these distributions account for these differences in predicted flood depths. This suggests that though the three distributions appear very similar over the range of the data, the choice of distribution would actually have significant impact on the larger predicted flood depths. It can also be noted from Table 4-15 that a trend in the predicted flood depths with increasing sample size is not apparent. As trends in the distribution parameters were not apparent with varying sample size, it was not expected that there would be any trend in the predicted flood depths.

As the sample size has increased from 10 to 100, the fitted distributions were observed to be better able to fit the sample, as would be expected. The larger samples better approximate and define the populations; thus, the fitted distributions can be considered to be more accurate when based on a larger sample. Though the increasing sample size resulted in a better ability to fit distributions to the total flood depths, the increase in sample size was not observed to result in any trends in the parameters of the fitted distributions. However, from the figures and the probability plot correlation coefficients calculated, it would appear that reasonably accurate estimates of the non-exceedance probabilities that correspond to total flood depths

could be obtained even when only a small sample size is available, at least for this particular scenario. Given the high variability typical in hydrologic data, this may not be the case for other scenarios.

4.2.5.2.2. Effects of Varying Sample Correlation

Distributions were also fitted to the total flood depths calculated for the scenarios with varying levels of correlation between the samples. These non-exceedance probabilities were compared to those calculated using the double integral procedure. The parameters calculated for the LP3, GEV, gamma, and normal distributions, which were fitted to the total flood depths for each scenario, are presented in Tables 4-16, 4-17, 4-18, and 4-19. These tables also present the probability plot correlation coefficients for each distribution in each scenario. The fitted distributions are graphically compared to the non-exceedance probabilities calculated using the double integral procedure in Figures 4-43, 4-44, and 4-45. Graphical comparisons are presented for correlations of 0.27, 0.37, and 0.60, but not for a correlation of 0.31 due to similarities to the correlation of 0.27. Table 4-20 provides flood depths calculated for each distribution for the 10%, 2%, and 1% annual chance events to compare the performance of the distributions. The results will be discussed after all tables and figures have been presented.

Table 4-16: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Correlation of 0.27

	LP3	GEV	Gamma	Normal
Shape	1.6795	0.4590	2.7523	N/A
Scale	0.4232	1.2511	1.8733	4.3639
Location	0.7370	3.3314	N/A	5.1560
Probability Plot Correlation Coefficient	0.9785	0.9816	0.9903	0.9309

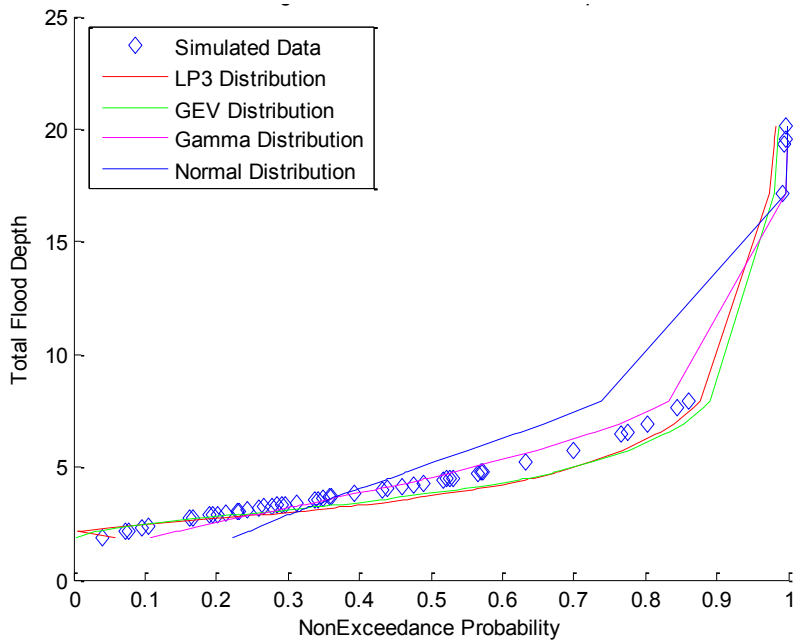

Figure 4-43: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample Correlation of 0.27, with Total Flood Depth Units in Feet

Table 4-17: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Correlation of 0.31

	LP3	GEV	Gamma	Normal
Shape	2.4181	0.4285	2.5514	N/A
Scale	0.3735	1.3829	1.9935	4.3708
Location	0.5149	3.2128	N/A	5.0861
Probability Plot Correlation Coefficient	0.9781	0.9816	0.9917	0.9292

Table 4-18: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Correlation of 0.37

	LP3	GEV	Gamma	Normal
Shape	3.6609	0.4824	2.3757	N/A
Scale	0.3209	1.4431	2.1250	4.4042
Location	0.2192	3.0585	N/A	5.0482
Probability Plot Correlation Coefficient	0.9765	0.9703	0.9945	0.9314

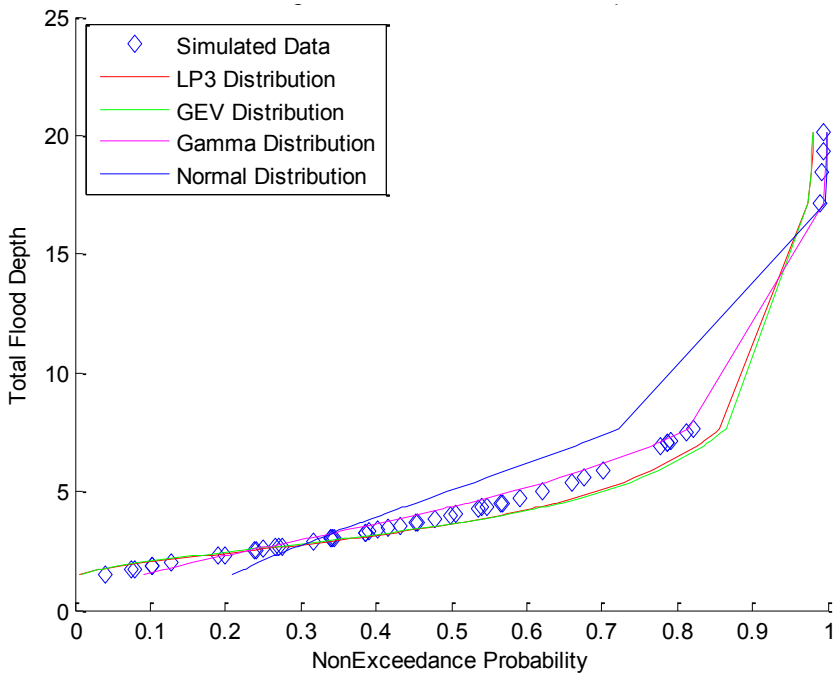

Figure 4-44: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample with Correlation Coefficient of 0.37, with Total Flood Depth Units in Feet

Table 4-19: Parameters and Probability Plot Correlation Coefficients Calculated for Total Flood Depths for a Sample Correlation of 0.60

	LP3	GEV	Gamma	Normal
Shape	9.1296	0.4878	2.0727	N/A
Scale	0.2258	1.6144	2.3989	4.4732
Location	-0.7174	2.8521	N/A	4.9721
Probability Plot Correlation Coefficient	0.9840	0.9712	0.9915	0.8929

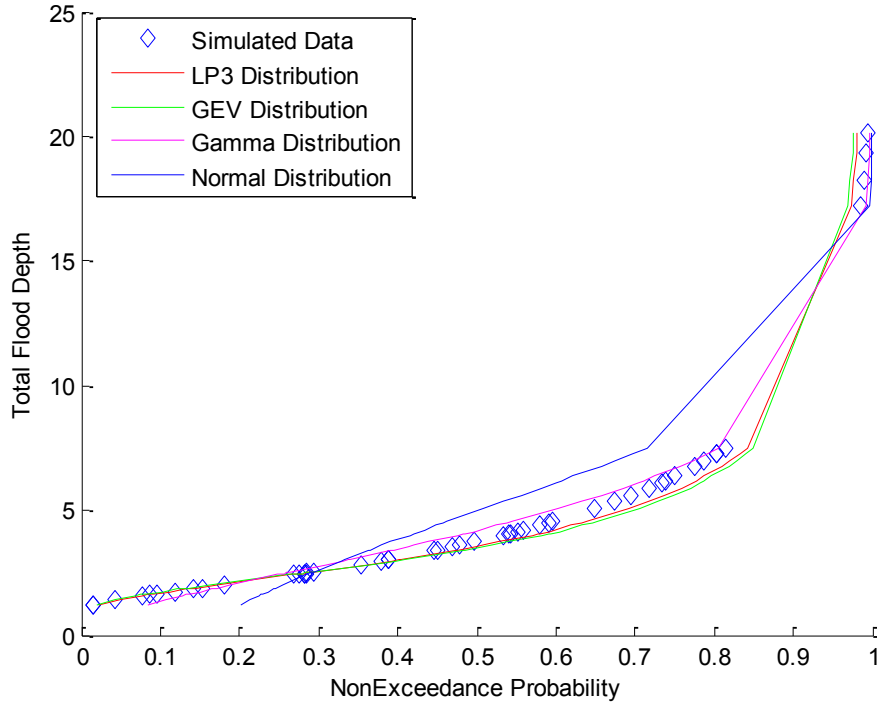

Figure 4-45: Comparing Non-Exceedance Probabilities Calculated for Total Flood Depths Using the Double Integral Procedure to Those Calculated by Fitting Distributions for a Sample with Correlation Coefficient of 0.60, where Total Flood Depth Units are in Feet

Table 4-20: Flood Depths in Feet Calculated for Exceedance Probabilities of 10%, 2%, and 1% for Each Distribution

Exceedance Probability	LP3	GEV	Gamma	Normal
	Sample Correlation 0.27			
0.1	8.8308	8.2629	9.3216	10.7486
0.02	19.3222	16.9474	13.325	14.1184
0.01	26.7977	23.121	14.9545	15.3079
	Sample Correlation 0.31			
0.1	8.9883	8.4504	9.3526	10.6875
0.02	19.3679	17.1635	13.5163	14.0626
0.01	26.5259	23.1544	15.2175	15.2541
	Sample Correlation 0.37			
0.1	9.1815	8.9253	9.4343	10.6924
0.02	19.429	19.7169	13.7791	14.0933
0.01	26.2512	27.5863	15.5607	15.2939
	Sample Correlation 0.60			
0.1	9.5105	9.4624	9.5896	10.7047
0.02	19.5668	21.7445	14.3019	14.1589
0.01	25.8622	30.7534	16.248	15.3783

To select the most appropriate distributions to represent the total flood depths for each scenario the probability plot correlation coefficients will first be examined. These suggest that the LP3, GEV, and gamma distributions would all adequately represent the total flood depths, while the normal distribution would typically not be able to represent the data. Differences in the probability plot correlation coefficients for the LP3, GEV, and gamma distributions do not appear to be significant. Based on the very similar probability plot correlation coefficients calculated for these distributions for all four scenarios, it would be expected that all three distributions should perform very similarly in representing the total flood depth samples.

Graphical comparisons of the fitted distributions to the non-exceedance probabilities calculated for the total flood depths using the double integral procedure were also evaluated to determine the most appropriate distribution to represent the total flood depths. For all four scenarios, these plots supported the conclusions drawn based on the probability plot correlation coefficients. The plots suggested that the LP3, GEV, and gamma distributions all provide excellent fits to the total flood depth data but the normal distribution does not. The LP3 and GEV distributions appear to perform almost identically, and both typically appear to match the curvature of the generated data slightly better than the gamma distribution. The LP3 distribution was chosen to represent the total flood depths for all scenarios based on the plots, the probability plot correlation coefficients, and the knowledge that this distribution was frequently observed to be the best fit in the other simulation studies.

As the level of correlation increased a trend was observed in the parameters of each of the fitted distributions. For the LP3 distribution, the shape parameter was

observed to increase and the scale and location parameters were observed to decrease as correlation increased. For the GEV distribution, the shape and scale parameters were observed to increase while the location parameter was observed to decrease. For the gamma distribution, the shape parameter was observed to decrease and the scale parameter was observed to decrease, and for the normal distribution, the scale parameter was observed to increase and the location parameter was observed to decrease. These trends in the parameter values indicate that the level of correlation between the riverine and tidal flood depth samples will influence the distributions fitted to the total flood depths. The level of correlation between the samples also influenced the total depths themselves, as the level of correlation controlled the tidal sample generated, and the riverine and tidal samples were summed to identify the total flood depths. It was expected that the level of correlation would have some impact on the distributions fitted to the total flood depths as the level of correlation was also observed on influence the joint pdfs. Because the purpose of the copula is to model dependence between variables, it is reasonable to expect that the level of dependence between the variables would impact the results of the copula and the results of probabilities corresponding to the total flood depths as well.

Despite the impact that the level of correlation was observed to have on the parameters of the fitted distributions, definite trends in the calculated flood depths corresponding to common exceedance probabilities in Table 4-20 were not observed. The flood depths predicted for each exceedance probability using the GEV and gamma distributions were observed to slightly increase with increasing correlation. Such trends were not evident when the LP3 and normal distributions were used to

predict the flood depths. This suggests that though trends were identified in the distribution parameters based on the level of correlation, those trends were not significant enough to impact the flood depths calculated using the distributions. Further examination of Table 4-20 also reveals the significant of the choice of distribution to represent the total flood depths, which was not so obvious over the range of the data. Though the LP3, GEV, and gamma distributions were observed to perform very similarly over the range of the data, Table 4-20 suggests that large differences in flood depths, especially for the larger flood events, would be possible depending on which distribution as chosen. The LP3 and GEV distribution consistently calculate larger flood depths corresponding to the 1% annual chance flood than does the gamma distribution. Thus, care must be taken in choosing the distribution best suited to represent the total flood depths.

4.2.6. Conclusions

The analyses presented examined the effect of sample size on the development of a joint cumulative distribution, making use of the copula method, and on the development of a comprehensive flood frequency curve that considered multiple flood sources. Though sample size has a clear impact on the fitting of marginal distributions, the impact of sample size on the joint distributions calculated by the copula method, and the fitting of distributions to represent the total flood depth, was observed to be minimal. As the sample size increased, more accurate distributions were fitted to the marginal samples, though trends were not observed in the fitted marginal distribution parameters as sample size varied. The joint distributions developed based on the copula were not observed to vary significantly

based on sample size. This suggests that a reasonable joint distribution can be developed based on a small sample. However, these studies were based on only one scenario, and higher degrees of variation may be apparent for other scenarios. Limited available data remains a significant challenge in hydrologic analyses due to the high variation typically evident.

The joint distributions were observed to be most strongly influenced by the calculated value of Kendall's τ, which was not observed to be strongly related to sample size. Since Kendall's τ is based solely on the ranks of the data, it would not be expected to vary systematically with sample size. Because the marginal distributions simulated for each sample size were generated for the same distributions with the same parameters, and Kendall's τ was not observed to vary significantly as the sample size varied, the joint distributions developed for each of the sample sizes were very similar. However, it can be expected that as sample size increased the joint distributions developed would be more accurate, as the sample better represents and defines the population. Based on the joint distributions, non-exceedance probabilities corresponding to total flood depths were calculated and distributions were fitted to represent the probabilities that corresponded to the total flood depths. Trends were not observed in the fitted distributions as the sample size varied. For instance, specific variations in the parameters of the fitted distributions were not observed. As sample size was not observed to significantly impact the joint distributions, it would not be expected to impact the distributions fitted to the total flood depths either. However, the distributions fitted to the larger sample sizes can be assumed to better represent the populations. These results indicate that even when the

sample size is small, reasonable joint distributions can be developed for the sample and non-exceedance probabilities can be calculated corresponding to the total flood depths.

The analyses presented also examined the effect of degree of correlation between the samples on the development of a joint cumulative distribution, making use of the copula method, and on the development of a comprehensive flood frequency curve that considered multiple flood sources. The change in correlation between the samples did have some impact on the fitting of distributions to the sample marginals, as the change in correlation between the simulated samples was achieved by varying the tidal sample while holding the riverine sample constant. However, specific trends were not evident in the fitted marginal distribution parameters as the level of correlation between the samples varied. The level of correlation between the samples was observed to impact the joint pdfs and cdfs, with a more visually obvious impact on the joint pdfs. The level of correlation between the samples describes the linear relationship and dependence between the samples. Thus, the joint distributions, which illustrate the dependence between the variables, should vary as the level of correlation varies. Within the joint pdfs, the peaks of the joint pdfs were observed to become narrower and steeper as the level of correlation between the samples increased. The level of correlation impacted the joint distributions because there was a relationship between correlation and Kendall's τ. Kendall's τ influenced the value of the copula parameter, and therefore influenced the joint distributions developed for a given set of samples.

Further, in scenarios for which positive values of Kendall's τ were calculated, which allowed all three copula families to be used to develop the joint distributions, the choice of copula family was observed to have more impact as the level of correlation increased. This result was expected as the primary difference between the copula families is how they model dependence between the variables. Differences were more visually evident in the joint pdfs between the three copula families as the level of correlation increased. This result suggests that the choice of copula family would have more impact on an analysis if the flood sources were strongly dependent, or had a high level of correlation between them. For lower levels of dependence, the results of this study indicate that the choice of copula family is less important and will have less impact on the results of the analysis.

The correlation also was observed to impact the non-exceedance probabilities corresponding to the total flood depths. As these probabilities were calculated based on the joint pdfs, it was expected that the level of correlation between the samples would influence the calculated probabilities. The correlation was also observed to impact the distributions fitted to the total flood depths. Trends were observed in the parameters of the distributions fitted to the total flood depths as the correlation between the samples increased. The level of correlation between the samples was expected to have a more significant impact on the joint distributions and the distributions fitting the total flood depths because the purpose of the copula is to model dependence between variables. As the correlation between the flood sources increases the dependence decreases, making the copulas more necessary to model the joint distributions. These results indicate that the level of correlation between the

samples will have a significant impact on the joint distributions developed and the calculation of non-exceedance probabilities that correspond to the total flood depths. More care must be taken in conducting analyses for highly dependent variables, to ensure that the most appropriate copula family is used to represent the joint distribution. This will also ensure that the best possible estimates of the non-exceedance probabilities that correspond to the total flood depths will be obtained.

4.3. Flood Risk Calculations

The previous sections have developed a process to assess the flood hazard for a location of interest based on multiple flood sources. The flood frequency assessments developed were based on joint distributions for the two flood sources calculated based on copulas. The flood hazard is only one part of flood risk. The other two components of flood risk are the vulnerability to flooding and the consequences that result from those flood events. Flood risk is calculated by multiplying the exceedance probability, the vulnerability, and the consequences for a given flood depth of interest.

The term vulnerability refers to the ability of the system in place to protect the location of interest from flooding. The system will include any structures near the location of interest designed to reduce the probability of flooding or any nonstructural measures taken at or near the location of interest to reduce either the probability or the consequences of flooding. Common examples of structural systems include levees or floodwalls. Nonstructural measures include building restrictions, elevation of structures, flood proofing structures, or creating additional room to store waters

naturally. Whatever the system may be that is identified as protecting the location of interest, failure is always a possibility. This uncertainty is incorporated into the vulnerability term of the risk equation.

To complete the flood risk calculations, the consequences of the flood event must be assessed. Consequences may be identified in economic terms, such as damages to structures and infrastructure, environmental damages, or loss of life or injuries as a result of the flood event, though economic terms are the most commonly used. For the purpose of providing an example of the method, economic consequences will be considered because they can be easily calculated and understood.

4.3.1. Description of Methods

The exceedance probabilities calculated for the total flood depths using the copula procedure were used as input to the flood risk calculations. Flood risk calculations were then made based on the other terms in the risk equation (exceedance probability multiplied by vulnerability multiplied by consequences), the vulnerability and the consequences. Vulnerability is a weight, that ranges between 0 and 1and indicates how well the in-place system to protect the location of interest from flood events is expected to perform. Therefore, vulnerability will vary with flood depth. A hypothetical vulnerability curve was created for use in experimental calculations, as shown in Figure 4-46, as detailed information about the vulnerability of the location of interest was not available. Vulnerability increases with depth in a linear fashion, which indicates that vulnerability is simply directly proportional to depth of flooding, which may be true for certain systems. Without specific

information about the system in place at the location of interest this linear curve seemed to be a reasonable approximation of vulnerability for the purpose of these calculations.

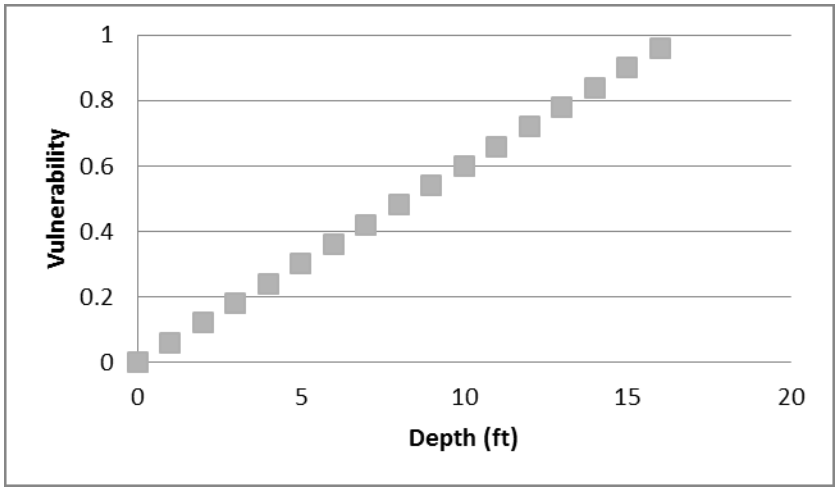

Figure 4-46: Illustration of the Vulnerability Curve Used in Flood Risk Assessments

Consequences, measured as damage to the structure and its contents due to the flood, will also vary with depth. Depth-damage curves, shown in Figure 4-47, obtained from USACE (2003), are used to determine the percentage of damage to the structure and to contents based on depth of flooding. These depth-percent damage curves are general curves that can be applied nationally, though it is also preferable to derive specific curves for a given location. For a hypothetical scenario using simulated data, the general curves were deemed to be appropriate. These percent damage values were then multiplied by the value of the structure and the value of the contents to determine the total monetary damage due to varying flood depths. For the hypothetical scenario, the two-story residential structure without a basement was assumed to be worth $150,000 and the contents were assumed to be worth $25,000. The resulting depth-monetary damage curves are provided in Figure 4-48.

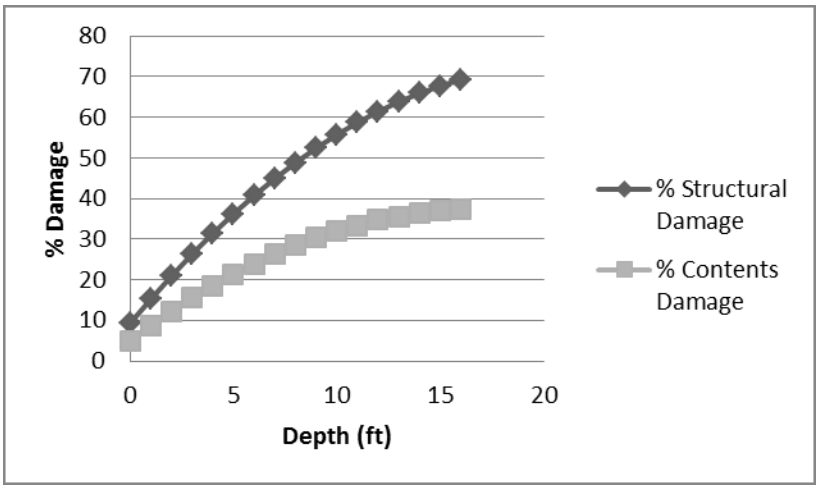

Figure 4-47: Depth-Percent Damage Curves Obtained from USACE (2003)

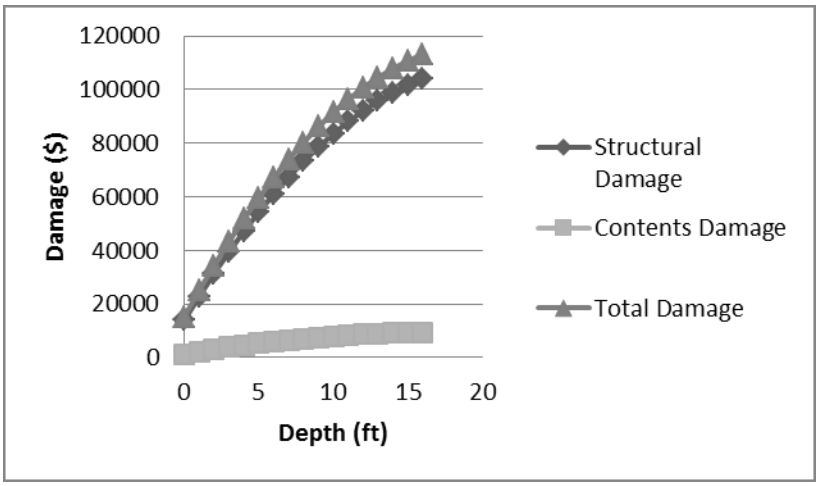

Figure 4-48: Depth-Monetary Damage Curves

Depending on the purpose of the risk analysis, it might be desirable in some circumstances to fit equations to the vulnerability and consequence curves, so that values could be determined for any flood depth of interest. However, the purpose of this analysis is simply to demonstrate the effect of the comprehensive probability assessment on risk calculations. Thus, the population that was fitted to the exceedance probabilities corresponding to the total flood depth values was used to determine the exceedance probabilities corresponding to various depths of flooding for which values of vulnerability and consequences had been obtained. These

exceedance probabilities were multiplied by the vulnerability values and consequence values for each flood depth to obtain flood risk values for each flood depth.

A total of eight scenarios were investigated using the simulation studies evaluating the sample size and sample correlation. Sample sizes of 10, 25, 50, and 100 were investigated, as were sample correlation coefficients of 0.27, 0.31, 0.37, and 0.60. For each of these scenarios, the exceedance probabilities corresponding to flood depths of 1 through 16 feet were first calculated using the probability distributions fit in the flood hazard assessment developed through the copula procedure. The exceedance probabilities corresponding to flood depths were used as input to the flood risk calculations for each simulation scenario. The exceedance probabilities were then multiplied by the vulnerability curve previously identified in Figure 4-46 and the consequences curve shown in Figure 4-48 to obtain calculations of flood risk for each simulation scenario.

4.3.2. Description of Results

Flood risk was calculated for each of the simulation scenarios developed to evaluate the impact of sample size and level of correlation between the samples. The exceedance probabilities for each scenario were calculated based on the probability distributions representing non-exceedance probabilities calculated for the total flood depths. The impacts of varying sample size and sample correlations were assessed.

4.3.2.1. Variation of Sample Size

Risk calculations were made for each of the four sample sizes considered to determine the impact of sample size on the estimation of flood risk for the location of

interest. Figure 4-49 provides the exceedance probabilities calculated based on the distribution fitted to total flood depths for each sample size. Figure 4-50 provides the flood risk calculations for each sample size, based on the probabilities presented in Figure 4-49, the vulnerability information presented in Figure 4-46, and the consequence information presented in Figure 4-48. The exceedance probabilities presented in Figure 4-49 suggest that sample size had minimal impact on the calculated probabilities. While there were variations in the parameters of the distributions previously fitted to the total flood depths for each sample size, particular trends or impacts on these fitted distributions corresponding to sample size were not observed; thus, significant difference in exceedance probabilities was not expected.

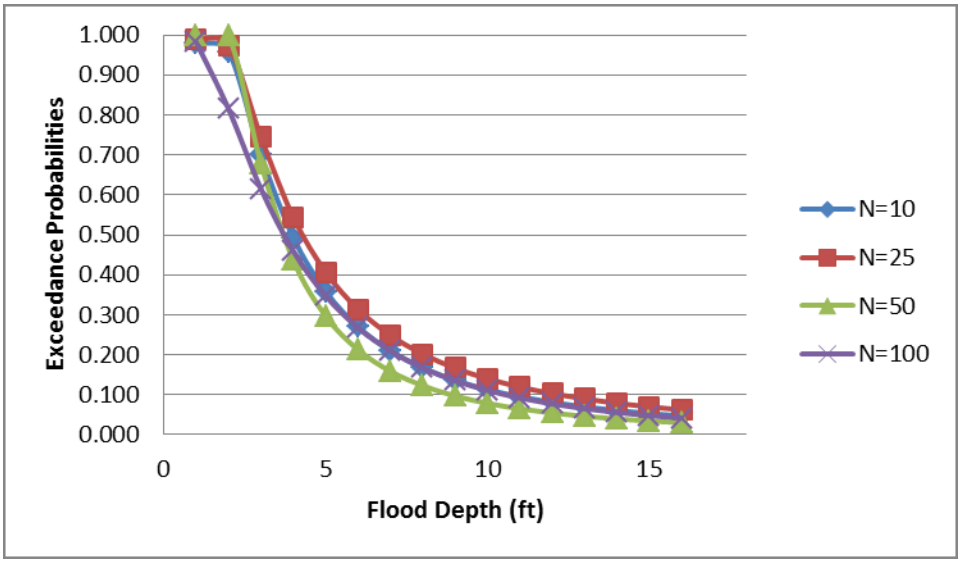

Figure 4-49: Exceedance Probabilities Corresponding to Total Flood Depths Calculated Using the Copula Procedure for All Four Sample Sizes

More variation is evident in the flood risk calculations in Figure 4-50 than in the non-exceedance probabilities. The flood risk calculations are the most similar for low flood depths where the exceedance probabilities were the most similar between the sample sizes. The small variations in exceedance probabilities are magnified by

the multiplication of the exceedance probabilities with the vulnerabilities and consequences. However, these variations do not seem to be related to sample size. It can be assumed that the flood risk calculations based on larger sample sizes would be more accurate, as the larger sample size should have better represented the population when distributions were fitted to the total flood depth values, and because the marginal distributions and copula fitted earlier in the process should also have become more accurate as sample size increased. Otherwise, trends cannot be identified in the flood risk calculations relating to the sample size. The exceedance probabilities steadily decreased with increasing flood depth, but the vulnerability and consequences increased with increasing flood depth, so that the highest flood risk was calculated for all sample sizes at flood depths between 4 and 7 feet. The flood risk for high flood depths is actually quite low because the exceedance probabilities become quite small as flood depth increases, which reduces the influence of the increasing vulnerability and consequence terms as the flood depth increases.

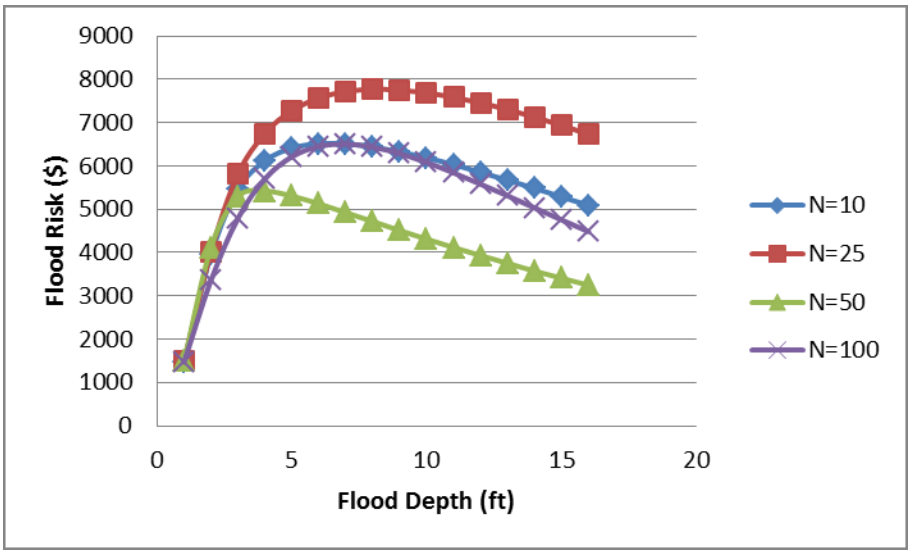

Figure 4-50: Flood Risk Calculations Corresponding to Total Flood Depths for All Four Sample Sizes

4.3.1.2. Variation of Sample Correlation

Flood risk calculations were also made for each of the simulation scenarios with different levels of correlation between the marginal samples. The level of correlation was previously observed to influence the joint distributions developed using the copula as well as the non-exceedance probabilities calculated corresponding to total flood depths, so it was believed that there may be some influence on the flood risk calculations as well. Figure 4-51 provides the exceedance probabilities calculated using the distribution fitted to the total flood depths for each level of correlation. The exceedance probabilities were observed to vary most significantly in the first three feet of flood depth and then to become very similar across all four scenarios. It was observed that the exceedance probability calculated for a given flood depth increased slightly as the level of correlation increased. Given the impact of the level or correlation on the distributions fitted to the total flood depths previously, an impact such as this on the exceedance probabilities calculated for the flood risk calculations was expected.

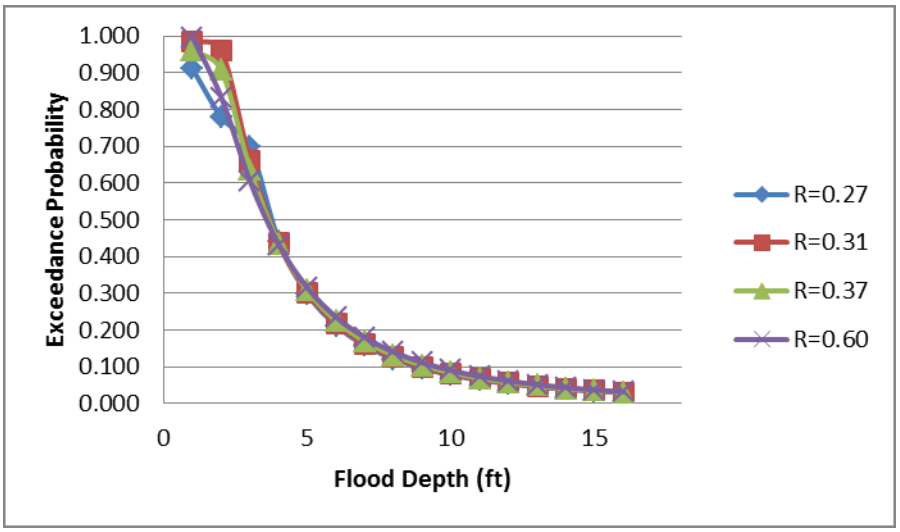

Figure 4-51: Exceedance Probabilities Corresponding to Total Flood Depths Calculated Using the Copula Procedure for All Four Levels of Correlation

The flood risk calculations for each of the four levels of correlation are presented in Figure 4-52. The minor differences in exceedance probabilities observed in Figure 4-51 have been amplified by the multiplication of the probabilities with the vulnerability and consequence values. The flood risk calculations were observed to vary the most in the mid-range flood depths. Though the exceedance probabilities were observed to decrease as flood depth increased, the increasing vulnerability and consequence curves result in the peak flood risk occurring between flood depths of 4 to 6 feet. Due to the decreasing exceedance probabilities with increasing flood depth, the calculated flood risk decreased as flood depth increased, despite the increased vulnerability and consequences as flood depth increases. As flood depth increased, the flood risk that corresponded to a given flood depth was observed to slightly increase as the level of correlation increases, due to the previously observed increase in exceedance probabilities with increasing levels of correlation. However, the differences in calculated flood risks between the correlation scenarios are overall minimal, suggesting that the trend may not be significant. These results do suggest that the level of correlation between the samples will have some influence on flood risk calculations, but the influence may not be significant.

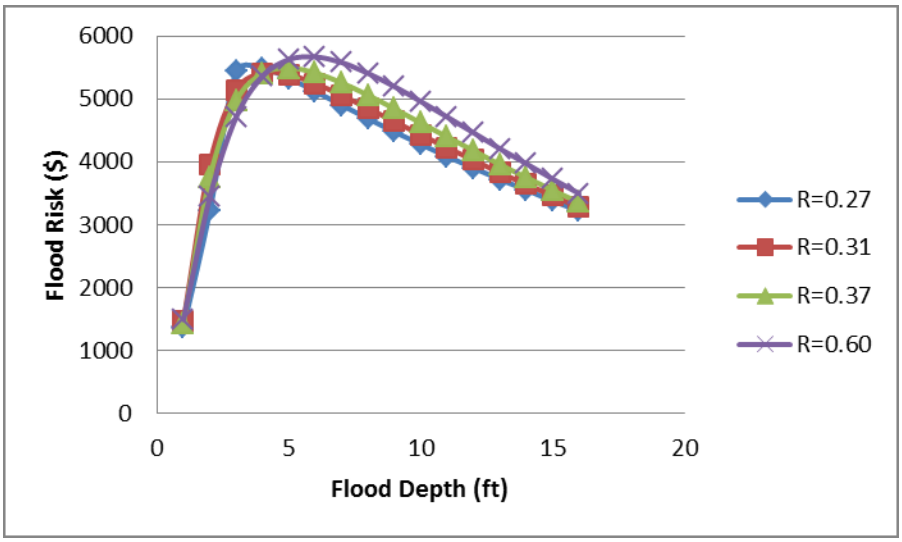

Figure 4-52: Flood Risk Calculations Corresponding to Total Flood Depths for All Four Levels of Correlation

4.3.3. Conclusions

Flood risks were calculated for each of the simulation scenarios to illustrate the effect that a comprehensive flood frequency analysis would have on flood risk assessments. These analyses also evaluated whether or not variations in sample size or the level of correlation between the riverine and tidal samples would influence the flood risk calculations. For two flood sources, the exceedance probabilities calculated based on the simulation scenarios were used to calculate flood risk. Variations in the sample size were not observed to result in impacts to the flood risk calculations. Given that the sample size was also not observed to impact the results of the joint distribution or the process of fitting distributions to the total flood depths to represent the probabilities of total flood depths, it was not expected that sample size would have a significant impact on flood risk calculations. It can be surmised that as sample size increased a better representation of the exceedance probability would be possible, as the population representing the flood probabilities would be better described, which would inherently result in more confidence in the flood risk

calculations; however, trends were not observed to support or counteract this supposition.

The level of correlation between the riverine and tidal flood depth samples was observed to have some influence on the results of the flood risk calculations. A general trend of increasing exceedance probabilities as the level of correlation increased was observed in these scenarios. This also resulted in slight increases in the flood risk calculations as the level of correlation increased, though there were some exceptions to this trend. It was expected that the level of correlation would have some impact on the flood risk calculations, as the correlation was observed to impact the joint distributions developed using the copula and a trend was observed in the parameters of distributions that were fitted to the total flood depths. These distributions were used to calculate the exceedance probabilities used in the flood risk calculations; thus, some impact was expected to be observed in the flood risk calculations. Overall, these results suggest that the level of correlation between the riverine and tidal samples has more influence on the results of the analysis than does the sample size, which was observed in previous steps of the analysis as well.

CHAPTER 5

TWO OBSERVED FLOOD SOURCES

5.1. Introduction

The simulation studies presented in Chapter 4 provide an understanding of the impact of sample size and sample correlation on the development of flood frequency assessment based on two flood sources. Because only a limited amount of observed data exist, these simulations studies were necessary to provide a context for results obtained from the observed data. The process of developing a comprehensive flood frequency analysis for a location in Florida, considering both a riverine and tidal flood source, will be discussed herein and results of this procedure will be presented.

5.2. Description of Experimental Location

A location of interest was identified close to both a river and the coast on the east coast of Florida. Discharge data (ft^3/s) from flow gage 02249007 on the Eau Gallie River at Heather Glen Circle at Melbourne were obtained. The tidal gage from which tidal height measurements (ft above NAVD 88 datum) were obtained was 8721604, located near Trident Pier. The tidal gage provided measurements of water height in feet. This tidal gage is approximately 20 miles from the location of interest, which is not ideal, but a closer location was not available.

5.3. Assessment of the Hazard

The first step to any risk assessment is to assess the hazard. The hazard is usually considered as the probability of a specific event occurring. The following sections will outline a methodology and present the results of assessing the flood hazard for a location of interest that could be impacted by two flood sources.

5.3.1. Fitting Marginal Distributions to Observed Annual Maximum Events

The first step to assessing the flood hazard when multiple flood sources must be considered was to understand the probability of flooding from each of the two flood sources. This first required calculating flood inundation depths based on observed gage measurements, and then identifying a series of annual maximum events. Finally, marginal distributions were fitted to the annual maximum flood depth events.

5.3.1.1. Calculation of Riverine-Caused Inundation Depths

To perform a flood frequency analysis for the location in Florida required observed riverine flow data. Discharge data were obtained from the United States Geological Society and information from this location, including elevations and channel characteristics, were obtained from ArcGIS and GoogleEarth for use in transforming the discharge data to flood depths at the location of interest. The process of transforming the discharges to inundation depths at the location of interest is explained here. The process involves two steps: (1) transposing the discharge from

the gaged site to a site on the river adjacent to the point of interest and (2) transforming the river discharge to an inland depth.

Prior to obtaining riverine-caused flood depths at the location of interest, it was necessary to obtain discharge measurements at a flow gage. Because the number of river flow gages is minimal, it is unlikely that a flow gage would be located at a point within the stream that is adjacent to the location of interest. Therefore, the discharges measured at the flow gage had to be transposed downstream to a point adjacent to the location of interest, which used the drainage area-ratio method (McCuen and Levy, 2000):

$$Q_u = Q_g \left(\frac{A_u}{A_g}\right)^n \tag{5-1}$$

where Q_u is the discharge at the ungaged location (cfs), Q_g is the discharge measured by the gage (cfs), A_u is the drainage area of the ungaged point (mi^2), A_g is the drainage area at the location of the flow gage (mi^2), and n is an empirical constant, with a value of 0.8 empirically derived. For the study location in Florida, A_g was 3.8 mi^2, as reported by the USGS, and A_u was 12.4 mi^2, as calculated in ArcGIS. The result of this step was a series of stream discharge estimates transposed downstream to the stream cross-section adjacent to the location of interest.

The channel cross-section information at the location of interest was not available; therefore, an estimate of the bankfull channel dimensions was required. The bankfull channel dimensions were first estimated at the location of the gage. Measured discharge and the corresponding gage heights were obtained from the gage.

The flow depth corresponding to each discharge was calculated by subtracting the gage datum, which was 11.39 feet, from the gage height measurements. Flow depths were plotted against corresponding discharges, and a power model was fitted to describe this relationship. The power model used to describe the relationship between discharge and flow depth was:

$$Discharge = 1.2212 * Depth^{4.6117} \quad (5\text{-}2)$$

Based on the observed gage data, the bankfull discharge at the location of the gage was identified to be 428 cfs. Using Equation 5-2, the bankfull depth was calculated to be 3.563 feet. Assuming a rectangular channel with a hydraulic radius approximately equal to the flow depth, Manning's equation was calculated to determine the corresponding channel width at bankfull flow. The bankfull channel width at the location of the gage was estimated to be 99.8 feet. This provided an estimate of the channel bankfull dimensions at the location of the gage.

The next step was to estimate the bankfull channel dimensions at the downstream location. The ratio of channel width to depth was assumed to remaine constant along the length of the channel. Using the downstream discharges calculated using Equation 5-1 and Manning's equation, bankfull channel width was solved. In doing so, channel depth was approximated as 0.0357 times the channel width, based on the ratio of channel depth to width at the gage location. This resulted in an estimated bankfull channel width of 123.5 feet at the location of interest. The ratio of channel depth to width then provided an estimate of the bankfull depth to be 4.409 feet.

Once bankfull channel dimensions had been estimated, the stream discharge values could be translated to flood inundation depths at the location of interest. Flow depths corresponding to the transposed discharges were calculated based on Manning's discharge equation. This was an iterative process, where depth was varied until the corresponding calculated discharge equaled the observed discharge rate. Manning's equation for discharge (ft^3/s) is:

$$Q = \frac{1.49}{n} * A_x * (\frac{A_x}{WP})^{2/3} * S_f^{1/2} \quad (5\text{-}3)$$

where n is the Manning's roughness coefficient, assumed to be 0.05, A_x is channel cross-sectional area (ft), WP is wetted perimeter (ft), and S_f is the channel slope, determined to be 0.0007 ft/ft. Both A_x and WP are dependent on channel geometry. A trapezoidal floodplain was assumed, with a slope of 49.2 feet horizontally for every one foot increase vertically. The values of channel and floodplain slope were determined using elevation data provided in ArcGIS for the Florida location. Cross-sectional area and wetted perimeter (see Figure 5-1) were calculated as:

$$A = W_b * d + z(d - D_b)^2 \quad (5\text{-}4a)$$

$$WP = W_b + 2 * D_b + 2 * (d - D_b) * \sqrt{1 + z^2} \quad (5\text{-}4b)$$

where d was flow depth and z was 49.2 feet/feet, determined based on floodplain slope. The channel geometry is illustrated in Figure 5-1. This iterative procedure identified the depth of flow corresponding to each discharge value in the sample obtained from the Florida flow gage.

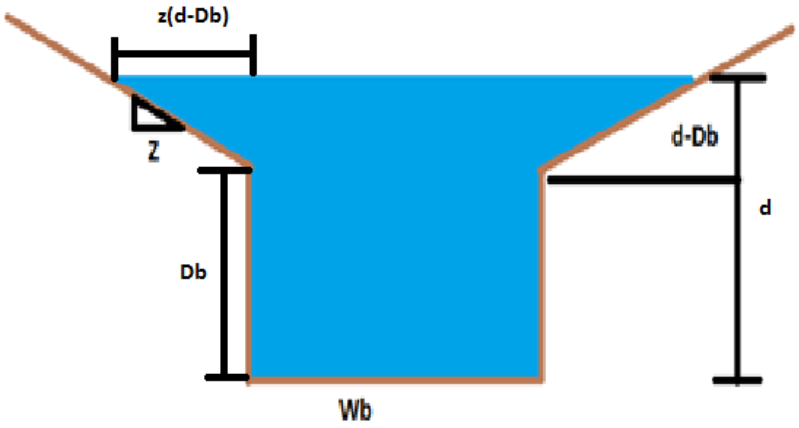

Figure 5-1: Assumed Channel Geometry

Once flow depths had been calculated, flood inundation depths at the location of interest could be determined. The next step in this process was to add the calculated flow depths to the channel bed elevation to obtain flow elevation. The channel bed elevation was assumed to be 0.4 feet above the North American Vertical Datum of 1988 (NAVD 88). This elevation was determined based on elevations provided by ArcGIS for the location in Florida. To determine the flood inundation depth at the location of interest, the elevation of the location of interest was then subtracted from the calculated flow elevation. The location of interest was assumed to be at an elevation of 4.2 feet above NAVD 88. This elevation was determined based on elevations provided by ArcGIS for the location in Florida. By following these steps, an estimate of the flood inundation depth at the location of interest can be calculated for each discharge value measured by the flow gage.

The procedure just explained was used to calculate flood inundation depths at a location of interest in Florida based on riverine discharge measurements taken by a local flow gage. Once flood inundation depths were calculated, the annual maximum events were identified and used as input to the remainder of the methodology.

5.3.1.2. Calculation of Tidally-Caused Inundation Depths

To determine the flood inundation depth at the location of interest due to tidal flooding, a procedure identified in a 1977 National Research Council report (NRC, 1977) to the Federal Insurance Administration that examined the effects of wave action on storm surge was used. The mean still-water elevation, which is the sum of astronomical tide and storm surge, the sum of which can be determined from tidal gage measurements, was determined at the location of the tidal gage. Next, the additional height caused by wave action was calculated and added to the still-water elevation. The shoreline geometry is illustrated in Figure 5-2. The equation to determine wave height is (NRC, 1977):

$$H = 0.78 * F * E_s \tag{5-5}$$

where H is the wave height (ft), F is a factor related to the length of fetch, which is the length of water over which the wind has blown, and E_s is the stillwater elevation (ft above NAVD88). For the Florida location, the fetch was taken as essentially unlimited, as the location opens to the ocean rather than a bay, which corresponded to a value of 1.0 for F. This provided the additional height of waves at the location of the tidal gage. In Table 5-1 the relationship is provided between fetch length and fetch factor, F, as indicated in NAS (1977).

Table 5-1: Relationship Between Fetch Length and Fetch Factor, F, Obtained from NAS (1977)

Fetch (miles)	F (Fetch Factor)
0.125	0.25
0.25	0.32
0.5	0.41
1	0.52
2	0.65
4	0.78
10	0.93
>20	1

The NRC procedure would next allow for the wave height past an obstruction, such as a dune or seawall, to be calculated. However, from GoogleEarth it was not possible to determine whether or not such an obstruction existed. Thus, it was assumed that an obstruction between the coastline and the location of interest did not exist.

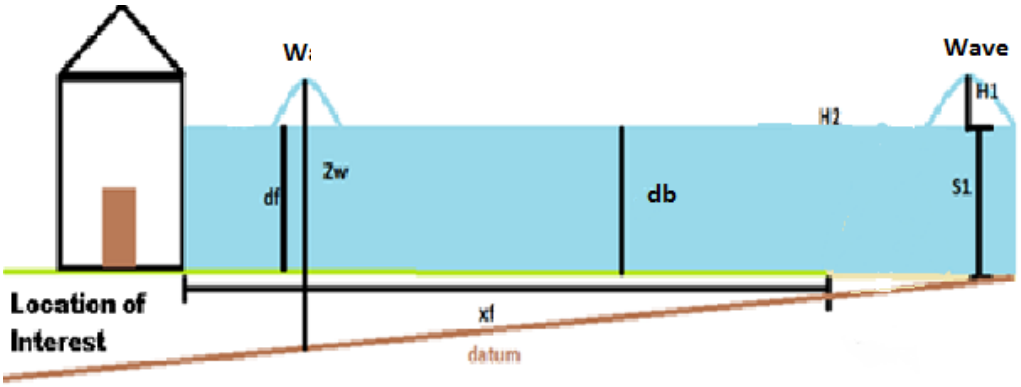

Figure 5-2: Shoreline Geometry, where S1 is Stillwater Elevation, H1 is Wave Height, db and df are Water Depths, H2 is Inland Wave Height, xf is Inland Fetch, and Zw is Flood Elevation, the Green Line is the Ground Surface, and the Brown Line is the Official Datum

The next step was to account for wave generation as the tidal surge moved inland between the dunes and the location of interest. This wave generation would primarily be caused by wind moving over the water surface. As the water travels over the land surface, additional waves may be generated, primarily through the

effects of wind. Wave height at the end of the inland fetch, H_f, was calculated as (NRC, 1977):

$$H_f = \sqrt{(G * d_f)^2 + H^2} \quad (5\text{-}6)$$

where G was a unitless value related to the inland fetch length, d_f was the average depth over the inland fetch (ft), and H was the wave height (ft) calculated behind the dunes. For the Florida location of interest, the inland fetch was determined to be 0.02 miles, estimated using GoogleEarth, corresponding to a G value of 0.2. The average depth, d_f, was calculated as the stillwater elevation minus the average elevation between the coastline and the location of interest, which was taken to be 1 foot above NAVD 88 based on measurements in ArcGIS.

Finally, the flood elevation at the location of interest, Z_w, was calculated using (NRC, 1977):

$$Z_w = E_s + 0.7 * H_f \quad (5\text{-}7)$$

where E_s was the stillwater elevation (ft) previously calculated based on tidal gage measurements and H_f was the wave height (ft) calculated at the location of interest (Equation 5-6). From Equation 5-7, the flood inundation at the location of interest was calculated by subtracting the elevation of the location of interest, which was 4.2 feet above NAVD 88 for the Florida location, from the Z_w values calculated.

The procedure just described was used to determine tidal flood inundation depths at the location of interest. Once the tidal flood inundation depths had been

calculated for the location of interest, the annual maximum events were identified. The annual maximum events were used as input to the remainder of the methodology.

5.3.1.3. Calculation of Flood Frequency Curves for Each Source Individually

To calculate a flood frequency curve for a single flood source requires identification of the annual maximum flood series. The annual maximum events were ranked from largest to smallest, with the largest magnitude assigned to a rank of 1 and the smallest magnitude a rank of n. After developing a ranked list of annual maximum flood inundation depths at the location of interest caused by an individual flood source, the Weibull probability equation was used to determine the exceedance probability of each event. Weibull's plotting position formula is:

$$P_i = \frac{i}{n+1} \tag{5-8}$$

where P_i is the exceedance probability, i is the rank of the given flood depth, and n is the number of sample values in the record. The flood depths were plotted against the non-exceedance probability, calculated as 1.0 minus the exceedance probability, which provides the probability of the flood depth not being equaled or exceeded in a given year. Based on the pairs of data (non-exceedance probability, flood depth), an equation was fitted that would provide non-exceedance probabilities for any flood inundation depth value.

Though Bulletin 17B recommends the Log-Pearson Type III distribution be used to model riverine flood flows, and the GEV distribution is frequently used to model tidal flood heights, clear evidence that either of these distributions provides

greater accuracy than that provided by other distributions is not available. Further, the variable of interest in this research was not discharge or tidal heights, as would be typical in a riverine or tidal flood analysis, but rather, was flood inundation depths at the location of interest. It was possible that the typical distributions would not be the most appropriate distributions to represent flood inundation depths. Thus, several distribution functions were fitted to the data to determine the most appropriate distribution for each data set. All tested distributions were plotted against the flood inundation depth data for comparison and probability plot correlation coefficients were calculated for each distribution to further assess the goodness-of-fit. For riverine-caused flood depth data, the distributions tested included the log-Pearson type III, lognormal, gamma, and Weibull distributions. For tidally-caused flood depth data, the distributions tested included the extreme value, generalized extreme value (GEV), and Rayleigh distributions.

5.3.1.4. Calculation of Flood Frequency Curves Considering Two Flood Sources

To develop a joint flood frequency curve requires the marginal distributions, or individual flood frequency curves that correspond to the two individual flood sources. The typical process for conducting a flood frequency analysis, outlined in section 5.3.1.3., was followed to determine the marginals, with one exception. When two potentially interacting flood sources must be considered, the maximum flood event experienced at the location of interest in a given year may not be caused by one individual source, but instead by a combination of the two sources. Thus, to identify the annual maximum flood depth at the location of interest, the riverine and tidally-caused flood inundation depths at the location of interest were summed for

corresponding time periods to identify the maximum total depth of flooding at the location of interest in each year. The assumption that the contributing flood depths from each source could be summed was a convenient assumption to make for the purpose of conducting this research; however, this methodology was not determined based on a technical analysis of the optimum approach for combining flood sources.

It was necessary to identify the annual maximum events while considering both potential flood sources. As each of the gages from which observed data were obtained took hourly measurements each day, the hourly measurements from each gage were summed and a single daily maximum flood depth at the location of interest was identified. Next, the annual maximum flood depth was identified based on the daily maximum events. This does not suggest that the annual maximum event could only occur through a combination of the two sources; this procedure could still identify an annual maximum event that was caused by only one source. The riverine contributions to the annual maximum flood inundation depth and the tidal contributions provided two flood series and flood frequency curves were developed from each. These flood frequency curves served as the marginal distributions used as input to the copula process.

5.3.1.5. Description of Results

In order to use the copula approach to develop joint distributions for the riverine and tidal data, the riverine and tidal flood depth samples must first have marginal distributions fitted. The log-Pearson Type III, lognormal, gamma, and Weibull distributions were fitted to the observed riverine sample. The extreme value, GEV, and Rayleigh distributions were fitted to the observed tidal sample. The fitted

parameters for each probability distribution, along with the probability plot correlation coefficient calculated for each, are presented in Table 5-2. The probability plot correlation coefficients calculated for the riverine distributions exhibit much more variation than was observed in the simulation studies presented in Chapter 4. This suggests that the choice of marginal distribution may have more impact on the observed data. For the tidal sample, all three of the distributions appear to provide reasonable goodness-of-fit, as all three distributions have probability plot correlation coefficients above 0.93. In the simulation studies, the probability plot correlation coefficients typically indicated that the GEV and Rayleigh distributions would result in reasonably good fits to the generated data but that the extreme value distribution would not result in an acceptable fit.

Table 5-2: Comparison of Parameters and Goodness-of-Fit Statistics for Four Marginal Distributions Fitted to Riverine and Tidal Observed Data

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	20.6015	-1.3118	0.4078	0.5271	N/A	-0.2224	N/A
Scale	0.4603	2.0893	3.3007	0.7652	1.1498	1.1038	1.3356
Location	-10.7949	N/A	N/A	N/A	2.0773	1.0476	N/A
Probability Plot Correlation Coefficient	0.7970	0.8402	0.9464	0.9150	0.9329	0.9648	0.9447

Figure 5-3 visually compares the four probability distributions of the observed riverine data. The riverine flood depth sample was based on only seven non-zero flood depths. In numerous years, as would be expected, the maximum discharge in the stream did not exceed the bankfull discharge, so no flooding occurred at the location of interest. In other years, though the river exceeded bankfull discharge at

some point in the year, the annual maximum total flood depth was caused by a large tidal event with no contribution from the river. Thus, it was necessary to attempt to fit a distribution to a sample in which more than half of the record contained zero-values. In order to consider multiple distributions to represent the riverine data, a very small value, which would not be large enough to impact the distribution of flood depths, was added to the zero-values. This is one strategy that can be employed in cases in which zero-flood years must be analyzed (Jennings and Benson, 1969). A value of approximately 1% of the maximum observed flood depth, or 0.05 feet, was added to each zero-value in the record, and all four distributions were fitted. The results of fitting the gamma distribution to this modified data set were compared to the results of fitting a gamma distribution to the original data set, including zero-values, to confirm that the addition of the small values did not significantly alter the probabilities calculated for the larger events. The addition of the small values was not found to influence the parameters of the fitted gamma distributions or the probabilities calculated using the gamma distribution; thus, the modified data set was used so that all four distributions could be considered to represent the riverine flood depth data.

The probability plot correlation coefficient in Table 5-2 suggests that the gamma distribution would by far be the best distribution to represent the riverine data set. Figure 5-3 suggests that none of the four distributions are able to fit the observed data well, which is not surprising given the large number of measurements within the sample that had the same value of 0.05 feet. However, the figure does agree with the probability plot correlation coefficients that the gamma distribution would be the best

distribution to represent the data, so it was selected for further use. That the distributions do not perfectly fit the observed data is not a surprise, given the number of zero-flood years and the small sample size. Only seventeen years of measured data were available, and it was previously observed in the simulation studies that the ability to fit marginal distributions to the generated samples improved as the sample size increased from a record of 10 values to a record of 100 values. Thus, it was expected that for a sample size of 17 it would be somewhat difficult to fit accurate marginal distributions.

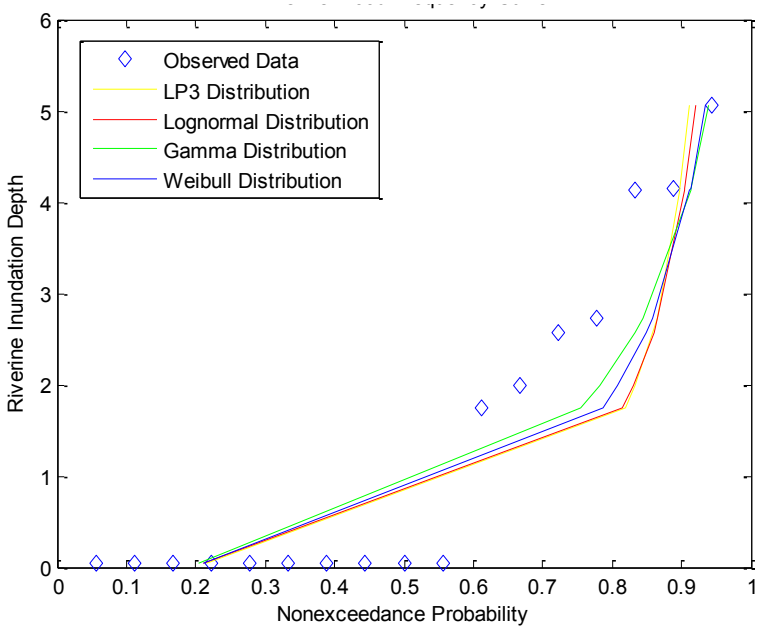

Figure 5-3: Comparison of Four Potential Marginal Distributions to Observed Riverine Flood Depth Data where Flood Depths are in Feet

Figure 5-4 visually compares the fitted distributions to the observed tidal flood depths. This figure confirms the conclusion based on evaluating the probability plot correlation coefficients, that all three distributions could provide a reasonable fit to the observed tidal flood depth data. Based on the probability plot correlation coefficient, the GEV distribution was selected as the marginal distribution

representing the observed tidal flood depths. The tidal data set contained four zero-values, which stemmed from years in which tidal flood did not occur or years in which the annual maximum total flood depth was caused by a riverine event without any tidal contribution. Though zero-values do not prevent any of three distributions considered from being used, the inclusion of 4 measurements of the same value (0 feet) in the sample decreases the ability to accurately fit any of the distributions. Between the zero-values and the small sample, it was expected that the marginal distribution would not provide as strong a fit to the observed sample.

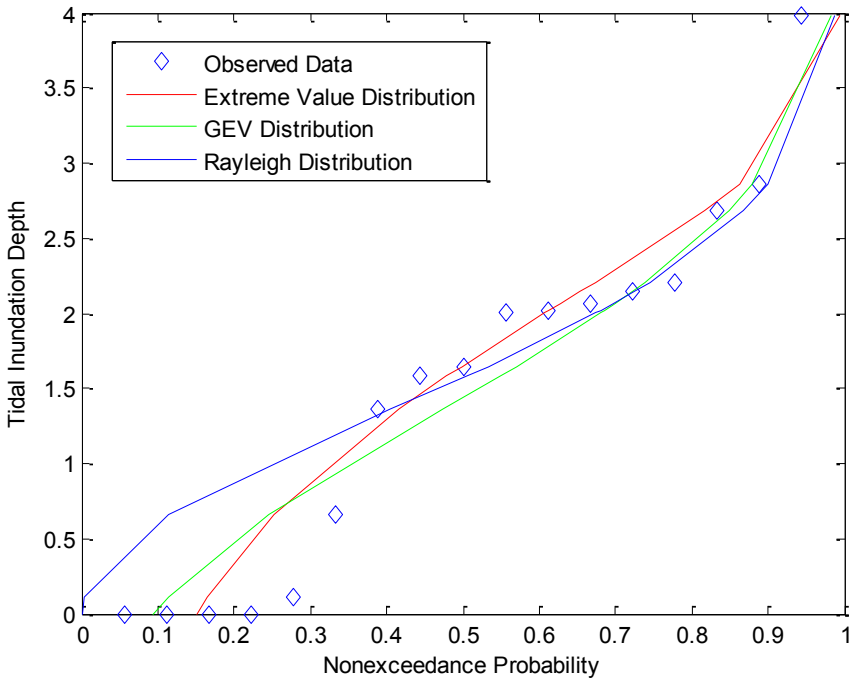

Figure 5-4: Comparison of Two Potential Marginal Distributions to Observed Tidal Flood Depth Data where Flood Depths are in Feet

5.3.2. Use of Copula Equations to Develop Joint Distributions

Once the marginals were developed, they were used as input to a copula equation to determine the combined probability of the two flood sources occurring.

The copula provided the probabilities of specific flood depths being caused by both flood sources simultaneously.

5.3.2.1. Description of Methods

Once the marginals had been developed based on the observed riverine and tidal flood inundation depths at the location of interest, a copula could be used to determine the joint probability of depths from the two flood sources. The first step to determining a joint distribution for the two flood sources was to calculate the appropriate dependence structure between the two flood sources. Kendall's τ is often used and is particularly convenient in bivariate scenarios because a direct relationship exists between τ and the copula parameter for Archimedean copulas. Kendall's τ was determined to serve as a measure of hydrologic dependence, which was defined as floods from multiple sources that occur simultaneously or occurring due to a common underlying cause. Kendall's τ measures the frequency with which both variables simultaneously experience high or low values; thus, it was determined that Kendall's τ could be used as a measure of hydrologic dependence between the flood sources.

Once Kendall's τ was calculated, the most appropriate family of copulas to represent the riverine and tidal samples had to be determined. Numerous families of copulas are available from which to choose. Several families were fitted and compared using Akaike's Information Criteria (AIC). The specific method of using AIC was described in Chapter 1 of this document. When AIC values are calculated, the smallest AIC value indicates the most appropriate copula family (Klein *et al.*, 2010). The Gumbel, Clayton, and Frank copula families were chosen for comparison as they are three of the more popular Archimedean families. For each

family, the copula parameter was calculated as a function of Kendall's tau, a measure of association between the two data sets. The equations relating tau and the copula parameters for the Gumbel-Hougaard, Clayton, and Frank families, respectively, are as follows:

$$\tau = 1 - \alpha_{GH}^{-1} \tag{5-9a}$$

$$\tau = \frac{\alpha_C}{\alpha_C + 2} \tag{5-9b}$$

$$\tau = 1 - \frac{4}{\alpha_F}\left(D_1(-\alpha_F) - 1\right) \tag{5-9c}$$

where α is the copula parameter, τ is Kendall's tau, and D_1 is the first-order Debye function which is calculated as:

$$D_1(-\alpha) = \frac{1}{-\alpha^1}\int_0^{-\alpha} \frac{t^1\, dt}{e^t - 1} \tag{5-10}$$

Next, the copula cumulative distribution functions for the Gumbel-Hougaard, Clayton, and Frank families, respectively, were calculated as:

$$C(X,Y) = \exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\} \tag{5-11a}$$

$$C(X,Y) = (u^{-\alpha} + v^{-\alpha} - 1) \tag{5-11b}$$

$$C(X,Y) = \frac{1}{\alpha}\ln\left[1 + \frac{(\exp(\alpha u) - 1)(\exp(\alpha v) - 1)}{\exp(\alpha) - 1}\right] \tag{5-11c}$$

where α is the copula parameter and u and v are the marginals, which were calculated from the cumulative distribution functions of the original random variables, riverine and tidal flood depths. The calculated copulas provided the cumulative joint distribution of combinations of riverine and tidally-caused flood inundation depths for the location of interest. Joint pdfs that correspond to the variables u and v were calculated by taking the second derivative of the joint cdf with respect to both u and v. To obtain joint pdf values that correspond to the riverine and tidal variables required that the joint pdfs calculated for the variables u and v then be multiplied by the marginal distribution pdf values (Wang *et al.*, 2009). The equations for the joint pdfs, expressed in terms of the riverine and tidal variables, are as follows:

$$f(x,y) = \frac{\partial^2 C(u,v)}{\partial u \partial v}\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} \quad (5\text{-}12)$$

where u and v represent the marginal distributions fitted to the riverine and tidal flood depths and x and y represent the riverine and tidal flood depths, respectively.

5.3.2.2. Description of Results

The following section will provide the results of using the copula approach to develop a joint distribution of riverine and tidal flood depths. Prior to using the copula approach, a copula parameter was first determined. Both a probability distribution and a cumulative distribution were developed using the copula approach and plots are provided.

5.3.2.2.1. Copula Fitting to Develop Joint Distributions

In order to calculate joint distributions for the observed data the copula parameter was calculated as a function of Kendall's τ. For the observed data from the riverine and tidal gages, the calculated value of τ was -0.4689. This value indicates that high flood inundation depths did not typically result from both sources simultaneously, but rather that high flood depths caused by one of the sources were most often accompanied by low depths caused by the other flood source. Because Kendall's τ had a negative value, only the Frank copula family could be used to develop the joint pdf and cdf. Based on the calculated copula parameter value, the Frank copula parameter was determined using Equation 5-9c to be -5.1978. Based on the simulation studies, a negative copula parameter was expected given the negative value of Kendall's τ. The simulation studies also revealed that the magnitude of the copula parameter increases as the magnitude of Kendall's τ increases; thus, a copula parameter on the order of -5 was expected given the magnitude of Kendall's τ calculated for the observed data. However, it should be noted that the numerous zero-values in the riverine and tidal samples may have impacted the copula parameter as well, since all of those values would have equal ranks. Thus, when Kendall's τ was calculated for the observed samples, it was based on a number of values with the same rank.

5.3.2.2.2. Plotting Joint Distributions through Copulas

The joint distributions calculated for the observed riverine and tidal flood depth samples were next investigated. The joint pdf calculated for the riverine and tidal samples is presented in Figure 5-5. The peak of the joint cdf was observed to be quite broad, as was observed in the simulation studies for scenarios with small sample

sizes and low levels of correlation between the samples. The location of the peak in the joint pdf also appears to be reasonable and rational given the observed riverine and tidal flood depth samples. The riverine sample has numerous very low values; thus, the peak occurs at a very low riverine flood depth values. The tidal flood depths ranged from approximately 0 to 4 feet, with the peak of approximately 2.5 feet.

Figure 5-6 shows the joint cdf for the riverine and tidal observed data sets. As would be expected given the riverine flood depth sample, a very steep rise was evident in the joint distribution parallel to the riverine flood depth axis. The riverine sample consisted of numerous zero flood depth values and otherwise ranged from approximately 1 foot to 5 feet. The tidal flood depth sample contained several low flood depth values, which resulted in a less steep slope parallel to the tidal flood depth axis than was observed in the simulation studies. Because the tidal flood depth sample had fewer zero flood depth values than the riverine flood depth sample, the slope parallel to the tidal flood depth axis was less steep than the slope along the riverine flood depth axis. The general shape of the joint cdf agrees with the shape of the joint cdfs developed for low sample size and low sample correlation within the simulation study.

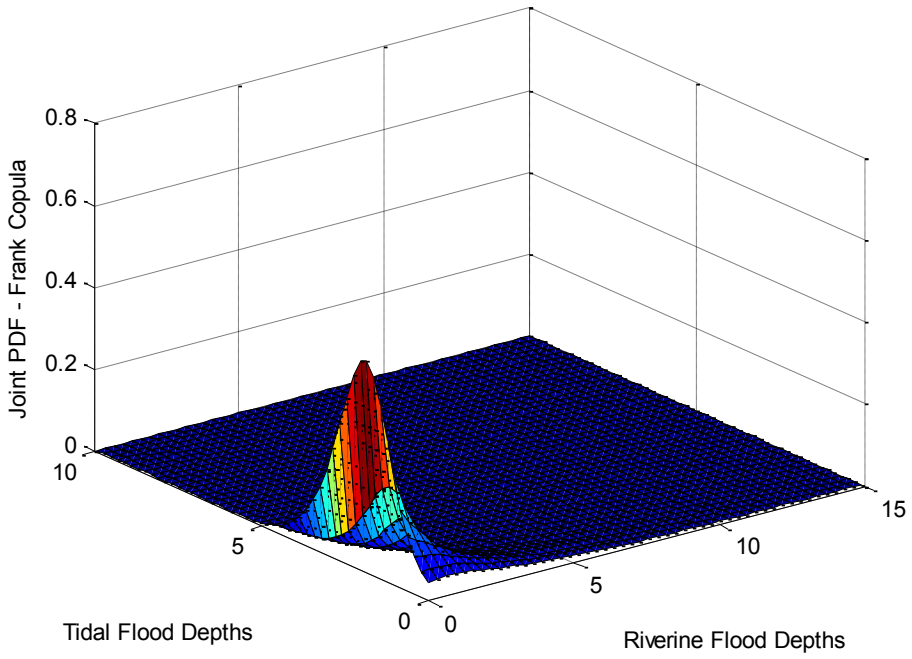

Figure 5-5: Joint Probability Distribution Developed by Copula for Observed Riverine and Tidal Flood Depths where Flood Depths are in Feet

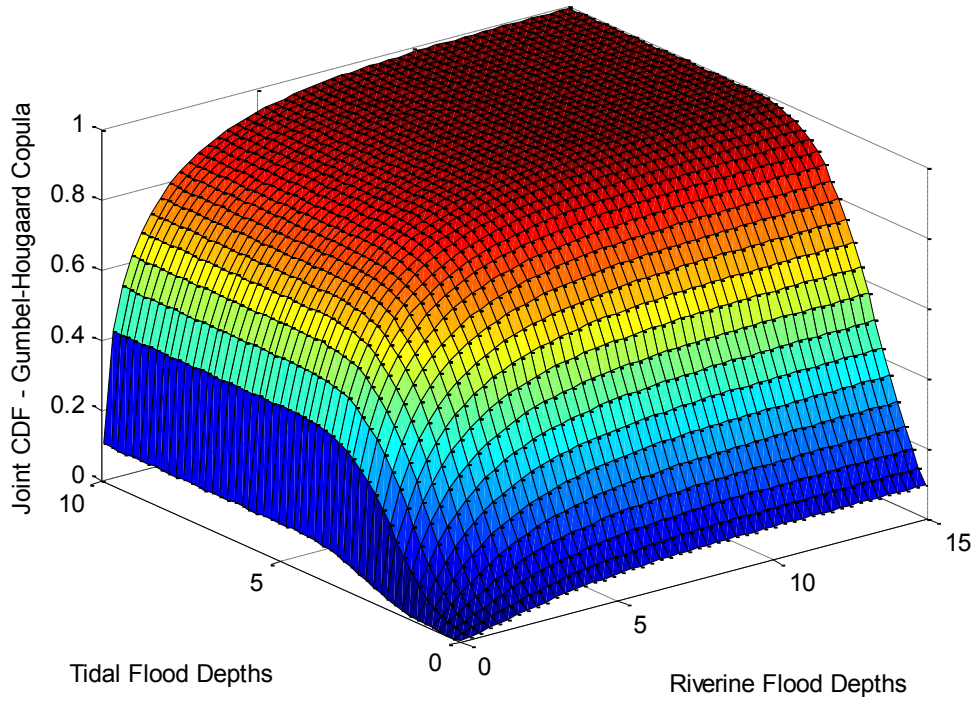

Figure 5-6: Joint Cumulative Probability Distribution Developed by Copula for Observed Riverine and Tidal Flood Depths where Flood Depths are in Feet

5.3.3. Calculation of Combined Flood Frequency Curves

The third step in developing a flood frequency assessment based on multiple flood sources used the joint pdf as input. While the joint pdf provided probabilities of given combinations of flood events, it did not provide probabilities of total flood depths with consideration of the fact that the flood depths could be caused by one or the other source individually or by some combination of the two sources. The final step of the procedure calculated the probability of total flood depths while considering the multiple ways in which those flood depths could occur.

5.3.3.1. Description of Methods

The results of the copula procedure described in section 5.3.2 were joint pdfs and cdfs, which can provide the probability that corresponds to a specific joint event. However, the primary interest is the probability of a certain flood depth occurring, when that flood depth might be due to either one of the sources individually, or some combination of the two sources, such that many possible combinations could result in the desired flood depth. Because the riverine and tidal contributions to flood inundation at the location of interest were summed in the process of identifying annual maximum events, flood depths were also summed to identify probabilities corresponding to total flood depths (D_T). Within the joint probability distribution, the area corresponding to a specific total flood depth was identified as bounded by a triangle of the riverine inundation depths (D_{RT}) and tidal inundation depths (D_{TT}) axes and a line connecting the points (0, tidal depth of interest) and (riverine depth of interest, 0). In other words, the points where the flood inundation depth of interest was caused only by one of the sources, either riverine or tidal, were identified and

connected by a line. All of the points on this line summed to a total depth equal to the depth of interest, and points inside the bounding triangle summed to total depths less than the depth of interest. An illustration of the procedure is shown in Figure 5-7.

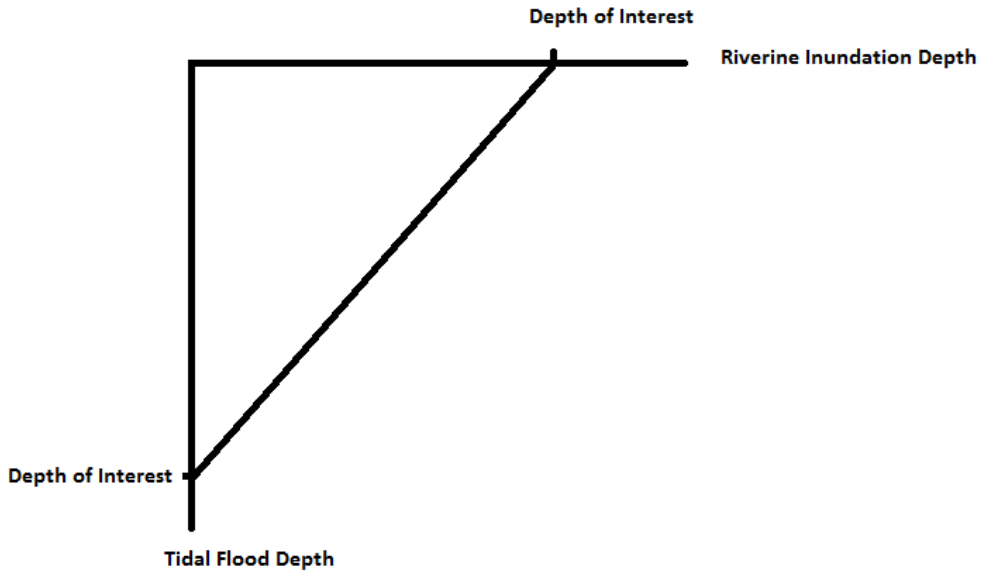

Figure 5-7: Graphical Illustration of Area Corresponding to Certain Flood Inundation Depth

The volume under the joint probability distribution within this bounding region represents the non-exceedance probability for a given flood depth. This was calculated using a double integral. The lower and upper bounds of the outer integral were 0 and total depth of interest (D_T), respectively. The lower and upper bounds (B_U) of the inner integral were 0 and:

$$B_U = D_T - D_{RT} \tag{5-13}$$

As an example, if the total depth of interest was 6.5 feet and the riverine flood depth was 3.5 feet, then the maximum tidal flood depth would be 3 feet. In this scenario, the outer integral lower and upper bounds would be 0 feet and 6.5 feet, respectively,

and the inner integral lower and upper bounds would be 0 and (6.5 feet – 3.5 feet), or 3 feet, respectively. The double integral was calculated in Matlab using the adaptive Simpson's method, a numerical evaluation method, which uses a recursive algorithm to approximate the integral based on the error in estimates calculated using Simpson's rule. The non-exceedance probability of each value of total flood depth was calculated using this double integral procedure, and then the probabilities were plotted against the total depth values.

The final task to determining the probability of total flood depths was to identify the appropriate distribution to represent the probabilities that correspond to the total flood depths. Four distributions were fitted to the total flood depths: the LP3, the GEV, the gamma, and the normal distributions. The GEV and gamma distributions were chosen for consideration because they were the two marginal distributions used to represent the simulated riverine and tidal data. The LP3 distribution was selected for consideration because it is commonly used to represent flood data. The normal distribution was selected for consideration because the purpose of the copula is to develop a joint distribution when the marginals are represented by different distributions; thus, it was believed that the results of the copula might not follow either of the marginal distributions.

These four distributions were fitted to the total flood depth samples using either Maximum Likelihood Estimation or Method of Moments and they were compared to the probabilities calculated for each copula family based on the double integral procedure. The fitted distributions were plotted against the total flood depths and compared to a plot of the non-exceedance probabilities calculated by double

integral versus the total flood depths. These plots, as well as probability plot correlation coefficients, were used to determine which of the four distributions could best represent the total flood depth populations for each simulated scenario. The result of this process was a population from which non-exceedance probabilities corresponding to total flood depths for the location of interest could be identified. From this distribution, the true nature of the flood hazard for the location of interest can be understood, which was the ultimate goal of this research.

5.3.3.2. Description of Results

The final step to develop a flood hazard analysis that considers multiple flood sources was to determine the non-exceedance probabilities that correspond to total flood depths. Four probability distributions were considered to represent these non-exceedance probabilities, the LP3, the GEV, the gamma, and the normal distributions. Table 5-4 presents the fitted parameters for each distribution as well as the probability plot correlation coefficients calculated to assess the fit of each distribution to the total flood depth values. As was observed in the simulation studies, the probability plot correlation coefficients indicate that the LP3, GEV, and gamma distributions would all be able to acceptably fit the observed data; however, the normal distribution would not fit the data so well. The probability plot correlation coefficients suggest that the LP3 and GEV distributions should perform essentially identically and that either could be selected to represent the total flood depths.

Table 5-3: Parameter Values and Probability Plot Correlation Coefficients for Distributions Fitted to Total Flood Depths

	LP3	GEV	Gamma	Normal
Shape	9.0005	0.3650	5.9509	N/A
Scale	0.1382	0.6796	0.4766	1.3063
Location	-0.2876	2.1417	N/A	2.8363
Probability Plot Correlation Coefficient	0.9913	0.9945	0.9773	0.9576

Table 5-4 provides the flood depths calculated for the 10%, 2%, and 1% annual chance flood event using each of the four distributions evaluated. For the 10% annual chance event little difference is observed in the calculated flood depths between any of the four distributions. However, for the larger flood events a wider difference in calculated flood depths is observed. For the 2% annual chance event the flood depths could range from 5.5 feet to 8 feet, depending on the distribution, and for the 1% annual chance flood event the depths could range from 5.9 to 10.3 feet. This differences in how each distribution perform in the tail are responsible for the differences in the calculated flood depths. These differences suggest that though the probability plot correlation coefficients suggested that the LP3, GEV, and gamma distribution would all be well able to represent the total flood depths, outside the range of the observed data the distributions would perform quite differently. Care must be taken in selecting the distribution used to represent the total flood depths in order to obtain an accurate understanding of the larger flood depths, which are typically of the most concern.

Table 5-4: Flood Depths in Feet Calculated for Each Distribution for the 10%, 2%, and 1% Annual Chance Events

Exceedance Probability	LP3	GEV	Gamma	Normal
0.1	4.5193	4.5131	4.3906	4.5104
0.02	7.0119	8.0152	5.6987	5.5191
0.01	8.3107	10.2606	6.2128	5.8752

The four fitted distributions plotted against the non-exceedance probabilities calculated using the double integral procedure for the total flood depths are shown in Figure 5-8. This figure suggests that the LP3 and GEV distribution adequately represent the total flood depths, approximately equally well, while the gamma and normal distributions are slightly less able to fit the observed data. Based on Figure 5-8, the LP3 distribution should be selected to represent the observed total flood depths. However, none of the four distributions are observed to fit the upper tail of the observed total flood depths well at all. Though the small sample size may have some role in this, the distributions fitted to the simulation studies were observed to fit the upper tails well overall, even for a sample size of 10, which suggests that further explanation of the results for the observed total depths is necessary.

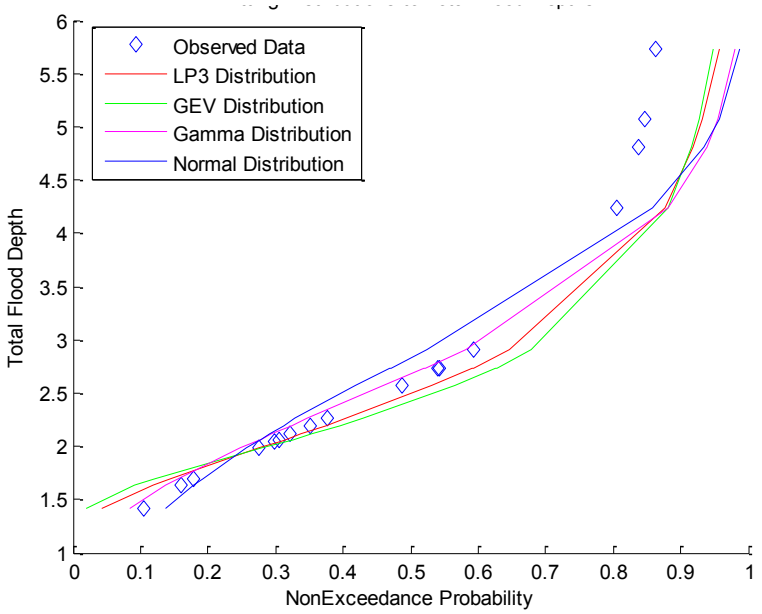

Figure 5-8: Comparison of Fitted Distributions to Non-Exceedance Probabilities Calculated for Observed Total Flood Depths where Flood Depths are in Feet

The non-exceedance probability calculated corresponding to the maximum observed total flood depth was 0.86, while the non-exceedance probability corresponding to the maximum generated total flood depth in the simulation studies was never below 0.97. The observed riverine and tidal data sets and the generated data sets used in the simulation studies differed in two ways. First, the riverine and tidal flood depth samples both contained numerous zero-values. The riverine depths sample contained 10 zero-values out of 17 values and the tidal depths sample contained 4 zero-values out of 17 values. Having so many zero-values made it very difficult to fit marginal distributions, which would influence the calculation of the copula pdf and cdf values. Also, having so many equivalent values in the flood depth samples would influence the calculation of Kendall's τ, as all the zero-values would have an equivalent rank. Thus, the zero-values may make it more difficult to define the joint distribution. The second major difference between the observed data sets and the generated data sets was the level of relationship between the riverine and tidal

flood depth samples. The annual maximum flood depth occurring at the location of interest was caused by a combination of the riverine and tidal flood sources in just 4 of the 17 years for which observed data was obtained. In the other 13 years of the record the annual maximum flood depth at the location of interest was caused by one of the flood sources alone. Thus, a joint distribution is being developed based on variables that did not frequently occur jointly. Stronger relationships between the simulated riverine and tidal flood depth samples were prescribed, such that the riverine and tidal flood depths did typically occur jointly in the simulation studies.

Several studies were conducted to verify the impact of the zero-values and the lack of a strong relationship between the observed riverine and tidal flood depth samples. First, the impact of the equivalent values in the riverine sample was assessed. The zero-values in the riverine sample had previously been replaced by a value of 0.05 feet such that the LP3 and lognormal distributions could be considered to represent the flood depth sample. In the first test, all of these values of 0.05 feet were replaced with varying values between 0.01 feet and 0.05 feet, which would influence the calculation of Kendall's τ. A second test was conducted with varying values between 0.1 and 0.5 to determine if the proximity to a value of 0 impacted the results. The first test, using varying small values between 0.01 and 0.05 was not observed to result in a significant increase in the non-exceedance probability calculated corresponding to the maximum total flood depth. The second test, adding slightly larger values instead of the zero-values was observed to slightly increase the maximum non-exceedance probability to 87%, but this was not result a significant improvement. A third and final test along these lines replaced the 4 zero-values in the

tidal sample with values ranging from 0.1 to 0.4, while keeping the varying values between 0.1 and 0.5 in the riverine sample. The maximum non-exceedance probability calculated for this scenario was above 0.9, which should be considered a significant change as compared to the original results. This suggests that the large number of zero-values in the riverine and tidal flood depth samples influences the joint distribution developed to represent the samples.

In the next study, the elevation of the location of interest was assumed to be lower than originally expected, such that deeper flood depths would be calculated and there would be fewer zero-values in both the riverine and tidal samples. When the location of interest was assumed to be at an elevation of 1 foot, as opposed to the original 4.2 feet above NAVD88, the riverine sample included no zero-values and the tidal sample included only 1 zero-value. This reduced the impact of zero-values on the analysis and also resulted in a stronger relationship between the riverine and tidal flood sources. In this case, the annual maximum total flood depth was caused by only one of the flood sources rather than by a combination of the two flood sources in only 1 year out of 17. The maximum non-exceedance probability calculated corresponding to the maximum total flood depth was above 0.91 for these flood depths. This suggests that the combination of zero-values in the flood depth samples and the lack of strong relationship between the two flood sources influence the development of joint distributions as well as calculations of the non-exceedance probabilities that correspond to total flood depths, calculated based on the joint pdf.

A final test confirmed that the marginal distributions themselves did not influence the development of joint distributions between the riverine and tidal

samples. A riverine sample and a tidal sample with a sample size of 17 were generated using the same populations that were fitted to the observed data. Marginal distributions were fitted to these samples and Kendall's τ was calculated. Based on a negative value of Kendall's τ, the corresponding parameter for the Frank copula family was calculated. The joint pdf and cdf were calculated for these samples, and then the double integral under the joint pdf was calculated to determine the non-exceedance probabilities corresponding to total flood depths. For this scenario, the maximum non-exceedance probability calculated was 0.95, which is much improved over the value of 0.86 calculated using the observed data. This suggests that the marginal distributions fitted to the observed riverine and tidal flood depth samples were not responsible for the low maximum non-exceedance probability calculated based on the observed data. Thus, it is believed that the high number of zero-values in both the riverine and tidal values, as well as the lack of a strong relationship between the riverine and tidal flood depth sources, were responsible for the unusual results calculated for the observed data sets.

5.3.4. Comparison of Flood Hazard Based On Different Assumptions

It is of interest to determine how the understanding of the flood hazard for the location of interest is changed by using the method that considers both flood sources and the dependencies between the two sources. Thus, the flood frequency assessment developed based on the joint distribution calculated by copula was compared to flood frequency assessments for both flood sources individually and to a second joint flood assessment that made the assumption of independent flood sources.

Annual maximum flood events were identified for both the riverine and tidal flood depth samples individually, without consideration of the total flood depths if flooding occurred from both sources simultaneously, in order to assess the assumption of independence. Table 5-5 provides the parameters fitted for each of the distributions used to represent the riverine and tidal observed, independent data, as well as the probability plot correlation coefficients calculated to assess the goodness-of-fit of each distribution. Figure 5-9 visually compares the four distributions to the observed, individual riverine flood depth sample. The results presented in Table 5-5 and Figure 5-9 can be compared to Table 5-2 and Figure 5-4, in which the annual maximum events were identified based on which events from both sources caused the maximum total flood depth at the location of interest, to assess the impact that consideration of the joint occurrences of the flood sources had on the identification of annual maximum riverine flood events. Small differences were observed in the parameters of the fitted distributions based on the independent and dependent riverine samples.

The number of zero-values in the riverine flood depth sample was the most significant difference based on the assumptions of dependence and independence. When identified based on the assumption of dependence, the annual maximum event was caused entirely by the tidal source in several years, so the dependent sample had a larger number of zero-values. Despite these additional zero-values in the dependent riverine sample, differences in the marginal distribution parameters calculated for each sample were fairly small between the independent and dependent samples. Further, a significant difference was not observed in the marginal non-exceedance

probabilities calculated for similar flood depths between the two samples. The assumption of independence or dependence between the flood sources is also not observed to influence which distribution fits the riverine flood depths best, based on the probability plot correlation coefficients. The gamma distribution was selected as the most appropriate distribution to represent the sample in both cases.

Table 5-5: Comparison of Parameters and Goodness-of-Fit Statistics for Four Marginal Distributions Fitted to Riverine and Tidal Observed, Independent Data

	Riverine				Tidal		
	LP3	Lognormal	Gamma	Weibull	Extreme Value	GEV	Rayleigh
Shape	113.6015	-0.7384	0.5415	0.6584	N/A	0.0824	N/A
Scale	-0.1770	1.8860	2.8120	1.1643	0.7784	0.3941	1.5825
Location	19.3636	N/A	N/A	N/A	2.4925	1.8982	N/A
Probability Plot Correlation Coefficient	0.8864	0.8670	0.9697	0.9546	0.8397	0.9533	0.9115

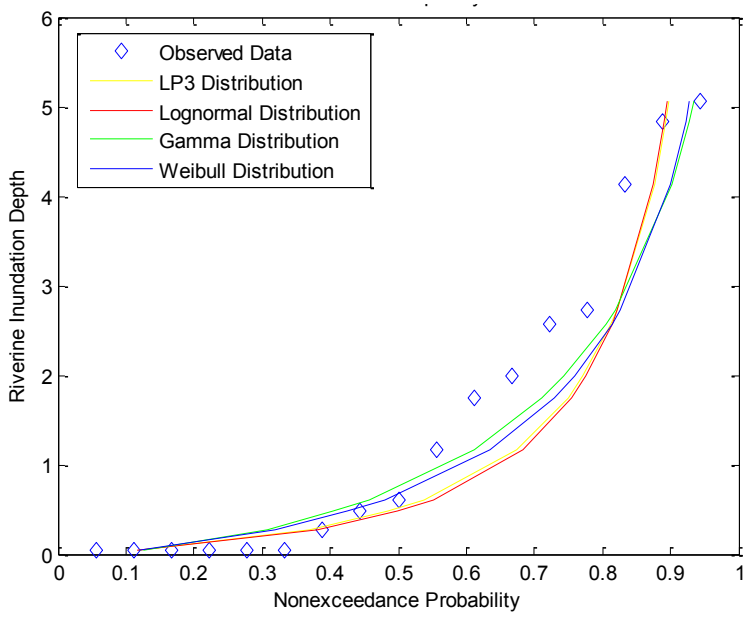

Figure 5-9: Comparison of Four Potential Marginal Distributions to Observed, Independent Riverine Flood Depth Data where Flood Depths are in Feet

Marginal distributions were also fitted to the annual maximum events from the tidal flood source that had been identified by considering only the tidal flood source rather than identifying the annual maximum flood depths at the location of interest which could be caused by joint occurrences of both flood sources. Table 5-5 provides the distribution parameters fitted to the individual tidal flood depths and the probability plot correlation coefficients calculated for each distribution, while Figure 5-10 provides a visual comparison of the three potential distributions to the observed data. In comparing these results to those presented in Table 5-3 and Figure 5-5, differences are again observed between the parameters fitted for each distribution. A primary difference in the annual maximum flood depth samples based on the assumptions of dependence and independence is again the number of zero-values in the sample. When the flood depths were considered to be dependent the annual maximum event was caused entirely by the riverine source in several years, which resulted in additional zero-values in the tidal flood record for that assumption. The differences in the non-exceedance probabilities calculated for tidal flood depths based on the assumptions of dependence and independence were most significant for the smaller flood depths, due to the additional zero-values in the dependent sample, but the differences for larger flood depths were minimal. The differences in the observed samples based on the assumptions of dependence and independence were not observed to influence the choice of marginal distribution used to fit the tidal sample. In both cases, the GEV distribution was observed to be the most appropriate distribution to represent the tidal sample.

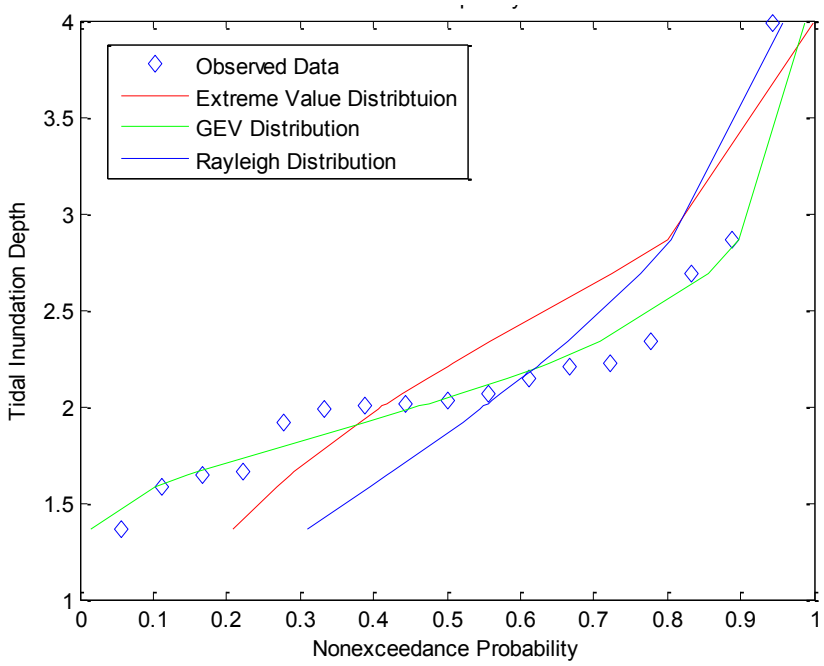

Figure 5-10: Comparison of Three Potential Marginal Distributions to Observed, Independent Tidal Flood Depth Data where Flood Depths are in Feet

Once differences in the marginal distributions based on the assumptions of dependence and independence had been identified, differences in the flood hazard assessment based on those assumptions were examined. The flood hazard assessments for the location of interest developed using four different approaches were compared. The flood hazard assessments based on the riverine source independently, the tidal source independently, the joint sources calculated based on an assumption of independence, and the joint sources calculated based on an assumption of dependence between the sources were all compared. The flood hazard assessments presented individually for each source were those calculated using the parameters of Table 5-5. The flood hazard assessment calculated based on the sources jointly with an assumption of independence between them were the exceedance probabilities calculated using the equation identified by Morris (1983) (see Equation 1-3). The flood hazard assessment calculated based on the sources

jointly with an assumption of dependence between the sources was the exceedance probabilities calculated for total flood depths, which were developed based on the copula. Figure 5-11 illustrates the exceedance probabilities corresponding to a sample of flood depths for each of the four methods used to calculate the flood hazard.

Based on these results, it is evident that considering only one of the flood sources independently could provide a much different understanding of the flood hazard for the location than considering both sources jointly. The riverine flood source, when considered independently, provides a significantly different understanding of the flood hazard for the location of interest than the tidal flood source when considered independently. The exceedance probabilities calculated using each method were compared over a range of flood depths from 0.5 feet to 10 feet. Due to the high number of zero-values in the independent riverine sample, the largest exceedance probabilities were calculated for flood depths less than 0.5 feet. By a depth of 0.5 feet, the exceedance probability calculated for the independent riverine sample had already decreased to approximately 0.60. The independent tidal sample; however, had exceedance probabilities of approximately 1.0 until the flood depths increased above 2 feet. If only the tidal flood source were considered, it would appear that a flood depth greater than 4 feet did not have a chance of occurring, which would not be the case if only the riverine independent sample were considered. This indicates that considering only one of the possible flood sources that could influence the location of interest would significantly impact the understanding of the flood hazard that was developed for the location.

Figure 5-11 can also be used to assess the assumption of independence or dependence between the two flood sources. In comparing the joint distributions calculated based on the assumptions of independence and dependence, only a small difference is observed in the exceedance probabilities. The difference in exceedance probabilities is most significant between 2 and 5 feet of flood depth, which is the range in which the independent tidal exceedance probabilities are decreasing most steeply. In this range, the assumption of dependence results in higher exceedance probabilities for a given flood depth, which means that the assumption of dependence results in a higher expectation of flooding at the location of interest than does the assumption of independence. Based on available data, it is not possible to say that one of the assumptions results in more "correct" exceedance probabilities than the other. Both joint distributions decrease more gradually than the independent tidal curve, due to the influence of the riverine curve. As the flood depths increase beyond approximately 6 feet, the four sets of exceedance probabilities converge to very low values. Based on the observed data, it was unlikely that a higher flood depth would occur. This analysis suggests that, for the location of interest, the assumption of dependence or independence between the riverine and tidal flood sources would have minimal impact on the final probabilities calculated to correspond to total flood depths. Given the lack of a strong relationship between the riverine and tidal flood sources previously discussed, this result was not unexpected.

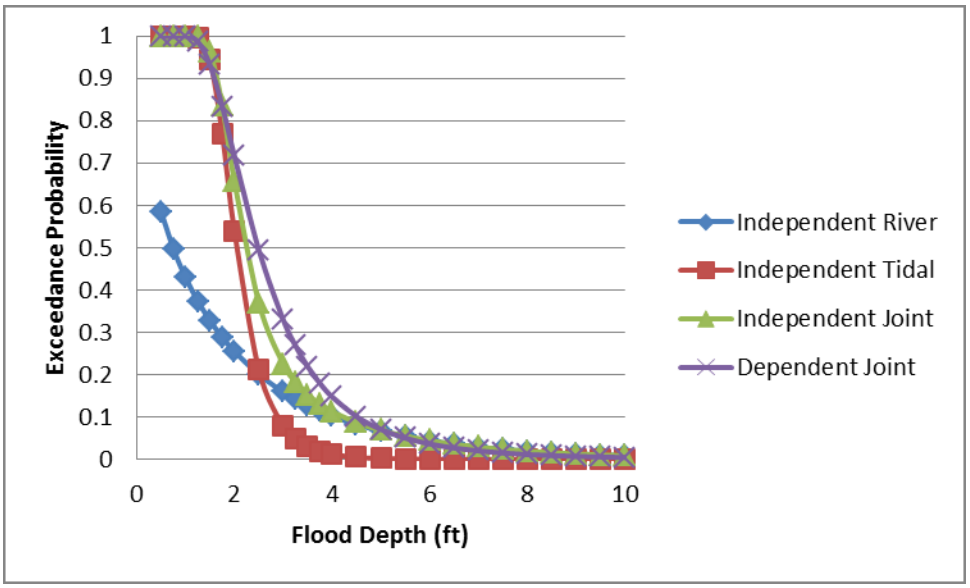

Figure 5-11: Comparison of Exceedance Probabilities Calculated Based on the Assumptions of Independence and Dependence Between the Riverine and Tidal Flood Sources

Though the assumption of independence or dependence between flood sources was not found to be important for this location, it is not believed that this would always be the case. For other locations, it is believed that this assumption could have great significance on the results of the analysis. An evaluation of the correlation between the riverine and tidal flood depth samples early in the analysis procedure could be of assistance in determining whether or not an assumption of dependence or independence should be made. If the data are found to be correlated, the assumption that the flood sources are dependent seems to be the more rational assumption to make. If the data are not found to be highly correlated, the assumption of independence would likely be reasonable, which would allow for a more simple procedure to be used to determine the probabilities that correspond to the total depths.

5.3.5. Conclusions

The process of developing a flood hazard assessment that considered both riverine and tidal flood sources was successful. The observed riverine and tidal flood

depth marginals were fitted to the gamma and GEV distributions, respectively. Fitting distributions to the riverine sample in particular was difficult due to the large number of zero-values in the sample. However, given that a river would not be expected to flood every year, these zero-values make sense from a hydrologic perspective. Due to a negative value of Kendall's τ between the two observed data sets, the Frank copula family was used to develop the joint pdf and cdf based on the marginal distributions.

The non-exceedance probabilities corresponding to total flood depths were calculated by taking the double integral under the joint pdf for each total flood depth value. This provided a prediction of non-exceedance probabilities for total flood depths for the location of interest in Florida that could be used as a general planning tool to determine the appropriate use of that land or to determine the best measures to take to protect against flooding. In fitting a distribution to the total flood depths to represent the non-exceedance probabilities, the fitted distributions agreed with the results of the double integral under the joint pdf more poorly than had been observed in the simulation studies. The non-exceedance probability that corresponded to the maximum observed total flood depth was only 0.86, while the fitted distribution estimated a non-exceedance probability of 0.95. Several factors were considered to be responsible for this discrepancy. First, the observed riverine and tidal flood depth samples had numerous zero-values, as would be expected since the location of interest should not be expected to flood on a yearly basis. These zeros interfered with the fitting of marginal distributions and the copula. Further, a strong relationship between the riverine and tidal flood sources was not evident. In only 4 of the 17

years was the annual maximum flood depth caused by a combination of the two flood sources. Thus, a joint distribution was being fit to data that did not typically occur jointly.

The assessment of the assumption of independence between samples provides several interesting findings. It is clear that if both flood sources are not considered in the process of calculating the flood hazard a different understanding of flood hazard will be obtained. The exceedance probabilities fitted to the independent riverine and tidal flood sources differed significantly for flood depths between 0.5 and 5 feet. Further, if a joint flood hazard assessment is made based on the assumption that the two flood sources are independent of each other, the resulting non-exceedance probabilities calculated will vary somewhat from those calculated based on the assumption that the two flood sources are dependent. However, calculation of exceedance probabilities based on both flood sources, based on assumptions of both independence and dependence between the sources, was found to produce very similar exceedance probabilities. The most difference was evident between flood depths of 2 and 5 feet. Because a strong relationship was not found to exist between the observed riverine and tidal flood sources, it was not expected that a significant difference would be found between the exceedance probabilities calculated based on the assumptions of independence and dependence in this case.

This conclusion, that the assumption of independence or dependence between flood sources does not make a significant difference to the predicted exceedance probabilities, may not necessarily hold true in all coastal locations though. While it typically cannot be stated that the analyses made based on one assumption or the

other are “correct”, whether or not a relation exists between the flood sources can at least be demonstrated that based on an analysis of the correlation between the two flood sources. A high correlation between the flood sources would suggest a strong relationship between them such that the assumption of dependence would reasonable. In these cases, the methodology developed based on the copula should be used to determine the joint probabilities and the probabilities corresponding to total flood depths. A low correlation would suggest that a strong relationship did not exist between the flood sources, such that the assumption of independence would be reasonable. In these cases, Equation 1-3 could be used to determine the probabilities corresponding to total flood depths, which is a much simpler and more straightforward method than the method developed based on copulas.

5.4. Flood Risk Calculations

The previous sections have developed a process to assess the flood hazard for a location of interest based on multiple flood sources. The flood frequency assessments developed were based on joint distributions for all flood sources calculated based on copulas. The flood hazard is only one part of flood risk. The other two components of flood risk are the vulnerability to flooding and the consequences that result from those flood events. Flood risk is calculated by multiplying the exceedance probability, the vulnerability, and the consequences for a given flood depth of interest.

The term vulnerability refers to the ability of the system in place to protect the location of interest from flooding. The system will include any structures near the

location of interest designed to reduce the probability of flooding or any nonstructural measures taken at or near the location of interest to reduce either the probability or the consequences of flooding. Common examples of structural systems include levees or floodwalls. Nonstructural measures include elevation of structures, flood proofing structures, or creating additional room to store waters naturally. Whatever the system may be that is identified as protecting the location of interest, failure is always a possibility. This uncertainty is incorporated into the vulnerability term of the risk equation.

To complete the flood risk calculations, the consequences of the flood event must be assessed. Consequences may be identified in economic terms, such as damages to structures and infrastructure, environmental damages, or loss of life or injuries as a result of the flood event, though economic terms are the most commonly used. For the purpose of providing an example of the method, economic consequences will be considered because they can be easily calculated and understood.

5.4.1. Description of Methods

The exceedance probabilities calculated for the total flood depths using the copula procedure were used as input to the flood risk calculations. Flood risk calculations also rely on measurements of vulnerability and consequences. Vulnerability is a weight, that ranges between 0 and 1and indicates how well the in-place system to protect the location of interest from flood events is expected to perform. Therefore, vulnerability will vary with flood depth. A hypothetical vulnerability curve was created for use in experimental calculations, as shown in

Figure 4-51. Vulnerability increases with depth in a linear fashion, which indicates that vulnerability is simply directly proportional to depth of flooding, which may be true for certain systems. Without specific information about the system in place at the location of interest this linear curve seemed to be a reasonable approximation of vulnerability for the purpose of these calculations. Vulnerability will vary significantly, depending on the system in place and on the condition of that system.

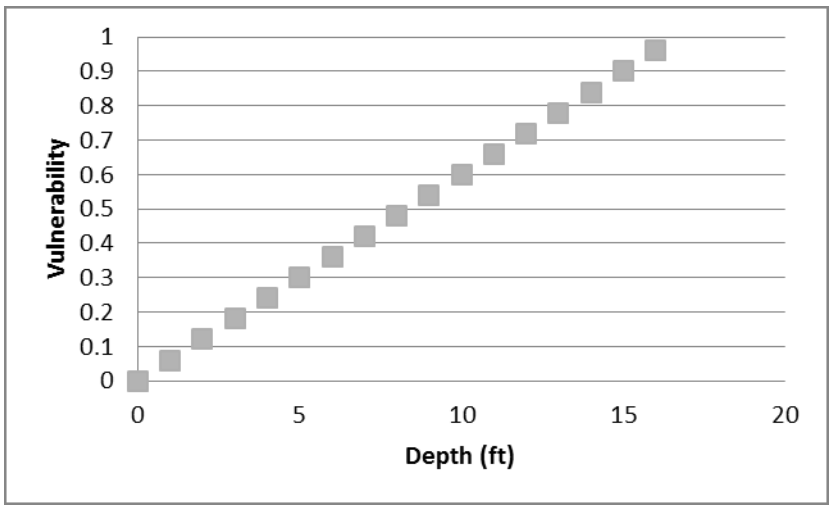

Figure 5-12: Illustration of the Vulnerability Curve Used in Flood Risk Assessments

Consequences, measured as damage to the structure and its contents due to the flood, will also vary with depth. Depth-damage curves, shown in Figure 5-12, obtained from USACE (2003), are used to determine the percentage of damage to the structure and to contents based on depth of flooding. These depth-percent damage curves are general curves that can be applied nationally, though it is also preferable to derive specific curves for a given location. For a hypothetical scenario using simulated data, the general curves were deemed to be appropriate. These percent damage values were then multiplied by the value of the structure and the value of the contents to determine the total monetary damage due to varying flood depths. For the

hypothetical scenario, the two-story residential structure without a basement was assumed to be worth $150,000 and the contents were assumed to be worth $25,000. The resulting depth-monetary damage curves are provided in Figure 5-13.

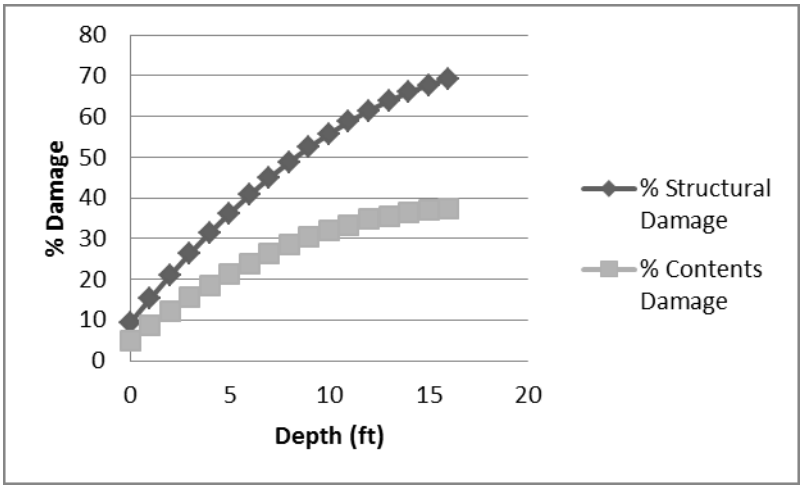

Figure 5-13: Depth-Percent Damage Curves Obtained from USACE (2003)

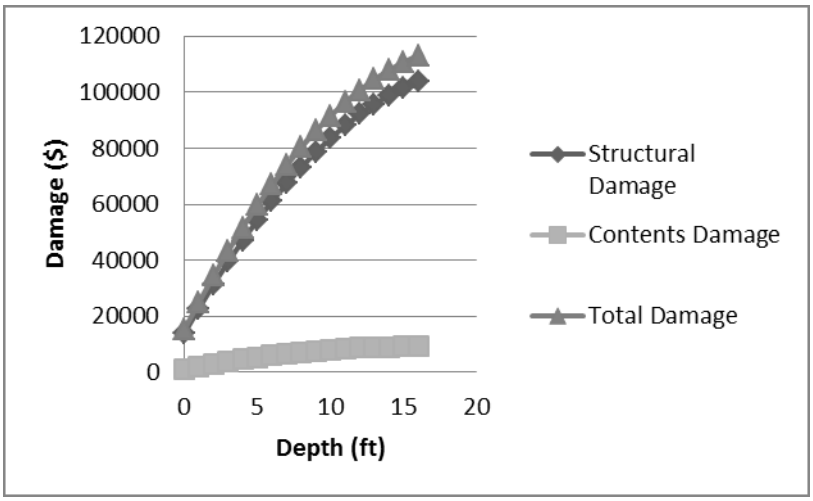

Figure 5-14: Depth-Monetary Damage Curves

Depending on the purpose of the risk analysis, it might be desirable in some circumstances to fit equations to the vulnerability and consequence curves, such that values could be determined for any flood depth of interest. However, the purpose of this analysis is simply to demonstrate the effect of the comprehensive probability

assessment on risk calculations. Thus, the population fitted to the exceedance probabilities corresponding to the total flood depth values was used to determine the exceedance probabilities corresponding to various depths of flooding for which values of vulnerability and consequences had been obtained. These exceedance probabilities were multiplied by the vulnerability values and consequence values for each flood depth to obtain flood risk values for each flood depth.

Several comparisons of flood risk were made using the observed data to understand the impact of multiple flood sources. First, the flood risk was calculated based on the exceedance probabilities corresponding to flood depths for each of the individual flood sources. For this, the probability distributions fitted to the annual maximum riverine and tidal flood depths, identified independently of each other, as is typically done in flood frequency analyses, were used. This analysis illustrated the impact of considering only one of the flood sources at a time on the understanding of flood risk at a specific location.

Flood risk calculations were also made using the exceedance probabilities calculated based on the copula, which considered both flood sources and the dependence between them. This was believed to provide a more complete understanding of the flood hazard, and thus should also provide a more complete understanding of the flood risk. These flood risk calculations were compared to those made using both of the individual flood source exceedance probabilities. The purpose of this analysis was to illustrate how the inclusion of both flood sources impacted flood risk, as well as how this method resulted in a varied understanding of flood risk for the location of interest.

Finally, flood risk was calculated using the exceedance probabilities calculated based on Equation 1-3. The results of this equation provided an understanding of the joint probability of flooding; however, the Equation 1-3 assumes that the two flood sources are independent. The purpose of this analysis was to assess the impact that this assumption of independence between flood sources would have on the understanding of flood risk for the location of interest.

5.4.2. Description of Results

Several analyses were made using the observed data to assess flood risk under various assumptions about the flood hazard at the location of interest. First, the flood risk calculated based on the probabilities of flooding of each source individually were assessed and compared. Next, the flood risk calculated based on the joint probabilities determined based on the copula, which accounted for dependence between the flood source, was compared to the risks calculated for each individual flood source. This illustrated the impact of considering both flood sources on the flood risk. Finally, the joint probabilities calculated using an assumption of independence between the flood sources were used to calculate flood risk. This assessed the impact of the assumption of independence between flood sources on the flood risk calculations.

Figure 5-15 provides the flood risk calculations for the observed flood depths, based on the assumptions of independence and dependence between the flood sources. Flood risk was calculated for the independent riverine flood depths, the independent tidal flood depths, and for total flood depths calculated based on assumptions of independence (using Equation 1-3 to calculate exceedance

probabilities) and dependence (using the copula procedure to calculate exceedance probabilities) between the flood sources. The exceedance probabilities are not presented as they were provided in Figure 5-11. The flood risk calculated for the joint distribution based on the assumption of dependence differed somewhat from the flood risk calculations based on the simulated studies. The increase and decrease in risk as the flood depth increased was steeper than observed in the simulated studies and the lowest values of flood risk for the observed joint distribution were lower than found for the simulation studies. The exceedance probabilities calculated for the higher flood depths were much lower than for the simulation studies. This can be attributed to the fact that the simulation studies typically contained a few large flood depths that were not included in the observed riverine and tidal samples. These lower exceedance probabilities resulted in lower flood risk for high flood depths, as would be expected.

As the exceedance probabilities calculated for either flood source independently were observed to differ, the flood risk calculations for each flood source independently also differed. The maximum exceedance probabilities for the riverine flood source occurred at very low flood depths, below a depth of 0.5 feet. Thus, the flood risk calculated for the riverine source independently was also lower than for the other scenarios in Figure 5-15. The tidal flood source had a very low probability of exceeding 4 feet of flood depth; thus, the flood risk decreases to very low values above 4 feet of flood depth. These results indicate that only considering one of these flood sources independently in evaluating flood risk would result in very different understandings of flood risk.

Flood risk calculations based on the assumptions of independence and dependence between the flood sources were also evaluated. The flood risk calculations based on these assumptions were not observed to differ significantly. The flood risk calculations differ the most between flood depths of 1 foot and 5 feet, which was similar to the range over which the exceedance probabilities were observed to vary the most as well. Because the exceedance probabilities were not observed to vary significantly based on the assumption of independence or dependence, it was expected that the flood risk calculations also would not vary significantly. However, it should be observed that the flood risk calculations based on both joint distributions vary significantly from the flood risk calculations based on the independent riverine and tidal flood sources. This would also be expected based on the differences in exceedance probabilities calculated. This suggests that considering only one of the potential flood sources that could impact a location of interest would result in a much different understanding of the flood risk for that location as well as a much different understanding of the flood hazard. However, because the flood hazard is a part of the flood risk calculation, this result would be expected.

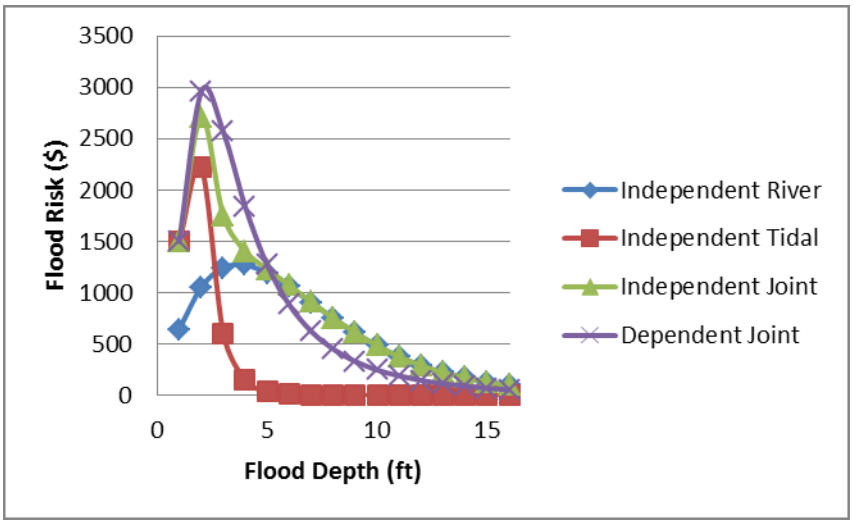

Figure 5-15: Flood Risk Calculated Based on Observed Riverine and Tidal Data with Assumptions of Independence and Dependence Between the Flood Sources

5.4.3. Conclusions

Flood risks were calculated for the observed data to illustrate the effect that a comprehensive flood frequency analysis would have on flood risk assessments. For two flood sources, the exceedance probabilities calculated based on the observed data were used to calculate flood risk. Calculations of flood risk based on the observed data were used to assess the impact of not considering both of the flood sources and the impact of assuming the flood sources were either independent or dependent. Considering either of the flood sources individually would result in quite different understandings of the flood risk for the location of interest, as would be expected. The exceedance probabilities, and therefore flood risk calculations, for the independent riverine flood source were highest at very low flood depths and then decreased gradually, while the exceedance probabilities and flood risk for the independent tidal flood source were highest at a slightly higher flood depth, but decreased very rapidly as flood depth increased. Considering the two flood sources jointly does provide a more complete understanding of the flood risk, regardless of

which method of calculating the exceedance probabilities is used. For this location, the assumption of independence or dependence between the flood sources was not observed to have significant impact on the flood risk calculations. However, it is believed that for other locations this assumption may in fact have significant impact. Thus, it will be important to assess the observed data and determine whether or not there are dependencies between the flood sources which must be considered. If so, the assumption of independence cannot be justified, though the method of calculating flood probabilities when independence is assumed is less complex.

CHAPTER 6

THREE SIMULATED FLOOD SOURCES

6.1. Introduction

Previous chapters have discussed the assessment of a flood hazard when two interacting flood sources must be considered. However, more than two potential flood sources could impact a given location. Thus, it is important to develop methodologies for assessing the flood hazard when more than two flood sources could impact a location. This chapter will develop a methodology for assessing the flood hazard when three flood sources must be considered, i.e., riverine, tidal, and pluvial, and will provide the results of such analyses.

As discussed previously, analyses based on data obtained from multiple gages will have a number of complicating factors. These include potential inaccuracies in one or more of the gage records, temporal, and spatial variations in hydrometeorological conditions, the use of different reference datum by each gage, and differences in the frequency of gage measurements, among others. The potential complicating factors will increase when the number of flood sources increases from two to three. Therefore, it is again beneficial to use a hypothetical scenario and simulated data as a preliminary step to verify the procedures developed and to better understand the expected results.

6.2. Assessment of the Flood Hazard

The first step of a risk assessment is to assess the hazard. In this case, the hazard was assessed by calculating the probability of certain flood inundation depths that occur at the location of interest. The following sections will describe the methods used to fit marginal distributions to each flood source, the use of the copula procedure to determine joint probabilities of given flood events, and the determination of the probabilities that correspond to the total flood depths, as well as presenting the results of each of these steps.

6.2.1. Generation of Correlated Random Variables

Generation of three correlated random variables with different distributions was somewhat more complicated than generation of two correlated random variables. The first step to this process was to generate three correlated random variables, all of which were normally distributed. Based on analyses of observed data for the location of interest in Florida, appropriate normal distribution parameters for all three flood sources were identified, along with levels of correlation between each flood source. The normal distributions representing the riverine, tidal, and pluvial flood depths, respectively, were calculated as follows, for i from 1 through sample size n:

$$R_i = \mu_r + z(0,1)\sigma_r \tag{6-1a}$$

$$T_i = \mu_t + \frac{\rho_{rt}\sigma_t}{S_r} * (R_i - \bar{R}) + z(0,1)\sigma_t(1-\rho_{rt}^2)^{0.5} \tag{6-1b}$$

$$P_i = \mu_p + 0.5\sigma_p\left[\rho_{rp}z_t + \rho_{tp}z_t\right] + z(0,1)\sigma_p\left[1 - 0.5\left(\rho_{rp} + \rho_{tp}\right)^2\right]^{0.5} \tag{6-1c}$$

where μ_j was the population mean for variable j, σ_j was the population standard deviation for variable j, $\bar{v}$ was the sample mean of random variable v (either R, T, or P), S_v was the sample standard deviation of random variable v (either, R, T, or P), V_i was a generated value of variable v (either R, T, or P), and ρ_{ij} was the population correlation between random variables i and j (where i and j could be any combination of R, T, and P). Within equations 6-1a through 6-1c, z_r was defined as $\frac{R_i - \bar{R}}{S_r}$ and z_t was defined as $\frac{T_i - \bar{T}}{S_t}$. The result of these equations was three sets of randomly generated values with the desired levels of correlation between each variable.

The desired random variables; however, were not all intended to be normally distributed. Thus, a transformation was needed from the normal distributions to the distributions of interest. For each random variable, the scale was set based on the minimum and maximum values of the random normal sample. The cumulative normal distribution that corresponded to each normal distribution was generated using the trapezoidal rule to integrate from the normal pdfs. The minimum and maximum values of the transformed variables were used to set the scale for each of the transformed random variables. The cumulative distribution functions that correspond to the transformed distributions were generated using the trapezoidal rule to integrate from the pdfs. The normal and transformed cdfs were then combined to create the transformation curve. Finally, the transformation curve was used to generate three correlated random variables with the desired distributions. This involved generating a normal random value from the population $N(\mu_j, \sigma_j)$, entering the normal cdf to obtain the cumulative normal probability that corresponded to the generated value, and then

using the cumulative probability identified from the normal distribution to generate a random variable from the desired distribution. For the riverine and pluvial flood sources, the desired distribution was the gamma distribution and for the tidal flood source the desired distribution was the generalized extreme value distribution. Thus, the result of the generation procedure was three correlated random samples represented by different probability distributions.

6.2.2. Overview of Simulation Scenarios

As with the scenarios in which two flood sources were investigated, two characteristics of the samples were investigated in this simulation study. The first characteristic was sample size and the second characteristic was the level of correlation between each of the flood depth samples. It was important to understand the effect of sample size because it is quite rare to have a long record of observed flood data at a particular location, especially for three flood sources. Four sample sizes, of 10, 25, 50, and 100 years, were evaluated. The minimum of 10 years was chosen because that is the smallest sample size recommended for use in Bulletin 17B, while the maximum of 100 years was chosen because very few longer observed records are in existence. The correlation between the samples was held roughly constant while samples of varying record length were generated, in order to understand the independent effect of sample size.

The relationship between the three flood depth samples has significant impact on the results of using copula equations. Thus, the effect of correlation between the generated samples on the results of the analyses was important to understand. This factor could not easily be assessed by using observed data because it would require

analyses in multiple locations, which would introduce unnecessary variables. Further, finding multiple locations of interest with a riverine, tidal, and rain gage close to each other is unlikely. For the simulation studies, two levels of correlation between the variables were assessed, moderately high correlation and relatively low correlation. Because many possible combinations of moderately high (correlation of approximately 0.3) and low (correlation of approximately 0.1) correlation between three sets of samples could arise, a total of six simulated data sets were developed to provide an understanding of the impact of the correlation between samples. Two of the generated sets of samples had similar levels of correlation between the three samples. One of these sets had relatively low correlation between all three samples (designated as RLA) and the other had moderately high correlation between all three samples (designated as RHA). The remaining four scenarios covered different combinations of moderately high and low correlation between the sample sets. Table 6-1 presents the levels of correlation between each set of samples for all six of the simulated scenarios. The names of the correlation scenarios presented in Table 6-1 indicate the level of correlation between each combination of variables. For instance, scenario RHLH indicates that the correlation between the Riverine and Tidal samples is high, the correlation between the Riverine and Pluvial samples is low, and the correlation between the Tidal and Pluvial samples is high.

Table 6-1: Explanation of Simulation Scenarios with Varying Correlation Between Samples, Where R Indicates Correlation, H indicates High Correlation Between Samples, L Indicates Low Correlation, and Position of R or L Indicates Combination of Variables for which the Designation Applies

	Correlation Between Samples		
Scenario	Riverine-Tidal	Riverine-Pluvial	Tidal-Pluvial
RHA	High	High	High
RLA	Low	Low	Low
RHLH	High	Low	High
RLHL	Low	High	Low
RLHH	Low	High	High
RHLL	High	Low	Low

6.2.3. Fitting Marginal Distributions to Annual Maximum Event Samples

The first step to the hazard assessment was to determine the marginal distributions. These were determined by fitting probability distributions to the generated riverine, tidal, and pluvial flood inundation depth samples. Several possible marginal distributions were considered to represent each flood source. The following section will describe the methods used to fit marginal distributions and then present the results.

6.2.3.1. Description of Methods

Though generated data represented riverine, tidal, and pluvial flood depths rather than riverine, tidal, and precipitation gage measurements, to determine the appropriate populations from which to generate samples it was necessary to determine the inundation depths for an observed location of interest based on measured gage data. The following sections describe the methods used to determine the appropriate populations from which to generate samples of flood depths from

each source and the methods used to fit the marginal distributions to the generated flood depth samples.

6.2.3.1.1. Calculation of Riverine-Caused Flood Inundation Depths

One variable of interest in this research was the flood inundation depth caused by a given riverine flood source; thus, one of the simulated random variables represented riverine flood inundation depths at the location of interest. Though the simulated data represented flood inundation depth at the location of interest, in order to determine the appropriate populations from which to simulate these flood inundation depths, a set of observed data from a location in Florida was assessed. The analysis of these data provided information as to the appropriate riverine flood depth population at the location in Florida. Discharge data from flow gage 02249007 on the Eau Gallie River at Heather Glen Circle at Melbourne was obtained. Information from this location, including elevations and channel characteristics, was also used to develop the simulation scenario. The population from which riverine flood depths were simulated was based on observed riverine flow gage discharge measurements translated to flood inundation depth at the location of interest. The process of transforming the discharges to inundation depths at the location of interest involves two steps: (1) transposing the discharge from the gaged site to a site on the river adjacent to the point of interest and (2) transforming the river discharge to an inland flood depth.

The procedure used to calculate flood inundation depths at the location of interest in Florida, based on riverine discharge measurements taken by a local flow gage, is explained in detail in Chapter 7, which discusses the analyses conducted for

three flood sources based on observed data. For the simulation study, based on information obtained from analysis of the location in Florida, generated samples represented flood inundation depths at the location of interest rather than gage measurements; thus, it was not necessary to transpose and transform simulated discharge measurements to flood inundation depth measurements. Once a sample of annual maximum flood inundation depths was generated, an appropriate distribution was fitted to the simulated data. Fitting a marginal distribution to the annual maximum events will be discussed in more detail in Section 6.2.3.1.4.

6.2.3.1.2. Calculation of Tidally-Caused Flood Inundation Depths

For the simulation studies, tidal flood inundation depths at the location of interest were simulated. In order to determine the appropriate distribution parameters to use as the basis for the simulation studies, a procedure was applied to calculate tidal flood depths based on tidal gage measurements for a location in Florida. The tidal gage from which measurements were obtained was 8721604, located near Trident Pier. This tidal gage is approximately 20 miles from the location of interest, which is not ideal, but a closer location was not available.

The procedure used to calculate flood inundation depths at the location of interest in Florida, based on tide height measurements taken by a local tidal gage, is explained in detail in Chapter 7, which discusses the analyses conducted using observed data from three flood sources. For the simulation study, based on information obtained from analysis of the location in Florida, the data generated represented flood inundation depths at the location of interest; thus, it was not necessary to transform tidal gage measurements to tidal flood inundation depths for

the simulation studies. Once a sample of annual maximum flood inundation depths was generated an appropriate population was fitted to the simulated data. Fitting a marginal distribution to the annual maximum events will be discussed in more detail in Section 6.2.3.1.4.

6.2.3.1.3. Calculation of Pluvially-Caused Flood Inundation Depths

For the simulation studies, pluvial flood inundation depths at the location of interest were simulated. In order to determine the appropriate distribution parameters to use as the basis for the simulation studies, a procedure was applied to precipitation gage measurements to obtain pluvial flood depths for a location in Florida. The precipitation gage from which measurements were obtained was 085612, located at the Melbourne Weather Forecasting Office. Precipitation depth measurements were obtained from the gage and translated into runoff depths, which were then translated into resulting discharge measurements. These discharge measurements were assumed to channelize in a street near the location of interest, and flood depths at the location of interest were calculated based on these discharge measurements.

The procedure used to calculate flood inundation depths at the location of interest in Florida, based on precipitation depth measurements taken by a local precipitation gage, is explained in detail in Chapter 7, which discusses the analyses conducted using observed data from three flood sources. For the simulation study, based on information obtained from analysis of the location in Florida, the data generated represented flood inundation depths at the location of interest; thus, it was not necessary to transform precipitation gage measurements to pluvial flood inundation depths for the simulation studies. Once a sample of annual maximum

flood inundation depths was generated an appropriate population was fitted to the simulated data. Fitting a marginal distribution to the annual maximum events will be discussed in more detail in Section 6.2.3.1.4.

6.2.3.1.4. Calculation of Flood Frequency Curves for Each Source Individually

To develop a flood frequency curve for a single flood source requires identification of the annual maximum flood series. The annual maximum events were ranked from largest to smallest, with the largest magnitude assigned to a rank of 1 and the smallest magnitude a rank of n. After developing a ranked list of annual maximum flood inundation depths at the location of interest caused by an individual flood source, the Weibull probability equation was used to determine the exceedance probability of each event. The Weibull plotting position formula is:

$$P_i = \frac{i}{n+1} \tag{6-2}$$

where P_i is the exceedance probability, i is the rank of the given flood depth, and n is the number of sample values in the record. The flood depths were plotted against the non-exceedance probability, calculated as 1.0 minus the exceedance probability, which provides the probability of the flood depth not being equaled or exceeded in a given year. Based on the pairs of data (non-exceedance probability, flood depth), a cumulative distribution function was fitted using either Maximum Likelihood Estimation or Method of Moments that would provide non-exceedance probabilities for any flood inundation depth value.

Though Bulletin 17B recommends the Log-Pearson Type III distribution be used to model riverine flood flows, the GEV distribution is frequently used to model

tidal flood heights, and the LP3 or gamma distribution are frequently used to model rainfall events, clear evidence that any one of these distributions provides greater accuracy than that provided by others does not exist. Further, given that the variable of interest in this research is flood depth at the location of interest, rather than discharge or tidal height; the typical distributions may not be the most appropriate distributions. Thus, several distribution functions were fitted to the data to determine the most appropriate distribution for each data set. All tested distributions were plotted against the simulated data for visual comparison and probability plot correlation coefficients were calculated to assess the goodness-of-fit of each distribution. For simulated riverine and pluvial data, the distributions tested included the log-Pearson type III, lognormal, gamma, and Weibull distributions. For simulated tidal data, the distributions tested included the extreme value, generalized extreme value, and Rayleigh distributions. This procedure was first conducted with the observed data from Florida, in order to determine appropriate populations from which to generate samples for the simulation study. It was also used in the simulation studies to calculate the distributions best fitting the generated samples.

6.2.3.1.5. Calculation of Flood Frequency Curves Considering Three Flood Sources

To develop a joint flood frequency curve requires calculation of the marginal distributions, or individual flood frequency curves that correspond to the three individual flood sources. The typical process for conducting a flood frequency analysis, which was outlined in section 6.2.3.1.4., was followed to determine the marginals, with one exception. When calculating the flood frequency curve for an individual flood source, the annual maximum flood event caused by the individual

source was identified. However, when three potentially interacting flood sources must be considered, it is possible that the maximum flood event experienced at the location of interest in a given year may not be caused by one individual source, but instead by a combination of the three sources. Thus, to identify the annual maximum flood depth at the location of interest, the simulated riverine, tidal, and pluvial flood inundation depths at the location of interest for corresponding time periods were summed. The assumption that the contributing flood depths from each source could be summed was a convenient assumption to make for the purpose of conducting this research; however, this methodology was not determined based on a technical analysis of the optimum approach for combining flood sources. This was completed first using the observed data from the location of interest in Florida, in order to identify the appropriate riverine, tidal, and pluvial flood frequency distributions. The same process was also used with the generated samples of riverine, tidal, and pluvial flood inundation depths at the location of interest, as part of the simulation studies.

The simulated data represented annual maximum flood depths; however, it was necessary to use the observed data from the location in Florida in order to determine the appropriate populations from which to generate the samples. Using the observed data, it was necessary to identify the annual maximum events while considering all three potential flood sources. As each of the gages from which observed data were obtained took hourly measurements each day, these measurements were summed and a single daily maximum flood depth at the location of interest was identified. Next, the annual maximum flood depth was identified based on the daily maximum events. This does not suggest that the annual maximum

event could only occur through some combination of the three sources; this procedure could still identify an annual maximum event that was caused by only one source individually or two sources jointly. The riverine, tidal, and pluvial contributions to the annual maximum flood inundation depth were separated into three flood series and flood frequency curves were developed for each. These flood frequency curves, based on the observed data from Florida, served to identify the appropriate populations from which samples of flood inundation depths should be generated for the simulation study. When the same process was used with generated samples, the flood frequency curves served as the marginals, which were inputs to the copula procedure.

6.2.3.2. Description of Results

The simulation study investigated both the effect of sample size and the effect of correlation between the samples. The following sections will present the results of fitting marginal distributions to generated riverine, tidal, and pluvial flood inundation depth samples of varying sample size and varying levels of correlation between the samples.

6.2.3.2.1. Effects of Varying Sample Size

Marginal distributions were fitted to the generated riverine, tidal, and pluvial samples of varying sample sizes in the first step. For each flood source sample for each sample size, the parameters of each potential distribution were calculated and the probability plot correlation coefficient was calculated to assess the fit of the distributions. To further assess the fit of the distributions each distribution was plotted against the sample to determine which distribution best agreed with the

generated sample. In Tables 6-2 through 6-4 and Figures 6-1 to 6-3 the results for a sample size of 10 were presented. For a sample size of 25 the results are presented in Tables 6-5 through 6-7 and Figures 6-4 through 6-6. Results for samples of size 50 are presented in Tables 6-8 through 6-10 and Figures 6-7 through 6-9 and results for samples of 100 are presented in Tables 6-11 through 6-13 and Figures 6-10 through 6-12. All results will be presented and then the results and implications will be discussed in detail.

Table 6-2: Fitted Parameters and Probability Plot Correlation Coefficients for the Riverine Samples of Size 10

	LP3	Lognormal	Gamma	Weibull
Shape	2.1117	-0.6463	0.5636	0.6905
Scale	-1.4707	2.1372	2.8114	1.2751
Location	2.4595	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9949	0.9684	0.9944	0.9962

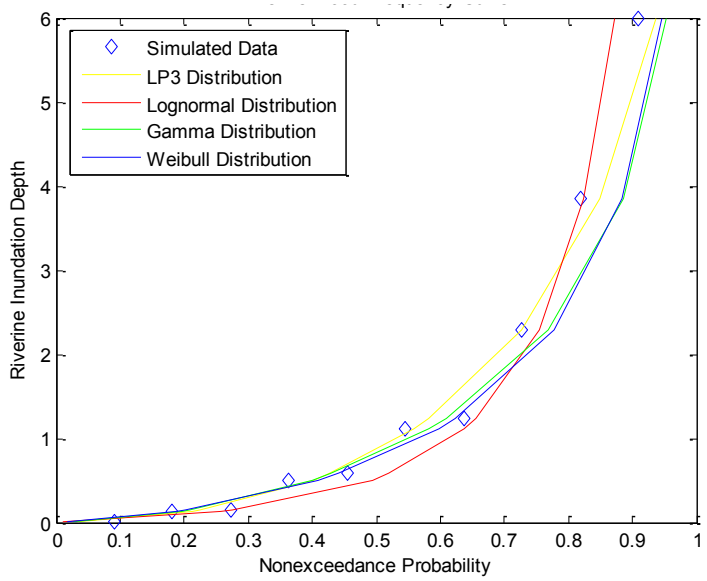

Figure 6-1: Comparison of Riverine Marginal Distribution Options for Sample Size of 10, where Flood Depths are in Feet

Table 6-3: Fitted Parameters and Probability Plot Correlation Coefficients for the Tidal Samples of Size 10

	Extreme Value	GEV	Rayleigh
Shape	N/A	-0.0296	N/A
Scale	1.0432	0.7032	1.4777
Location	2.3754	1.5131	N/A
Probability Plot Correlation Coefficient	0.8987	0.9496	0.9391

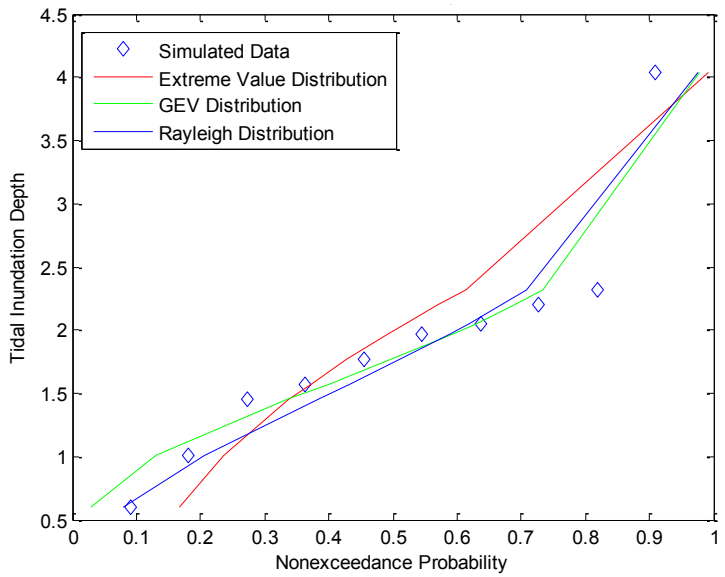

Figure 6-2: Comparison of Tidal Marginal Distribution Options for Sample Size of 10, where Flood Depths are in Feet

Table 6-4: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Samples of Size 10

	LP3	Lognormal	Gamma	Weibull
Shape	1.7405	-0.5578	1.3665	1.2873
Scale	-0.8869	1.1701	0.6303	0.9255
Location	0.9859	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9828	0.9465	0.98	0.9817

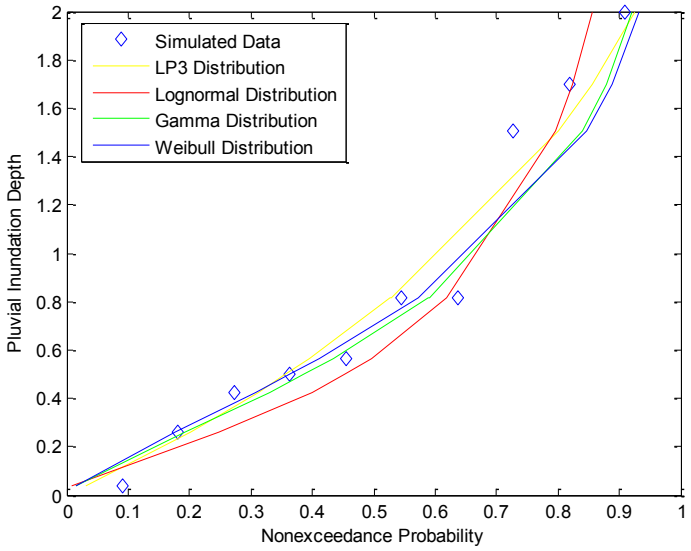

Figure 6-3: Comparison of Pluvial Marginal Distribution Options for Sample Size of 10, where Flood Depths are in Feet

Table 6-5: Fitted Parameters and Probability Plot Correlation Coefficients for the Riverine Sample of 25

	LP3	Lognormal	Gamma	Weibull
Shape	2.6544	-0.9135	0.4808	0.6171
Scale	-1.5464	2.5194	3.1526	1.165
Location	3.1911	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9454	0.8153	0.9634	0.9484

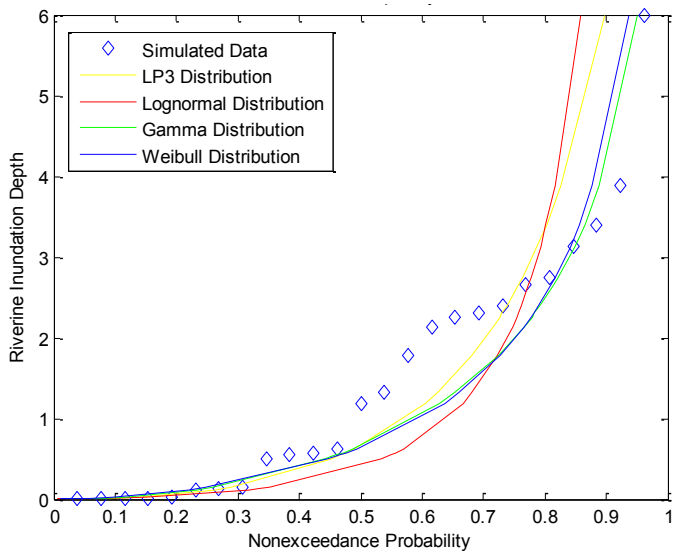

Figure 6-4: Comparison of Riverine Marginal Distribution Options for Sample Size of 25, where Flood Depths are in Feet

Table 6-6: Fitted Parameters and Probability Plot Correlation Coefficients for Tidal Sample of Size 25

	Extreme Value	GEV	Rayleigh
Shape	N/A	-0.1628	N/A
Scale	1.0166	0.8661	1.2885
Location	2.0469	1.1864	N/A
Probability Plot Correlation Coefficient	0.9495	0.9785	0.9705

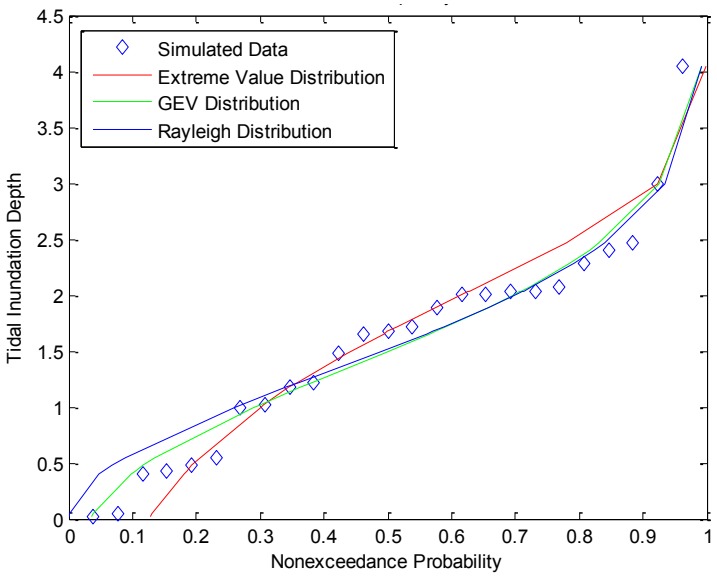

Figure 6-5: Comparison of Tidal Marginal Distribution Options for Sample Size of 25, where Flood Depths are in Feet

Table 6-7: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample of Size 25

	LP3	Lognormal	Gamma	Weibull
Shape	1.5137	-0.8546	1.1658	1.1602
Scale	-1.0397	1.2792	0.5937	0.7259
Location	0.7193	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9846	0.9278	0.9864	0.9888

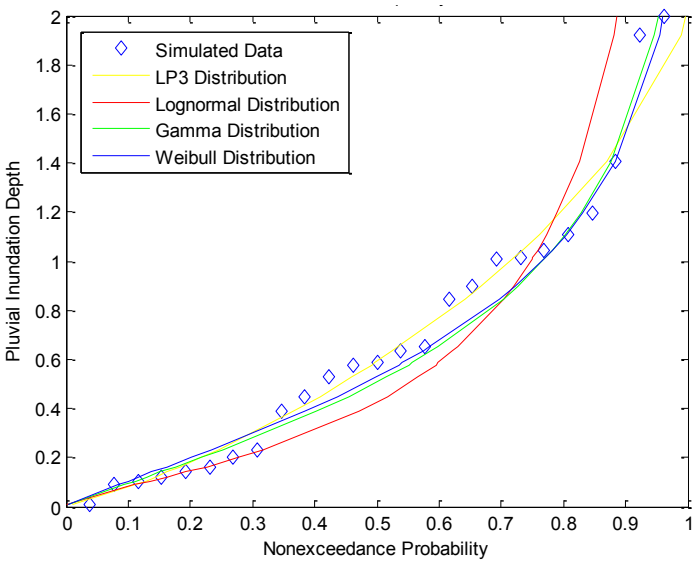

Figure 6-6: Comparison of Tidal Marginal Distribution Options for Sample Size of 25, where Flood Depths are in Feet

Table 6-8: Fitted Parameters and Probability Plot Correlation Coefficients for the Riverine Sample of Size 50

	LP3	Lognormal	Gamma	Weibull
Shape	2.5418	-0.6447	0.568	0.7022
Scale	-1.3388	2.1344	2.7664	1.3097
Location	2.7582	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9707	0.7867	0.9712	0.9591

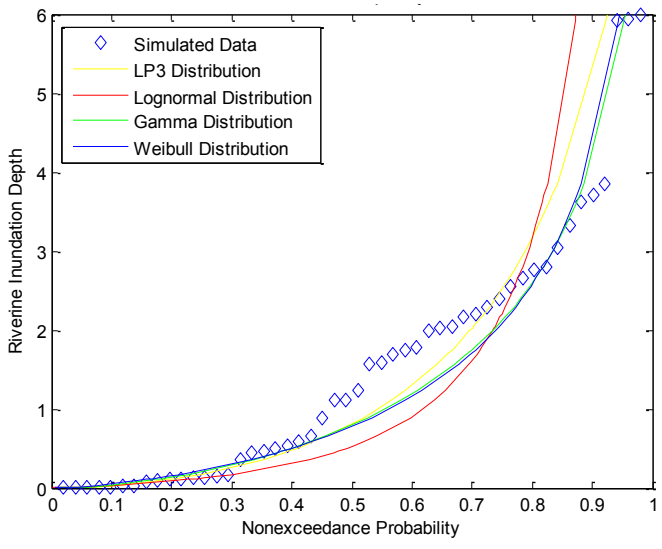

Figure 6-7: Comparison of Riverine Marginal Distribution Options for Sample Size of 50, where Flood Depths are in Feet

Table 6-9: Fitted Parameters and Probability Plot Correlation Coefficients for the Tidal Sample of Size 50

	Extreme Value	GEV	Rayleigh
Shape	N/A	-0.0764	N/A
Scale	2.0339	0.7803	1.2756
Location	0.9972	1.1529	N/A
Probability Plot Correlation Coefficient	0.9347	0.9958	0.9878

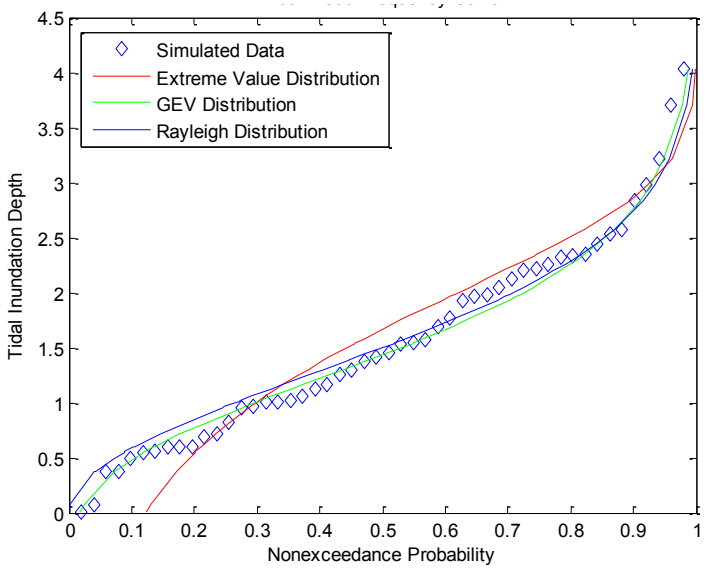

Figure 6-8: Comparison of Tidal Marginal Distribution Options for Sample Size of 50, where Flood Depths are in Feet

Table 6-10: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample of Size 50

	LP3	Lognormal	Gamma	Weibull
Shape	1.8678	-1.0876	0.7993	0.9063
Scale	-1.2291	1.6798	0.8861	0.6806
Location	1.2081	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9833	0.7699	0.9542	0.9526

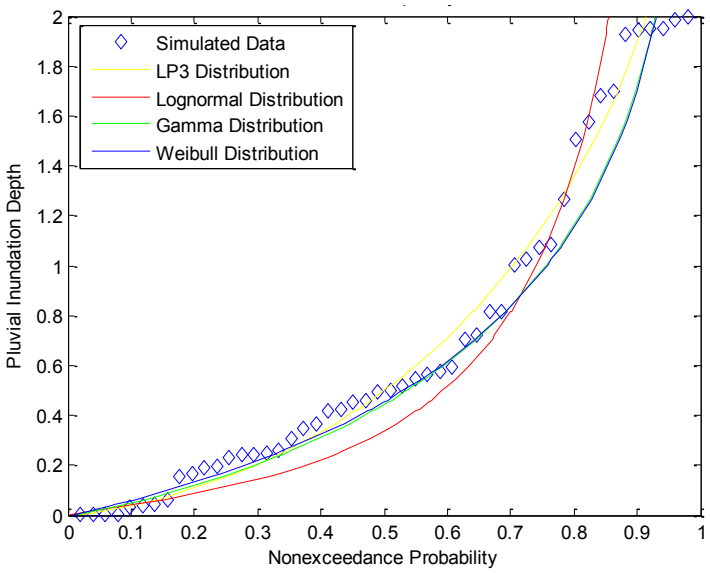

Figure 6-9: Comparison of Pluvial Marginal Distribution Options for Sample Size of 50, where Flood Depths are in Feet

Table 6-11: Fitted Parameters and Probability Plot Correlation Coefficients for the Riverine Sample of Size 100

	LP3	Lognormal	Gamma	Weibull
Shape	2.7435	-0.6238	0.5677	0.6974
Scale	-1.2501	2.0707	2.8287	1.3183
Location	2.8059	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9691	0.7011	0.9611	0.9405

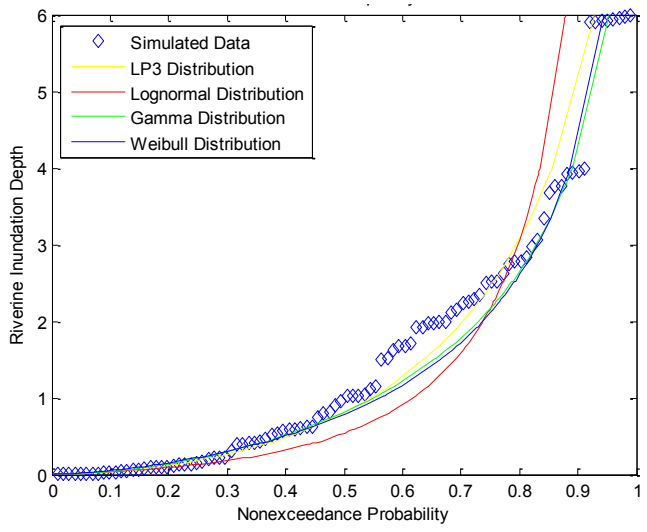

Figure 6-10: Comparison of Riverine Marginal Distribution Options for Sample Size of 100, where Flood Depths are in Feet

Table 6-12: Fitted Parameters and Probability Plot Correlation Coefficients for the Tidal Sample of Size 100

	Extreme Value	GEV	Rayleigh
Shape	N/A	0.0575	N/A
Scale	1.9657	0.743	1.2381
Location	0.992	0.986	N/A
Probability Plot Correlation Coefficient	0.91	0.9739	0.9751

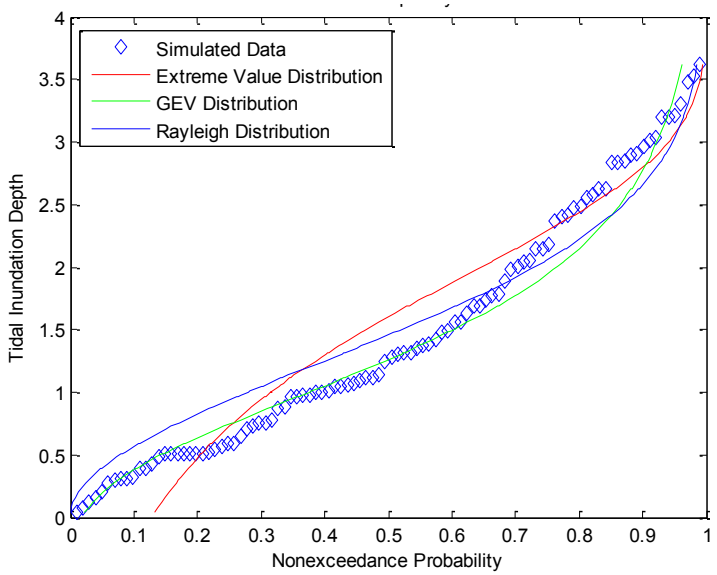

Figure 6-11: Comparison of Tidal Marginal Distribution Options for Sample Size of 100, where Flood Depths are in Feet

Table 6-13: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample of Size 100

	LP3	Lognormal	Gamma	Weibull
Shape	1.8097	-0.896	0.9949	1.0282
Scale	-1.0235	1.3768	0.7332	0.7372
Location	0.9562	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9861	0.7786	0.9491	0.952

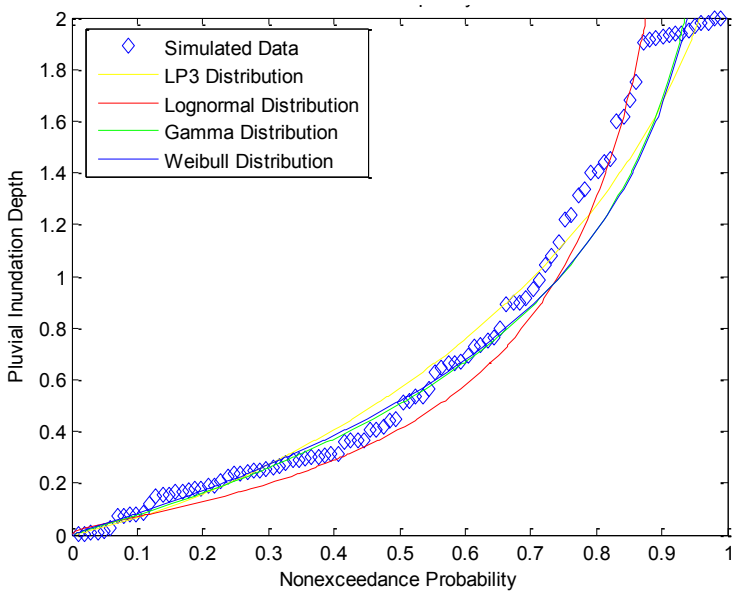

Figure 6-12: Comparison of Pluvial Marginal Distribution Options for Sample Size of 100, where Flood Depths are in Feet

For each sample size, distributions were first fitted to the riverine samples. For a sample size of 10, the probability plot correlation coefficients suggest that the LP3, gamma, and Weibull distributions would all adequately represent the sample, as the correlation coefficients were above 0.99 for all three distributions. However, from this information, it would not appear that the lognormal distribution would be able to represent the sample as well as the other three distributions. The plot comparing the four distributions to the generated sample confirmed that the LP3, gamma, and Weibull distributions all fit the sample well and slightly better than the lognormal distribution. Though the results will not be presented until Chapter 7, distributions had already been fitted to the observed data in order to determine the populations from which to generate data; thus, it was known that the gamma distribution was observed to provide the best fit to the observed data. In the interest of having the simulated samples consistent with the observed data, the gamma

distribution was selected for use in representing this sample, though it was not clearly superior to the LP3 or Weibull distributions.

For a sample of 25, the probability plot correlation coefficient suggested that the gamma distribution would be the best choice to represent the riverine sample, which was confirmed by the plot comparing the distributions to the generated sample. An examination of the probability plot correlation coefficients and the plot for a riverine sample of 50, the gamma distribution again appears to be the most appropriate distribution to represent the sample. Finally, for a riverine sample of 100, the probability plot correlation coefficients suggested that the LP3 and gamma distributions should both be able to adequately represent the sample and should perform better than either the lognormal or Weibull distributions. The plot comparing the distributions to the generated sample of 100 suggested that the gamma distribution would be most appropriate to represent the sample. Thus, the gamma distribution was selected to represent each generated riverine sample of varying size.

For each sample size distributions were next fitted to the generated tidal flood depth samples. For a sample size of 10 the probability plot correlation coefficients suggested that the GEV distribution would be slightly better than the Rayleigh distribution to represent the sample, and both distributions would better represent the sample than the extreme value distribution. The plot comparing the three distributions to the sample agreed with these findings and suggested that the GEV distribution was the most appropriate distribution to represent this sample. For a sample size of 25, the probability plot correlation coefficients suggest that the GEV and Rayleigh distributions should perform approximately equally in representing the

sample and somewhat better than the extreme value distribution. The plot comparing the distributions to the sample agrees that the GEV and Rayleigh distributions are superior to the extreme value distribution and that the GEV distribution is the best choice to represent the sample. Based on the probability plot correlation coefficients, for a sample size of 50 the GEV distribution would be the best choice to represent the sample, though the Rayleigh distribution would also adequately represent the sample. The plot agreed that either distribution could adequately represent the sample; thus, the GEV distribution was selected. For a sample of 100, the probability plot correlation coefficients suggest that the GEV and Rayleigh distributions should both adequately and approximately equally represent the sample and the extreme value distribution should not be able to represent the sample. The plot comparing the distributions to the generated sample actually suggested that the GEV distribution would be superior to the Rayleigh distribution. Thus, the GEV distribution was selected to represent all four of the tidal samples of varying size, which was consistent with the observed data, which was known to be best represented by the GEV distribution.

The final set of samples to be evaluated for each sample size was the pluvial flood depth sample. For a sample of 10, the probability plot correlation coefficients suggest that the LP3, gamma, and Weibull distributions should all be well able to represent the sample, though the lognormal distribution would not provide as accurate representation. The plot comparing the four distributions to the generated sample agreed that the LP3, gamma, and Weibull distributions could all satisfactorily represent the sample. Because the observed data was previously determined to be

best represented by the gamma distribution, which will be discussed in Chapter 7, the gamma distribution was selected to represent this sample. Similar results were observed for a sample of 25; thus, the gamma distribution was also selected for a sample of 25. For a sample of 50, the probability plot correlation coefficients suggested that the LP3 distribution would be the best distribution to represent the sample, the gamma and Weibull distribution would provide slightly less accurate, but approximately equal, representations of the sample, and the lognormal distribution would not be able to adequately represent the sample. The plot comparing the four distributions to the sample suggested that the gamma and Weibull distributions could adequately fit the sample. To be consistent with the observed sample, the gamma distribution was selected to represent the sample of 50 pluvial flood depth values. Similar results were observed for the probability plot correlation coefficients for a sample of 100 pluvial flood depth values. The plot comparing the four distributions to the sample again suggested that the gamma and Weibull distributions could adequately represent the sample despite the slightly lower probability plot correlation coefficient; thus, the gamma distribution was selected to represent this sample along with the three smaller pluvial flood depth samples.

The results for each sample size must be compared to determine whether or not trends can be attributed to sample size. In evaluating the parameters fitted to each sample size, trends were not apparent. Though trends were not identified in the marginal distribution parameters, the sample size did impact the marginal distributions. The distributions appear to more accurately represent the data as sample size increases, which was an expected result. Because larger sample sizes

better represent the underlying population, the distributions fitted to those samples are expected to be more accurate.

6.2.3.2.2. Effects of Varying Sample Correlation

Marginal distributions were also fitted to six sets of generated samples with varying levels of correlation between them. Generally speaking, the samples did exhibit the desired correlation structures. The same riverine sample was generated for all six scenarios and the tidal and pluvial samples were varied in order to reach the desired level of correlation. Thus, the results of fitting marginal distributions to the riverine samples will be presented only once, in Table 6-14 and Figure 6-13. The probability plot correlation coefficients suggest that the LP3 and gamma distributions should both adequately represent the riverine sample. The Weibull distribution should also be reasonably able to represent the sample, though the lognormal distribution could not be expected to provide an accurate fit to the sample. The plot comparing the distributions to the sample agree that the LP3, gamma, and Weibull distributions could all adequately represent the sample, though the gamma distribution appears to be slightly superior. Because the gamma distribution appears to be superior based on the probability plot correlation coefficient and the plot, and because this distribution was previously observed to provide the best fit to the observed data, it was selected to represent the riverine samples for the varying correlation scenarios.

Table 6-14: Fitted Parameters and Probability Plot Correlation Coefficients for the Riverine Sample Used for All Six Varying Correlation Scenarios

	LP3	Lognormal	Gamma	Weibull
Shape	2.5418	-0.6447	0.568	0.7022
Scale	-1.3388	2.1344	2.7664	1.3097
Location	2.7582	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9707	0.7867	0.9712	0.9591

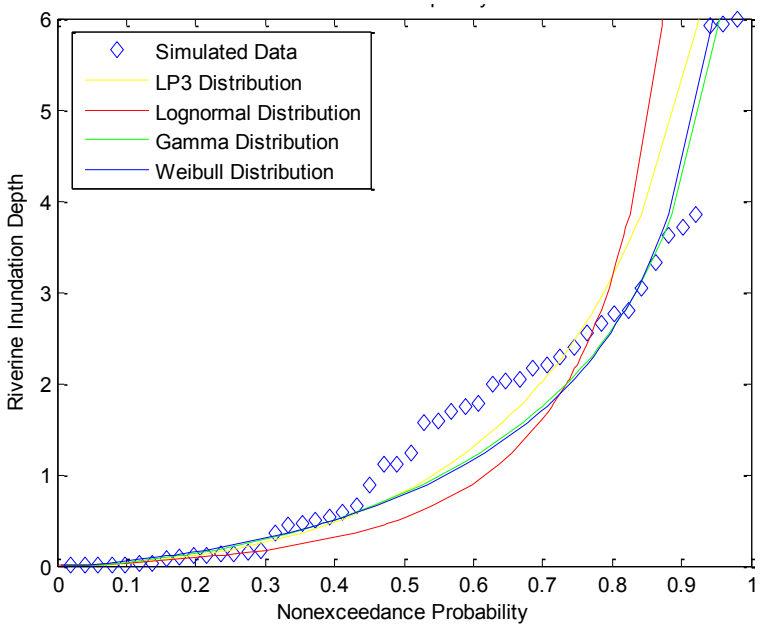

Figure 6-13: Comparison of Riverine Marginal Distribution Options for Scenario RHA

Marginal distributions were next fitted to the tidal samples. For all samples with a high level of correlation between the riverine and tidal samples, including scenarios RHA, RHLH, and RHLL, the same tidal sample was generated. A different tidal sample was generated for all three of the scenarios with a low level of correlation between the riverine and tidal samples, including scenarios RLA, RLHH, and RLHL. Thus, the results of fitting marginal distributions to two tidal samples will be discussed. The marginal distribution parameters and probability plot correlation coefficients for scenarios RHA, RHLH, and RHLL are presented in Table

6-15 and Figure 6-14 compares the three fitted distributions to the generated sample. The probability plot correlation coefficients suggest that the GEV distribution would be the best distribution to represent the sample, though the Rayleigh distribution would also be able to adequately represent the sample. The plot agrees that the GEV and Rayleigh distributions adequately represent the generated sample though the extreme value distribution does not. The GEV distribution was selected to represent the tidal sample for these three scenarios. The results for scenarios RLA, RLHH, and RLHL are presented in Table 6-16 and Figure 6-15. The probability plot correlation coefficients and the plot comparing the distributions to the generated sample again suggest that either the GEV or the Rayleigh distribution could adequately represent the sample, though the extreme value distribution could not. The GEV distribution was selected to represent this sample as well due to the probability plot correlation coefficient and because the GEV distribution was previously identified as the most appropriate distribution to represent the observed tidal sample.

Table 6-15: Fitted Parameters and Probability Plot Correlation Coefficients for the Tidal Sample Used for Scenarios RHA, RHLH, and RHLL

	Extreme Value	GEV	Rayleigh
Shape	N/A	-0.0286	N/A
Scale	2.0499	0.7301	1.2818
Location	0.9942	1.1704	N/A
Probability Plot Correlation Coefficient	0.9248	0.9958	0.9883

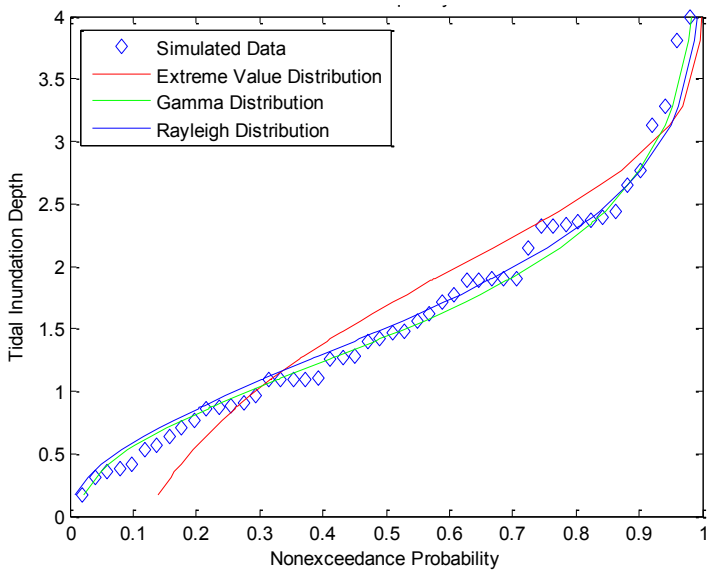

Figure 6-14: Comparison of Tidal Marginal Distribution Options for Scenarios RHA, RHLH, and RHLL

Table 6-16: Fitted Parameters and Probability Plot Correlation Coefficients for the Tidal Sample Used for Scenarios RLA, RLHH, and RLHL

	EV	GEV	Rayleigh
Shape	N/A	-0.0764	N/A
Scale	2.0339	0.7803	1.2756
Location	0.9972	1.1529	N/A
Probability Plot Correlation Coefficient	0.9347	0.9958	0.9878

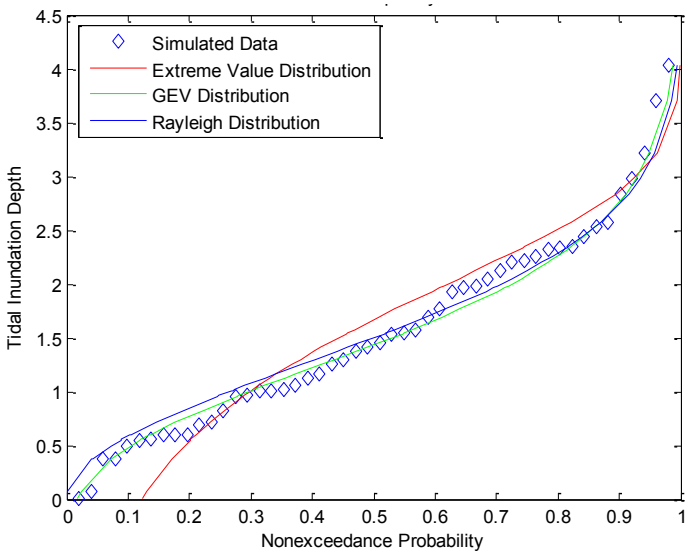

Figure 6-15: Comparison of Tidal Marginal Distribution Options for Scenarios RLA, RLHH, and RLHL

The pluvial samples generated were different for all six correlation scenarios. The pluvial sample needed to be varied for each scenario in order to obtain the desired levels of correlation between the riverine and pluvial and tidal and pluvial samples. Table 6-17 and Figure 6-16 provide the results for scenario RHA, Table 6-18 and Figure 6-17 provide the results for scenario RHLH, and Table 6-19 and Figure 6-18 provide the results for scenario RHLL. Table 6-20 and Figure 6-19 provide the results for scenario RLA, Table 6-21 and Figure 6-20 provide the results for scenario RLHH, and Table 6-22 and Figure 6-21 provide the results for scenario RLHL. The results and implications for each scenario will be discussed after all results have been presented.

Table 6-17: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RHA

	LP3	Lognormal	Gamma	Weibull
Shape	1.147	-1.1497	0.7507	0.8762
Scale	-1.7367	1.86	0.9365	0.6665
Location	0.8422	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9869	0.74	0.9546	0.9519

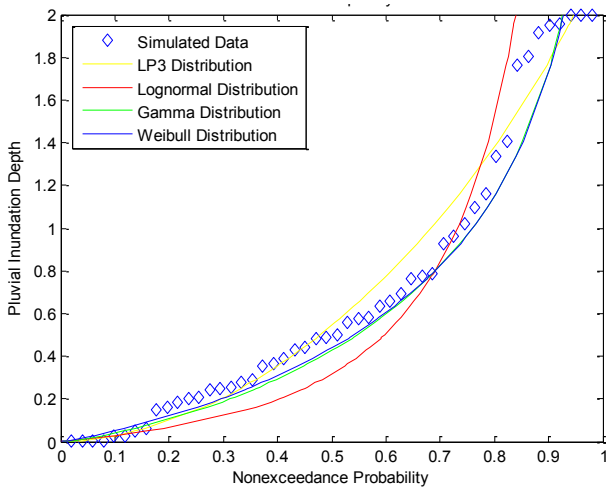

Figure 6-16: Comparison of Pluvial Marginal Distribution Options for Scenarios RHA, where Flood Depths are in Feet

Table 6-18: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RHLH

	LP3	Lognormal	Gamma	Weibull
Shape	1.0842	-1.1397	0.7428	0.8737
Scale	-1.833	1.9086	0.9651	0.6795
Location	0.8476	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9889	0.7181	0.9494	0.9464

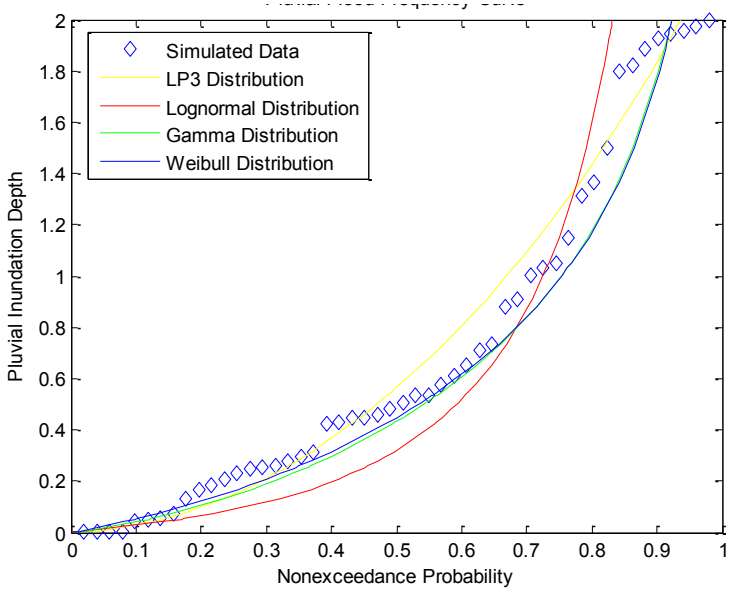

Figure 6-17: Comparison of Pluvial Marginal Distribution Options for Scenarios RHLH, where Flood Depths are in Feet

Table 6-19: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RHLL

	LP3	Lognormal	Gamma	Weibull
Shape	1.7002	-1.0569	0.8106	0.9168
Scale	-1.2856	1.6763	0.8904	0.6972
Location	1.1288	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9849	0.7617	0.9508	0.9498

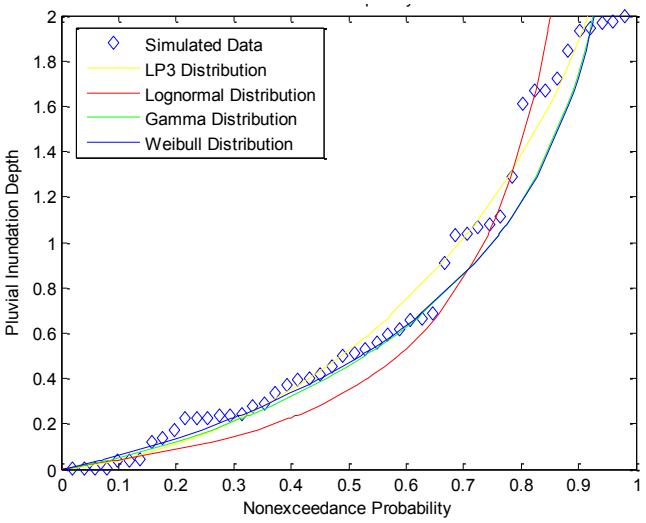

Figure 6-18: Comparison of Pluvial Marginal Distribution Options for Scenarios RHLL, where Flood Depths are in Feet

Table 6-20: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RLA

	LP3	Lognormal	Gamma	Weibull
Shape	1.1659	-1.0942	0.775	0.8951
Scale	-1.6732	1.8066	0.9322	0.6913
Location	0.8566	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9894	0.7337	0.9487	0.947

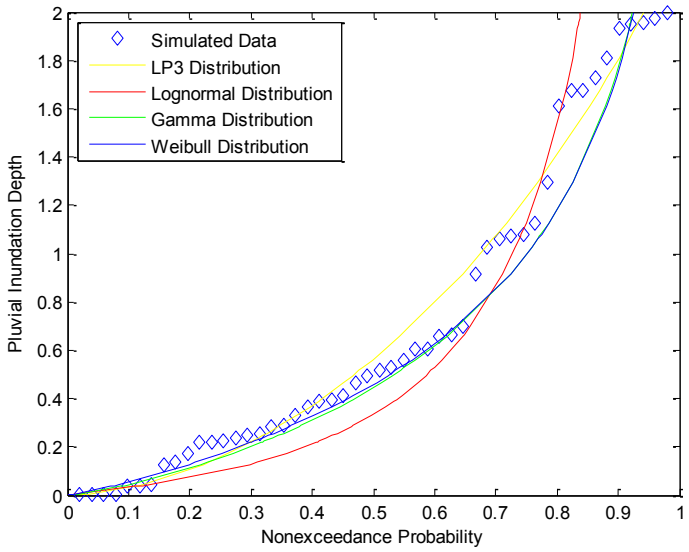

Figure 6-19: Comparison of Pluvial Marginal Distribution Options for Scenarios RLA, where Flood Depths are in Feet

Table 6-21: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RLHH

	LP3	Lognormal	Gamma	Weibull
Shape	1.1424	-1.1462	0.75	0.8764
Scale	-1.7473	1.8676	0.9413	0.6696
Location	0.8499	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9867	0.7325	0.9514	0.9488

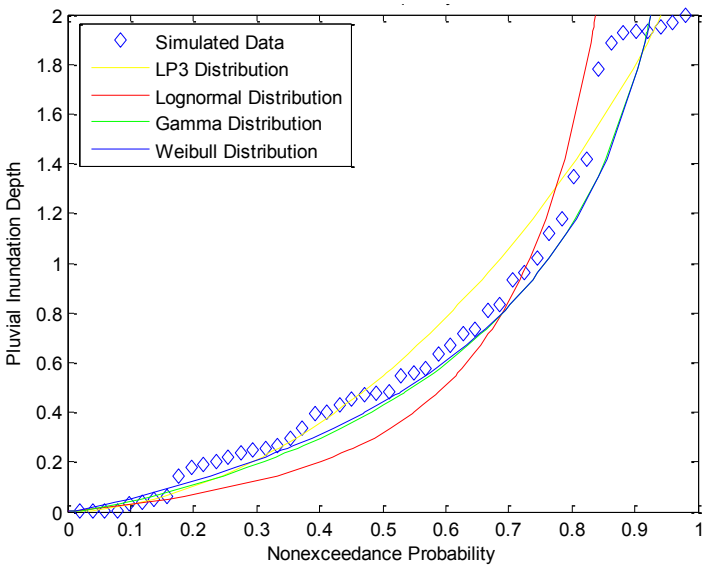

Figure 6-20: Comparison of Pluvial Marginal Distribution Options for Scenarios RLHH, where Flood Depths are in Feet

Table 6-22: Fitted Parameters and Probability Plot Correlation Coefficients for the Pluvial Sample for Scenario RLHL

	LP3	Lognormal	Gamma	Weibull
Shape	1.8678	-1.0876	0.7993	0.9063
Scale	-1.2291	1.6798	0.8861	0.6806
Location	1.2081	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.9833	0.7699	0.9542	0.9526

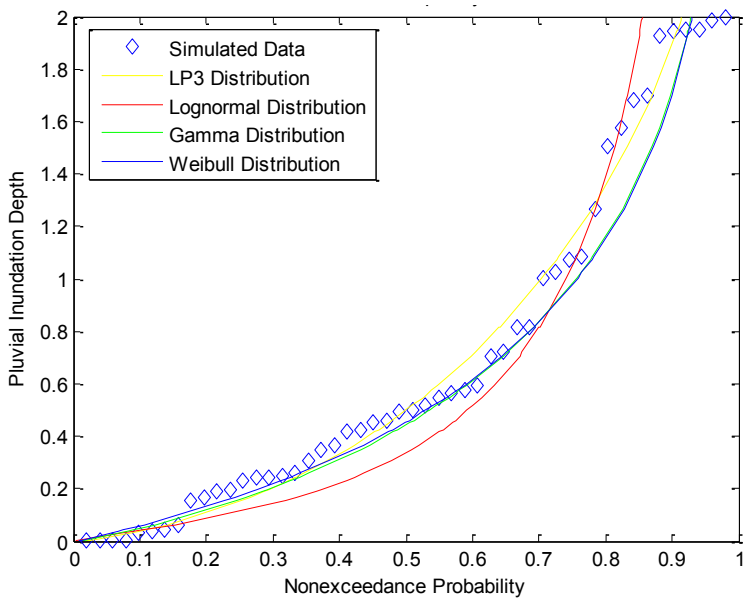

Figure 6-21: Comparison of Pluvial Marginal Distribution Options for Scenarios RLHL, where Flood Depths are in Feet

To determine which marginal distribution best represented the pluvial flood depth sample for each correlation scenario the probability plot correlation coefficients and the plot comparing the fitted distributions to the sample were examined. For each of the six scenarios, the probability plot correlation coefficients suggested that the LP3 distribution would be the best distribution to represent the sample, and slightly better than either the gamma or Weibull distributions. The probability plot correlation coefficients also suggested that the lognormal distribution would not be able to accurately represent the sample in any scenario. An examination of the plots comparing the fitted distributions to the samples agreed that the lognormal distribution was unable to represent the samples; however, the plots suggested that the gamma and Weibull distributions would be able to accurately represent the distributions despite their slightly lower probability plot correlation coefficients. Because the gamma distribution was previously identified as the most appropriate

distribution to represent the observed data, the gamma distribution was selected for use in representing all six pluvial flood depth samples.

6.2.4. Using Copula Equations to Develop Joint Distributions

Next in the process, joint distributions were developed for flood inundation depths from the riverine, tidal, and pluvial sources. The joint distributions provided probabilities of specific combinations of riverine, tidal, and pluvial flood depths. The next sections will discuss the methods used to develop the joint distributions and present the results.

6.2.4.1. Description of Methods

Once the marginals had been developed based on the generated riverine, tidal, and pluvial flood inundation depths at the location of interest, a copula could be used to determine the joint probability of depths for the three flood sources. The first step was to calculate the appropriate dependence structure between the three flood sources. As with the scenario in which two flood sources interacted, the dependence structure was based on the ranks of the data; however, for three flood sources, Kendall's τ was not appropriate. The method used involved maximum pseudo-likelihood estimation, which is an alternative to the method of moments (Genest and Favre, 2007). This method determines the value of the parameter, α, based exclusively on the ranks of the riverine, tidal, and pluvial sample data, R, S, and T respectively:

$$l(\alpha) = \frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} \frac{\dot{c_\theta}(\frac{R_i}{n+1}, \frac{S_i}{n+1}, \frac{T_i}{n+1})}{c_\theta(\frac{R_i}{n+1}, \frac{S_i}{n+1}, \frac{T_i}{n+1})} = 0 \qquad (6\text{-}3)$$

Where l(α) is the likelihood function of the copula parameter, c(α) is the copula pdf, $\dot{c}(\alpha)$ is the derivative of the copula pdf with respect to the copula parameter, n is the sample size, and R_i, S_i, and T_i are the ranks of the riverine, tidal, and pluvial sample values, respectively. This equation was solved based on the generated samples of riverine, tidal, and pluvial flood depths for each of the simulation scenarios in order to determine the copula parameter value for each of the three copula families. To solve Equation 6-3, the value of α was varied by a step size of 0.0001 over the range of expected values and the value that resulted in the equation being most closely equal to zero was selected as the copula parameter.

Once the measure of dependence was calculated, the appropriate family of copulas to fit with the data had to be determined. Numerous families of copulas are available from which to choose. Several families were fitted and compared using Akaike's Information Criteria (AIC). The specific method of using the AIC was described in Chapter 1 of this document. When AIC values are calculated, the minimum AIC value indicates the most appropriate copula family (Klein *et al.*, 2010). Three Archimedean copula families, the Gumbel-Hougaard, Clayton, and Frank families, were evaluated. The equations for these three families, respectively, are as follows:

$$C(u,v,w) = \exp(-[(-\ln u)^{\alpha} + (-\ln v)^{\alpha} + (-\ln w)^{\alpha}]^{\frac{1}{\alpha}}) \quad (6\text{-}4a)$$

$$C(u,v,w) = (u^{-\alpha} + v^{-\alpha} + w^{-\alpha} - 2)^{-1/\alpha} \quad (6\text{-}4b)$$

$$C(u,v,w) = \frac{-1}{\alpha}\left(\log\left(1 + \frac{(e^{-\alpha*u}-1)(e^{-\alpha*v}-1)(e^{-\alpha*w}-1)}{(e^{-\alpha}-1)^2}\right)\right) \quad (6\text{-}4c)$$

where u, v, and w were the cumulative distributions fitted to the riverine, tidal, and pluvial samples, respectively, and α was the copula parameter calculated for each family.

Finally, the joint pdfs were calculated for each copula family. Joint pdfs that correspond to the variables u, v, and w were calculated by taking the second derivative of the joint cdf with respect to u, v, and w. To obtain joint pdf values that correspond to the riverine, tidal, and pluvial variables required that the joint pdfs calculated for the variables u, v, and w then be multiplied by the marginal distribution pdf values (Wang *et al.*, 2009). The equations for the joint pdfs, expressed in terms of the riverine, tidal, and pluvial variables, are as follows:

$$f(x,y,z) = \frac{\partial^3}{\partial u \partial v \partial w} C_\alpha(u,v,w) \frac{\partial u}{\partial x}\frac{\partial v}{\partial y}\frac{\partial w}{\partial z} \quad (6\text{-}5)$$

where u, v, and w were the marginal cdf values corresponding to the riverine, tidal, and pluvial flood depth samples, respectively, and x, y, and z represent the riverine, tidal, and pluvial flood depth samples, respectively. Graphical displays of the joint pdfs and cdfs could not be produced for these scenarios as four-dimensional plots would have been required.

6.2.4.2. Description of Results

The effects of both sample size and correlation between marginal samples were investigated in the simulation studies. Copulas were used to develop joint distributions for each of the four sample sizes and each of the six sample correlations that were investigated in the simulation study.

6.2.4.2.1. Effects of Varying Sample Size

For the four sample sizes considered, the most appropriate copula family was first identified and then the identified copula families were used to develop joint distributions. The results of these steps are presented in the following sections.

6.2.4.2.1.1. Copula Fitting to Develop Joint Distributions

The first step to using a copula to develop a joint distribution was to calculate the copula parameters for each copula family and then determine which copula family was the most appropriate to represent each data set. The most appropriate copula family was chosen based on Akaike's Information Criteria (AIC) values, as discussed in Chapters 4 and 5. The copula parameters calculated for each sample size for each of the three potential families, as well as the AIC values calculated for each potential family are presented in Table 6-23. For each sample size, the calculated parameters met the constraint that the Gumbel-Hougaard parameter must be greater than 1.0 and the Clayton and Frank parameters must be greater than 0.0.

Generally speaking, as the sample size increased, the copula parameter values were observed to decrease, though for the Gumbel-Hougaard and Frank families, the parameters for a sample of 50 were actually slightly higher than for a sample of 25. The sample of 10 had the most significantly different parameter values, which suggests that as the sample size increases beyond this low sample size the parameters and also the joint distributions tend to stabilize. When two flood sources were examined; however, the copula parameters were observed to have a fairly wide standard error, which would suggest that the differences in copula parameters in Table 6-23 would not be significant. The joint distributions were calculated for the

sample size of 10 using the parameters calculated based on the sample of 100 values and these joint distributions were compared to the joint distributions originally calculated for a sample of 10. The joint cumulative distribution values varied from approximately 1% to 6% based on the different copula parameters. This suggests that the four sets of copula parameters calculated for the different sample sizes were not significantly different, such that the four samples of varying size could all be represented by the same joint distributions.

In the process of attempting to calculate copula parameters for samples of varying size, some impact from the level of correlation between the samples was also observed. The original set of samples generated for different sample sizes had strongly negative levels of correlation between the samples, particularly between the riverine and tidal and the tidal and pluvial samples. For these samples, it was not possible to determine a copula parameter for any of the three families that resulted in Equation 6-3 approaching a value of 0. When a second set of samples was generated with more positive levels of correlation between the samples, copula parameters could be calculated for all three families. The level of correlation between the samples would influence the ranks of the three sets of samples generated, and these ranks were used to calculate the copula parameter. This suggests that it would be more difficult to represent a set of samples with strongly negative levels of correlation between them using the three Archimedean copulas considered in this research. When only two flood sources were considered, the Gumbel-Hougaard and Clayton families were not able to represent samples with negative values of Kendall's τ, which was found to be related to the level of correlation between the samples, so it

does not seem unreasonable that the copula families might also have difficulty representing samples with negative levels of correlation between them when three flood sources were considered.

Table 6-23: Copula Family Parameters and AIC Values Calculated for Three Simulated Sources of Varying Sample Sizes

N	Gumbel-Hougaard Alpha	Gumbel-Hougaard AIC	Clayton Alpha	Clayton AIC	Frank Alpha	Frank AIC
10	1.2635	3.5674	0.4266	5.8091	1.2828	2.0633
25	1.0765	1.7691	0.1881	2.6517	0.6867	1.4221
50	1.0952	-1.0454	0.1288	2.1756	0.6882	0.015
100	1.0202	1.5455	0.0108	2.1134	0.2101	1.7703

The AIC values must be examined in order to determine the most appropriate copula family to represent each sample. For samples of 10 and 25, the Frank family produces the lowest AIC values. For these samples then, the Frank family was expected to be the most appropriate family to represent the joint distributions. For samples of 50 and 100 the Gumbel-Hougaard family produces the lowest AIC value. Thus, the Gumbel-Hougaard family was expected to be the most appropriate family to represent these samples. However, for all four scenarios, all three copula families will be used to calculate the joint distributions, to determine how significant the choice of copula family could be on the development of joint distributions.

6.2.4.2.1.2. Calculation of Joint Distributions

Once the copula parameters were calculated and the most appropriate copula family determined for each scenario, the joint pdf and cdf were calculated. Though the AIC values identified the most appropriate copula family for each sample sizes,

all three copula families were used to calculate the joint pdf and cdf for each sample size to illustrate the impact of the choice of copula.

Table 6-24 provides the joint pdfs and cdfs calculated based on the Gumbel-Hougaard, Clayton, and Frank families for a sample size of 10. The joint pdfs and cdfs calculated based on each family were compared and minimal differences in the joint distributions were observed. The differences in the joint cdfs were not observed to exceed 2-3%. This suggests that the choice of copula family had very little impact on the joint distributions developed for this sample of flood depths. In evaluating the joint cdf, it is observed that the maximum joint cdf value for the sample did not exceed 0.6. A maximum sample value closer to 1.0 would have been expected for a cdf, which suggests that the sample of 10 values may not have been representative of the population. Sampling variation could also have been responsible for the low maximum joint cdf value. This was confirmed by calculating the joint cdf value for a combination of the largest riverine, tidal, and pluvial flood depths, which could be expected if high correlation existed between the three flood sources. For this combination of flood depths the joint cdf would exceed 0.86, which is closer to expectation. The joint pdfs were necessary in order to calculate probabilities corresponding to total flood depths, which is the next step of the procedure, but they do not provide valuable information in and of themselves.

Table 6-24: Joint PDFs and CDFs Calculated for a Sample Size of 10

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.5744	0.5649	0.1002	0.1598	0.0932	0.1634	0.0883	0.1477
0.5794	4.0352	0.4975	0.0050	0.2097	0.0096	0.2103	0.0080	0.2000
2.2926	2.1996	0.4259	0.0307	0.2335	0.0336	0.2242	0.0293	0.2180
0.0038	1.7635	0.8139	0.8475	0.0150	0.3969	0.0205	0.8679	0.0129
5.9868	2.0517	1.9953	0.0013	0.5931	0.0016	0.5669	0.0016	0.5709
1.2359	2.3238	1.6990	0.0170	0.4642	0.0176	0.4296	0.0167	0.4351
0.1463	1.4513	0.8172	0.2464	0.0750	0.2212	0.0897	0.2240	0.0679
0.1306	1.0073	1.5051	0.0545	0.0390	0.0565	0.0542	0.0575	0.0344
0.505	1.9667	0.0360	0.0883	0.0078	0.0338	0.0121	0.0939	0.0065
3.8553	0.6013	0.2585	0.0038	0.0097	0.0030	0.0172	0.0046	0.0079

An analysis for a sample of 25 values was next made. The joint distributions calculated for a sample of 25 were very similar to the joint distributions calculated for a sample of 10, so a table of the results for a sample of 25 was not presented. For a sample of 25, the choice of copula family was again observed to have minimal impact on the joint distributions. The joint cdfs were observed to vary by no more than 2-3% based on the three different copula families. The maximum joint cdf value for this sample again did not exceed 0.6. If a combination of the largest riverine, tidal, and pluvial flood depth samples were to occur within this population, the maximum joint cdf value would exceed 0.89. This suggests that due to sampling variation a sample of 25 may still be too small to fully define the population of the joint distribution. Trends in the joint pdfs and cdfs were not observed based on the tables of values; however, it is possible that because the distributions cannot be visualized existing trends cannot be discerned.

Joint distributions were next calculated for a sample of 50 riverine, tidal, and pluvial flood depths. Table 6-25 provides the joint distributions for this scenario.

The impact of the choice of copula families on the joint distributions was again observed to be minimal. The differences in joint cdf values based on the copula families were not observed to exceed 3%. As the sample size has increased the maximum joint cdf value has also been observed to increase. For a sample of 50, the maximum joint cdf value has increased to above 0.75, which is closer to the results that would be expected from a cumulative distribution. The joint distributions developed for each sample size were visually compared to determine whether there were any trends in the joint cdfs or pdfs for similar combinations of riverine, tidal, and pluvial flood depths. Such trends were not observed; however, because the distributions could not be visualized, it is possible that there may be trends that are not evident in tables of values. Based on the results observed for two flood sources; however, the sample size was not expected to significantly impact the joint distributions.

Table 6-25: Joint PDFs and CDFs Calculated for a Sample Size of 50

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.5744	0.5649	0.0601	0.2189	0.0595	0.2093	0.0584	0.2187
0.5794	4.0352	0.4975	0.0041	0.2479	0.0055	0.2417	0.0050	0.2480
2.2926	2.1996	0.4259	0.0220	0.3247	0.0229	0.3099	0.0224	0.3201
0.0038	1.7635	0.8139	0.5076	0.0147	0.4188	0.0161	0.4724	0.0145
5.9868	2.0517	1.9953	0.0009	0.6746	0.0007	0.6598	0.0009	0.6642
1.2359	2.3238	1.6990	0.0075	0.4745	0.0074	0.4562	0.0081	0.4655
0.1463	1.4513	0.8172	0.1369	0.0892	0.1334	0.0897	0.1285	0.0901
0.1306	1.0073	1.5051	0.0543	0.0626	0.0561	0.0656	0.0532	0.0633
0.5050	1.9667	0.0360	0.2283	0.0301	0.2091	0.0332	0.2169	0.0303
3.8553	0.6013	0.2585	0.0133	0.0525	0.0137	0.0555	0.0130	0.0529
1.6877	0.3688	0.3059	0.0288	0.0257	0.0260	0.0288	0.0274	0.0258
0.0105	2.3460	0.3493	0.4078	0.0211	0.3713	0.0235	0.3897	0.0210
0.0161	0.9608	1.9868	0.0681	0.0199	0.0697	0.0232	0.0701	0.0198

0.4584	0.3734	1.5774	0.0119	0.0304	0.0117	0.0333	0.0117	0.0305
0.0028	1.9336	0.1655	1.7485	0.0058	1.4917	0.0078	1.7331	0.0056
2.5557	0.4966	0.7074	0.0105	0.0632	0.0104	0.0639	0.0096	0.0630
0.1181	0.5485	0.0031	1.2003	0.0005	1.4005	0.0014	1.2266	0.0004
2.0325	0.7162	0.1531	0.0552	0.0417	0.0528	0.0457	0.0533	0.0423
2.2027	1.2558	0.2441	0.0517	0.1289	0.0519	0.1271	0.0494	0.1300
0.0334	2.1223	0.1882	0.4662	0.0259	0.4374	0.0296	0.4517	0.0260
0.5331	2.5339	1.0237	0.0216	0.2961	0.0237	0.2843	0.0223	0.2923
2.1586	3.2144	1.9531	0.0015	0.6812	0.0011	0.6656	0.0015	0.6702
3.0516	1.6912	0.4613	0.0205	0.2923	0.0215	0.2793	0.0209	0.2893
3.3222	2.2125	0.4165	0.0127	0.3524	0.0136	0.3377	0.0133	0.3469
2.6521	1.5382	0.0045	0.0881	0.0086	0.0696	0.0097	0.0842	0.0084
1.7418	2.2556	0.0410	0.0544	0.0586	0.0542	0.0591	0.0497	0.0583
5.9349	1.5450	0.7219	0.0038	0.3683	0.0043	0.3549	0.0043	0.3640
2.7991	0.6912	0.3678	0.0216	0.0742	0.0217	0.0758	0.0205	0.0748
0.8908	2.3390	1.2676	0.0152	0.3861	0.0159	0.3696	0.0160	0.3795
1.7783	2.5741	1.9451	0.0035	0.5905	0.0031	0.5718	0.0037	0.5791
2.0397	0.5936	0.0050	0.1170	0.0026	0.1080	0.0042	0.1169	0.0024
0.6548	2.8419	0.2303	0.0392	0.1557	0.0437	0.1530	0.0396	0.1564
1.5845	1.3740	0.5491	0.0468	0.2095	0.0468	0.2008	0.0454	0.2093
2.3893	1.2985	0.1957	0.0503	0.1190	0.0508	0.1179	0.0480	0.1200
0.0317	0.8275	0.4913	0.4557	0.0150	0.4126	0.0190	0.4565	0.0151
1.9830	0.6023	0.2434	0.0404	0.0437	0.0383	0.0472	0.0386	0.0442
0.1724	1.4101	0.2477	0.3418	0.0507	0.3143	0.0545	0.3327	0.0520
2.7613	0.0744	1.0835	0.0017	0.0177	0.0015	0.0183	0.0016	0.0174
0.4493	1.9860	1.9517	0.0119	0.2736	0.0138	0.2641	0.0126	0.2708
0.0737	2.4420	0.5955	0.1052	0.0827	0.1088	0.0829	0.0985	0.0827
5.9161	0.9561	1.9255	0.0006	0.2534	0.0008	0.2482	0.0007	0.2506
1.1116	1.0163	1.0728	0.0285	0.1602	0.0294	0.1558	0.0272	0.1604
3.6203	3.7034	1.6821	0.0008	0.7805	0.0004	0.7667	0.0005	0.7698
0.3609	1.1603	0.0040	0.7512	0.0028	0.6252	0.0043	0.7696	0.0027
1.5643	0.0048	0.5155	0.0073	0.0084	0.0057	0.0098	0.0070	0.0083
3.7035	2.9762	0.4539	0.0040	0.4339	0.0045	0.4204	0.0046	0.4270
0.0009	0.5526	0.0609	5.7175	0.0004	7.3480	0.0011	5.7572	0.0003
0.1282	1.1263	1.0026	0.1163	0.0649	0.1134	0.0675	0.1108	0.0658
0.1062	1.0124	0.5755	0.2443	0.0408	0.2252	0.0448	0.2372	0.0416
0.0914	1.0541	0.0303	1.0546	0.0059	0.9987	0.0092	1.1049	0.0058

The final scenario evaluated for varying sample sizes involved a sample size of 100. The joint distributions were similar to those presented for a sample size of 50 so a table of the joint distributions for a sample size of 100 was not presented. For a sample size of 100 the copula family again had minimal impact on the joint distributions. Variations in the joint cdf values of only 2-3% were observed based on the different copula families. The maximum joint cdf value further increased as the sample size increased. For a sample size of 100 the maximum joint cdf value was above 0.8. This suggests that as the sample size increased the sample was a better representation of the joint distribution population. The joint pdf and cdf values were compared across sample sizes for similar combinations of riverine, tidal, and pluvial flood depths to determine whether or not the sample size influenced the joint distribution values. However, trends in the joint distributions related to sample size were not observed. Trends may have existed that were not apparent in tables of the joint distributions; however, trends were not expected given the results observed for two flood sources. An indirect impact on the joint distributions may be attributed to sample size, in that the cdfs calculated for each of the three flood source samples become more accurate as the sample size increases. This should suggest that the joint cdfs and pdfs calculated also become more accurate as sample size increases.

6.2.4.2.2. Effects of Varying Sample Correlation

For the six levels of sample correlation considered, the most appropriate copula family was first identified and then the identified copula families were used to develop joint distributions. The results of these steps are presented in the following sections.

6.2.4.2.2.1. Copula Fitting to Develop Joint Distributions

For the samples with varying levels of correlation, the copula parameters and the AIC values are presented in Table 6-27. For all six scenarios, the parameters met the requirements that the Gumbel-Hougaard parameter be greater than 1.0 and that the Clayton and Frank parameters be greater than 0.0. It was observed that a decreased level of correlation between the riverine and pluvial, the tidal and pluvial, or the riverine and tidal samples resulted in decreased copula parameters, while increased levels of correlation between the samples resulted in increased copula parameters. The parameters for scenario RHA, which had high correlations between all three flood sources, were the highest parameter values, while the parameters for scenario RLA, which had low correlation between all three flood sources, were the lowest parameter values. A similar trend in copula parameters was identified when only two flood sources were examined. The level of correlation between the samples was expected to influence the copula parameters because the parameters were calculated based on the ranks of the data in the three flood source samples, which would be influenced by the level of correlation between the samples.

Though trends were identified in the copula parameter variations due to the level of sample correlation, the differences between the copula parameters was quite small from one scenario to the other. Thus, it was not expected that the differences in copula parameters would result in significant differences in the joint distributions. The joint distributions were calculated using the copula parameters from scenario RLA and the samples from scenario RHA and the joint distributions were compared. The joint cdfs based on the two different parameters varied between 1% and 4%,

which was not a significant difference. When two flood sources were examined, it was determined that the level of correlation could result in significant differences between the samples; however, a broader range of levels of correlation was examined in that case. For higher levels of correlation between the samples, it can be expected that a more significant difference between the copula parameters would be apparent for three flood sources as well because the higher correlation between the flood sources represents a different relationship between them that must be modeled by the copula. The AIC values were examined to determine the most appropriate family to represent each scenario. The Gumbel-Hougaard family produced the lowest AIC value for each scenario, suggesting that it was the most appropriate family to use in representing all six scenarios.

Table 6-26: Copula Family Parameters and AIC Values Calculated for Three Simulated Sources with Varying Levels of Correlation between Samples

	Gumbel-Hougaard Alpha	Gumbel-Hougaard AIC	Clayton Alpha	Clayton AIC	Frank Alpha	Frank AIC
RHA	1.1964	-6.9299	0.2559	2.0004	1.4392	-5.3409
RHLH	1.1707	-5.205	0.2276	2.3404	1.2307	-3.4997
RHLL	1.1356	-2.8447	0.1695	3.3645	0.9761	-1.2641
RLA	1.075	-0.089	0.0843	2.4599	0.5388	0.8291
RLHH	1.1317	-2.2034	0.1593	2.6769	0.955	-0.8578
RLHL	1.0952	-1.0454	0.1288	2.1756	0.6882	0.015

6.2.4.2.2.2. Calculation of Joint Distributions

Based on the parameter values previously calculated, joint pdfs and cdfs were calculated for the various correlation scenarios. Though the AIC values indicated that the Gumbel-Hougaard copula family was the most appropriate family to use in representing all of the correlation scenarios, joint pdfs and cdfs were developed using

all three copula families, to determine how significant a difference the choice of copula family could have on the results.

The first correlation scenario to be assessed was scenario RHA, which had relatively high levels of correlation between all three flood source samples. Table 6-28 provides the joint distributions calculated for correlation scenario RHA. The choice of copula family was observed to have minimal impact on the joint distributions, as was observed for the scenarios with varying sample size. The joint cdfs were observed to vary by no more than 3% based on the different copula families. The maximum joint cdf value was above 0.8, which may be due to the fact that the sample size was 50.

Table 6-27: Joint PDFs and CDFs Calculated for Correlation Scenario RHA

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.6230	0.5802	0.0640	0.2629	0.0608	0.2413	0.0623	0.2663
0.5794	3.9886	0.6907	0.0028	0.3128	0.0048	0.2971	0.0041	0.3133
2.2926	2.3254	0.4884	0.0190	0.3985	0.0194	0.3675	0.0200	0.3917
0.0038	1.2669	0.7623	0.6781	0.0126	0.4725	0.0160	0.6153	0.0128
5.9868	2.3713	1.9955	0.0014	0.7664	0.0006	0.7390	0.0011	0.7474
1.2359	2.3523	1.7646	0.0073	0.5053	0.0073	0.4728	0.0083	0.4928
0.1463	1.2803	0.7845	0.1545	0.0897	0.1458	0.0917	0.1394	0.0943
0.1306	1.0939	1.1618	0.0884	0.0763	0.0897	0.0804	0.0811	0.0800
0.5050	1.8974	0.0479	0.2342	0.0498	0.2106	0.0547	0.2081	0.0520
3.8553	0.8682	0.2500	0.0164	0.0982	0.0189	0.1005	0.0158	0.1023
1.6877	0.5270	0.2753	0.0375	0.0391	0.0328	0.0444	0.0335	0.0407
0.0105	1.8907	0.3639	0.5728	0.0238	0.4494	0.0277	0.5004	0.0243
0.0161	0.6359	1.8014	0.0612	0.0123	0.0684	0.0186	0.0636	0.0122
0.4584	0.3764	1.4060	0.0112	0.0266	0.0103	0.0311	0.0102	0.0272
0.0028	1.4030	0.1625	2.9282	0.0062	2.1104	0.0101	3.0046	0.0062
2.5557	0.7074	0.6550	0.0134	0.1017	0.0142	0.1003	0.0113	0.1035
0.1181	0.4131	0.0050	1.4194	0.0009	1.8693	0.0035	1.5369	0.0007
2.0325	0.8856	0.1487	0.0671	0.0673	0.0647	0.0726	0.0620	0.0711

2.2027	1.4158	0.2559	0.0523	0.1829	0.0532	0.1734	0.0492	0.1880
0.0334	1.7773	0.2011	0.6339	0.0316	0.5473	0.0380	0.5876	0.0331
0.5331	2.4443	1.0994	0.0199	0.3209	0.0233	0.2991	0.0206	0.3173
2.1586	3.2799	1.9566	0.0018	0.6971	0.0012	0.6697	0.0018	0.6800
3.0516	1.8840	0.5005	0.0189	0.3751	0.0194	0.3454	0.0198	0.3704
3.3222	2.3959	0.4851	0.0101	0.4364	0.0108	0.4064	0.0112	0.4271
2.6521	1.7154	0.0003	0.1139	0.0020	0.0392	0.0023	0.1078	0.0019
1.7418	2.3369	0.0618	0.0384	0.1006	0.0426	0.0987	0.0319	0.1014
5.9349	1.8980	0.7728	0.0035	0.5009	0.0038	0.4699	0.0043	0.4907
2.7991	0.9079	0.3532	0.0262	0.1207	0.0278	0.1194	0.0240	0.1250
0.8908	2.3224	1.3345	0.0146	0.4159	0.0154	0.3859	0.0157	0.4077
1.7783	2.6434	1.9958	0.0038	0.6196	0.0030	0.5871	0.0041	0.6026
2.0397	0.7684	0.0043	0.1671	0.0050	0.1318	0.0085	0.1645	0.0048
0.6548	2.7650	0.2918	0.0346	0.2070	0.0417	0.1970	0.0357	0.2111
1.5845	1.4810	0.5576	0.0505	0.2637	0.0487	0.2424	0.0492	0.2667
2.3893	1.4691	0.2097	0.0487	0.1728	0.0506	0.1646	0.0454	0.1773
0.0317	0.5649	0.4275	0.4431	0.0101	0.4019	0.0169	0.4615	0.0102
1.9830	0.3521	0.1870	0.0246	0.0184	0.0205	0.0236	0.0228	0.0188
0.1724	1.2574	0.2444	0.4160	0.0575	0.3605	0.0642	0.4096	0.0623
2.7613	0.3145	0.9637	0.0025	0.0328	0.0022	0.0338	0.0020	0.0327
0.4493	1.9022	1.9146	0.0122	0.2840	0.0158	0.2671	0.0132	0.2829
0.0737	2.1517	0.6326	0.1160	0.0916	0.1185	0.0911	0.0970	0.0935
5.9161	1.5659	1.9953	0.0008	0.5239	0.0010	0.5034	0.0011	0.5138
1.1116	1.0907	1.0225	0.0324	0.1916	0.0338	0.1808	0.0302	0.1953
3.6203	3.8012	1.9476	0.0009	0.8143	0.0003	0.7916	0.0006	0.7974
0.3609	1.0961	0.0025	1.1729	0.0033	0.7852	0.0061	1.2585	0.0032
1.5643	0.1709	0.4412	0.0092	0.0119	0.0060	0.0144	0.0081	0.0119
3.7035	3.1338	0.5743	0.0030	0.5262	0.0033	0.4996	0.0038	0.5143
0.0009	1.0921	0.0275	14.0654	0.0009	14.6188	0.0031	14.9743	0.0008
0.1282	0.9630	0.9286	0.1265	0.0595	0.1209	0.0655	0.1168	0.0628
0.1062	1.1000	0.3893	0.3878	0.0479	0.3353	0.0550	0.3776	0.0516
0.0914	0.8650	0.0282	1.4882	0.0071	1.4329	0.0137	1.7056	0.0071

Scenarios RHLH and RHLL resulted in very similar joint distributions so the results of RHLL only are presented. Table 6-28 presents the joint distributions for this scenario. As was observed for scenario RHA, the choice of copula family had minimal impact on the joint distributions. The joint cdf values based on the different copula families did not vary by more than 3%. The joint distributions were also

compared to those developed for scenarios with higher levels of correlation between the samples, including scenarios RHA and RHLH. As the level of correlation between the flood depth samples decreased the joint cdf values were observed to decrease as well. However, the differences in joint cdfs based on the different levels or correlation were still fairly small, on the order of approximately 5%. However, the difference in levels of correlation between these scenarios was only moderate. For a more significant change in levels of correlation between the samples a more significant impact on the joint distributions would be expected, based on the results observed when two flood sources were previously evaluated.

Table 6-28: Joint PDFs and CDFs Calculated for Correlation Scenario RHLL

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.6230	0.5598	0.0640	0.2355	0.0622	0.2197	0.0619	0.2360
0.5794	3.9886	0.5090	0.0042	0.2552	0.0061	0.2451	0.0054	0.2558
2.2926	2.3254	0.3932	0.0196	0.3269	0.0210	0.3071	0.0202	0.3218
0.0038	1.2669	1.0336	0.4943	0.0121	0.4009	0.0143	0.4662	0.0120
5.9868	2.3713	1.9951	0.0011	0.7577	0.0005	0.7370	0.0009	0.7426
1.2359	2.3523	1.7230	0.0075	0.4895	0.0073	0.4642	0.0082	0.4782
0.1463	1.2803	0.9069	0.1347	0.0842	0.1305	0.0847	0.1248	0.0861
0.1306	1.0939	1.6695	0.0459	0.0742	0.0501	0.0764	0.0450	0.0757
0.5050	1.8974	0.0401	0.2280	0.0336	0.2053	0.0369	0.2116	0.0341
3.8553	0.8682	0.2226	0.0189	0.0793	0.0206	0.0813	0.0185	0.0810
1.6877	0.5270	0.2901	0.0383	0.0340	0.0347	0.0375	0.0358	0.0345
0.0105	1.8907	0.4519	0.5347	0.0227	0.4545	0.0252	0.4882	0.0228
0.0161	0.6359	1.9679	0.0557	0.0108	0.0610	0.0148	0.0591	0.0105
0.4584	0.3764	1.6698	0.0090	0.0247	0.0088	0.0276	0.0087	0.0248
0.0028	1.4030	0.2454	2.2813	0.0061	1.7855	0.0086	2.2969	0.0060
2.5557	0.7074	0.6658	0.0151	0.0938	0.0157	0.0923	0.0136	0.0942
0.1181	0.4131	0.0029	1.0921	0.0003	1.4639	0.0012	1.0894	0.0003
2.0325	0.8856	0.1391	0.0682	0.0529	0.0654	0.0566	0.0648	0.0543
2.2027	1.4158	0.2224	0.0544	0.1461	0.0551	0.1406	0.0512	0.1480
0.0334	1.7773	0.2392	0.5874	0.0287	0.5235	0.0327	0.5585	0.0292

0.5331	2.4443	1.0658	0.0222	0.3033	0.0248	0.2861	0.0229	0.2990
2.1586	3.2799	1.9324	0.0016	0.6867	0.0011	0.6655	0.0016	0.6726
3.0516	1.8840	0.4172	0.0198	0.3140	0.0210	0.2946	0.0202	0.3100
3.3222	2.3959	0.3708	0.0105	0.3487	0.0119	0.3301	0.0112	0.3425
2.6521	1.7154	0.0046	0.0656	0.0097	0.0491	0.0107	0.0604	0.0096
1.7418	2.3369	0.0366	0.0427	0.0552	0.0436	0.0553	0.0375	0.0550
5.9349	1.8980	0.6177	0.0039	0.4308	0.0044	0.4092	0.0046	0.4236
2.7991	0.9079	0.3334	0.0289	0.1032	0.0299	0.1024	0.0273	0.1051
0.8908	2.3224	1.2922	0.0154	0.3968	0.0160	0.3734	0.0164	0.3888
1.7783	2.6434	1.9733	0.0035	0.6071	0.0028	0.5820	0.0037	0.5926
2.0397	0.7684	0.0043	0.1385	0.0032	0.1163	0.0051	0.1386	0.0030
0.6548	2.7650	0.2344	0.0394	0.1606	0.0451	0.1553	0.0396	0.1622
1.5845	1.4810	0.5298	0.0512	0.2344	0.0504	0.2190	0.0496	0.2347
2.3893	1.4691	0.1752	0.0510	0.1327	0.0524	0.1285	0.0476	0.1342
0.0317	0.5649	0.5952	0.3140	0.0093	0.2882	0.0135	0.3172	0.0091
1.9830	0.3521	0.2267	0.0235	0.0169	0.0205	0.0203	0.0224	0.0169
0.1724	1.2574	0.2806	0.3692	0.0513	0.3280	0.0552	0.3610	0.0535
2.7613	0.3145	1.0359	0.0029	0.0314	0.0027	0.0319	0.0025	0.0310
0.4493	1.9022	1.9454	0.0128	0.2732	0.0155	0.2597	0.0138	0.2707
0.0737	2.1517	0.6846	0.1213	0.0863	0.1240	0.0852	0.1082	0.0867
5.9161	1.5659	1.8470	0.0010	0.5104	0.0012	0.4932	0.0013	0.5011
1.1116	1.0907	1.0806	0.0315	0.1805	0.0327	0.1716	0.0298	0.1813
3.6203	3.8012	1.6117	0.0009	0.7763	0.0004	0.7572	0.0006	0.7622
0.3609	1.0961	0.0040	0.8030	0.0028	0.6285	0.0046	0.8394	0.0026
1.5643	0.1709	0.5000	0.0095	0.0109	0.0072	0.0126	0.0088	0.0108
3.7035	3.1338	0.4006	0.0030	0.4075	0.0038	0.3915	0.0037	0.4004
0.0009	1.0921	0.1209	6.3821	0.0016	5.5135	0.0032	6.7585	0.0014
0.1282	0.9630	1.1153	0.1013	0.0555	0.0995	0.0589	0.0962	0.0568
0.1062	1.1000	0.6606	0.2361	0.0514	0.2158	0.0549	0.2239	0.0530
0.0914	0.8650	0.0437	1.0495	0.0064	0.9927	0.0107	1.1442	0.0063

The next scenario to be used to develop joint pdfs and cdfs was scenario RLA, which had low correlation between all three flood source samples. Table 6-29 presents the joint distributions calculated for this scenario. As with the other scenarios, the choice of copula family had minimal impact on the joint distributions. The joint cdfs typically varied by no more than 2% based on the copula family used in their calculation. The joint distributions were also compared to those calculated for

different levels of correlation in order to determine the impact of the correlation between the samples. The joint cdf values were observed to decrease by up to 6% as compared to the joint distributions for scenario RHA, which had high levels of correlation between all three flood sources. When two flood sources were evaluated, the level of correlation was observed to have a more significant impact on the joint distributions. However, for two flood sources a wider range of correlations were evaluated. It is expected that if the difference in the levels of correlation between the samples for these two scenarios was more significant the impact on the joint distributions would also be more significant. It is also possible that the full impact of the levels of correlation on the joint distributions was not evident based on the tables of values. Since the joint distributions for three flood sources could not be visualized, it is possible that trends in the joint distributions were not fully evident.

Table 6-29: Joint PDFs and CDFs Calculated for Correlation Scenario RLA

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.5744	0.5593	0.0583	0.2107	0.0575	0.2004	0.0567	0.2101
0.5794	4.0352	0.5156	0.0041	0.2476	0.0052	0.2404	0.0049	0.2471
2.2926	2.1996	0.3907	0.0221	0.3027	0.0228	0.2896	0.0224	0.2990
0.0038	1.7635	1.0607	0.3621	0.0154	0.3330	0.0159	0.3431	0.0152
5.9868	2.0517	1.9947	0.0008	0.6679	0.0006	0.6552	0.0008	0.6591
1.2359	2.3238	1.7260	0.0071	0.4680	0.0069	0.4518	0.0076	0.4603
0.1463	1.4513	0.9167	0.1172	0.0892	0.1162	0.0873	0.1114	0.0895
0.1306	1.0073	1.6772	0.0447	0.0612	0.0467	0.0618	0.0443	0.0616
0.5050	1.9667	0.0409	0.2276	0.0330	0.2164	0.0341	0.2186	0.0331
3.8553	0.6013	0.2181	0.0146	0.0457	0.0151	0.0468	0.0145	0.0459
1.6877	0.3688	0.2870	0.0296	0.0236	0.0276	0.0250	0.0286	0.0236
0.0105	2.3460	0.4662	0.3325	0.0236	0.3198	0.0244	0.3180	0.0235
0.0161	0.9608	1.9573	0.0766	0.0189	0.0788	0.0205	0.0781	0.0188
0.4584	0.3734	1.6741	0.0111	0.0294	0.0112	0.0306	0.0109	0.0293
0.0028	1.9336	0.2565	1.3748	0.0071	1.2209	0.0082	1.3518	0.0070

2.5557	0.4966	0.6585	0.0115	0.0591	0.0116	0.0585	0.0108	0.0589
0.1181	0.5485	0.0035	1.1620	0.0005	1.2585	0.0011	1.1781	0.0005
2.0325	0.7162	0.1372	0.0570	0.0374	0.0550	0.0389	0.0556	0.0377
2.2027	1.2558	0.2200	0.0531	0.1173	0.0532	0.1138	0.0511	0.1179
0.0334	2.1223	0.2464	0.4041	0.0294	0.3880	0.0308	0.3913	0.0294
0.5331	2.5339	1.0725	0.0201	0.2940	0.0216	0.2826	0.0207	0.2904
2.1586	3.2144	1.9330	0.0015	0.6730	0.0011	0.6595	0.0014	0.6637
3.0516	1.6912	0.4127	0.0212	0.2682	0.0219	0.2564	0.0214	0.2658
3.3222	2.2125	0.3670	0.0130	0.3216	0.0138	0.3091	0.0135	0.3173
2.6521	1.5382	0.0046	0.0989	0.0092	0.0882	0.0098	0.0956	0.0090
1.7418	2.2556	0.0365	0.0592	0.0545	0.0605	0.0540	0.0553	0.0542
5.9349	1.5450	0.6062	0.0043	0.3320	0.0047	0.3203	0.0047	0.3290
2.7991	0.6912	0.3286	0.0231	0.0671	0.0233	0.0668	0.0222	0.0674
0.8908	2.3390	1.2961	0.0144	0.3804	0.0147	0.3653	0.0150	0.3746
1.7783	2.5741	1.9741	0.0033	0.5843	0.0028	0.5682	0.0034	0.5747
2.0397	0.5936	0.0043	0.1288	0.0023	0.1223	0.0032	0.1287	0.0022
0.6548	2.8419	0.2365	0.0389	0.1555	0.0425	0.1508	0.0393	0.1556
1.5845	1.3740	0.5273	0.0466	0.1986	0.0463	0.1892	0.0453	0.1981
2.3893	1.2985	0.1730	0.0521	0.1070	0.0525	0.1043	0.0501	0.1075
0.0317	0.8275	0.6063	0.3601	0.0154	0.3329	0.0173	0.3579	0.0154
1.9830	0.6023	0.2227	0.0417	0.0397	0.0401	0.0409	0.0403	0.0400
0.1724	1.4101	0.2846	0.3055	0.0526	0.2852	0.0532	0.2974	0.0534
2.7613	0.0744	1.0252	0.0020	0.0168	0.0019	0.0171	0.0019	0.0166
0.4493	1.9860	1.9481	0.0125	0.2679	0.0140	0.2585	0.0131	0.2653
0.0737	2.4420	0.6970	0.0910	0.0859	0.0958	0.0841	0.0864	0.0855
5.9161	0.9561	1.8127	0.0008	0.2478	0.0010	0.2426	0.0009	0.2453
1.1116	1.0163	1.0790	0.0282	0.1550	0.0288	0.1491	0.0272	0.1548
3.6203	3.7034	1.6089	0.0007	0.7647	0.0004	0.7528	0.0005	0.7556
0.3609	1.1603	0.0005	1.2878	0.0005	1.0730	0.0008	1.3155	0.0005
1.5643	0.0048	0.4953	0.0077	0.0078	0.0066	0.0086	0.0074	0.0077
3.7035	2.9762	0.3971	0.0041	0.3966	0.0046	0.3856	0.0046	0.3914
0.0009	0.5526	0.1280	3.8780	0.0005	4.1232	0.0010	3.9302	0.0005
0.1282	1.1263	1.1264	0.0974	0.0642	0.0968	0.0642	0.0939	0.0647
0.1062	1.0124	0.6642	0.2054	0.0410	0.1934	0.0422	0.1997	0.0414
0.0914	1.0541	0.0453	0.9153	0.0076	0.8483	0.0095	0.9470	0.0075

Scenarios RLHH and RLHL were observed to have similar joint distributions so only the results of scenario RLHH were presented. Table 6-30 provides the joint distributions for scenario RLHH. The choice of copula family was not observed to

lead to significant differences in the joint distributions. The joint cdfs were observed to vary by only 2-3% based on the copula family. More significant impact based on the copula family was observed in some cases for two flood sources; however, the copula family was observed to have the most impact on the joint distributions when the level of correlation between the flood depth samples was quite high. However, a wider range of levels of correlation were assessed for two flood sources than for three, so it is possible that a more significant difference in joint distributions would be evident for a higher level of correlation. The joint distributions were also compared to those developed for other levels of correlation. As compared to scenario RLA, which had low correlation between all three samples, the joint cdf values were observed to increase. This reinforces the conclusion previously drawn that for decreasing levels of correlation between the samples the value of the joint distributions decreased and for increasing levels of correlation between the samples the values of the joint cdfs also increased.

Table 6-30: Joint PDFs and CDFs Calculated for Correlation Scenario RLHH

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
1.1092	1.5744	0.5784	0.0590	0.2351	0.0573	0.2189	0.0571	0.2358
0.5794	4.0352	0.7154	0.0027	0.3059	0.0040	0.2932	0.0037	0.3046
2.2926	2.1996	0.4795	0.0210	0.3641	0.0212	0.3405	0.0216	0.3579
0.0038	1.7635	0.8307	0.4385	0.0158	0.3578	0.0171	0.3938	0.0157
5.9868	2.0517	1.9954	0.0010	0.6787	0.0007	0.6587	0.0010	0.6656
1.2359	2.3238	1.7790	0.0070	0.4865	0.0069	0.4616	0.0078	0.4757
0.1463	1.4513	0.8119	0.1319	0.0952	0.1284	0.0935	0.1211	0.0970
0.1306	1.0073	1.1801	0.0822	0.0633	0.0824	0.0654	0.0780	0.0648
0.5050	1.9667	0.0495	0.2211	0.0460	0.2066	0.0484	0.2048	0.0469
3.8553	0.6013	0.2341	0.0130	0.0548	0.0139	0.0573	0.0126	0.0558
1.6877	0.3688	0.2654	0.0301	0.0266	0.0269	0.0297	0.0282	0.0270

0.0105	2.3460	0.4009	0.3318	0.0248	0.3046	0.0267	0.3054	0.0248
0.0161	0.9608	1.8842	0.0724	0.0213	0.0756	0.0246	0.0736	0.0213
0.4584	0.3734	1.4193	0.0138	0.0319	0.0135	0.0346	0.0131	0.0322
0.0028	1.9336	0.1876	1.5810	0.0072	1.2882	0.0094	1.5446	0.0071
2.5557	0.4966	0.6340	0.0106	0.0641	0.0108	0.0639	0.0094	0.0643
0.1181	0.5485	0.0036	1.5493	0.0009	1.7027	0.0023	1.6366	0.0008
2.0325	0.7162	0.1415	0.0567	0.0457	0.0542	0.0491	0.0540	0.0468
2.2027	1.2558	0.2468	0.0508	0.1428	0.0512	0.1370	0.0482	0.1451
0.0334	2.1223	0.2196	0.4149	0.0323	0.3889	0.0355	0.3911	0.0328
0.5331	2.5339	1.1223	0.0182	0.3135	0.0205	0.2962	0.0190	0.3089
2.1586	3.2144	1.9694	0.0017	0.6864	0.0011	0.6655	0.0017	0.6726
3.0516	1.6912	0.4851	0.0198	0.3161	0.0206	0.2956	0.0203	0.3127
3.3222	2.2125	0.4711	0.0121	0.3950	0.0126	0.3719	0.0129	0.3874
2.6521	1.5382	0.0004	0.1616	0.0019	0.0929	0.0023	0.1596	0.0018
1.7418	2.2556	0.0601	0.0500	0.0905	0.0533	0.0884	0.0444	0.0905
5.9349	1.5450	0.7351	0.0036	0.3826	0.0042	0.3631	0.0042	0.3776
2.7991	0.6912	0.3375	0.0215	0.0771	0.0219	0.0775	0.0200	0.0784
0.8908	2.3390	1.3499	0.0138	0.4017	0.0143	0.3786	0.0147	0.3938
1.7783	2.5741	1.9350	0.0039	0.5967	0.0032	0.5711	0.0042	0.5825
2.0397	0.5936	0.0049	0.1375	0.0036	0.1254	0.0057	0.1371	0.0034
0.6548	2.8419	0.2981	0.0335	0.1977	0.0387	0.1887	0.0346	0.1989
1.5845	1.3740	0.5497	0.0464	0.2228	0.0457	0.2081	0.0448	0.2236
2.3893	1.2985	0.2011	0.0492	0.1343	0.0502	0.1295	0.0466	0.1364
0.0317	0.8275	0.4538	0.4859	0.0165	0.4280	0.0207	0.4883	0.0168
1.9830	0.6023	0.1756	0.0465	0.0406	0.0439	0.0441	0.0441	0.0416
0.1724	1.4101	0.2543	0.3401	0.0584	0.3059	0.0609	0.3279	0.0608
2.7613	0.0744	0.9348	0.0018	0.0177	0.0016	0.0183	0.0016	0.0174
0.4493	1.9860	1.9244	0.0119	0.2798	0.0143	0.2657	0.0127	0.2770
0.0737	2.4420	0.6684	0.0843	0.0918	0.0912	0.0900	0.0764	0.0921
5.9161	0.9561	1.9305	0.0005	0.2559	0.0008	0.2488	0.0006	0.2528
1.1116	1.0163	1.0194	0.0295	0.1656	0.0305	0.1574	0.0278	0.1667
3.6203	3.7034	1.9496	0.0008	0.8062	0.0003	0.7892	0.0005	0.7930
0.3609	1.1603	0.0017	1.1344	0.0022	0.8578	0.0035	1.1811	0.0021
1.5643	0.0048	0.4279	0.0079	0.0085	0.0059	0.0100	0.0074	0.0084
3.7035	2.9762	0.5612	0.0038	0.5029	0.0039	0.4817	0.0044	0.4926
0.0009	0.5526	0.0366	8.3622	0.0004	11.0560	0.0013	8.4549	0.0003
0.1282	1.1263	0.9594	0.1185	0.0689	0.1151	0.0702	0.1109	0.0706
0.1062	1.0124	0.3958	0.3359	0.0382	0.2978	0.0422	0.3302	0.0397
0.0914	1.0541	0.0316	1.1966	0.0081	1.0882	0.0121	1.2786	0.0080

6.2.5. Calculation of Combined Flood Frequency Curve

The final step of the procedure required using a triple integral to calculate the area under the joint distributions corresponding to total flood depths. This would provide the non-exceedance probabilities that correspond to total flood depths that were influenced by all three flood sources. This would provide information about the flood hazard when three dependent sources were considered to impact the location. The following sections will explain the methods used to develop the flood frequency assessment for total flood depths.

6.2.5.1. Description of Methods

The results of the copula procedure described in section 6.2.4 were joint pdfs and cdfs, which can provide the probability that corresponded to a specific joint event. However, the primary interest is the probability of a certain flood depth occurring, when that flood depth might be due to either of the sources individually, or some combination of the three sources, such that many possible combinations could result in the desired flood depth. Because the riverine, tidal, and pluvial contributions to flood inundation at the location of interest were summed in the process of identifying annual maximum events, flood depths were also summed to identify probabilities that corresponded to total flood depths. Within the joint probability distribution, the region that corresponded to a specific total flood depth was identified as bounded by the riverine inundation depths, tidal inundation depths, and pluvial inundation depths axes and a triangular plane connecting the points (0, tidal depth of interest, 0), (riverine depth of interest, 0, 0), and (0, 0, pluvial depth of interest). In other words, the points where the flood inundation depth of interest was caused only

by one of the sources, riverine, tidal, or pluvial, were identified and connected by a plane. All of the points on this plane summed to a total depth equal to the depth of interest to be used in flood risk assessments.

The mass under the joint probability distribution within this bounding region represents the non-exceedance probability for a given flood depth. This was calculated using a triple integral. Within the triple integral, the lower and upper bounds of the outer integral were 0 and total depth of interest, respectively. The lower and upper bounds of the middle integral were 0 and:

$$Upper\ Bound = Total\ Depth\ of\ Interest - Riverine\ Contribution\ to\ Total\ Depth \quad (6\text{-}6)$$

The lower and upper bounds of the inner integral were 0 and:

$$Upper\ Bound = Total\ Depth\ of\ Interest - River\ Contribution\ to\ Total\ Depth - Tidal\ Contribution\ to\ Total\ Depth \quad (6\text{-}7)$$

As an example, if the riverine flood depth was 3.5 feet, the corresponding tidal flood depth was 3 feet, and the pluvial flood depth was 2 feet, the total flood depth of interest would be 8.5 feet. In this scenario, the outer integral lower and upper bounds would be 0 feet and 8.5 feet, respectively, the middle integral bounds would be 0 and (8.5 feet – 3.5 feet), or 5 feet, and the inner integral bounds would be 0 feet and (8.5 feet – 3.5 feet – 3 feet), or 2 feet. The triple integral was calculated in Matlab using the adaptive Simpson's method, a numerical evaluation method, which uses a recursive algorithm to approximate the integral based on the error in estimates calculated using Simpson's rule. The non-exceedance probability of each value of

total flood depth was calculated using this triple integral procedure, and then the probabilities were plotted against the total depth values.

The final task to determining the probability of total flood depths was to identify the appropriate distribution to represent the population. Four distributions were fitted to the total flood depths; the LP3, the GEV, the gamma, and the normal distributions. The GEV and gamma distributions were chosen for consideration because they were the two marginal distributions used to represent the simulated riverine, pluvial and tidal data. The LP3 distribution was selected for consideration because it is commonly used to represent flood data. The normal distribution was selected for consideration because the purpose of the copula is to develop a joint distribution when the marginals are represented by different distributions; thus, it was believed that the results of the copula might not follow either of the marginal distributions.

These four distributions were fitted to the total flood depth samples using either MLE or MOM and they were compared to the probabilities calculated for each copula family based on the triple integral procedure. The fitted distributions were plotted against the total flood depths and compared to a plot of the non-exceedance probabilities calculated by triple integral versus the total flood depths. These plots, as well as probability plot correlation coefficients, were used to determine which of the four distributions could best represent the total flood depth populations for each simulated scenario. The result of this process was a population from which non-exceedance probabilities corresponding to total flood depths for the location of interest could be identified. From this distribution, the true nature of the flood hazard

for the location of interest could be understood, which was the ultimate goal of this research.

6.2.5.1. Description of Results

For each of the sample sets of varying size and varying levels of correlation, the cumulative probabilities that correspond to total flood depths were calculated and then various distributions were fitted to represent those probabilities. The results of these efforts are presented in the following sections.

6.2.5.1.1. Effects of Varying Sample Size

Non-exceedance probabilities were calculated for the total flood depths by fitting four distributions to each total flood depth sample. The accuracy of these fitted distributions was assessed by comparing the resulting non-exceedance probabilities to those calculated by taking the triple integral under the joint distribution calculated using the previously identified most appropriate copula family. For a sample size of 10, Table 6-31 provides the parameters fitted to the sample for each distribution as well as the probability plot correlation coefficients, while Figure 6-22 compares each fitted distribution to the total depth sample. The fitted distributions were compared to the non-exceedance probabilities calculated by taking the triple integral under the joint pdf calculated using the Frank family. The probability plot correlation coefficients suggest that the LP3 and gamma distributions would both be adequate, and approximately equally, able to represent the sample. The normal distribution would be less able to accurately represent the sample, and the GEV distribution would not be expected to be able to represent the sample at all. The plot agrees that both the LP3 and gamma distributions represent the sample

reasonably well while the other two distributions do not. The gamma distribution was selected to represent the total flood depths for this sample.

Table 6-31: Parameters and Probability Plot Correlation Coefficients Calculated for LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for a Sample Size of 10

	LP3	GEV	Gamma	Normal
Shape	6.4642	1.4078	4.9412	N/A
Scale	0.1819	0.5827	0.879	2.3229
Location	0.1881	2.7587	N/A	4.3433
Probability Plot Correlation Coefficient	0.9962	0.8921	0.9932	0.9726

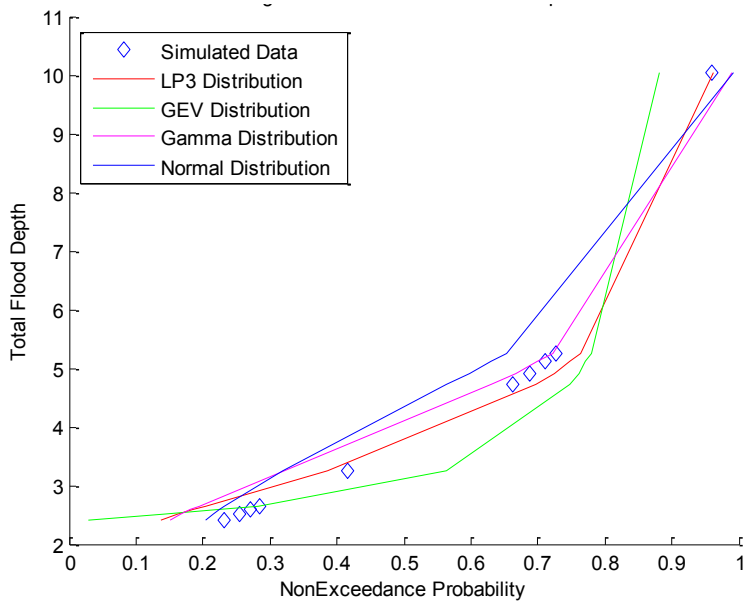

Figure 6-22: Comparison of the LP3, GEV, Gamma, and Normal Distributions to Non-Exceedance Probabilities Calculated for Total Flood Depths for a Sample Size of 10

For the larger samples of 25, 50, and 100 the results were very similar. Thus, the tables of parameters will be presented for all three sets of samples but the plots comparing the fitted distributions to the total flood depth samples will be presented only for samples of 25 and 100. The results and implications will be discussed after all of the results have been presented. Table 6-32 and Figure 6-23 provide the results of fitting distributions to the total flood depth sample of 25. Table 6-33 provides the

distribution parameters and probability plot correlation coefficients for a sample of 50. Table 6-34 and Figure 6-24 provide the results for a sample of 100 total flood depth values. For a sample size of 25, as with the sample of 10, the fitted distributions were compared to the non-exceedance probabilities calculated based on the Frank family. For the larger sample sizes the fitted distributions were compared to the non-exceedance probabilities calculated using the Gumbel-Hougaard copula family.

Table 6-32: Parameters and Probability Plot Correlation Coefficients Calculated for LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for a Sample Size of 25

	LP3	GEV	Gamma	Normal
Shape	8.6493	-0.0643	3.5331	N/A
Scale	-0.2012	1.6164	1.0677	1.9532
Location	2.9201	2.9281	N/A	3.7723
Probability Plot Correlation Coefficient	0.9825	0.9778	0.9854	0.9465

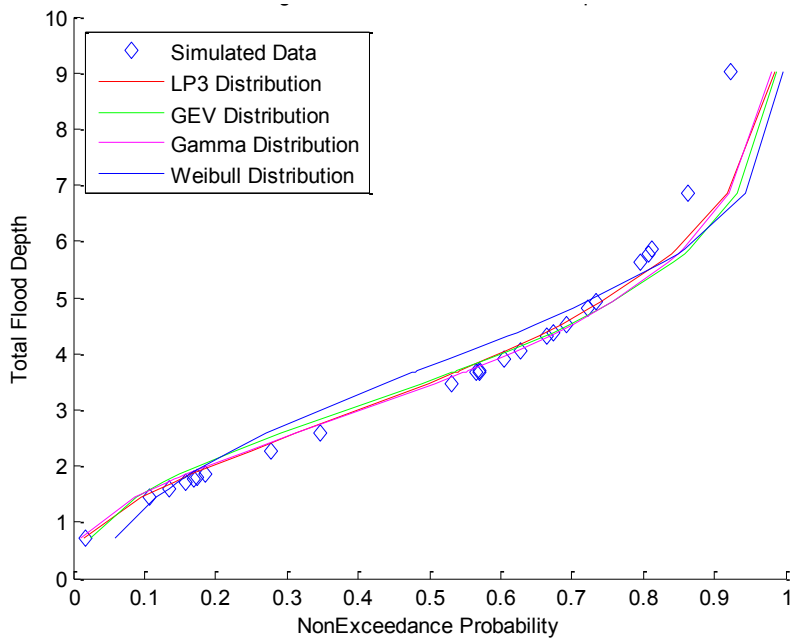

Figure 6-23: Comparison of LP3 Distribution to Non-Exceedance Probabilities Calculated for Total Flood Depths for a Sample Size of 50

Table 6-33: Parameters and Probability Plot Correlation Coefficients Calculated for LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for a Sample Size of 50

	LP3	GEV	Gamma	Normal
Shape	12.3858	0.065	3.3292	N/A
Scale	-0.1689	1.5443	1.1515	2.1525
Location	3.2783	2.8341	N/A	3.8337
Probability Plot Correlation Coefficient	0.9864	0.9909	0.9869	0.9324

Table 6-34: Parameters and Probability Plot Correlation Coefficients Calculated for LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for a Sample Size of 100

	LP3	GEV	Gamma	Normal
Shape	12.265	0.0317	3.2295	N/A
Scale	-0.1728	1.5907	1.1757	2.1165
Location	3.2911	2.825	N/A	3.797
Probability Plot Correlation Coefficient	0.9864	0.9875	0.9868	0.931

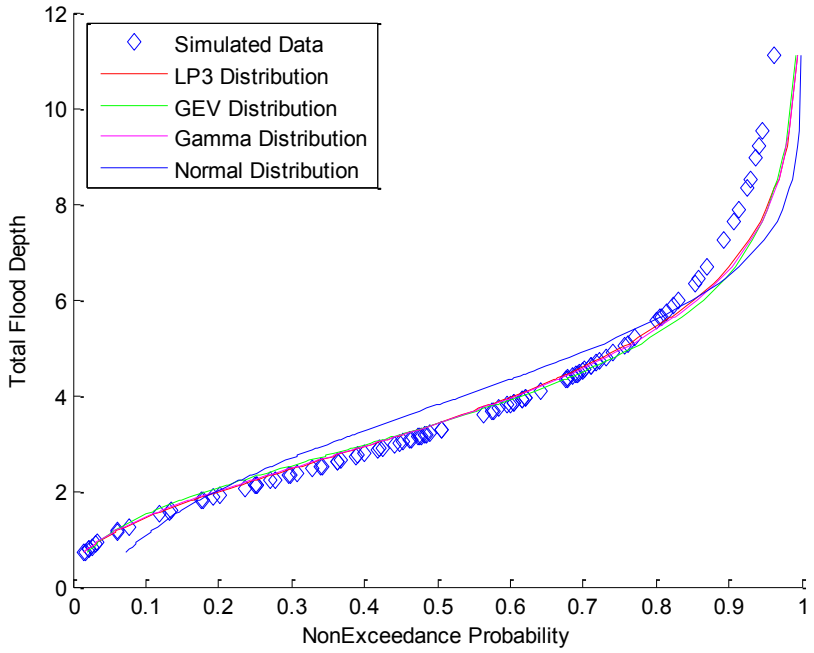

Figure 6-24: Comparison of LP3 Distribution to Non-Exceedance Probabilities Calculated for Total Flood Depths for a Sample Size of 100

For all three of the larger sample sizes assessed, the probability plot correlation coefficients and plots comparing the distributions to the samples were

evaluated. The probability plot correlation coefficients in all cases suggested that the normal distribution could not adequately represent the samples, but all three of the other distributions could adequately and approximately equally represent the total depth samples. The plots show the LP3, GEV, and gamma distributions as essentially identical and all able to represent the sample very well. However, in all cases, the normal distribution is observed to fit the samples more poorly than the others. For all three samples, the gamma distribution was selected to represent the total flood depths. As the sample size increased, the gamma shape parameter was observed to slightly decrease and the scale parameter was observed to increase. This suggests that the sample size may have some influence on the probabilities calculated corresponding to the total flood depths; however, the impact is fairly minimal. The trend in parameters may be due to the fact that as the sample size increased the distributions became more accurate representations of the population.

The impact of the choice of distribution to represent each sample was next evaluated. Table 6-35 provides the predicted flood depths corresponding to common exceedance probabilities for each distribution for each sample size. For a sample of 10, there is significant difference in the predicted flood depths based on the chosen distribution. The GEV distribution in particular predicts unreasonably high flood depths for the larger events. This was expected given the poor fit of the GEV distribution observed in Figure 6-22. For the larger sample sizes, there was minimal difference in the predicted flood depths for the smallest flood events. For the larger flood events, there was minimal difference in the predicted flood depths between the LP3, GEV, and gamma distributions; however, there tended to be more variation in

the predicted flood depths based on the normal distributions. Based on the similarities in the performance of the three distributions previously observed for the larger sample sizes, these results were expected.

Table 6-35: Predicted Flood Depths (in Feet) Corresponding to Common Exceedance Probabilities for Each Distribution for Each Sample Size

Exceedance Probability	LP3	GEV	Gamma	Normal
N=10				
0.1	7.2558	12.1799	6.9594	7.3202
0.02	12.1285	102.9333	9.2242	9.114
0.01	14.8325	271.1479	10.1214	9.7472
N=25				
0.1	6.5678	6.3147	6.4632	6.2754
0.02	8.7652	8.5062	8.9292	7.7837
0.01	9.5578	9.3649	9.9209	8.3161
N=50				
0.1	6.7051	6.5764	6.6509	6.5922
0.02	9.2043	9.6929	9.2609	8.2544
0.01	10.1556	11.1145	10.3135	8.8412
N=100				
0.1	6.6921	6.5354	6.63	6.5094
0.02	9.2328	9.432	9.2696	8.1438
0.01	10.2023	10.7029	10.3357	8.7207

6.2.5.1.2. Effects of Varying Sample Correlation

Distributions were next fitted to the total flood depth samples for the varying levels of correlation. These non-exceedance probabilities were compared to the probabilities calculated by taking the triple integral under the joint pdf calculated using the Gumbel-Hougaard copula for each scenario. The results for all six scenarios were quite similar, so the results and implications will be discussed after all results have been presented. Table 6-36 and Figure 6-25 present the results for scenario RHA. The results of scenario RHLH were very similar, so the plot

comparing the distributions to the sample was not presented. The fitted distribution parameters and the probability plot correlation coefficients for this scenario are presented in Table 6-37. The results of scenario RHLL are presented in Table 6-38 and Figure 6-26. For scenario RLA, results are presented in Table 6-39 and Figure 6-27. For scenario RLHH, the results are presented in Table 6-40 and Figure 6-28. The results of the final scenario, RLHL, were very similar, so the plot comparing the distributions to the sample was not presented. Table 6-40 provides the fitted distribution parameters and probability plot correlation coefficients for this scenario.

Table 6-36: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RHA

	LP3	GEV	Gamma	Normal
Shape	26.3279	0.1262	2.9301	N/A
Scale	-0.1229	1.5818	1.3143	2.3224
Location	4.404	2.7225	N/A	3.851
Probability Plot Correlation Coefficient	0.9925	0.9952	0.9876	0.9334

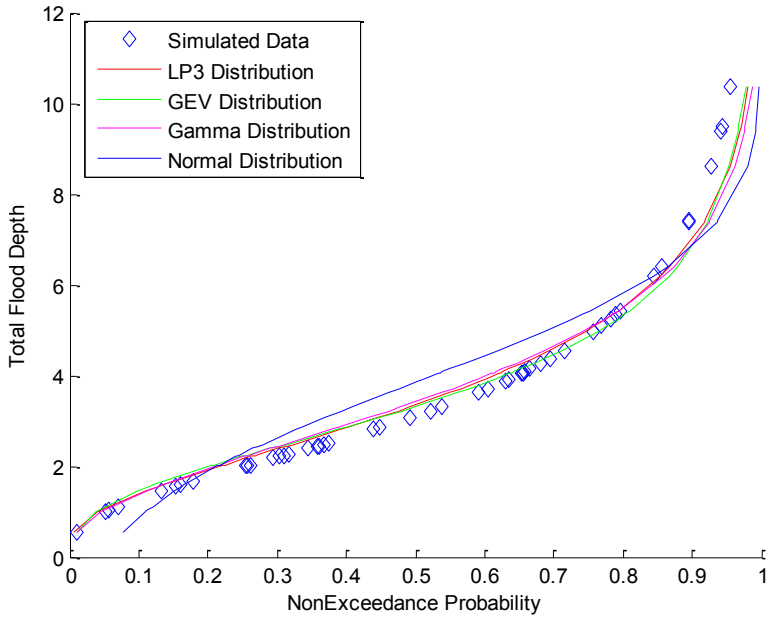

Figure 6-25: Comparison of LP3 Distribution to Non-Exceedance Probabilities Calculated for Total Flood Depths for Scenario RHA

Table 6-37: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RHLH

	LP3	GEV	Gamma	Normal
Shape	22.0677	0.1109	3.0694	N/A
Scale	-0.1311	1.5697	1.2592	2.2789
Location	4.0724	2.7708	N/A	3.8649
Probability Plot Correlation Coefficient	0.9914	0.9947	0.9876	0.934

Table 6-38: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RHLL

	LP3	GEV	Gamma	Normal
Shape	19.2497	0.0978	3.1845	N/A
Scale	-0.1377	1.5589	1.2152	2.2405
Location	3.8389	2.8053	N/A	3.8698
Probability Plot Correlation Coefficient	0.9907	0.9945	0.988	0.934

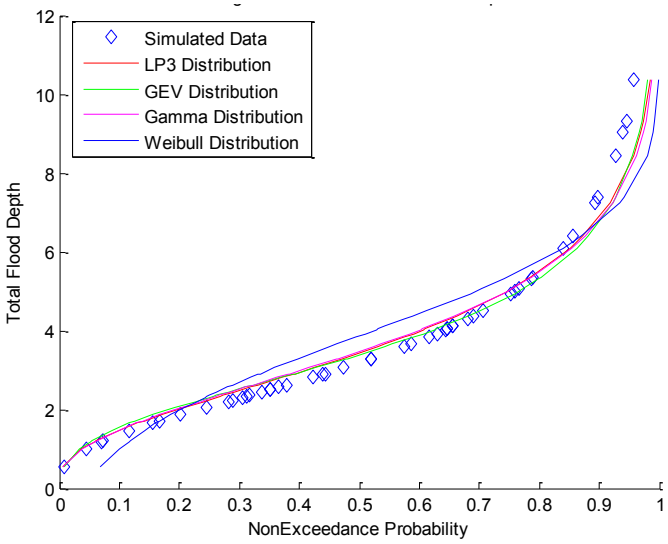

Figure 6-26: Comparison of LP3 Distribution to Non-Exceedance Probabilities Calculated for Total Flood Depths for Scenario RHLL

Table 6-39: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RLA

	LP3	GEV	Gamma	Normal
Shape	12.849	0.0574	3.485	N/A
Scale	-0.1615	1.5276	1.1041	2.1163
Location	3.2726	2.8716	N/A	3.8479
Probability Plot Correlation Coefficient	0.9866	0.9912	0.9872	0.9345

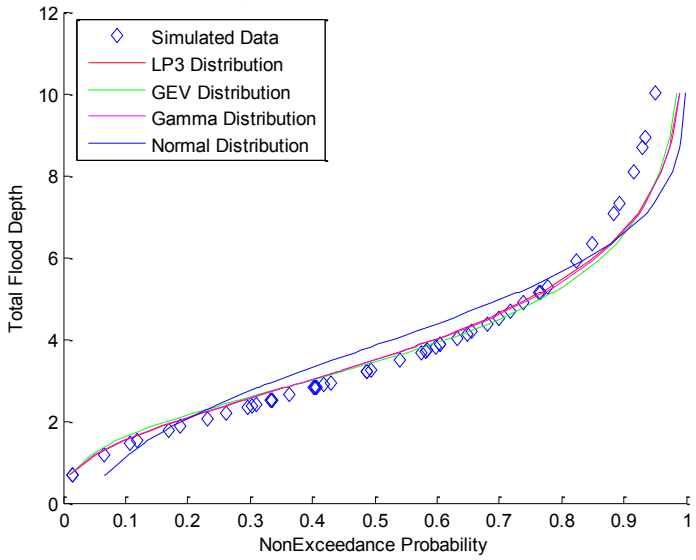

Figure 6-27: Comparison of LP3 and Normal Distributions to Non-Exceedance Probabilities Calculated for Total Flood Depths for Scenario RLA

Table 6-40: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RLHH

	LP3	GEV	Gamma	Normal
Shape	13.367	0.0757	3.2086	N/A
Scale	-0.1658	1.558	1.1941	2.1909
Location	3.3952	2.805	N/A	3.8315
Probability Plot Correlation Coefficient	0.9869	0.991	0.9864	0.9317

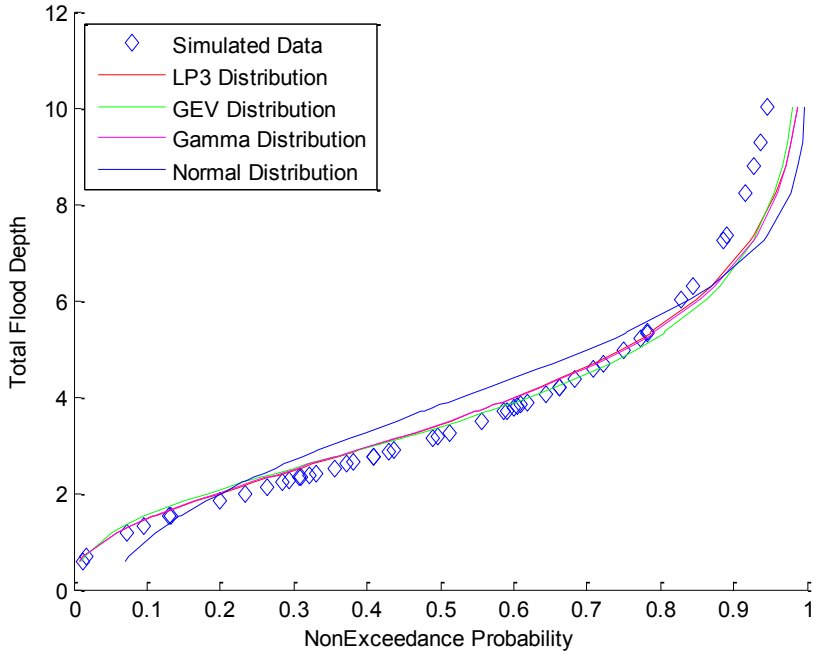

Figure 6-28: Comparison of LP3 and Normal Distributions to Non-Exceedance Probabilities Calculated for Total Flood Depths for Scenario RLHH

Table 6-41: Parameters and Probability Plot Correlation Coefficients Calculated for the LP3, GEV, Gamma, and Normal Distributions Corresponding to Total Flood Depths for Scenario RLHL

	LP3	GEV	Gamma	Normal
Shape	12.3858	0.065	3.3292	N/A
Scale	-0.1689	1.5443	1.1515	2.1525
Location	3.2783	2.8341	N/A	3.8337
Probability Plot Correlation Coefficient	0.9864	0.9909	0.9869	0.9324

For all six scenarios, the fitted distributions were compared to the results of the triple integral procedure using the probability plot correlation coefficients and the plots comparing the distributions to the samples. In all cases, the probability plot correlation coefficients indicated that the LP3, GEV, and gamma distributions would all be able to adequately, and approximately equally, represent the total flood depth samples. They also suggested that the normal distribution would not be able to represent the samples as well as the other distributions. These findings were

confirmed by the plots that compared the distributions to the samples. In all cases, the LP3, GEV, and gamma distributions were essentially identical and agreed very well with the generated sample. The normal distribution; however, was not observed to agree so well with the simulated samples. Thus, for all cases, the gamma distribution was selected to represent the total flood depth samples. As the level of correlation between the samples decreased, the gamma shape parameter was observed to increase and the scale parameter was observed to decrease. This suggests that the level of correlation between the samples does have some influence on the probabilities that correspond to the total flood depths. This was expected given the impact that the level of correlation was observed to have when two flood sources were considered.

The impact of the choice of distribution to represent the total flood depths was next examined. Table 6-42 provides the predicted flood depths for several common exceedance probabilities for each distribution in each scenario. In all cases, minimal differences in the predicted flood depths can be observed between the LP3, GEV, and gamma distributions for any magnitude event. However, more variation between these three distributions is evident than was evident for the larger sample size scenarios. The differences in the predicted flood depths for the larger magnitude events appear to be more significant in the scenarios with higher levels of correlation between the samples. This suggests that the distribution selected to represent the total flood depth samples would have more impact when higher levels of correlation exist between the samples. Table 6-42 demonstrates the difference in predicted flood depths based on the normal versus the other three distributions. As expected, the

normal distribution would predict quite different flood depths, especially for the larger magnitude events. This confirms that the normal distribution should not be used to represent the total flood depths but the choice between the other three distributions would have minimal impact on the results of the analyses.

Table 6-42: Predicted Flood Depths (in Feet) Corresponding to Common Exceedance Probabilities for Each Distribution for Each Level of Correlation between the Samples

Exceedance Probability	LP3	GEV	Gamma	Normal
RHA				
0.1	6.9978	6.839	6.8673	6.8273
0.02	10.2596	10.6977	9.7304	8.6206
0.01	11.6211	12.5868	10.8921	9.2537
RHLH				
0.1	6.9255	6.7831	6.823	6.7854
0.02	9.9707	10.4346	9.6057	8.5452
0.01	11.2152	12.1912	10.7322	9.1664
RHLL				
0.1	6.8691	6.7294	6.7776	6.7411
0.02	9.7472	10.2115	9.494	8.4712
0.01	10.9033	11.8614	10.5917	9.082
RLA				
0.1	6.661	6.5412	6.6115	6.56
0.02	9.0907	9.5525	9.1505	8.1942
0.01	10.0151	10.9139	10.1722	8.7712
RLHH				
0.1	6.7647	6.6929	6.6995	6.6393
0.02	9.3938	9.9429	9.375	8.3311
0.01	10.4085	11.4438	10.4558	8.9283
RLHL				
0.1	6.7051	6.5764	6.6509	6.5922
0.02	9.2043	9.6929	9.2609	8.2544
0.01	10.1556	11.1145	10.3135	8.8412

6.2.6. Conclusions

The procedure developed provided a probability of flooding when flooding could be caused by any combination of three sources. Cumulative probability

distributions were fitted to the observed flood inundation depths from the riverine, tidal, and pluvial flood sources. Once the marginal distributions had been obtained, the copula parameters were calculated for the three copula families and then the most appropriate copula family was identified based on calculations of Akaike's Information Criteria. Once the most appropriate copula families had been identified for each simulation scenario, the joint pdfs and cdfs were calculated using the copula equations. The joint pdfs were used to calculate non-exceedance probabilities that corresponded to total flood depths, and then four distributions were fitted to the total flood depths to represent these non-exceedance probabilities. This procedure provided an assessment of the flood hazard for a given location of interest when three flood sources could potentially impact the location.

The procedure developed was used in a series of simulation studies in order to understand the effects of sample size and levels of correlation between the flood source samples. Specific trends in marginal distribution parameters were not observed due to either sample size or level of correlation. This is likely due to sampling variation. As would be expected, as the marginal sample size increased, the probability distributions better fit the simulated data sets. As the sample size increased, the copula parameters were observed to generally decrease. The copula parameters were also observed to decrease as the levels of correlation between the three flood depth samples decreased. However, the copula parameters calculated for both the varying sample size and varying sample correlation scenarios were observed to have a fairly wide standard error, as was also observed when two flood sources were examined. The variations in copula parameters calculated for these scenarios

did not result in widely different joint distributions. Thus, though trends in the copula parameters were identified in the copula parameters, they were not significant. However, it is expected that a wider variation in the levels of correlation between the generated samples may result in a more significant difference in the copula parameters, as was observed when two flood sources were assessed.

Based on the copula parameters identified, joint pdfs and cdfs were calculated for each scenario. Trends in the joint distributions were not identified based on either the sample size or the level of correlation between the flood depth samples. It is possible that because the joint distributions could not be visualized as they were when only two flood sources were considered trends due to either factor may have been overlooked. However, because the differences in copula parameters were observed to be so minimal, it is expected that there would not have been trends in the joint distributions due to these factors. The choice of copula family used to represent each scenario was also not observed to significantly impact the joint distributions. For two flood sources, the choice of copula family was not observed to impact the joint distributions developed for varying sample size but the choice of copula family was observed to impact the joint distributions for varying levels of correlation between the samples. It is possible that the choice of copula family would also have a more significant impact when three flood sources were considered if a wider range of levels of correlation were examined.

The next step in the procedure was to calculate the cdfs that corresponded to total flood depths using a triple integral under the joint pdfs. Four distributions were fitted to each total flood depth sample and compared to the non-exceedance

probabilities calculated using the triple integral procedure. The log-Pearson Type III, Generalized Extreme Value, and gamma distributions were typically observed to fit the total flood depths quite well, though the normal distribution was not observed to fit the samples as accurately. Based on a comparison of predicted flood depths that corresponded to several common exceedance probabilities between the four distributions considered, it was determined that the choice of distribution among the LP3, GEV, and gamma distributions had minimal impact. The parameters of the fitted distributions were observed to vary slightly depending on the sample size, which indicated that the sample size may have some impact on the fitted distributions. This may simply be due to the fact that the fitted distributions typically become more accurate as the sample size increases. A small trend in the fitted distribution parameters was also observed for varying levels of sample correlation. This was likely due to the fact that the level of correlation influenced the generated tidal and pluvial samples and therefore influenced the total flood depth samples. However, these changes in the fitted distribution parameters were likely not significant, given the minimal impact that either sample size or sample correlation had on the joint distributions. The probability distributions that represented the non-exceedance probabilities of the total flood depths could then be used to assess the flood risk at the location of interest. The results of these simulation studies will help to understand the results obtained using observed data.

6.3. Flood Risk Calculations

The previous sections have developed a process to assess the flood hazard for a location of interest based on multiple flood sources. The flood frequency

assessments developed were based on joint distributions for all three flood sources calculated based on copulas. However, the flood hazard is only one part of flood risk. The other two components of flood risk are the vulnerability to flooding and the consequences that result from those flood events. Flood risk is calculated by multiplying the exceedance probability, the vulnerability, and the consequences for a given flood depth of interest.

The term vulnerability refers to the ability of the system in place to protect the location of interest from flooding. The system will include any structures near the location of interest designed to reduce the probability of flooding or any nonstructural measures taken at or near the location of interest to reduce either the probability or the consequences of flooding. Common examples of structural systems include levees or floodwalls. Nonstructural measures include elevation of structures, flood proofing structures, or creating additional room to store waters naturally. Whatever the system may be that is identified as protecting the location of interest, failure is always a possibility. This uncertainty is incorporated into the vulnerability term of the risk equation.

To complete the flood risk calculations, the consequences of the flood event must be assessed. Consequences may be identified in economic terms, such as damages to structures and infrastructure, environmental damages, or loss of life or injuries as a result of the flood event, though economic terms are the most commonly used. For the purpose of providing an example of the method, economic consequences will be considered because they can be easily calculated and understood.

6.3.1. Description of Methods

The exceedance probabilities calculated for the total flood depths using the copula procedure were used as input to the flood risk calculations. Flood risk calculations were then made based on the other terms in the risk equation, the vulnerability and the consequences. Vulnerability is a weight that ranges between 0 and 1 and indicates how well the in-place system to protect the location of interest from flood events is expected to perform. Therefore, vulnerability will vary with flood depth. A hypothetical vulnerability curve was created for use in experimental calculations, as shown in Figure 6-29. Vulnerability increases with depth in a linear fashion, which indicates that vulnerability is simply directly proportional to depth of flooding, which may be true for certain systems. Without specific information about the system in place at the location of interest this linear curve seemed to be a reasonable approximation of vulnerability for the purpose of these calculations.

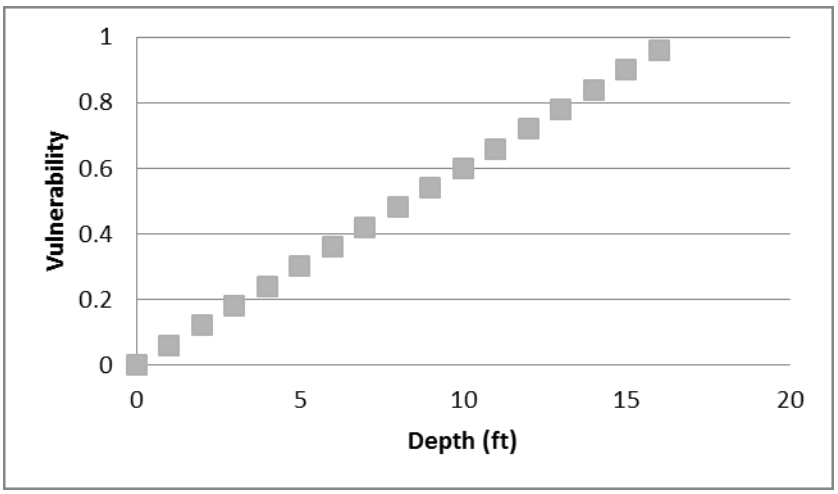

Figure 6-29: Illustration of the Vulnerability Curves Used in Flood Risk Assessments

Consequences, which are measured as damage to the structure and its contents due to the flood, will also vary with depth. Depth-damage curves (e.g., see Figure 6-

30) obtained from USACE (2003) are used to determine the percentage of damage to the structure and to contents based on the depth of flooding. These depth-percent damage curves are general curves that can be applied nationally, though it is also preferable to derive specific curves for a given location. For a hypothetical scenario using simulated data, the general curves were deemed to be appropriate to show the methodology. These percent damage values were then multiplied by the value of the structure and the value of the contents to determine the total monetary damage due to varying flood depths. For the hypothetical scenario, the structure was assumed to be a two-story residential structure without a basement, assumed to be worth $150,000 with contents assumed to be worth $25,000. The resulting depth-monetary damage curves are provided in Figure 6-31.

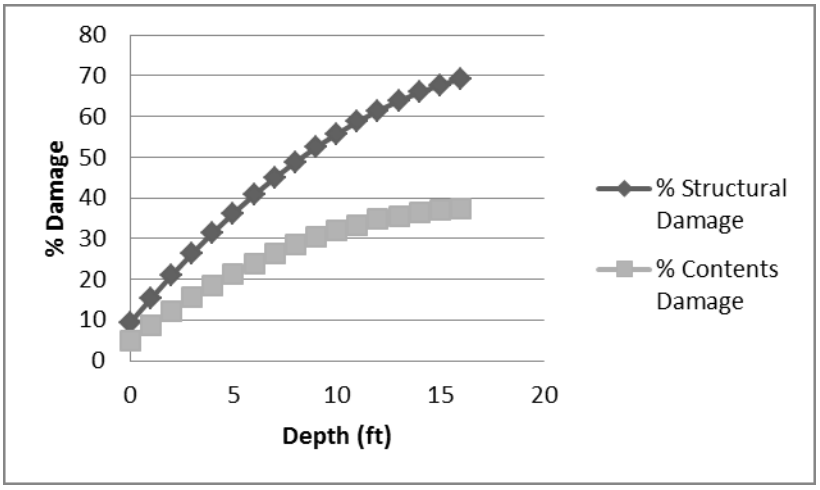

Figure 6-30: Depth-Percent Damage Curves Obtained from USACE (2003) for 2-Story Residential Home Without Basement

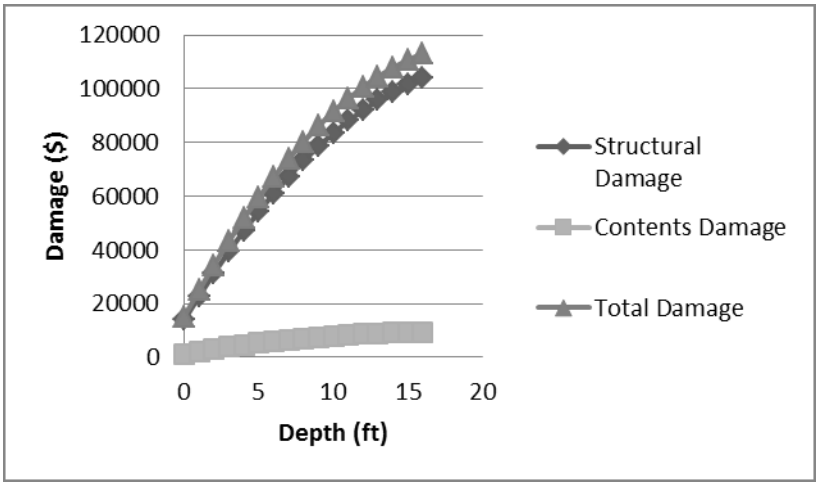

Figure 6-31: Depth-Monetary Damage Curves for 2-Story Residential Home Without Basement

Depending on the purpose of the risk analysis, it might be desirable in some circumstances to fit equations to the vulnerability and consequence curves, such that values could be determined for any flood depth of interest. However, the purpose of this analysis is simply to demonstrate the effect of the comprehensive probability assessment on risk calculations. Thus, the population fitted to the exceedance probabilities that corresponded to the total flood depth values was used to determine the exceedance probabilities corresponding to various depths of flooding for which values of vulnerability and consequences had been obtained. These exceedance probabilities were multiplied by the vulnerability values and consequence values for each flood depth to obtain flood risk values for each flood depth.

A total of 10 scenarios were investigated using the simulation study. Sample sizes of 10, 25, 50, and 100 were investigated, as were sample correlation scenarios RHA, RHLH, RHLL, RLA, RLHH, and RLHL. For each of these scenarios, the exceedance probabilities that correspond to flood depths of 1 through 16 feet were first calculated using the probability distributions fit in the flood hazard assessment. The exceedance probabilities corresponding to flood depths were used as input to the

flood risk calculations for each simulation scenario. The exceedance probabilities were then multiplied by the vulnerability curve previously identified in Figure 6-29 and the consequences curve shown in Figure 6-31 to obtain calculations of flood risk for each simulation scenario.

6.3.2. Description of Results

For each simulation scenario in which a flood hazard assessment was made, resulting in non-exceedance probabilities that corresponded to total flood depths, flood risk calculations were then made. The following sections present the results of these risk analyses for both varying sample size and varying levels of correlation.

6.3.2.1. Varying Sample Size

Prior to calculating the flood risk for flood depths of 1.0 to 16.0 feet, the exceedance probabilities for that range of flood depths were calculated for each of the four sample sizes as described in Section 6.2.5.1.1. These exceedance probabilities are presented in Figure 6-32. This information was used to calculate the corresponding flood risks, which are presented in Figure 6-33. As expected given the results presented when the distributions were fitted to the total flood depths, the exceedance probabilities calculated for a sample of 10 differed somewhat from the exceedance probabilities calculated for the larger sample sizes. Once the sample size was greater than 25 the exceedance probabilities varied minimally, indicating that sample size had minimal impact on the probabilities calculated for given flood depths of interest. It can be assumed that the smallest sample size could not be as accurately represented as the larger sample sizes.

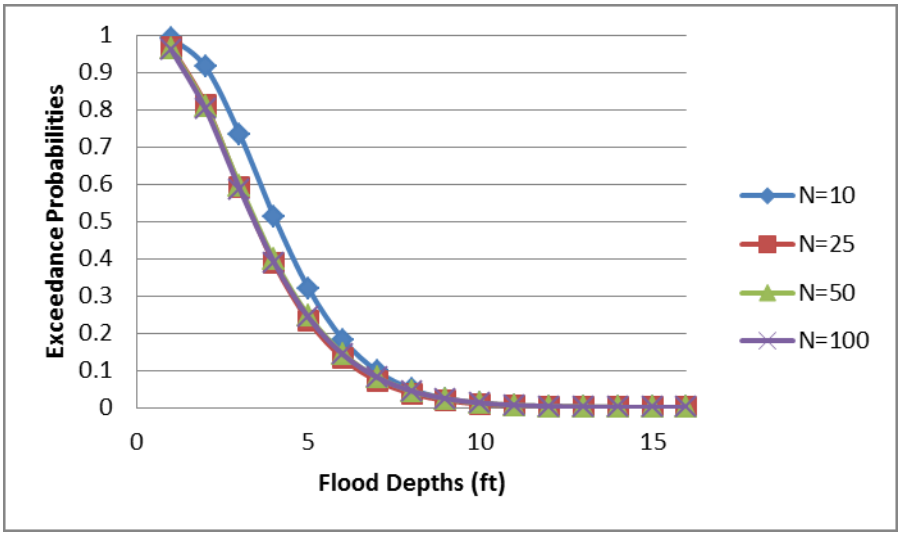

Figure 6-32: Exceedance Probabilities Calculated for Each Sample Size

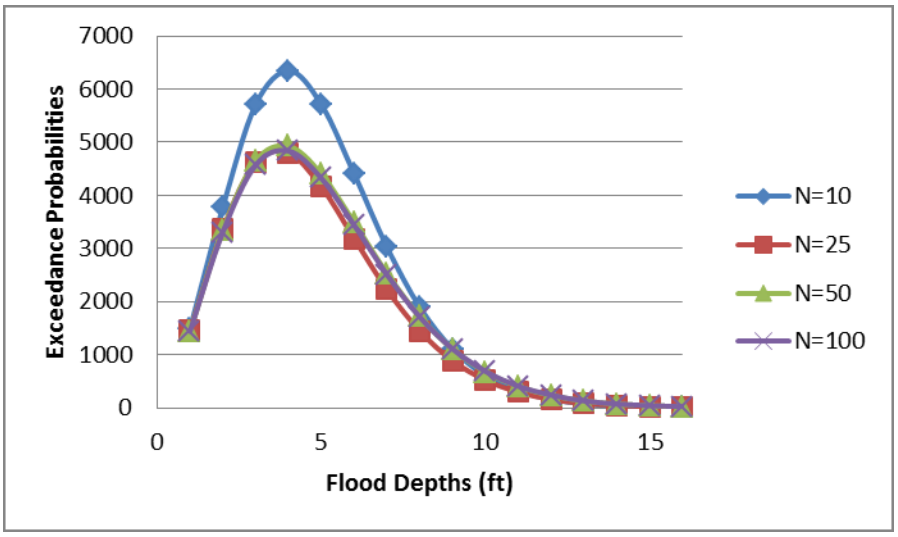

Figure 6-33: Flood Risk Calculations for Each Sample Size

Though the differences in exceedance probabilities calculated for a sample of 10 as compared to the larger samples did not appear to be extremely large, the difference in flood risk calculations for a sample of 10 does seem to differ widely from those calculated for the larger samples. Because the exceedance probabilities calculated for a sample of 10 were larger than those calculated for the larger samples the sample of 10 also produced larger estimates of flood risk than the larger samples. As expected, once the sample size reached 25, the flood risks calculated were

essentially identical. This suggests that once the sample size was large enough to provide an accurate representation of the underlying population, the sample size had minimal impact on the understanding of the flood risk for the location of interest. In examining the flood risk curves, it should be noted that the highest flood depths do not have the highest calculated flood risk values. This is because the probability of flooding decreased with increasing flood depths, though vulnerability and consequences increased with increasing flood depths. Because flood risk calculations are made by multiplying probability, vulnerability, and consequence values for each flood depth this result was expected.

6.3.2.2. Varying Sample Correlation

Flood risk calculations were also made for the six scenarios with varying levels of correlation between the samples. Exceedance probabilities calculated based on distributions fitted to the total flood depths for each correlation scenario were first compared, as presented in Figure 6-34. It would appear that there is virtually no impact on the exceedance probabilities based on the level of correlation between the samples, as would be expected given the results observed when distributions were fitted to the total flood depths for each scenario.

Based on these exceedance probabilities, flood risk calculations were made for each scenario, which are presented in Figure 6-35. The calculated flood risks for each scenario were observed to peak at a flood depth of approximately 4 feet. Though the exceedance probabilities decreased continuously with increasing flood depth, the vulnerabilities and consequences increased with increasing flood depth, leading to the location of the peak flood risk calculations. As there was virtually no

difference in the exceedance probabilities based on the level of correlation, there was also virtually no difference in the flood risk calculations based on the level of correlation between the samples either. In Figure 6-35, a very slight trend can be identified that as the level of correlation between the samples increased the calculated flood risk also increased, but for the range of correlations examined in these scenarios the differences in calculated flood risk are minimal. This suggests that the level of correlation between the samples has minimal impact on the understanding of the flood risk for the location of interest, at least for the range of correlations examined in this study. Based on the slight trend observed, it is believed that for higher levels of correlation between the samples, a more significant difference in the flood risk calculations would exist. The results observed for an analysis of two flood sources suggest that the level of correlation between the samples does have the ability to significantly impact the results of the analysis; however, a wider range of correlations were examined for two flood sources.

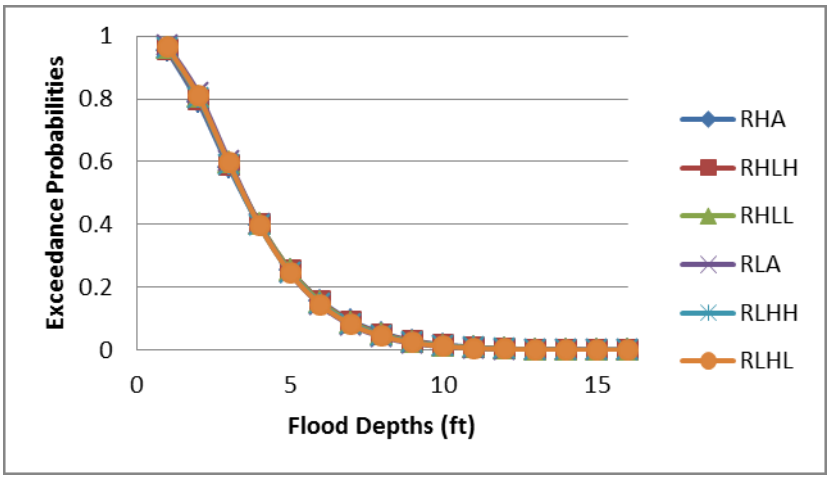

Figure 6-34: Exceedance Probabilities Calculated for Each Level of Sample Correlation

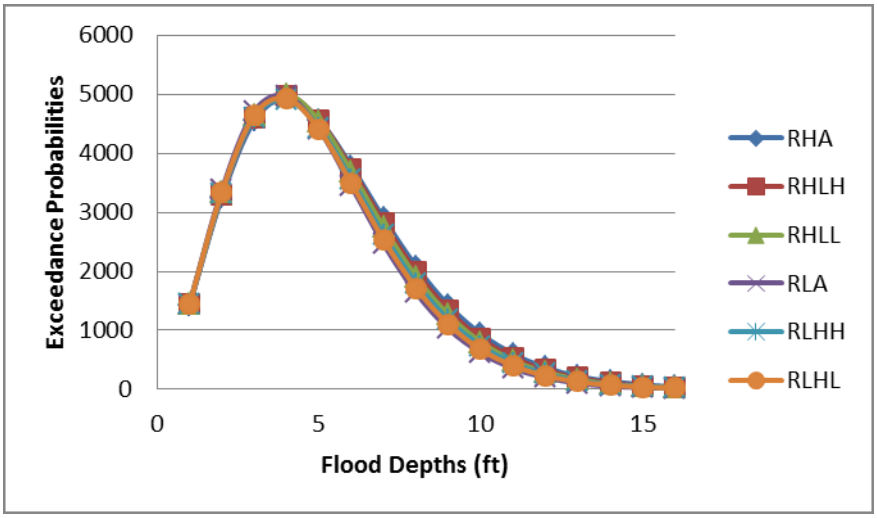

Figure 6-35: Flood Risk Calculations for Each Level of Sample Correlation

6.3.3. Conclusions

Flood risk was calculated for a series of simulated flood depths from three flood sources impacting a location of interest. The flood risk calculations assessed the impact of sample size and level of correlation between the samples on the risk calculations. Exceedance probabilities for each simulation scenario were calculated using the copula procedure to develop joint distributions based on three flood sources. As the results indicated, neither sample size nor sample correlation were observed to produce any particular impacts on the flood risk. A difference was observed in both exceedance probabilities and flood risk calculations for the smallest sample size of 10, but once the sample size reached 25 an impact on the probabilities and risk calculation was not observed. This suggests that once the sample size was large enough to be reasonably accurate the sample size did not significantly impact the results of the analysis. Significant trends were not observed in the flood risk calculations based on the level of correlation between the samples. A very slight trend of increased flood risk calculations with increased levels of correlation between

the samples was observed, but for the range of sample correlations evaluated in these studies the impact on the results of the analyses was minimal. For larger levels of correlation a more significant impact on the calculated flood risks would be expected based on these minor trends as well as based on the more significant impact that level of correlation was observed to have on the results when only two flood sources were considered. These flood risk calculations will be helpful in providing some understanding of future risk calculations made using observed data.

CHAPTER 7

THREE OBSERVED FLOOD SOURCES

7.1. Introduction

The simulation studies presented in Chapter 6 provide an understanding of the impact of sample size and sample correlation on the development of flood frequency assessment based on three flood sources. Because only a limited amount of observed data exist, these simulations studies were necessary to provide a context for results obtained from the observed data. In addition, because the mathematics of combining three flood sources using the copula procedure are quite complex, the simulation studies assisted in determining that the results obtained from the developed procedures were rational and logical. The process of developing a comprehensive flood frequency analysis for a location in Florida, considering a riverine, tidal, and pluvial flood source, will be discussed herein and results will be presented.

7.2. Description of Experimental Location

A location of interest was identified close to both a river and the coast on the east coast of Florida. Discharge data (ft^3/s) from flow gage 02249007 on the Eau Gallie River at Heather Glen Circle at Melbourne were obtained. The tidal gage from which tidal height measurements (ft above NAVD 88 datum) were obtained was 8721604, located near Trident Pier. The tidal gage provided measurements of water height in feet. This tidal gage is approximately 20 miles from the location of interest, which is not ideal, but a closer location was not available. Precipitation depths (in

inches) were obtained from rain gage 085612, located at the nearby Melbourne Weather Forecast Office. The watershed area was 12.4 mi^2 and the elevations over the watershed ranged from approximately 0 to 35 feet above MSL.

7.3. Assessment of the Hazard

The first step to conducting a risk assessment is to assess the hazard. The hazard is usually considered as the probability of a specific event occurring. The following sections will outline a methodology and present the results of assessing the flood hazard for a location which could be impacted by three flood sources.

7.3.1. Fitting Marginal Distributions to Observed Annual Maximum Events

When three flood sources are considered, the first step to assessing the flood hazard was to understand the probability of flooding from each of the flood sources. This required calculating flood inundation depths based on observed gage measurements, and then identifying a series of annual maximum events. Finally, marginal distributions were fitted to the depths of the annual maximum flood events from each flood source.

7.3.1.1. Description of Methods

Observed riverine, tidal, and pluvial samples were obtained and transformed into flood inundation depths at the location of interest. Marginal distributions were then fitted to these inundation depth samples. The following sections will describe the methods used to calculate flood inundation depths from each source and to fit marginal distributions to those samples.

7.3.1.1.1. Calculation of Riverine-Caused Flood Inundation Depths

To perform a flood frequency analysis for the location in Florida required observed riverine flow data. Discharge data were obtained from the United States Geological Society and information from this location, including elevations and channel characteristics, were obtained from ArcGIS and GoogleEarth for use in transforming the discharge data to flood depths at the location of interest. The process of transforming the discharges to inundation depths at the location of interest is explained here. The process involves two steps: (1) transposing the discharge from the gaged site to a site on the river adjacent to the point of interest and (2) transforming the river discharge to an inland depth.

Prior to obtaining riverine-caused flood depths at the location of interest, it was necessary to obtain discharge measurements at a flow gage. Because the number of river flow gages is minimal, it is unlikely that a flow gage would be located at a point within the stream that is adjacent to the location of interest. Therefore, the discharges measured at the flow gage had to be transposed downstream to a point adjacent to the location of interest, which used the drainage area-ratio method (McCuen and Levy, 2000):

$$Q_u = Q_g \left(\frac{A_u}{A_g}\right)^n \tag{7-1}$$

where Q_u is the discharge at the ungaged location (cfs), Q_g is the discharge measured by the gage (cfs), A_u is the drainage area of the ungaged point (mi^2), A_g is the drainage area at the location of the flow gage (mi^2), and n is an empirical constant, with a value of 0.8 empirically derived. For the study location in Florida, A_g was 3.8

mi^2, as reported by the USGS, and A_u was 12.4 mi^2, as calculated in ArcGIS. The result of this step was a series of stream discharge estimates transposed downstream to the stream cross-section adjacent to the location of interest.

The channel cross-section information at the location of interest was not available; therefore, an estimate of the bankfull channel dimensions was required. The bankfull channel dimensions were first estimated at the location of the gage. Measured discharge and the corresponding gage heights were obtained from the gage. The flow depth corresponding to each discharge was calculated by subtracting the gage datum, which was 11.39 feet, from the gage height measurements. Flow heights were plotted against corresponding discharges, and a power model was fitted to describe this relationship. The power model used to describe the relationship between discharge and flow depth was:

$$Discharge = 1.2212 * Depth^{4.6117} \quad (7\text{-}2)$$

Based on the observed gage data, the bankfull discharge at the location of the gage was identified to be 428 cfs. Using Equation 7-2, the bankfull depth was calculated to be 3.563 feet. Assuming a rectangular channel with a hydraulic radius was approximately equal to the flow depth, Manning's equation was calculated to determine the corresponding channel width at bankfull flow. The bankfull channel width at the location of the gage was estimated to be 99.8 feet. This provided an estimate of the channel bankfull dimensions at the location of the gage.

The next step was to estimate the bankfull channel dimensions at the downstream location. A uniform channel was assumed, such that the ratio of channel

width to depth remained constant along the length of the channel. Using the downstream discharges calculated using Equation 5-1 and Manning's equation, bankfull channel width was solved. In doing so, channel depth was approximated as 0.0357 times the channel width, based on the ratio of channel depth to width at the gage location. This resulted in an estimated bankfull channel width of 123.5 feet at the location of interest. The ratio of channel depth to width then provided an estimate of the channel's bankfull depth to be 4.409 feet.

Once bankfull channel dimensions had been estimated, the stream discharge values could be translated to flood inundation depths at the location of interest. Flow depths that correspond to the transposed discharges were calculated based on Manning's discharge equation. This was an iterative process, where depth was varied until the corresponding calculated discharge equaled the observed discharge rate. Manning's equation for discharge (ft^3/s) is:

$$Q = \frac{1.49}{n} * A_x * \left(\frac{A_x}{WP}\right)^{2/3} * S_f^{1/2} \quad (7\text{-}3)$$

where n is the Manning's roughness coefficient, assumed to be 0.05; A_x is channel cross-sectional area (ft); WP is wetted perimeter (ft); and S_f is the channel slope, determined to be 0.0007 ft/ft. Both A_x and WP are dependent on channel geometry. A trapezoidal floodplain was assumed, with a slope of 49.2 feet horizontally for every one foot increase vertically. The values of channel and floodplain slope were determined using elevation data provided in ArcGIS for the location of interest in Florida. Cross-sectional area and wetted perimeter (see Figure 7-1) were calculated as:

$$A = W_b * d + z(d - D_b)^2 \quad (7\text{-}4a)$$

$$WP = W_b + 2 * D_b + 2 * (d - D_b) * \sqrt{1 + z^2} \quad (7\text{-}4b)$$

where d was flow depth and z was 49.2 feet/feet, determined based on floodplain slope. The channel geometry is illustrated in Figure 7-1. This iterative procedure identified the depth of flow corresponding to each discharge value in the sample obtained from the Florida flow gage.

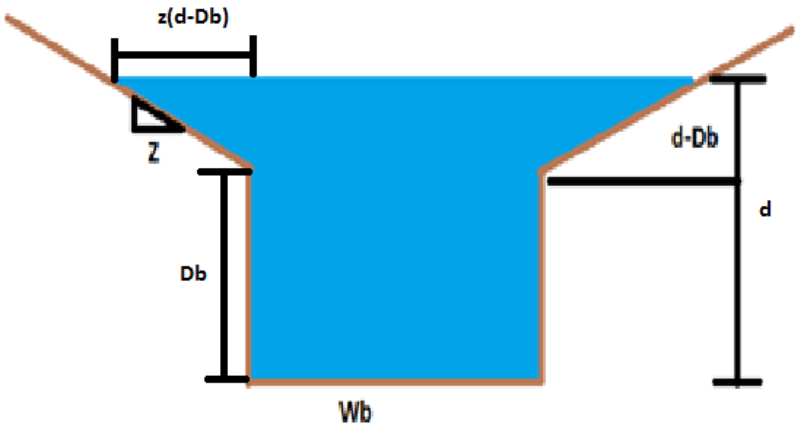

Figure 7-1: Assumed Channel Geometry

Once flow depths had been calculated, flood inundation depths at the location of interest could be determined. The next step in this process was to add the calculated flow depths to the channel bed elevation to obtain flow elevation. The channel bed elevation was assumed to be 0.4 feet above the North American Vertical Datum of 1988 (NAVD 88). This elevation was determined based on elevations provided by ArcGIS for the location in Florida. To determine the flood inundation depth at the location of interest, the elevation of the location of interest was then subtracted from the calculated flow elevation. The location of interest was assumed to be at an elevation of 4.2 feet above NAVD 88. This elevation was determined

based on elevations provided by ArcGIS for the location in Florida. By following these steps, an estimate of the flood inundation depth at the location of interest can be calculated for each discharge value measured by the flow gage.

The procedure just explained was used to calculate flood inundation depths at a location of interest in Florida, based on riverine discharge measurements taken by a local flow gage. Once flood inundation depths were calculated, the annual maximum events were identified and used as input to the remainder of the methodology.

7.3.1.1.2. Calculation of Tidally-Caused Flood Inundation Depths

To determine the flood inundation depth at the location of interest due to tidal flooding, a procedure identified in a 1977 National Research Council report (NRC, 1977) to the Federal Insurance Administration examining the effects of wave action on storm surge was used. The mean stillwater elevation, which is the sum of astronomical tide and storm surge, the sum of which can be determined from tidal gage measurements, was determined at the location of the tidal gage. Next, the additional height caused by wave action was calculated and added to the stillwater elevation. The shoreline geometry is illustrated in Figure 7-2. The equation to determine wave height is (NRC, 1977):

$$H = 0.78 * F * E_s \tag{7-5}$$

where H is the wave height (ft), F is a factor related to the length of fetch, which is the length of water over which the wind has blown, and E_s is the stillwater elevation (ft above NAVD88). For the Florida location, the fetch was taken as essentially unlimited, as the location opens to the ocean rather than a bay, which corresponded to

a value of 1.0 for F. This provided the additional height of waves at the location of the tidal gage. In Table 7-1 the relationship is provided between fetch length and fetch factor, F, as indicated in NAS (1977).

Table 7-1: Relationship Between Fetch Length and Fetch Factor, F (NAS, 1977)

Fetch (miles)	F (Fetch Factor)
0.125	0.25
0.25	0.32
0.5	0.41
1	0.52
2	0.65
4	0.78
10	0.93
>20	1

The NRC procedure would allow for the wave height past an obstruction, such as a dune or seawall, to be calculated. However, from GoogleEarth it was not possible to determine whether or not such an obstruction existed. Thus, it was assumed that an obstruction between the coastline and the location of interest did not exist.

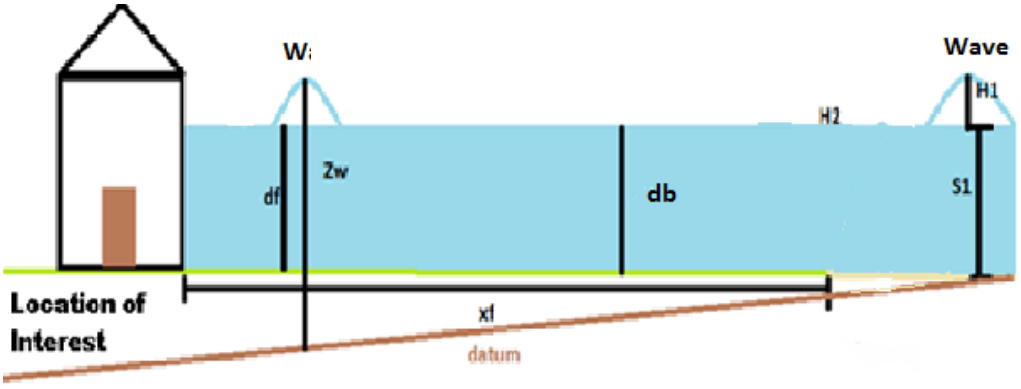

Figure 7-2: Shoreline Geometry, where S1 is Stillwater Elevation, H1 is Wave Height, db and df are Water Depths, xf is Inland Fetch, H2 is Inland Wave Height, and Zw is Flood Elevation, the Green Line is the Ground Surface, and the Brown Line is the Official Datum

The next step was to account for wave generation as the tidal surge moved inland between the dunes and the location of interest. This wave generation would

primarily be caused by wind moving over the water surface. As the water travels over the land surface, additional waves may be generated, primarily through the effects of wind. Wave height at the end of the inland fetch, H_f, was calculated as (NRC, 1977):

$$H_f = \sqrt{(G * d_f)^2 + H^2} \quad (7\text{-}6)$$

where G was a unitless value related to the inland fetch length, d_f was the average depth over the inland fetch (ft), and H was the wave height (ft) calculated behind the dunes. For the Florida location of interest, the inland fetch was determined to be 0.02 miles, estimated using GoogleEarth, corresponding to a G value of 0.2. The average depth, d_f, was calculated as the stillwater elevation minus the average elevation between the coastline and the location of interest, which was taken to be 1 foot above NAVD 88 based on measurements in ArcGIS.

Finally, the flood elevation at the location of interest, Z_w, was calculated using (NRC, 1977):

$$Z_w = E_s + 0.7 * H_f \quad (7\text{-}7)$$

where E_s was the stillwater elevation (ft) previously calculated based on tidal gage measurements and H_f was the wave height (ft) calculated at the location of interest (Equation 7-6). From Equation 7-7, the flood inundation at the location of interest was calculated by subtracting the elevation of the location of interest, which was 4.2 feet above NAVD 88 for the Florida location, from the Z_w values calculated.

The procedure just described was used to determine tidal flood inundation depths at the location of interest. Once the tidal flood inundation depths had been calculated for the location of interest, the annual maximum events were identified. The annual maximum events were used as input to the remainder of the procedure.

7.3.1.1.3. Calculation of Pluvially-Caused Flood Inundation Depths

Pluvial flooding is typically overland flow caused by quantities of rainfall that exceed the drainage capacity. Pluvial flooding is commonly considered as an urban phenomenon, in which the capacity of the sewer system is overwhelmed. However, in rural and other areas where sewer systems do not exist, the infiltration capacity of the land may also be exceeded, leading to overland flow. Pluvial flood depths are calculated as depth of runoff caused by a given depth of rainfall.

Calculation of pluvial flood depths is a multi-step process that was necessary to determine the population from which to generate pluvial flood depth samples. First, total rainfall depths that correspond to individual storm events are identified. It was assumed that the annual maximum pluvial flood depths would occur on the same date and time as the annual maximum riverine flood depths. The annual maximum riverine flood depth occurrences were previously identified in conjunction with tidal flood depths. Thus, rainfall storm events that occur on the dates of the previously identified annual maximum riverine flood depths were identified and then the rainfall depths for the duration of the storm were summed from the hourly rain gage measurements. Once storm rainfall depths had been calculated, the Natural Resources Conservation Service (NRCS) runoff depth estimation method was used to calculate corresponding runoff depths. Runoff depth, Q, was calculated as:

$$Q = \begin{Bmatrix} \frac{(P-0.2S)^2}{P+0.8S}, for\ P > 0.2S \\ 0, for\ P < 0.2S \end{Bmatrix} \quad (7\text{-}8)$$

where P is the rainfall depth, in inches, and S is the retention. The retention, S, was calculated as a function of the curve number (CN), which was based on land use and soil type. The retention was calculated as:

$$S = \frac{1000}{CN} - 10 \quad (7\text{-}9)$$

For the location of interest in Florida, a weighted curve number was required. The location of interest was determined to be in a fully developed urban area in fair condition with soil from soil group A, containing deep sand, which would have a curve number of 49. Based on aerial photographs of the watershed, a fraction of impervious surface was estimated to be 0.5. The weighted curve number was calculated as:

$$CN_w = CN_p(1-f) + f(98) \quad (7\text{-}10)$$

where CN_p is the curve number for pervious areas of the watershed, or 49, and f was the fraction of impervious area (McCuen, 2005). The weighted curve number was calculated to be 73. Based on this information, the runoff depth corresponding to every measured rainfall value measured by the rain gage was calculated.

Once equations 7-8 through 7-10 had been used to calculate runoff depths at correspond to the total storm depths, a peak discharge over the watershed resulting from these runoff depths could be calculated. From McCuen (2005), the following equation was identified to calculate peak discharges:

$$q_p = q_{um} A_m Q \tag{7-11}$$

where q_p is the peak discharge (ft^3/s), q_{um} is a unit peak discharge ($ft^3/sec/mi^2/in.$), A_m is the watershed area for the pluvial runoff, estimated to be 0.75 mi^2 in ArcGIS, and Q is the runoff depth (in.). The unit peak discharge, q_{um}, was identified from a graph based on the time of concentration (t_c) and the ratio of initial abstraction to precipitation depth (I_a/P). To determine the ratio of I_a/P, I_a must be calculated as 0.2*S. In order to calculate the time of concentration for the watershed, the following equation was used:

$$t_c = 0.00526 L^{0.8} \left(\frac{1000}{CN} - 9\right)^{0.7} S^{-0.5} \tag{7-12}$$

where L is the watershed length (ft), CN is the curve number, and S is the watershed slope (ft/ft). Using ArcGIS the watershed length was estimated to be 4,232 ft and the watershed slope was estimated to be 0.004 ft/ft. This resulted in a time of concentration of 3.26 hours. Using Equations 7-10 through 7-15, the maximum discharge caused by rainfall events was estimated for the watershed.

Based on the calculated peak discharges, pluvial flood depths could be calculated. It was assumed that the surface runoff ultimately drained through a roadway near the location of interest. This roadway was estimated to be 20 feet wide, with a curb height of 0.75 feet. There was also a space 10 feet wide between the edge of the road and the nearest building. The geometry of this street is illustrated in Figure 7-3. The pluvial discharge was conveyed through the roadway and the sidewalk-area between the road and the buildings. Using Manning's equation, the flow depth corresponding to each measured discharge was computed. The area and

wetted perimeter were calculated based on the geometry of the roadway, and Manning's coefficient of roughness was assumed to be 0.012. This provided the pluvial flood depth values, which were evaluated to obtain the annual maximum flood depth series. This information provided information on the observed population, from which pluvial flood depth samples were then generated.

Figure 7-3: Geometry of Street Through Which Pluvial Flow is Conveyed

7.3.1.1.4. Calculation of Flood Frequency Curves for Each Source Individually

To calculate a flood frequency curve for a single flood source requires identification of the annual maximum flood series. The annual maximum events were ranked from largest to smallest, with the largest magnitude assigned to a rank of 1 and the smallest magnitude a rank of n. After developing a ranked list of annual maximum flood inundation depths at the location of interest caused by an individual flood source, the Weibull probability equation was used to determine the exceedance probability of each event. Weibull's plotting position formula is:

$$P_i = \frac{i}{n+1} \tag{7-13}$$

where P_i is the exceedance probability, i is the rank of the given flood depth, and n is the number of sample values in the record. The flood depths were plotted against the

non-exceedance probability, calculated as one minus the exceedance probability, which provides the probability of the flood depth not being equaled or exceeded in a given year. Based on the pairs of data (non-exceedance probability, flood depth), a cumulative distribution function was fitted using either Maximum Likelihood Estimation or Method of Moments that would provide non-exceedance probabilities for any flood inundation depth value.

Though Bulletin 17B recommends the Log-Pearson Type III distribution be used to model riverine flood flows, the GEV distribution is frequently used to model tidal flood heights, and the LP3 or gamma distribution are frequently used to model rainfall events, clear evidence that any one of these distributions provides greater accuracy than that provided by others does not exist. Further, the variable of interest in this research is flood depth at the location of interest, rather than discharge or tidal height, as would be typical in a riverine, tidal, or pluvial analysis. The watershed processes are known to modify the distributions of the events, so it is possible that the typical distributions may not be the most appropriate distributions. Thus, several distribution functions were fitted to the data to determine the most appropriate distribution for each data set. All tested distributions were plotted against the observed data for visual comparison, and probability plot correlation coefficients were calculated to assess the goodness-of-fit of each distribution. For simulated riverine and pluvial data, the distributions tested included the log-Pearson type III, lognormal, gamma, and Weibull distributions. For simulated tidal data, the distributions tested included the extreme value, generalized extreme value, and Rayleigh distributions. This procedure was first conducted with the observed data

from Florida, in order to determine appropriate populations from which to generate samples for the simulation study. It was also used in the simulation studies to calculate the distributions best fitting the generated samples.

7.3.1.1.5. Calculation of Flood Frequency Curves Considering Three Flood Sources

To develop a joint flood frequency curve requires the marginal distributions or individual flood frequency curves that correspond to the three individual flood sources. The typical process for conducting a flood frequency analysis, outlined in section 7.3.1.1.4., was followed to determine the marginals, with one exception. When three potentially interacting flood sources must be considered, the maximum flood event experienced at the location of interest in a given year may not be caused by one individual source, but instead by a combination of the three sources. Thus, to identify the annual maximum flood depth at the location of interest, the riverine, tidal, and pluvial flood inundation depths at the location of interest were summed for corresponding time periods to identify the maximum total depth of flooding at the location of interest in each year. The assumption that the contributing flood depths from each source could be summed was a convenient assumption to make for the purpose of conducting this research; however, this methodology was not determined based on a technical analysis of the optimum approach for combining flood sources.

It was necessary to identify the annual maximum events while considering all potential flood sources. As each of the gages from which observed data were obtained took hourly measurements each day, the hourly measurements from each gage were summed and a single daily maximum flood depth at the location of interest was identified. Next, the annual maximum flood depth was identified based on the

daily maximum events. This does not suggest that the annual maximum event could only occur through a combination of the three sources; this procedure could still identify an annual maximum event that was caused by only one source or by a combination of two of the sources. The riverine contributions to the annual maximum flood inundation depth, the tidal contributions, and the pluvial contributions were separated into three flood series and flood frequency curves were developed from each. These flood frequency curves served as the marginal distributions used as input to the copula process.

7.3.1.2. Description of Results

In order to develop a comprehensive flood frequency assessment, the first step was to fit marginal distributions to the observed riverine, tidal, and pluvial flood inundation depth samples. Tables 7-2 through 7-4 provides the fitted marginal parameters and the calculated probability plot correlation coefficients for each potential marginal distribution for the riverine, tidal, and pluvial flood depths, respectively. In Figures 7-4 through 7-6, the possible marginal distributions are compared to the observed riverine, tidal, and pluvial flood depth samples, respectively. Based on the probability plot correlation coefficients and Figure 7-4, it appears that the gamma distribution would be the distribution best able to represent the riverine data. Due to the large number of zero-flood years in the observed data, it was difficult to fit any of the distributions and none were as accurate as would be desired. The small sample size also likely contributed to the difficulty in fitting the distributions to the observed riverine sample, as the simulation studies demonstrated

that distributions could more accurately represent large sample sizes than small sample sizes.

For the tidal data, the GEV distribution results in the highest probability plot correlation coefficient and Figure 7-5 shows that the GEV distribution most closely agrees with the observed tidal flood depths. Either of the other distributions would also provide acceptable fits to the observed data, though the GEV distribution would be superior. Thus, the GEV distribution was chosen to represent the tidal flood source. The observed tidal flood depth sample was fitted more accurately than the riverine sample because the tidal flood depth sample contained fewer zero-flood years than the riverine flood depth sample. However, the fit of the distributions to the observed tidal sample was still less accurate than desired, due to the small sample size. The simulation studies demonstrated that distributions could more accurately represent large samples than small samples.

For the pluvial observed data, the probability plot correlation coefficients and Figure 7-6 suggests that the gamma distribution would be the most appropriate distribution to represent the observed data. Thus, the gamma distribution was selected for use in representing the pluvial flood depths, though none of the distributions considered fit the observed data as well as desired due to the large number of zero-flood years. In addition to the zero-flood years, another factor influencing the accuracy of the fitted distributions was the sample size. The simulation studies demonstrated that smaller samples could not be as accurately represented by the potential marginal distributions as larger samples.

Table 7-2: Fitted Parameters and Probability Plot Correlation Coefficients for the Observed Dependent Riverine Sample

	LP3	Lognormal	Gamma	Weibull
Shape	20.6015	-1.3118	0.4078	0.5271
Scale	0.4603	2.0893	3.3007	0.7652
Location	-10.7949	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.797	0.8402	0.9464	0.915

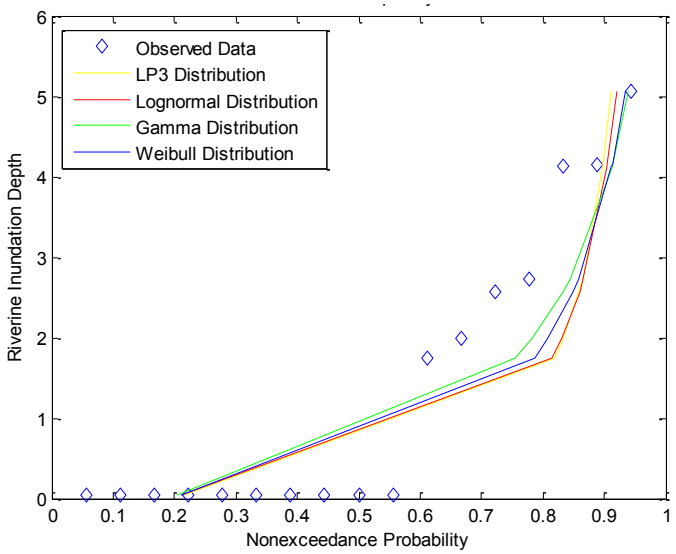

Figure 7-4: Comparison of Riverine Marginal Distribution Options for Observed Data, Where Flood Depths are in Feet

Table 7-3: Fitted Parameters and Probability Plot Correlation Coefficients for the Observed Dependent Tidal Sample

	Extreme Value	GEV	Rayleigh
Shape	N/A	-0.2224	N/A
Scale	1.1498	1.1038	1.3356
Location	2.0773	1.0476	N/A
Probability Plot Correlation Coefficient	0.9328	0.9648	0.9447

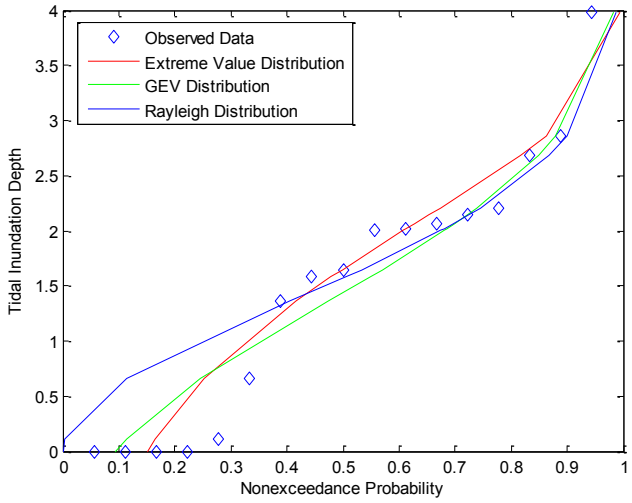

Figure 7-5: Comparison of Tidal Marginal Distribution Options for Observed Data, Where Flood Depths are in Feet

Table 7-4: Fitted Parameters and Probability Plot Correlation Coefficients for the Observed Dependent Pluvial Sample

	LP3	Lognormal	Gamma	Weibull
Shape	12.7906	-2.8569	0.3716	0.4949
Scale	0.5992	2.1429	0.9278	0.1695
Location	-10.5207	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.8148	0.8682	0.9598	0.932

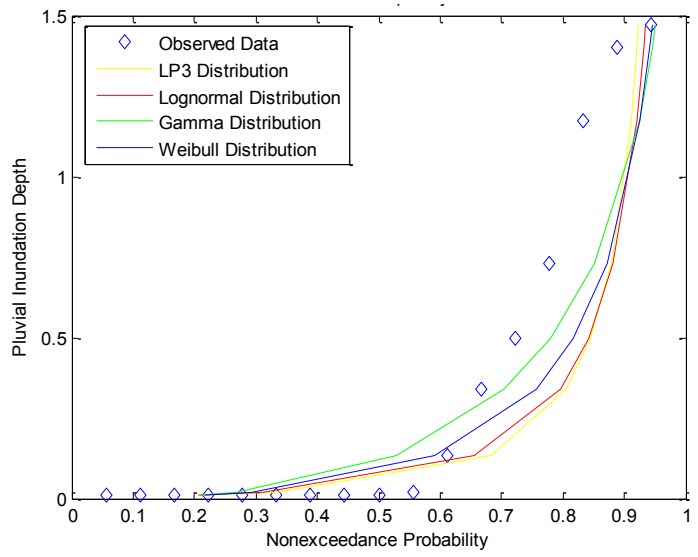

Figure 7-6: Comparison of Pluvial Marginal Distribution Options for Observed Data, Where Flood Depths are in Feet

7.3.2. Using Copula Equations to Determine Joint Distributions

Joint distributions were developed for flood inundation depths from the riverine, tidal, and pluvial sources. The joint distributions provided probabilities of specific combinations of riverine, tidal, and pluvial flood depths. The next sections will discuss the methods used to develop the joint distributions and present the results.

7.3.2.1. Description of Methods

Once the marginals had been developed based on the observed riverine, tidal, and pluvial flood inundation depths at the location of interest, a copula could be used to determine the joint probability of depths for the three flood sources occurring. The first step was to calculate the appropriate dependence structure between the three flood sources. As with the scenario in which two flood sources interacted, the dependence structure was based on the ranks of the data; however, for three flood sources, Kendall's τ was not appropriate. The method used involved maximum pseudo-likelihood estimation, which is an alternative to the method of moments (Genest and Favre, 2007). This method determines the value of the parameter, α, based exclusively on the ranks of the riverine, tidal, and pluvial sample data, R, S, and T, respectively.

$$l(\alpha) = \frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} \frac{\dot{c_\theta}(\frac{R_i}{n+1}, \frac{S_i}{n+1}, \frac{T_i}{n+1})}{c_\theta(\frac{R_i}{n+1}, \frac{S_i}{n+1}, \frac{T_i}{n+1})} = 0 \qquad (7\text{-}14)$$

where c(α) is the copula pdf, $\dot{c}(\alpha)$ is the derivative of the copula pdf, n is the sample size, and R_i, S_i, and T_i are the ranks of the riverine, tidal, and pluvial sample values,

respectively. This equation was solved based on the observed samples of riverine, tidal, and pluvial flood depths in order to determine the copula parameter value.

Once the measure of dependence was calculated, the appropriate family of copulas to fit with the data had to be determined. Numerous families of copulas are available from which to choose. Several families were fitted and compared using Akaike's Information Criteria (AIC). The specific method of using the AIC was described in Chapter 1. When AIC values are calculated, the smallest AIC value indicates the most appropriate copula family (Klein *et al.*, 2010). Three Archimedean copula families, the Gumbel-Hougaard, Clayton, and Frank families, were evaluated. The equations for these three families, respectively, are as follows:

$$C(u,v,w) = \exp(-[(-\ln u)^{\alpha} + (-\ln v)^{\alpha} + (-\ln w)^{\alpha}]^{\frac{1}{\alpha}}) \quad (7\text{-}15a)$$

$$C(u,v,w) = (u^{-\alpha} + v^{-\alpha} + w^{-\alpha} - 2)^{-1/\alpha} \quad (7\text{-}15b)$$

$$C(u,v,w) = \frac{-1}{\alpha}(\log\left(1 + \frac{(e^{-\alpha*u}-1)(e^{-\alpha*v}-1)(e^{-\alpha*w}-1)}{(e^{-\alpha}-1)^2}\right)) \quad (7\text{-}15c)$$

where u, v, and w were the cumulative distributions fitted to the riverine, tidal, and pluvial samples, respectively, and α was the copula parameter calculated for each family.

Finally, the joint pdfs were calculated for each copula family. Joint pdfs that correspond to the variables u, v, and w were calculated by taking the second derivative of the joint cdf with respect to u, v, and w. To obtain joint pdf values that correspond to the riverine, tidal, and pluvial variables required that the joint pdfs calculated for the variables u, v, and w then be multiplied by the marginal distribution

pdf values (Wang *et al.*, 2009). The equations for the joint pdfs, expressed in terms of the riverine, tidal, and pluvial variables, are as follows:

$$f(x,y,z) = \frac{\partial^3}{\partial u \partial v \partial w} C_\alpha(u,v,w) \frac{\partial u}{\partial x}\frac{\partial v}{\partial y}\frac{\partial w}{\partial z} \quad (7\text{-}16)$$

where u, v, and w were the marginal cdf variables that correspond to the riverine, tidal, and pluvial flood depth samples, respectively, and x, y, and z represent the riverine, tidal, and pluvial flood depth variables, respectively.

7.3.2.2. Description of Results

As was done in the simulation scenarios, the first step to determine the joint distributions was to calculate the parameters for each copula family and identify the most appropriate copula family for further use. Based on the observed data, the copula parameters calculated were 1.0232, 0.0738, and 0.1945 for the Gumbel-Hougaard, Clayton, and Frank families, respectively. These parameters met the requirements of trivariate copulas, in that the parameter for the Gumbel-Hougaard family is above 1.0 and the parameters for the Clayton and Frank families are above 0.0. These parameters tend to be a bit lower than the parameters calculated in the simulation studies evaluating various samples sizes. Because the copula parameters are related to the ranks of the data in the riverine, tidal, and pluvial flood depth samples, it was not expected that the sample size would greatly influence the parameters. The copula parameters calculated for the observed data were similar overall to the parameters calculated for the varying correlation scenarios in which the correlations between samples tended to be low. In particular, the copula parameters were most similar to those calculated for scenarios RLA and RLHL. These scenarios

were the most similar to the observed data. The level of correlation between the flood source samples was expected to impact the copula parameters because the copula parameters were based on the relative ranks of the three data sets. The level of correlation would be expected to impact these relative ranks.

Based on the calculated copula parameters, the AIC values were then calculated for each of the three copula families in order to draw conclusions about the most appropriate copula family for further use. The AIC values calculated were 35.9917, 17.3744, and 17.3369, for the Gumbel-Hougaard, Clayton, and Frank copula families, respectively. These values are quite a bit higher than the AIC values calculated in the simulation studies. This suggests that it was more difficult to fit the copulas to the observed data than the simulation study data, which would be expected because the observed data contained numerous zero-flood years that the observed data did not. The most appropriate copula family appears to be the Frank family. Though the AIC value suggests that the Frank family is the best family to represent the observed data, all three copula families will be used to calculate copula cdfs and pdfs to determine the impact of the choice of copula family.

Table 7-5 provides the joint pdfs and cdfs calculated using both the Gumbel-Hougaard, Clayton, and Frank copula families for the observed flood inundation depth data. The joint pdfs are not valuable in themselves, but were necessary for further use in determining the non-exceedance probabilities corresponding to the total flood depths. The joint distributions, both the pdfs and the cdfs, were very similar for the three families. This suggests that the choice of copula family would not have a significant impact on the development of the joint distributions. The representation

of the dependence between the flood depths is the primary difference between the copula families. This suggests that the level of dependence between the three sets of flood depths was such all three copula families could represent the data equally.

It should be noted that the maximum cumulative probability identified for the sample in Table 7-5 was only 0.58. However, if the largest riverine, tidal, and pluvial flood depths had occurred jointly, the maximum joint cumulative probability would have been 0.88. This suggests that the joint distribution was a complete cdf; however, the small sample did not fully represent the joint population. The copula parameter was calculated based on the ranks of the observed riverine, tidal, and pluvial flood depths, and for each sample, all of the zero-flood years received equal ranks. This may have skewed the resulting copula parameter, such that the observed samples were not representative of the joint distribution. The maximum joint cdf value for the observed sample based on two flood sources was also lower than expected for a cdf; however, in that case the distribution could be plotted, which allowed the full distribution to be visualized. Since the distribution cannot be visualized for three flood sources, it is more difficult to understand the full population of the joint cdf. The low maximum joint cumulative probability may also have been influenced by both the sample size and the level of correlation between the samples, based on observations from the simulation studies. The maximum cdf value for the samples of size 10 and 25 were also observed to be lower than expected for a cdf, but for larger sample sizes the maximum joint cumulative probability calculated for the sample was observed to increase. Further, lower joint cumulative probabilities for the samples were observed for simulated samples with lower levels of correlation

between the samples than for samples with high levels of correlation. The observed data had low correlation between the riverine and tidal and tidal and pluvial samples; thus, lower joint cdf values could be expected.

Table 7-5: Joint PDFs and CDFs Calculated for the Observed Flood Depth Data

Riverine Flood Depth (ft)	Tidal Flood Depth (ft)	Pluvial Flood Depth (ft)	Gumbel-Hougaard pdf	Gumbel-Hougaard cdf	Clayton pdf	Clayton cdf	Frank pdf	Frank cdf
2.5685	0.0000	0.3400	0.0069	0.0573	0.0065	0.0602	0.0067	0.0574
5.0699	0.0000	1.1750	0.0004	0.0827	0.0004	0.0834	0.0004	0.0825
0.0500	2.2087	0.0100	3.3338	0.0337	3.2328	0.0392	3.3042	0.0341
0.0500	1.5846	0.0100	4.3147	0.0254	4.1998	0.0309	4.3110	0.0258
0.0500	2.8648	0.1350	0.3365	0.0985	0.3362	0.1032	0.3317	0.0990
4.1238	0.1125	0.7310	0.0014	0.0904	0.0014	0.0919	0.0014	0.0903
0.0500	1.6486	0.0100	4.2495	0.0263	4.1318	0.0318	4.2407	0.0267
0.0500	2.6888	0.0100	2.3414	0.0383	2.2844	0.0438	2.3232	0.0387
0.0500	2.0007	0.0100	3.7299	0.0312	3.6156	0.0367	3.7029	0.0316
1.9846	0.0000	0.0200	0.0837	0.0212	0.0797	0.0255	0.0830	0.0214
1.7504	3.9849	0.4970	0.0020	0.5846	0.0021	0.5821	0.0021	0.5823
0.0500	2.1447	0.0100	3.4602	0.0329	3.3545	0.0385	3.4309	0.0333
0.0500	2.0167	0.0100	3.7012	0.0314	3.5877	0.0369	3.6738	0.0318
2.7329	0.0000	0.0100	0.0856	0.0178	0.0822	0.0218	0.0853	0.0179
4.1497	0.6565	1.4000	0.0007	0.2143	0.0008	0.2147	0.0007	0.2137
0.0500	2.0647	1.4730	0.0302	0.1388	0.0312	0.1416	0.0302	0.1390
0.0500	1.3606	0.0100	4.4531	0.0220	4.3570	0.0274	4.4700	0.0224

7.3.3. Calculation of Combined Flood Frequency Curve

The final step of the procedure was to use the joint probability distribution developed for the observed data to determine the non-exceedance probabilities corresponding to total flood depths. This provided information about the flood hazard when three dependent sources were considered to impact the location. The following sections will explain the methods used to develop the final flood frequency assessment for total flood depths.

7.3.3.1. Description of Methods

The results of the copula procedure described in section 7.3.2. were joint pdfs and cdfs, which provide the probability that corresponded to a specific joint event. However, the primary interest was the probability of a certain flood depth occurring, when that flood depth might be due to one of the sources individually, or some combination of the three sources, such that many possible combinations could result in the desired flood depth. Because the riverine, tidal, and pluvial contributions to flood inundation at the location of interest were summed in the process of identifying annual maximum events, flood depths were also summed to identify probabilities that corresponded to total flood depths. Within the joint probability distribution, the region that corresponded to a specific total flood depth was identified as bounded by the riverine inundation depths, tidal inundation depths, and pluvial inundation depths axes and a triangular plane connecting the points (0, tidal depth of interest, 0), (riverine depth of interest, 0, 0), and (0, 0, pluvial depth of interest). In other words, the points where the flood inundation depth of interest was caused only by one of the sources, riverine, tidal, or pluvial, were identified and connected by a plane. All of the points on this plane summed to a total depth equal to the depth of interest to be used in flood risk assessments.

The mass under the joint probability distribution within this bounding region represents the non-exceedance probability for a given flood depth. This was calculated using a triple integral. Within the triple integral, the lower and upper bounds of the outer integral were 0 and total depth of interest, respectively. The lower and upper bounds of the middle integral were 0 and:

$$Upper\ Bound = Total\ Depth\ of\ Interest - Riverine\ Contribution\ to\ Total\ Depth \quad (7\text{-}17)$$

The lower and upper bounds of the inner integral were 0 and:

$$Upper\ Bound = Total\ Depth\ of\ Interest - River\ Contribution\ to\ Total\ Detph - Tidal\ Contribution\ to\ Total\ Depth \quad (7\text{-}18)$$

As an example, if the riverine flood depth was 3.5 feet, the corresponding tidal flood depth was 3 feet, and the pluvial flood depth was 2 feet, the total flood depth of interest would be 8.5 feet. In this scenario, the outer integral lower and upper bounds would be 0 feet and 8.5 feet, respectively, the middle integral bounds would be 0 and (8.5 feet – 3.5 feet), or 5 feet, and the inner integral bounds would be 0 feet and (8.5 feet – 3.5 feet – 3 feet), or 2 feet. The triple integral was calculated in Matlab using the adaptive Simpson's method, a numerical evaluation method, which uses a recursive algorithm to approximate the integral based on the error in estimates calculated using Simpson's rule. The non-exceedance probability of each value of total flood depth was calculated using this triple integral procedure, and then the probabilities were plotted against the total depth values.

The final task to determining the probability of total flood depths was to identify the appropriate distribution to represent the population. Four distributions were fitted to the total flood depths: the LP3, the GEV, the gamma, and the normal distributions. The GEV and gamma distributions were chosen for consideration because they were the two marginal distributions used to represent the simulated riverine, pluvial and tidal data. The LP3 distribution was selected for consideration

because it is commonly used to represent flood data. The normal distribution was selected for consideration because the purpose of the copula is to develop a joint distribution when the marginals are represented by different distributions; thus, it was believed that the results of the copula might not follow either of the marginal distributions.

These four distributions were fitted to the total flood depth samples using either Maximum Likelihood Estimation or Method of Moments and they were compared to the probabilities calculated for each copula family based on the triple integral procedure. The fitted distributions were plotted against the total flood depths and compared to a plot of the non-exceedance probabilities calculated by triple integral versus the total flood depths. These plots, as well as probability plot correlation coefficients, were used to determine which of the four distributions could best represent the total flood depth populations for each simulated scenario. The result of this process was a population from which non-exceedance probabilities corresponding to total flood depths for the location of interest could be identified. From this distribution, the true nature of the flood hazard for the location of interest could be understood, which was the ultimate goal of this research.

7.3.3.2. Description of Results

Using the joint distributions developed based on the observed riverine, tidal, and pluvial flood depth data, non-exceedance probabilities were calculated that corresponded to total flood depths. The LP3, GEV, gamma, and normal distributions were fitted and compared to the non-exceedance probabilities calculated for total flood depths. Table 7-8 provides the parameters calculated for each distribution for

the observed total flood depths and the probability plot correlation coefficients calculated for each distribution. The probability plot correlation coefficients indicate that the LP3, GEV, and gamma distributions would all fit the observed total flood depths well, and should perform approximately equally well. As was observed in the simulation studies, the normal distribution does not appear to be able to fit the total flood depths as well as the other three distributions, based on the probability plot correlation coefficients.

Figure 7-7: Comparison of LP3 and Normal Distributions Fitted to the Observed Total Flood Depths

	LP3	GEV	Gamma	Normal
Shape	12.1435	0.459	4.4992	N/A
Scale	0.138	0.8221	0.707	1.6761
Location	-0.6333	2.2406	N/A	3.181
Probability Plot Correlation Coefficient	0.9985	0.9997	0.9916	0.9779

Based on the probability plot correlation coefficients, it would appear that either the LP3, GEV, or gamma distributions could be selected to represent the non-exceedance probabilities corresponding to the observed total flood depths. Figure 7-9 compares the fitted distributions to the observed total flood depths and the non-exceedance probabilities calculated using the triple integral procedure for the Frank copula equation, which was identified as the most appropriate copula to represent the observed data. This figure suggests that the LP3, GEV, and gamma distributions perform approximately equally. However, none of the four distributions fit the upper tail of the observed data, based on the non-exceedance probabilities calculated using the triple integral. Based on probability plot correlation coefficients and Figure 7-9,

the gamma distribution was selected to represent the total flood depths. This distribution appeared to better match the curvature of the observed data in Figure 7-9.

In the simulation studies discussed in Chapter 6 the distributions were observed to provide excellent fits overall to the total flood depths. However, that none of the four distributions was found to fit the observed total flood depths was also observed when only two flood sources were considered. For two flood sources, the poor fit to the upper tail of the observed data was determined to be primarily related to the large number of zero-flood years in the flood depth samples. Because there were numerous zero-flood years in the observed data for three flood sources as well, it was expected that this remained a primary cause of the poor fit between the non-exceedance probabilities calculated using the triple integral and the distributions fitted to the total flood depths. When the poor fit in the upper tail was investigated for two flood sources, it was also determined that the lack of a strong relationship between the two flood sources was also an important factor. A strong relationship was also not evident between the three flood sources. To determine the potential impact of this factor on the non-exceedance probabilities calculated for the total flood depths influenced by three flood sources the largest riverine, tidal, and pluvial flood depths were assumed to occur jointly. The non-exceedance probability calculated for this combination of flood depths was 0.89, which was greatly improved over the maximum non-exceedance probability calculated by the triple integral in Figure 7-8. This suggests that, if a stronger relationship between three flood sources existed, the triple integral procedure would result in larger non-exceedance probabilities as would be expected for a cumulative distribution.

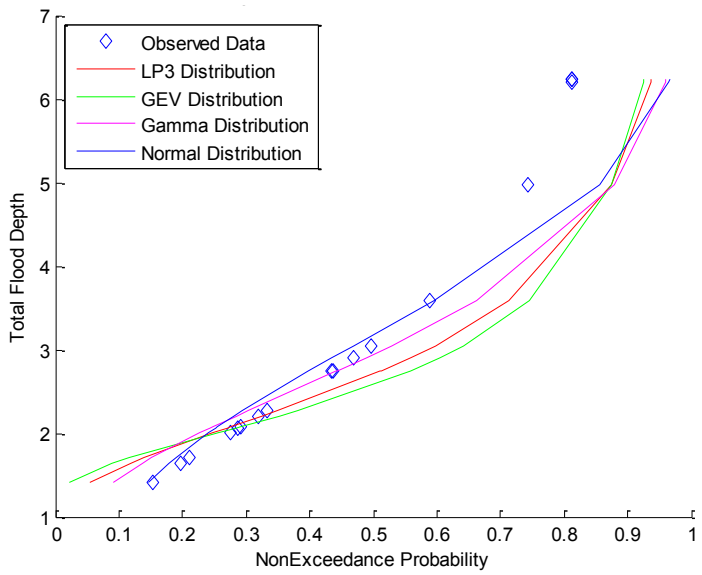

Figure 7-8: Comparison of LP3 and Normal Distributions to Non-Exceedance Probabilities Calculated for Total Flood Depths for the Observed Data

The impact of the choice of distributions to represent the total flood depths on the predicted flood depths for several common exceedance probabilities was also investigated. Flood depths were calculated for using each of the four distributions for exceedance probabilities of 0.1, 0.02, and 0.01. For the exceedance probability of 0.1 there was very little difference in the predicted flood depths, which agreed with the results presented in Figure 7-9. For the larger events, exceedance probabilities of 0.02 and 0.01, the choice of distribution would have more influence on the predicted flood depths. For an exceedance probability of 0.01, the predicted flood depths could vary from approximately 7 feet to 15.2 feet, which was likely not an insignificant difference. Thus, though the four distributions appear to perform similarly over the range of the observed data, giving the impression that the choice of distribution is not terribly important, the prediction of flood depths for larger flood events suggests that this decision could greatly influence the understanding of the flood hazard.

Table 7-6: Flood Depths (in Feet) Calculated For Common Exceedance Probabilities Using Each Distribution

Exceedance Probability	LP3	GEV	Gamma	Normal
0.1	5.3695	5.4811	5.1899	5.329
0.02	8.7676	11.1877	6.9557	6.6233
0.01	10.578	15.2444	7.658	7.0802

7.3.4. Comparison of Flood Hazard Based on Different Assumptions

The results presented previously have been derived based on the assumptions that all three flood sources must be considered together and that the three flood sources are not independent of each other. These results can be compared to flood frequency curves obtained using each of the three sources individually in order to draw conclusions about the importance of considering all three flood sources together. These results can also be compared to flood frequency curves computed based on all three flood sources but with an assumption that the flood sources are independent of each other.

Annual maximum flood depths for each independent flood source were identified without consideration of the other flood sources (independently of each other) and flood frequency assessments were made. Based on the identified annual maximum flood depths, distributions were fitted and used to provide an understanding of the likelihood of flooding from each source. The marginal distribution parameters fitted to the riverine, tidal, and pluvial flood depth data are presented in Table 7-7 through 7-9, respectively, along with the probability plot correlation coefficients computed to determine the most appropriate distribution to use to represent each flood source. Figures 7-10 through 7-12 compare each possible probability distribution to the observed independent flood sources. Based on the

agreement of the fitted distributions to the observed flood depths in these figures and the probability plot correlation coefficients, the gamma distribution was chosen to represent the independent riverine and pluvial flood depths and the GEV distribution was chosen to represent the independent tidal data. As was observed for the dependent flood sources, numerous zero-flood years in the riverine and pluvial samples made it difficult to fit distributions to these flood depth samples.

Table 7-7: Marginal Distribution Parameters and Calculated Probability Plot Correlation Coefficients for Independent Observed Riverine Flood Sources

	LP3	Lognormal	Gamma	Weibull
Shape	113.6015	-0.7384	0.5415	0.6584
Scale	-0.177	1.886	2.812	1.1643
Location	19.3636	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.8864	0.867	0.9697	0.9546

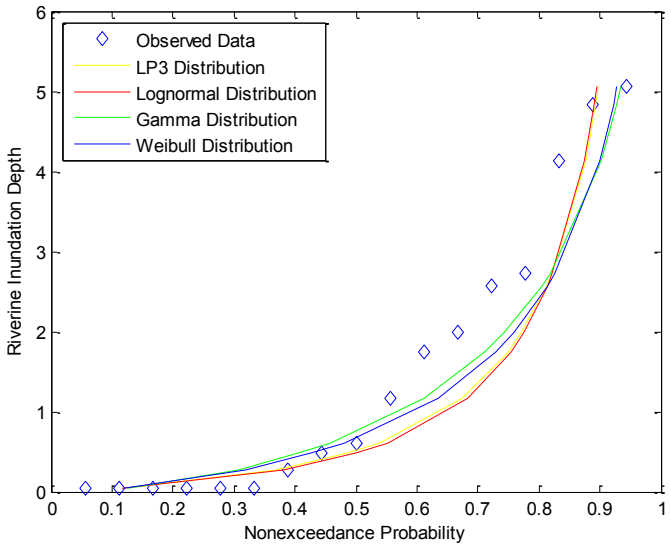

Figure 7-9: Marginal Distributions Fitted to the Independent Observed Riverine Flood Depths, where the Flood Depths were in Feet

Table 7-8: Marginal Distribution Parameters and Calculated Probability Plot Correlation Coefficients for Independent Observed Tidal Flood Sources

	Extreme Value	GEV	Rayleigh
Shape	N/A	0.0824	N/A
Scale	2.4925	0.3941	1.5825
Location	0.7784	1.8982	N/A
Probability Plot Correlation Coefficient	0.8397	0.9533	0.9115

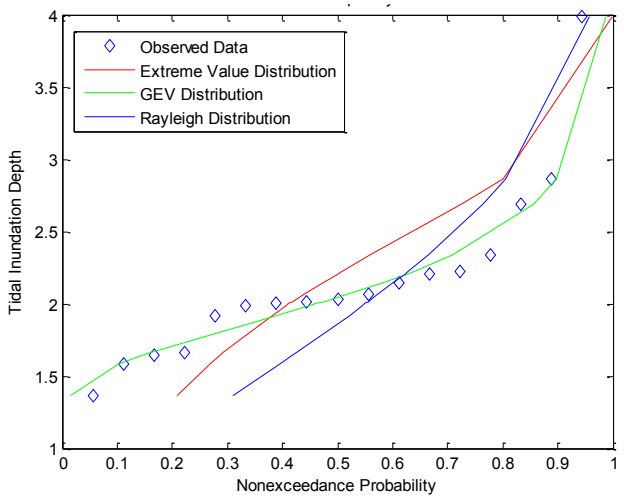

Figure 7-10: Marginal Distributions Fitted to the Independent Observed Tidal Flood Depths, where the Flood Depths were in Feet

Table 7-9: Marginal Distribution Parameters and Calculated Probability Plot Correlation Coefficients for Independent Observed Pluvial Flood Sources

	LP3	Lognormal	Gamma	Weibull
Shape	7.0022	-3.0935	0.3698	0.4953
Scale	0.7779	2.0584	0.7435	0.1303
Location	-8.5404	N/A	N/A	N/A
Probability Plot Correlation Coefficient	0.8973	0.9449	0.9837	0.976

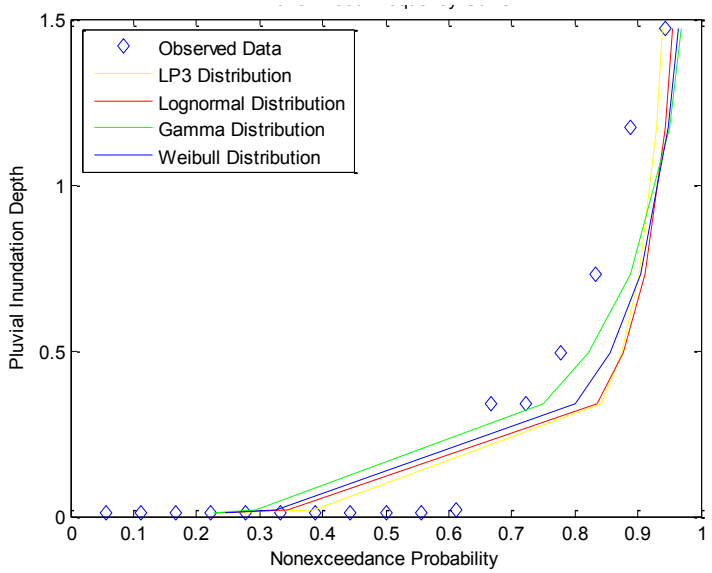

Figure 7-11: Marginal Distributions Fitted to the Independent Observed Pluvial Flood Depths, where the Flood Depths were in Feet

Based on comparison of Figures 7-10 through 7-12, it appears that consideration of only one of the potential flood sources would result in quite different understandings of the likelihood of flooding at the location of interest. When the riverine flood source was considered independently, the maximum flood depth did not change significantly, but the curvature of the frequency curve did vary, which would impact the predicted flood depths beyond the range of the data. The same held true for the tidal and pluvial flood depth samples. When all three flood sources were considered independently there were also fewer zero-flood years in each flood depth sample. Consideration of the three flood sources dependently resulted in more frequent occurrences of zero-flood years because the three flood sources do not always occur simultaneously and in some years the annual maximum flood event is caused by only one or two flood sources. These differences in expected flood probabilities based on which flood source was considered would certainly impact precautions taken to protect against flood events.

It is also important to compare the flood frequency curves to assess the importance of the assumption of independence. The marginal distribution parameter values in Tables 7-7, 7-8, and 7-9 can also be compared to those presented in Tables 7-2, 7-3, and 7-4, which were developed based on the understanding that the three flood sources were dependent. When the annual maximum flood depths caused by each flood source were determined independently of each other, the riverine gamma shape parameter increased and the scale parameter decreased. When the annual maximum flood depths were calculated independently the GEV distribution shape and location parameters increased, while the scale parameter decreased. When the annual maximum flood depths were calculated independently, the gamma distribution shape parameter increased and the scale parameter decreased. This suggests that consideration of the flood sources independently or dependently would influence the annual maximum events chosen, and therefore also the marginal distributions fitted to the annual maximum flood depth data.

It is also possible to calculate the joint probability of a given flood depth occurring while making the assumption that the flood sources are independent of each other. For three flood sources, the equation to do so, expanded from Equation 1-3, would be:

$$P_{joint} = P_{river} + P_{tide} + P_{pluvial} - P_{river} * P_{tide} - P_{river} * P_{pluvial} - P_{tide} * P_{pluvial} + P_{river} * P_{tide} * P_{pluvial} \quad (7\text{-}19)$$

where P indicates the exceedance probability corresponding to a given flood depth and the subscript indicates the flood source. Figure 7-12 provides the exceedance

probabilities corresponding to total flood depths for the independent riverine, tidal, and pluvial scenarios, the joint distribution calculated based on the assumption of independence, and the joint distribution based on the assumption of dependence.

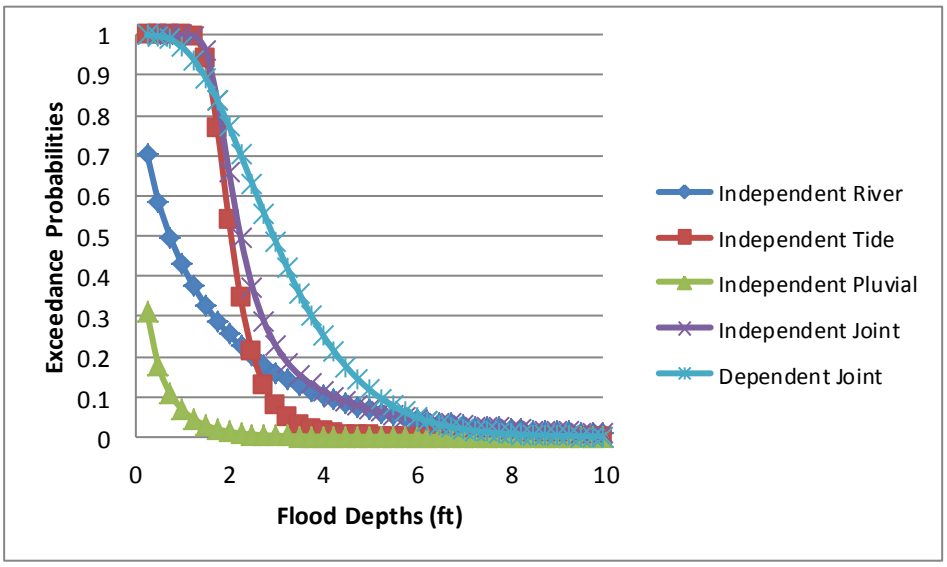

Figure 7-12: Exceedance Probabilities Calculated for the Independent Riverine, Tidal, and Pluvial Flood Sources and the Independent and Dependent Joint Distributions

From Figure 7-12 it was clear that the consideration of only one of the flood sources independently would result in a very different understanding of the flood hazard. The exceedance probabilities calculated for the independent tidal flood source decreased much more steeply as flood depth increased than either the riverine or pluvial flood sources. The riverine and pluvial flood sources both had more zero-flood years than the tidal flood source; thus, the highest exceedance probabilities for these flood sources occurred at very low flood depths. Because of the zero-flood years, the riverine and pluvial flood samples had much lower exceedance probabilities for flood depths below approximately 3 feet than did the tidal flood source. The riverine flood source then had the highest exceedance probabilities for

flood depths greater than 3 feet, which was expected given that the riverine depth samples had greater maximum flood depths than did the tidal or pluvial samples.

The importance of the assumption of independence or dependence between the flood sources was assessed based on Figure 7-12. The joint distributions result in fairly similar exceedance probabilities overall. Both differ from the three independent flood sources, as would be expected given that they are based on all three of the flood sources. It can be expected that consideration of all three flood sources, either independently or dependently will result in a more complete understanding of the flood hazard for the location of interest. Based on Figure 7-12, it does appear that the assumption of dependence or independence could alter the expected probability of a given flood depth, but only for flood depths in the mid-depth ranges. The most significant differences would occur for the mid-range flood depths, where the most difference exists in the exceedance probabilities calculated for each flood source. Through these flood depths, the exceedance probabilities calculated using the assumption of dependence decrease more gradually, leading to a higher expectation of these flood depths occurring when the three flood depths are considered to be dependent.

Less significant difference between the joint dependent and joint independent exceedance probabilities were observed for two flood sources, where it was conjectured that the lack of a strong relationship between the flood sources resulted in the minimal difference in exceedance probabilities calculated based on the assumptions of dependence and independence. For three flood sources, a strong relationship did not exist between the riverine and tidal or the tidal and pluvial flood

depth samples, but a strong relationship was evident between the riverine and pluvial samples. This relationship may explain why the assumption of independence or dependence between the flood sources seems to have more impact when three flood sources were considered than when two flood sources were considered.

7.3.5. Conclusions

A procedure was successfully developed to calculate the probability of flooding when three flood sources could potentially impact a location of interest. The procedure required identifying the annual maximum flood events caused by each flood source based on first identifying the annual maximum flood depths caused by all three flood sources. Marginal distributions were then fitted to these annual maximum event data sets. The gamma distribution was selected to represent both the riverine and the pluvial flood depths and the GEV distribution was selected to represent the tidal flood depths.

Once the marginal distributions were identified, joint distributions were determined using copula families. The copula parameter was calculated and the copula family most appropriate to represent the observed data was chosen. The Frank family was determined to be the most appropriate copula family to represent the observed data, as was the case in several simulation studies. However, minimal difference was observed between the joint distributions calculated based on all three of the copula families, as was also observed in the simulation studies. The joint probability distributions were then used to determine the probability of total flood depths by taking the triple integral under the joint distribution. These non-exceedance probabilities were then compared to cdfs calculated by fitting four

distributions to the total flood depths. The gamma distribution was selected as the best distribution to represent the total flood depths, as was observed in the simulation studies, based on comparison to the results of the triple integral procedure. These probabilities were used to represent the flood hazard for the location of interest based on the three potential flood sources.

The importance of the assumption of independence or dependence between flood sources was also assessed for three flood sources. Differences were observed between the calculated probabilities that corresponded to riverine, tidal, or pluvial flood sources individually and between the joint probabilities when considering the three flood sources as dependent or independent. This indicates that it is very important to consider all possible flood sources in order to obtain a full understanding of the likelihood of flooding at a location of interest. Further, the two assumptions were observed to result in different probability estimates for mid-range flood depths. This suggests that the assumption of independence or dependence has some impact on the understanding of the flood hazard at the location of interest. The assumption of independence or dependence was not observed to have a significant impact on the flood hazard assumption when only two flood sources were considered. However, a strong relationship did not exist between the two flood sources, and for three flood sources a strong relationship was apparent between the riverine and pluvial flood depth samples. This relationship likely explains the more significant impact of the assumption of independence or dependence observed for three flood sources.

7.4. Flood Risk Calculations

The previous sections have developed a process to assess the flood hazard for a location of interest based on multiple flood sources. The flood frequency assessments developed were based on joint distributions for all three flood sources calculated based on copulas. However, the flood hazard is only one part of flood risk. The other two components of flood risk are the vulnerability to flooding and the consequences that result from those flood events. Flood risk is calculated by multiplying the exceedance probability, the vulnerability, and the consequences for a given flood depth of interest.

The term vulnerability refers to the ability of the system in place to protect the location of interest from flooding. The system will include any structures near the location of interest designed to reduce the probability of flooding or any nonstructural measures taken at or near the location of interest to reduce either the probability or the consequences of flooding. Common examples of structural systems include levees or floodwalls. Nonstructural measures include elevation of structures, flood proofing structures, or creating additional room to store waters naturally. Regardless of the system designed to protect the location of interest, failure is always a possibility. This potential is incorporated into the vulnerability term of the risk equation.

To complete the flood risk calculations, the consequences of the flood event must be assessed. Consequences may be identified in economic terms, such as damages to structures and infrastructure, environmental damages, or loss of life or

injuries as a result of the flood event, though economic terms are the most commonly used. For the purpose of providing an example of the method, economic consequences will be considered because they can be easily calculated and understood.

7.4.1. Description of Methods

The exceedance probabilities calculated for the total flood depths using the copula procedure were used as input to the flood risk calculations. Flood risk calculations also rely on measurements of vulnerability and consequences. Vulnerability is a weight, that ranges between 0 and 1and indicates how well the in-place system to protect the location of interest from flood events is expected to perform. Therefore, vulnerability will vary with flood depth. A hypothetical vulnerability curve was created for use in experimental calculations, as shown in Figure 7-13. Vulnerability increases with depth in a linear fashion, which indicates that vulnerability is simply directly proportional to depth of flooding, which may be true for certain systems. Without specific information about the system in place at the location of interest this linear curve seemed to be a reasonable approximation of vulnerability for the purpose of these calculations. Vulnerability will vary significantly, depending on the system in place and on the condition of that system.

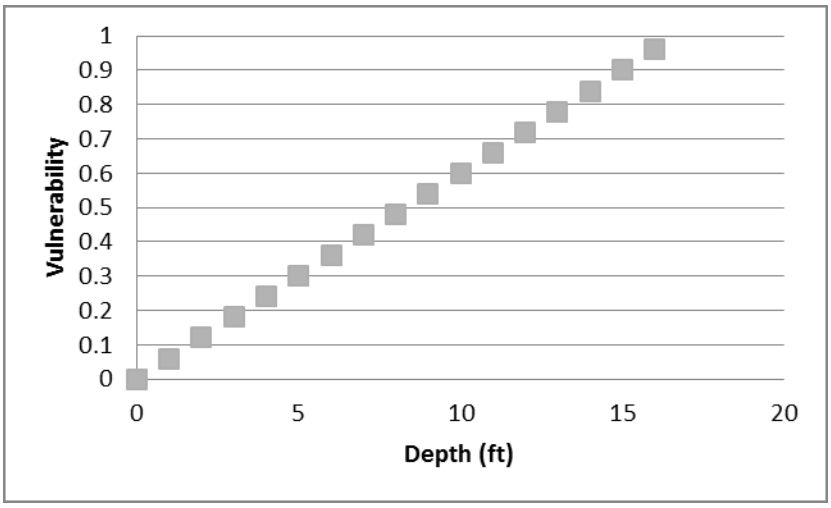

Figure 7-13: Illustration of the Vulnerability Curves Used in Flood Risk Assessments

Consequences, which are measured as damage to the structure and its contents due to the flood, will also vary with depth. Depth-damage curves (e.g., see Figure 7-13) obtained from USACE (2003) are used to determine the percentage of damage to the structure and to contents based on the depth of flooding. These depth-percent damage curves are general curves that can be applied nationally, though it is also preferable to derive specific curves for a given location. For a hypothetical scenario using simulated data, the general curves were deemed to be appropriate. These percent damage values were then multiplied by the value of the structure and the value of the contents to determine the total monetary damage due to varying flood depths. For the hypothetical scenario, the structure was assumed to be a two-story residential structure without a basement, assumed to be worth $150,000 with contents assumed to be worth $25,000. The resulting depth-monetary damage curves are provided in Figure 7-14.

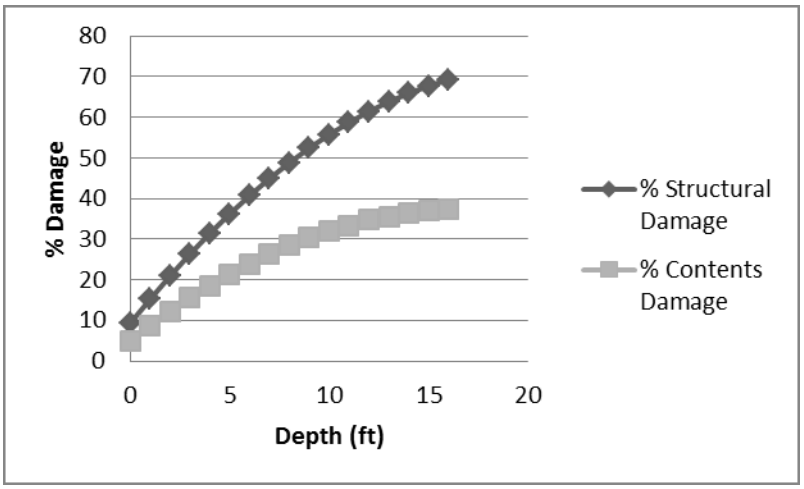

Figure 7-14: Depth-Percent Damage Curves Obtained from USACE (2003)

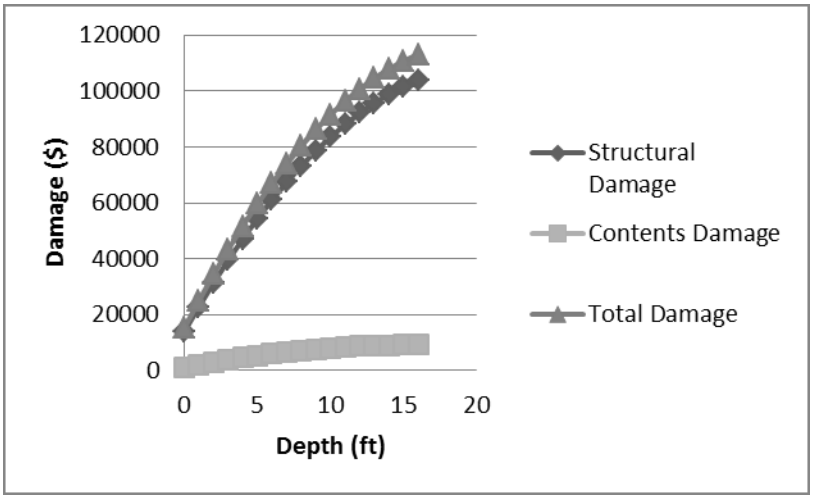

Figure 7-15: Depth-Monetary Damage Curves

Depending on the purpose of the risk analysis, it might be desirable in some circumstances to fit equations to the vulnerability and consequence curves, such that values could be determined for any flood depth of interest. However, the purpose of this analysis is simply to demonstrate the effect of the comprehensive probability assessment on risk calculations. Thus, the population fitted to the exceedance probabilities that corresponded to the total flood depth values was used to determine the exceedance probabilities corresponding to various depths of flooding for which values of vulnerability and consequences had been obtained. These exceedance

probabilities were multiplied by the vulnerability values and consequence values for each flood depth to obtain flood risk values for each flood depth.

Several comparisons of flood risk were made using the observed data to understand the impact of multiple flood sources. First, the flood risk was calculated based on the exceedance probabilities corresponding to flood depths for each of the individual flood sources. For this, the probability distributions fitted to the annual maximum riverine and tidal flood depths, identified independently of each other, as is typically done in flood frequency analyses, were used. This analysis illustrated the impact of considering only one of the flood sources at a time on the understanding of flood risk at a specific location.

Flood risk calculations were also made using the exceedance probabilities calculated based on the copula, which considered all three flood sources and the dependence between them. This was believed to provide a more complete understanding of the flood hazard, and thus should also provide a more complete understanding of the flood risk. These flood risk calculations were compared to those made using each of the individual flood source exceedance probabilities. The purpose of this analysis was to illustrate how the inclusion of all flood sources impacted flood risk, as well as how this method resulted in a varied understanding of flood risk for the location of interest.

Finally, flood risk was calculated using the exceedance probabilities calculated based on Equation 7-19. The results of this equation provided an understanding of the joint probability of flooding; however, the Equation 7-19

assumes that the three flood sources are independent. The purpose of this analysis was to assess the impact that this assumption of independence between the flood sources would have on the understanding of flood risk for the location of interest.

7.4.2. Description of Results

Based on the exceedance probabilities calculated based on all three flood sources, considered to be dependent, the vulnerability curve, and the consequence curve, flood risk calculations were made. Figure 7-16 provides the exceedance probabilities and Figure 7-17 provides the flood risk calculated based on the assumption that all three flood sources should be considered to be dependent. The exceedance probabilities decreased steeply over a range of flood depths from approximately 1 to 6 feet. The flood risk peaked at a flood depth of approximately 4 feet and then decreased gradually as flood depth increases. Though the exceedance probability calculated for this scenario peaked at a lower flood depth than the flood risk, the multiplication of the exceedance probabilities with the vulnerabilities and consequences resulted in this slightly higher peak. The gradual decline in flood risk does mirror the gradual design in exceedance probabilities as the flood depths increased.

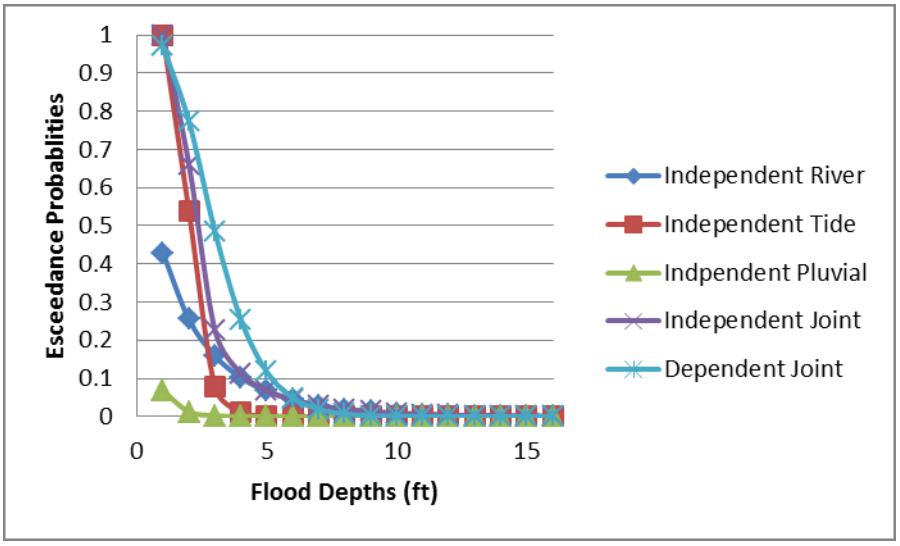

Figure 7-16: Exceedance Probabilities Calculated for the Independent Riverine, Tidal, and Pluvial Flood Sources and the Independent and Dependent Joint Distributions

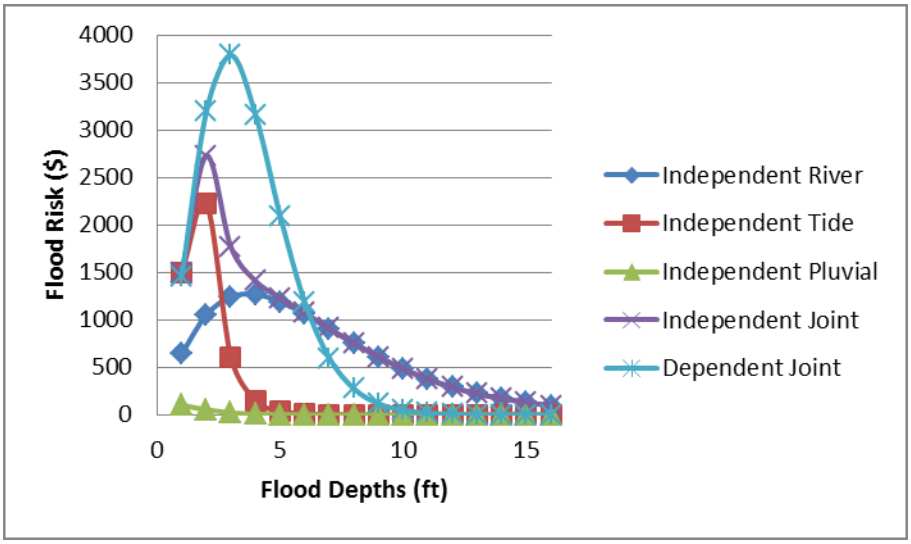

Figure 7-17: Flood Risks Calculated for the Independent Riverine, Tidal, and Pluvial Flood Sources and the Independent and Dependent Joint Distributions

Comparison of the flood risks calculated based on each of the three flood sources independently suggests that a much different understanding of flood risk would be developed if only one of the flood sources were considered, as would be expected given the variations in the exceedance probabilities calculated for each flood source. For instance, the independent tidal flood depth sample resulted in the highest flood risk for low flood depths, because the exceedance probabilities for this sample

were highest for low flood depths. The pluvial flood depth sample resulted in very low flood risk, which would be expected given the very low flood depths in this sample. The riverine flood risk calculated was greater than the pluvial flood risk but less than the tidal flood risk for low flood depths, and greater than both for larger flood depths. The riverine flood risk also peaked at a higher flood depth than the riverine exceedance probabilities, based on the multiplication of the exceedance probabilities by the vulnerabilities and consequences.

The flood risk, when calculated based on the joint probabilities, appears to differ widely when calculated based on the assumptions of dependence and independence. The multiplication of the exceedance probabilities with the vulnerabilities and consequences exacerbated the differences in the exceedance probabilities. The assumption of dependence between the flood sources results in higher flood risks calculated for given flood depths. This was not unexpected, given the typically higher exceedance probabilities observed for a given flood depth when the assumption of dependence was made. This suggests that the difference in exceedance probabilities based on the assumption of dependence and independence carried over into the flood risk calculations, such that this assumption would influence the understanding of the flood risk for the location of interest. The use of either joint distribution also differs widely from the flood risks calculated based on the individual flood sources, though they align most closely with the riverine flood risk calculations. This can also be explained by the differences in understanding of probabilities of flood events for all of these scenarios. This suggests that the

consideration of the flood sources jointly, either independently or dependently, would result in a more complete understanding of the flood risk for the location of interest.

7.5.3. Conclusions

Flood risk calculations were made using the probabilities of flooding calculated based on all three flood sources. The primary purpose of this assessment was to determine how consideration of all three flood sources would influence understanding of the flood risk at a location of interest. Flood risk calculations were made based on joint probabilities calculated under the assumptions of independence and dependence and based on probabilities calculated for riverine, tidal, and pluvial flood sources independently. These results suggest that only considering one of the flood sources does not provide an understanding of the full risk of flooding at a location. Further, consideration of the joint probabilities calculated based on the assumption of dependence typically resulted in larger exceedance probabilities, and therefore larger flood risk, than based on the assumption of independence. Neither set of probabilities can be considered "correct", but based on the correlation between the observed flood sources; an argument for the rationality of the assumption of dependence can be made. This variation in calculated probabilities of flooding and flood risk will influence the planning steps taken to prepare for and reduce damages from flood events, so the decision as to what method to use in assessing the flood hazard and flood risk should be made with careful thought and consideration.

CHAPTER 8

UNCERTAINTY ANALYSES FOR MULTIPLE FLOOD SOURCES

8.1. Introduction

In addition to developing and testing a method of assessing the impact of multiple flood sources on flood hazard and flood risk assessments, another objective of this research was to quantify the various sources of uncertainty that could be associated with each of the potential sources of flooding and show the effect of the uncertainties on flood risk estimates. When considering multiple potential sources of flood waters, such as riverine flow, coastal water, surface water flow, or groundwater, numerous sources of uncertainty are associated with each individual source. In addition to the uncertainty associated with each individual flood source, the process of developing a flood hazard assessment based on all possible sources of flooding may increase or otherwise alter the level of uncertainty.

Singh *et al.* (2007) defined uncertainty as a measure of the imperfect knowledge or probable error that could occur in an analysis. This error could be introduced through data collection, modeling and analysis of engineering systems, and prediction of a random process. Categories and sources of uncertainty identified included natural, model, parameter, data, computational, and operational. Based on these definitions, the uncertainty analysis was defined as the study of the uncertainty

in a model's output as a function of an inventory of the sources of uncertainty associated with the model inputs.

Many sources of uncertainty enter into flood risk assessments. For instance, the exact location of a flood event cannot be specified with certainty, nor can the timing or intensity of the event (Golding, 2009). Uncertainty in measured flow data influences the computed inundation depth in the surrounding area, as depth-discharge relations are imprecise. This can be estimated based on knowledge of the topography, but these estimates will always involve some level of uncertainty. In addition to the physical processes that influence flooding, uncertainty results from the simplification of the hydrologic and hydraulic models used. Numerous parameters used throughout the flood frequency analysis procedure are estimated based on uncertain data, which can lead to uncertain model results. Understanding the uncertainties in both the physical processes and the modeling of floods will be valuable in improving the use of flood estimates. By quantifying the uncertainties that may be associated with a flood frequency analysis and determining which of these uncertainties are most significant, an understanding can be developed of where to focus future research efforts to improve the accuracy of flood predictions.

8.2. Two Flood Sources

The level of uncertainty was first assessed for consideration of two potential flood sources. The following sections describe the methods used to assess the uncertainty in the analysis when riverine and tidal flooding are of interest, though the methodology would be applicable to other combinations of two flood sources as well.

The results provide a quantification of the level of uncertainty in the assessment of the flood hazard for the location of interest in Florida.

8.2.1. Description of Methods

The purpose of this uncertainty analysis was to identify the impact that uncertainties or errors throughout the analysis procedure had on the flood depths calculated for desired exceedance probabilities. The typical method of conducting an uncertainty or error analysis requires identifying the magnitude of errors that can reasonably be expected for given parameters or coefficients in order to determine the magnitude of expected impacts (McCuen, 2003). The analysis conducted for the observed data at the location of interest, presented in Chapter 5, included: (1) the calculation of flood inundation depths based on gage measurements; (2) the fitting of marginal distributions to the flood inundation depth records; (3) the fitting of joint distributions using the copula equations; (4) the calculation of non-exceedance probabilities that correspond to total flood depths using the double integral procedure; and (5) the fitting cumulative distributions to represent the probabilities associated with total flood depths. However, steps 2 through 4 were used only to assess the goodness-of-fit of several distributions fitted to the total flood depths.

In order to determine the cumulative probabilities of total flood depths, the steps that were required included: (1) the calculation of riverine inundation depths based on gage measurements; (2) the calculation of tidal inundation depths based on gage measurements; (3) the calculation of total flood depths by summing the corresponding riverine and tidal inundation depths; and (4) the fitting of distributions to the total flood depth data series. Within this procedure, uncertainties were present

in the processes used to calculate flood inundation depths based on the gage data. Further uncertainty may have been introduced through the choice of distribution fitted to the total flood depths, as clear guidance for choosing one distribution or copula family over the other possibilities does not exist.

8.2.1.1. Generation of Simulated Gage Measurements and Calculation of Flood Depths

The first step of the uncertainty analysis for two flood sources was to randomly generate the riverine and tidal gage data. In order to make the results of the uncertainty analysis comparable to the results of the observed data, samples representing 17 annual maximum flood depths were generated for each flood source. In order to understand the error associated with the procedure, a total of 3,000 sets of samples were generated and used in the assessment of uncertainty. The riverine samples were generated using both a log-Pearson Type III population (shape parameter of 196.4906, scale parameter of 0.1477, and location parameter of -24.5884) while the tidal samples were generated using a GEV population (shape parameter of -0.6427, scale parameter of 1.4606, and location parameter of 2.8305). These parameters were estimated based on the observed gage measurements. The simulated riverine gage measurements were then transposed downstream, as described in Chapter 5. This process involved one uncertain parameter. The value of the transposition parameter for each run was randomly generated from a log-normal population with a mean of 0.8 and a standard deviation of 0.28 (log-mean of -0.2809 and log-standard deviation of 0.34 for the log-normal distribution), where the mean value was identified from the observed data and the standard deviation was

approximated by multiplying the mean value by the expected error, which was believed to be approximately 35%. The expected error was identified based on experience working with hydrologic data.

Next, riverine flood inundation depths were calculated based on the transposed discharge data using the process described in Chapter 5. The parameters and coefficients of Manning's equation were sources of uncertainty in this step. For each simulation run, each parameter value was generated randomly from lognormal populations, with mean values identified as the value calculated for the observed data set and standard deviations identified as the expected error in those parameters. For Manning's roughness, the mean, based on the observed data, was 0.05 and the standard deviation was 0.015 (for the log-normal distribution the log-mean was -3.0388 and the log-standard deviation was 0.2936), which was approximated based on an expected error in Manning's roughness coefficient of 30%. The expected error in all parameters was expected based on experience working with hydrologic data. For the channel slope the mean was 0.0007, based on the observed data, and the standard deviation was 0.000105 (log-mean of -7.2756 and log-standard deviation of 0.1492), which was approximated based on an expected error of 15%. The hydraulic radius could not be randomly generated, though this coefficient could have significant uncertainty associated with it, because both wetted perimeter and area are a function of depth. Because discharge was given, the hydraulic radius had to be calculated in order to determine inundation depth. Prior to using Manning's equation to calculate inundation depths corresponding to the each discharge value, the discharge values were compared to the estimated bankfull discharge for the given channel slope and

roughness coefficient. If the discharge value was less than the bankfull discharge, the flood inundation depth at the location of interest was equal to zero. If the discharge was greater than bankfull discharge then Manning's equation was used to determine the corresponding inundation depth at the location of interest.

Flood inundation depths also had to be calculated for the tidal flood source based on the generated tidal gage measurements. In this process, the coefficients were determined based on the observed data. The coefficients were F and G, which were based on the fetch length and inland fetch length, respectively. These parameters were discussed in more detail in Chapter 5, and were determined from tables which related the fetch and inland fetch lengths to F and G, respectively. Because the location of interest was along the open ocean, rather than along a harbor or bay, the fetch length was not randomly varied, and the value of F was a constant, unitless value of 1.0 for all simulations. The length of the inland fetch was randomly generated from a lognormal population with the mean identified as the observed data value and the standard deviation identified based on the expected error. The mean of the inland fetch was 0.02 and the standard deviation was approximated to be 0.003 (log-mean of -3.9231 and log-standard deviation of 0.1492 for the lognormal distribution), based on an expected error of 15%. Based on this information, the method and equations presented in Chapter 5 were used to compute tidal flood inundation depths corresponding to the generated tidal gage measurements.

8.2.1.2. Fitting Distributions to Total Flood Depths

Once the riverine and tidal flood depths that correspond to the generated gage measurements had been determined, the total flood depths that impacted the location

of interest could be calculated. As explained in Chapters 4 and 5, the observed riverine and tidal flood depths that occurred together were summed in order to determine the total flood depths. For each simulation in the uncertainty analysis, the generated riverine and generated tidal flood depths were summed to calculate the total flood depths.

Once the total flood depths were calculated, a distribution was required to provide the non-exceedance probabilities corresponding to each flood depth. The analysis procedure examined the suitability of the following four distributions: the LP3, the GEV, the gamma, and the normal distributions. The LP3, GEV, and gamma distributions were all typically observed to perform fairly well, while the normal distribution was not typically found to provide a good fit to the data. Therefore, the uncertainty analysis examined the potential distinctions between the LP3, GEV, and gamma distributions. The three distributions were fitted to each set of total flood depth values calculated in the uncertainty analysis. Then, each distribution was used to calculate the total flood depths corresponding to the 2-year, 10-year, 25-year, 50-year, and 100-year events (exceedance probabilities of 0.5, 0.1, 0.04, 0.02, and 0.01, respectively). This provided a measure of the uncertainty that could be caused in the final flood depth estimates due to uncertainty in the parameters used to calculate the flood inundation depths based on the observed gage data and due to the uncertainty in the choice of the distribution used to model the non-exceedance probabilities of the total flood depths.

8.2.1.3. Analysis of Simulation Results

The objective of this analysis was to demonstrate the range of flood depths that could result from different parameter values and different choices of distributions fitted to the total flood depths. For each of the three distributions evaluated, several flood depths were identified. These flood depths corresponded to the 2-year, 10-year, 25-year, 50-year, and 100-year events. For each flood depth, from the 3,000 simulation runs, the minimum depth, the 5^{th} percentile, the 10^{th} percentile, the 25^{th} percentile, the 50^{th} percentile, the 75^{th} percentile, the 90^{th} percentile, the 95^{th} percentile, and the maximum flood depth were identified. This provided information similar to a box-and-whisker plot about the likely range of flood depths that could be predicted, depending on the level of uncertainty in the parameters and distributions. This analysis also allowed for comparisons to be made across distributions, to determine whether certain combinations of distributions resulted in more or less uncertainty in the resulting flood depths. The same range of flood depths was also identified based on the entire set of simulated data as well, in order to identify the full range of uncertainty in flood depths while taking into account the possible variations due to parameters and choice of distributions. This information was provided in table form and graphically. The results of the observed data set were plotted against the full set of uncertainty analyses to place the results of the observed data into context.

8.2.2. Results and Discussion

The flood depths that correspond to the 2-year, 10-year, 25-year, 50-year, and 100-year return periods were identified from each simulation run. For each of the three distributions fitted to the total flood depths, a range of flood depths were

identified that corresponded to each return period of interest. The range of flood depths predicted for each distribution are presented and then the combined results for all three distributions are compared to the observed flood depths.

8.2.2.1. Log-Pearson Type III Distribution

The first distribution to be evaluated to represent the total flood depths was the LP3 distribution. This analysis assessed the uncertainty caused by the parameters used to calculate the riverine and tidal flood depths, and the parameters of the fitted LP3 distribution. Table 8-1 and Figure 8-1 provide the range of flood depths calculated for various exceedance probabilities. Based on these results, it is evident that for all exceedance probabilities, a wide range of flood depths might be calculated based on the combination of parameters used in the analysis. Further, the level of uncertainty appears to increase as the exceedance probability increases. For the 2-year flood (exceedance probability of 0.5) the flood depths range from 0.74 to 6.50 feet, which is a much smaller range than observed for the 100-year flood (exceedance probability of 0.01). For the larger return period, flood depths range from 4.35 to 51.97 feet. This level of uncertainty in the larger return periods is significant because they correspond to the larger flood events of most concern.

This wide range of flood depths calculated for each return period suggests the importance of obtaining accurate information in order to calculate the riverine and tidal flood depths. Potentially inaccurate flood depth parameters can be observed to have very large impact on the estimated total flood depths corresponding to different return periods. For instance, an estimate of the 2-year flood depth of 0.74 feet versus 6.50 feet would make significant difference in the preparations taken for the 2-year

flood. The range of flood depths estimated for the 100-year flood event are even larger and so would have an even more significant impact on preparations taken, based on whether a flood depth of 4.35 feet or 51.97 feet were expected.

Table 8-1: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the LP3 Distribution is Used to Represent the Total Flood Depths

	Exceedance Probability				
Flood Depth	0.5	0.1	0.04	0.02	0.01
Minimum	0.74	3.79	4.18	4.33	4.35
5th Percentile	1.20	5.32	6.68	7.31	7.76
10th Percentile	1.32	5.84	7.57	8.41	8.98
25th Percentile	1.59	6.85	9.39	10.96	12.12
50th Percentile	1.94	8.18	11.67	14.09	16.14
75th Percentile	2.38	9.69	14.52	18.16	21.95
90th Percentile	2.83	11.27	17.08	21.60	26.35
95th Percentile	3.18	12.39	18.56	23.88	29.77
Maximum	6.50	20.77	28.72	39.91	51.97

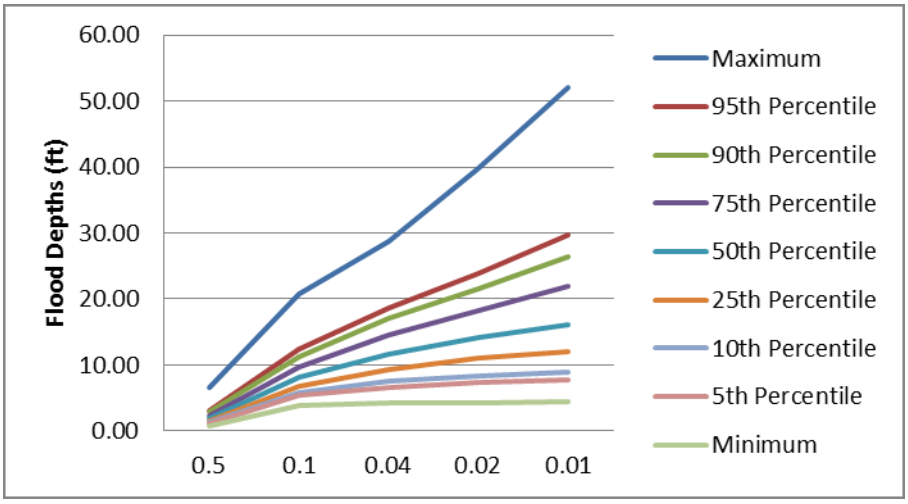

Figure 8-1: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the LP3 Distribution Fitted to Total Flood Depths

The maximum flood depths identified, particularly for the larger return periods, do not seem particularly reasonable. A predicted flood depth of 52 feet corresponding to the 1% annual chance flood depth is clearly unreasonable for this watershed. In examining the results of the uncertainty analyses, it was determined that the particularly large predicted flood depths for the larger return periods were the

result of a poor fit of the distribution to the sample of total flood depths. When the maximum non-exceedance probability determined by the fitted distribution was low, for instance approximately 0.9, then the fitted distribution had to predict the flood depths that corresponded to the larger return periods. The behavior of the distribution in the upper tail then controlled the larger predicted flood depths. The distribution was observed to poorly fit the total flood depth samples in this way when the sample consisted of numerous very low flood depth values and 1 or 2 comparatively large flood depths. This was primarily controlled by the generated riverine flood depth sample, which frequently had a large number of zero-flood years. When the riverine sample generated consisted of numerous zero-flood years and 1 or 2 large flood depths, the total flood depth sample typically consisted of numerous very low flood depths and 1 or 2 high flood depths. The LP3 distribution typically could not easily fit these total flood depth samples. The standard deviations of the distributions used to randomly generate the riverine and tidal flood depth calculation parameters also was observed to have some influence on the predicted flood depths for the return periods of interest. When smaller standard deviations were used it was less likely that particularly large riverine flood depths would be generated. Thus, the final distribution fitted to the total flood depths were able to fit more accurately, and more reasonable flood depths were predicted for the 1% annual chance flood event.

8.2.2.2. Generalized Extreme Value Distribution

The GEV distribution was next evaluated for its ability to fit the total flood depths. Table 8-2 and Figure 8-2 illustrate the level of uncertainty in flood depths calculated for each exceedance probability using the GEV distribution to fit the total

flood depths. Based on the parameters used to calculate the riverine and tidal flood depths, a wide range of flood depths was again predicted to correspond to each return period evaluated. For instance, the 2-year flood depth was determined to vary from 0.66 to 6.12 feet while the 100-year flood depth was determined to vary from 3.46 to 81.14 feet. Another consistent trend identified is that the level of uncertainty increases as the return period increases. The larger flood events corresponding to larger return periods are typically the more serious floods with more significant consequences, so they are of more concern; however, predictions of the expected flood depths corresponding to large returns periods appear to be inherently less certain. This again suggests the importance of obtaining accurate parameters for use in calculating the riverine and tidal flood depths, as these parameters will ultimately have a significant impact on the calculation of total flood depths and the calculation of flood depths corresponding to return periods of interest.

Table 8-2: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the GEV Distribution is Used to Represent the Total Flood Depths

	Exceedance Probability				
Flood Depth	0.5	0.1	0.04	0.02	0.01
Minimum	0.66	2.97	3.24	3.37	3.46
5th Percentile	1.20	4.61	6.31	7.63	9.02
10th Percentile	1.35	4.97	7.04	8.99	10.77
25th Percentile	1.62	5.94	9.14	12.14	15.89
50th Percentile	1.92	7.14	12.21	17.60	24.83
75th Percentile	2.28	8.65	16.27	25.80	40.72
90th Percentile	2.68	9.91	20.12	34.74	59.13
95th Percentile	3.06	11.00	21.91	38.78	68.36
Maximum	6.12	18.21	33.73	52.27	81.14

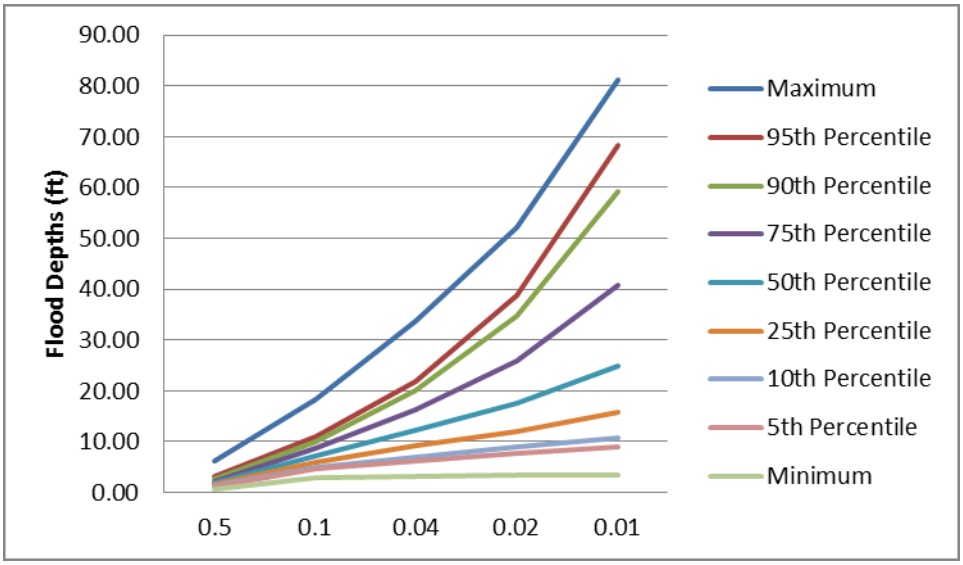

Figure 8-2: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the GEV Distribution Fitted to Total Flood Depths

The flood depths calculated for each exceedance probability should also be compared to the flood depths calculated based on the fitted LP3 distribution. A large difference in range of flood depths calculated for each return period is observed between the GEV and LP3 distributions. The minimum flood depths were observed to decrease when the GEV distribution was fitted to the total flood depths as compared to the LP3 distribution. For the smaller events, the 2-year and 10-year return periods, the maximum flood depths were also observed to be slightly smaller when the GEV distribution was fitted to the total flood depths as compared to the LP3 distribution. However, for the larger return periods (25-year, 50-year, and 100-year return periods) the GEV distribution was observed to result in larger maximum flood depths than the LP3 distribution. This suggests that fitting the GEV distribution to the total flood depths results in more uncertainty in the flood depths calculated for each exceedance probability than fitting the LP3 distribution. The behavior of the GEV distribution in the upper tail results in larger predicted flood depths than for the LP3 distribution. The choice of distribution fitted to the total flood depths clearly has

quite a bit of influence on the predicted flood depths corresponding to exceedance probabilities of interest.

As was observed when the results based on the LP3 distribution were examined, the maximum flood depths identified seem highly unreasonable. A predicted flood depth of 81 feet corresponding to the 1% annual chance flood depth is clearly unreasonable for this watershed. The results of the uncertainty analyses were examined and it was determined that the particularly large predicted flood depths for the larger return periods were the result of a poor fit of the distribution to the total flood depths. When the maximum non-exceedance probability determined by the fitted distribution was low, for instance approximately 0.9, then the fitted distribution had to predict the flood depths that corresponded to the larger return periods. The behavior of the distribution in the upper tail then controlled the larger predicted flood depths. The behavior of the GEV distribution in the upper tail was observed to result in much larger predicted flood depths than either the LP3 or gamma distributions. The distribution was observed to poorly fit the total flood depth samples in this way when the sample consisted of numerous very low flood depth values and 1 or 2 comparatively large flood depths. This was primarily controlled by the generated riverine flood depth sample, which frequently had a large number of zero-flood years. When the riverine sample generated consisted of numerous zero-flood years and 1 or 2 large flood depths, the total flood depth sample typically consisted of numerous very low flood depths and 1 or 2 high flood depths. The GEV distribution typically could not easily fit these total flood depth samples. The standard deviations of the distributions used to randomly generate the riverine and tidal flood depth calculation

parameters also were observed to have some influence on the predicted flood depths for the return periods of interest. When smaller standard deviations were used it was less likely that particularly large riverine flood depths would be generated. Thus, the final distribution fitted to the total flood depths were able to fit more accurately, and more reasonable flood depths were predicted for the 1% annual chance flood event.

8.2.2.3. Gamma Distribution

The gamma distribution was also evaluated as a potential marginal distribution to represent the total flood depths. Table 8-3 provides the range of flood depths calculated for each exceedance probability evaluated, where an exceedance probability of 0.5 corresponds to the 2-year flood, an exceedance probability of 0.1 corresponds to the 10-year flood, etc. The range of flood depths calculated for each return period is also provided graphically in Figure 8-3. As observed when using the LP3 and GEV distributions, based on the parameters used to calculate the riverine and tidal flood depths, a wide range of flood depths was calculated corresponding to each return period. The flood depths calculated for an exceedance probability of 0.5 were determined to range from 0.85 feet to 5.05 feet, while the flood depths calculated for an exceedance probability of 0.01 were observed to vary from 5.90 feet to 41.96 feet. This is a fairly wide range of flood depths based on the choice of parameters used to calculate the riverine and tidal flood depths, suggesting how important it is to have accurate information to use in calculating these flood depths. Further, the range of flood depths calculated increased for the larger events (100-year event) versus the smaller events (2-year event).

The flood depth calculations made based on fitting the gamma distribution to the total flood depths were compared to those calculations made based on fitting the LP3 and GEV distributions. Overall, the flood depths corresponding to each return period were slightly smaller when the gamma distribution was used than when the LP3 or GEV distributions were used. The differences in the smaller return periods (2-year and 10-year events) were fairly small, but for the larger return periods the choice of distribution had more significant impact. The 50-year (exceedance probability of 0.02) and 100-year (exceedance probability of 0.01) flood depths calculated based on the gamma distribution were smaller than either the LP3 or GEV distribution, though they were more similar to the flood depths calculated for the LP3 distribution. The choice of distribution can obviously have serious impact, but none of these three distributions can be considered to be the "correct" distribution to represent the total flood depths. Thus, it is important to compare the results of the fitted distribution to the non-exceedance probabilities calculated using the copula method, or Equation 1-3 if the flood sources are determined to be independent, to ensure that the results of the fitted distribution are reasonably accurate.

Table 8-3: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the Gamma Distribution is Used to Represent the Total Flood Depths

	Exceedance Probabilities				
Flood Depth	0.5	0.1	0.04	0.02	0.01
Minimum	0.85	3.45	4.72	5.32	5.90
5th Percentile	1.30	4.92	6.95	8.40	9.85
10th Percentile	1.42	5.28	7.46	9.10	10.76
25th Percentile	1.63	6.13	8.71	10.68	12.68
50th Percentile	1.94	7.30	10.49	12.90	15.35
75th Percentile	2.27	8.70	12.56	15.60	18.58
90th Percentile	2.70	10.11	14.74	18.29	22.03
95th Percentile	3.01	11.24	16.32	20.35	24.45
Maximum	5.05	18.96	27.96	34.92	41.96

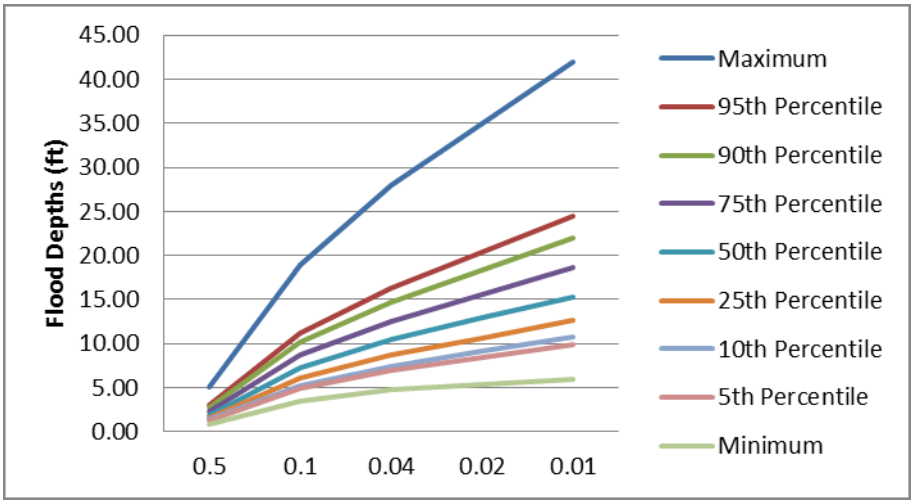

Figure 8-3: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the Gamma Distribution Fitted to Total Flood Depths

The maximum flood depths predicted for the larger return periods using the gamma distribution were more reasonable than was observed using the LP3 and GEV distributions. However, a predicted flood depth of 42 feet corresponding to the 1% annual chance flood depth is still clearly unreasonable for this watershed. The results of the uncertainty analyses were examined and it was determined that the particularly large predicted flood depths for the larger return periods were the result of a poor fit of the distribution to the total flood depths. When the maximum non-exceedance probability determined by the fitted distribution was low, for instance approximately 0.9, then the fitted distribution had to predict the flood depths that corresponded to the larger return periods. The behavior of the distribution in the upper tail then controlled the larger predicted flood depths. The behavior of the gamma distribution in the upper tail was observed to result in lower predicted flood depths than either the LP3 or GEV distributions. The distribution was observed to poorly fit the total flood depth samples in this way when the sample consisted of numerous very low flood depth values and 1 or 2 relatively large flood depths. This was primarily controlled by the generated riverine flood depth sample, which frequently had a large number of

zero-flood years. When the riverine sample generated consisted of numerous zero-flood years and 1 or 2 large flood depths, the total flood depth sample typically consisted of numerous very low flood depths and 1 or 2 high flood depths. The gamma distribution typically could not easily fit these total flood depth samples. The standard deviations of the distributions used to randomly generate the riverine and tidal flood depth calculation parameters also was observed to have some influence on the predicted flood depths for the return periods of interest. When smaller standard deviations were used it was less likely that particularly large riverine flood depths would be generated. Thus, the final distribution fitted to the total flood depths were able to fit more accurately, and more reasonable flood depths were predicted for the 1% annual chance flood event.

8.2.2.4. Consolidated Results of Uncertainty Analyses

The previous analyses evaluated the uncertainty in analyses based on specific choices of distributions to represent the non-exceedance probabilities corresponding to the total flood depths. These analyses evaluated the influence of the parameters used to calculate the riverine and tidal flood depths in addition to the influence of choice of distribution representing the total flood depths. The next step was to evaluate the full range of flood depths that could be calculated based on variations in the flood depth calculation parameters and on all possible distributions fitted to the total flood depths. The flood depths calculated for each return period based on the observed data was also compared to the range of flood depths calculated in the uncertainty analysis. Table 8-4 and Figure 8-4 illustrate the range of flood depths calculated for each exceedance probability based on the various sets of parameters

used in the calculations and based on the various possible choices of marginal distributions and copula families.

As would be expected, the range of flood depths calculated for each return period varies widely. The range of flood depths calculated for the smaller return periods (for example, the 2-year event) is smaller than the range of flood depths calculated for the larger return periods (for example, the 100-year event). Because the larger magnitude events, such as the 100-year event, are typically the flood events of most concern, this is an important point to keep in mind. The parameters used to calculate the riverine and tidal flood depths cannot be known with absolute certainty; however, as much care as is reasonably possible should be taken in determining the values of each of these parameters so that the analyses are based on as accurate flood depths as possible. In choosing a distribution to represent the non-exceedance probabilities corresponding to the total flood depths, none of the three evaluated can be considered the "correct" distribution. However, it is clear that the choice of distribution will have a strong influence on the calculated flood depths for each return period, especially for the larger return periods of the most concern. Thus, it is important to compare the fitted distributions to the results of either the copula or Equation 1-3 in the process of conducting an analysis in order to choose the most appropriate distribution.

As discussed for each distribution individually, the maximum flood depths predicted for the larger return periods are not reasonable. These particularly large predicted flood depths were the result of a poor fit of the distribution to the total flood depths. When the maximum non-exceedance probability determined by the fitted

distribution was low, for instance approximately 0.9, then the fitted distribution had to predict the flood depths that corresponded to the larger return periods. The behavior of the distribution in the upper tail then controlled the larger predicted flood depths. The GEV distribution had the most difficulty in fitting the distributions typically, followed by the LP3 distribution, while the gamma distribution had the least difficulty fitting the total flood depths. The distribution was observed to poorly fit the total flood depth samples in this way when the sample consisted of numerous very low flood depth values and 1 or 2 relatively large flood depths. This was primarily controlled by the generated riverine flood depth sample, which frequently had a large number of zero-flood years. When the riverine sample generated consisted of numerous zero-flood years and 1 or 2 large flood depths, the total flood depth sample typically consisted of numerous very low flood depths and 1 or 2 high flood depths. The standard deviations of the distributions used to randomly generate the riverine and tidal flood depth calculation parameters also was observed to have some influence on the predicted flood depths for the return periods of interest. When smaller standard deviations were used it was less likely that particularly large riverine flood depths would be generated. Thus, the final distribution fitted to the total flood depths were able to fit more accurately, and more reasonable flood depths were predicted for the 1% annual chance flood event. These results suggest that caution must be used in applying this method to observed data in which the riverine sample in particular consists of a large number of low flood depths or zero-flood years and a few high flood depths, as the final distribution fitted to the total flood depths may not provide an accurate fit.

The flood depths calculated for each return period based on the observed data were next compared to the results of the uncertainty analysis to evaluate the amount of variation that could be possible in the observed data. For all exceedance probabilities, the observed flood depths were determined to fall within the range predicted from the uncertainty analysis. For the 2-year return period, the observed flood depth fell approximately in the middle of the range predicted by the uncertainty analysis, while for the larger return periods, the observed depths tended to be approximately equal to the 5^{th} percentile flood depths calculated in the uncertainty analysis. The observed riverine and tidal flood depths were both fairly low, which explains the low flood depths corresponding to the larger flood depths. Also, the maximum predicted flood depths in the uncertainty analysis were known to be unreasonable due to the inability of the distributions to fit the total flood depth samples under conditions in which numerous zero-flood years occurred. The uncertainty analysis revealed that variations in the parameters used to calculate the riverine and tidal flood depths could result in much larger calculated flood depths. The larger riverine and tidal flood depth samples that could be calculated based on these different parameter sets then resulted in larger total flood depths, which influenced the distribution fitted to the total flood depths and the flood depths predicted by these distributions to correspond to the return periods of interest. This suggests that if there were any inaccuracies in the parameters used to calculate the observed flood depths, the results of the analysis based on the observed data could have varied quite widely.

Table 8-4: Range of Flood Depths (in Feet) Calculated for Various Exceedance Probabilities for All Distributions

	Exceedance Probabilities				
Flood Depth	0.5	0.1	0.04	0.02	0.01
Minimum	0.66	2.97	3.24	3.37	3.46
5th Percentile	1.24	4.85	6.58	7.74	8.78
10th Percentile	1.37	5.31	7.38	8.88	10.16
25th Percentile	1.62	6.25	9.04	11.12	12.99
50th Percentile	1.93	7.50	11.29	14.22	17.22
75th Percentile	2.30	9.01	14.31	18.79	24.03
90th Percentile	2.75	10.56	17.53	25.65	37.41
95th Percentile	3.09	11.61	19.94	31.65	51.71
Maximum	6.50	20.77	33.73	52.27	81.14
Observed	2.49	4.52	5.86	7.01	8.31

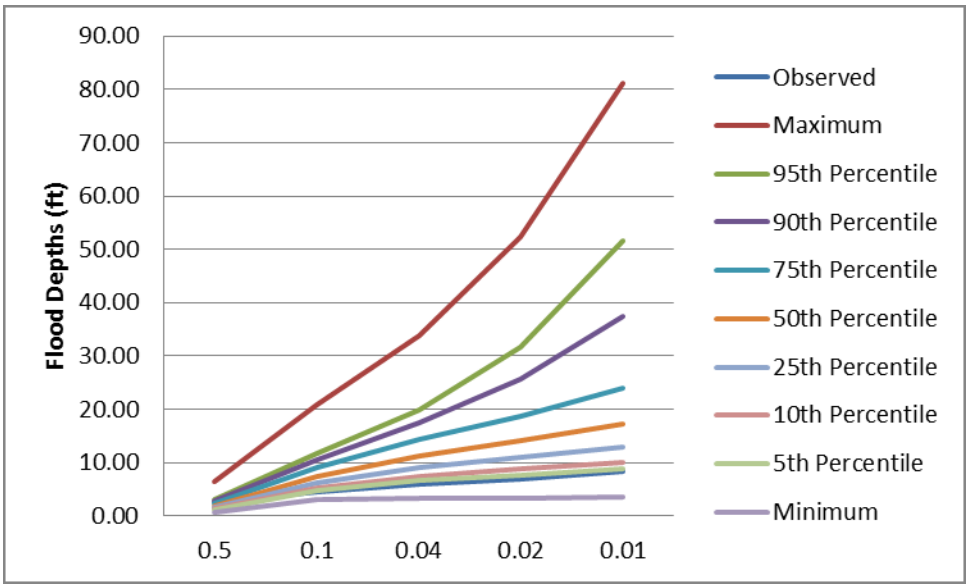

Figure 8-4: Plot Illustrating Range of Flood Depths for Various Exceedance Probabilities Calculated for All Distributions

8.2.2.5. Further Investigation of the Effect of Zero-Flood Years

The unreasonably high maximum flood depths predicted by the uncertainty analyses were further assessed using several additional studies. The high values were believed to be due to the number of zero-flood years in the riverine and tidal flood depth samples influencing the resulting total flood depth samples such that the distributions could not accurately fit the total flood depth samples. The first study

generated 25,000 samples of 10 flood depth values each with 0, 1, 2, 3, and 4 zero-flood years within the record. The samples were generated from a log-Pearson Type III distribution with a log-mean of 0.4152, log-standard deviation of 0.18, and a skew of 0.6666. These parameters were the LP3 distribution parameters fitted to the observed total flood depth sample, which was influenced by zero-flood years in the riverine and tidal flood depth samples. From this population, the known 2-year, 10-year, and 100-year flood depths were 2.50, 4.51, and 8.15 feet, respectively.

In order to determine the effect of the zero-flood years, for each simulation run, the 2-year, 10-year, and 100-year flood depths were predicted and compared to the known values. The relative error, calculated by subtracting the population depth from the computed depth, and dividing by the population depth, was used to determine the impact of the zero-flood years on the predicted flood depths. Table8-5 provides the relative errors in predicted flood depths for the 2-year, 10-year, and 100-year flood depths. The predicted flood depths based on no zero-flood years were not equivalent to the population values due to sampling variation. The relative errors for the 2-year and 10-year flood depths were generally negative because the zero-values cause the mean to decrease, while the relative errors for the 100-year flood depths are generally positive because the zero-values increase the sample skew, which causes the 100-year computed flood depths to increase. As the number of zero-flood years in the sample increased, the relative error was observed to increase in magnitude, suggesting that the larger number of zero-values in the samples, the more inaccuracy there would be in the predicted flood depths. This demonstrates that the zero-flood years in the sample do significantly influence the predicted flood depths.

Table 8-5 Relative Errors in 2-Year, 10-Year, and 100-Year Predicted Flood Depths Based on Zero-Flood Years

	Relative Errors		
Number of Zeros	2-Year	10-Year	100-Year
0	0.0269	0.0020	0.0128
1	-0.0324	-0.0091	-0.0412
2	-0.1247	-0.0309	0.0022
3	-0.2201	-0.0769	0.0459
4	-0.3054	-0.1325	0.0944

An investigation was conducted to confirm that the zero-flood years were responsible for the unrealistically high predicted flood depths that correspond to the larger return periods observed in the uncertainty analyses. The zero-flood years were removed from both the observed riverine and tidal flood depth samples, and distributions were fitted to these revised samples. While this will distort the exceedance probabilities, it is intended to show the unrealistic effect of including zero-flood values on the computed flood depths. Then, these distribution parameters were used to generate samples of riverine and tidal data that were similar to the observed data without the presence of zero-values. The generated riverine and tidal flood depth samples were summed to create samples of total flood depths and the LP3, GEV, and gamma distributions were fitted to these total flood depth samples. A total of 3,000 simulations were conducted, as had been done in the original uncertainty analysis. For each simulation, the three distributions were used to calculated flood depths corresponding to the 2-year, 10-year, 25-year, 50-year, and 100-year flood events. The minimum, 5^{th} percentile, 10^{th} percentile, 25^{th} percentile, 50^{th} percentile, 75^{th} percentile, 90^{th} percentile, 95^{th} percentile, and maximum predicted flood depths for each return period were examined and compared to the results of the original uncertainty analyses.

The predicted flood depths calculated using the LP3 distribution were first examined. Table 8-6 presents the predicted flood depths that correspond to the return period of interest based on this distribution. The maximum flood depths predicted for all five return periods of interest have decreased. For the 100-year flood depth, the maximum flood depth has decreased from nearly 45 feet to 14 feet, which is a much more reasonable flood depth value to predict for the location of interest. The predicted flood depths calculated using the gamma distribution are presented in Table 8-7. For this distribution, the results were very similar to the results observed using the LP3 distribution. The maximum flood depth predicted for the 100-year event was reduced from approximately 43 feet to 13 feet using the methodology presented above. The simulated riverine and tidal flood depth data using these distribution parameters would have been fairly similar to the observed data set, except that there would not have been zero-values in the samples. These results do suggest that when zero-values are not a factor, the distributions fitted to the total flood depth samples result in much more reasonable predicted flood depths. However, zero-flood years are a reality, but a method of properly adjusting for them is lacking.

Table 8-6: Predicted Flood Depths for Various Exceedance Probabilities Calculated Using the LP3 Distribution

	Exceedance Probabilities				
Flood Depths (ft)	0.5	0.1	0.04	0.02	0.01
Minimum	3.86	5.32	5.49	5.56	5.61
5th Percentile	4.44	6.21	6.75	7.01	7.21
10th Percentile	4.55	6.39	6.98	7.31	7.56
25th Percentile	4.77	6.72	7.43	7.87	8.26
50th Percentile	5.03	7.10	7.95	8.52	9.08
75th Percentile	5.28	7.50	8.57	9.37	10.08
90th Percentile	5.53	7.90	9.12	10.03	10.99
95th Percentile	5.67	8.12	9.41	10.49	11.56
Maximum	6.47	8.99	11.00	12.49	14.09

Table 8-7: Predicted Flood Depths for Various Exceedance Probabilities Calculated Using the Gamma Distribution

	Exceedance Probabilities				
Flood Depths (ft)	0.5	0.1	0.04	0.02	0.01
Minimum	3.75	5.44	6.00	6.31	6.58
5th Percentile	4.45	6.21	6.84	7.27	7.64
10th Percentile	4.55	6.39	7.09	7.54	7.94
25th Percentile	4.77	6.68	7.45	7.97	8.45
50th Percentile	4.99	7.08	7.95	8.55	9.11
75th Percentile	5.23	7.44	8.44	9.13	9.79
90th Percentile	5.45	7.85	8.97	9.73	10.48
95th Percentile	5.60	8.09	9.27	10.10	10.91
Maximum	6.40	9.00	10.61	11.73	12.81

The GEV distribution provided slightly different results than the LP3 and GEV distribution. These results are presented in Table 8-8. The maximum flood depths are greatly reduced using the method described above. The maximum flood depth predicted for the 100-year flood event has been reduced from approximately 83 feet to 50 feet. However, a maximum flood depth of 50 feet is still not reasonable for the location of interest. However, the minimum predicted flood depths through the 95th percentile predicted flood depths for each return period do look fairly reasonable. This suggests that in only a small number of the simulations did the GEV distribution predict the flood depths for the larger return periods outside the range of the sample due to poor fits between the sample and the distribution. In the original uncertainty analyses, this happened regularly for all three distributions, but it was noted that the GEV distribution had the most difficultly in accurately fitting the total flood depth samples. These results suggest that the GEV distribution may still occasionally not be able to accurately fit the total flood depth samples, but when zero-values are not a factor; the GEV distribution generally is able to predict reasonable flood depths for all return periods. The results presented in Tables 8-6 through 8-8 suggest that the

presence of zero-values in the generated riverine and tidal values were responsible for the unreasonably large predicted flood depths present in the original uncertainty analyses.

Table 8-8: Predicted Flood Depths for Various Exceedance Probabilities Calculated Using the GEV Distribution

	Exceedance Probabilities				
Flood Depths (ft)	0.5	0.1	0.04	0.02	0.01
Minimum	3.82	5.31	5.57	5.68	5.76
5th Percentile	4.43	6.13	6.53	6.68	6.79
10th Percentile	4.57	6.27	6.68	6.89	7.02
25th Percentile	4.79	6.61	7.18	7.48	7.68
50th Percentile	5.08	6.98	7.71	8.13	8.51
75th Percentile	5.36	7.37	8.34	9.03	9.72
90th Percentile	5.63	7.77	8.95	9.88	10.73
95th Percentile	5.79	7.97	9.40	10.71	12.06
Maximum	6.65	8.70	14.91	26.42	49.53

8.3. Three Flood Sources

The level of uncertainty was next assessed for consideration of three potential flood sources. The following sections describe the methods used to assess the uncertainty in the analysis when riverine, tidal, and pluvial flooding are of interest, though the methodology would be applicable to other combinations of three flood sources as well. The results provide a quantification of the level of uncertainty in the assessment of the flood hazard for the location of interest in Florida.

8.3.1. Description of Methods

The purpose of this uncertainty analysis was to identify the impact that uncertainties or errors throughout the analysis procedure had on the flood depths calculated for desired exceedance probabilities. The typical method of conducting an uncertainty or error analysis requires identifying the magnitude of errors that can

reasonably be expected for given parameters or coefficients in order to determine the magnitude of expected impacts (McCuen, 2003). The analysis conducted for the observed data at the location of interest, presented in Chapter 7, included: (1) the calculation of flood inundation depths based on gage measurements; (2) the fitting of marginal distributions to the flood inundation depth records; (3) the fitting of joint distributions using the copula equations; (4) the calculation of non-exceedance probabilities that correspond to total flood depths using the triple integral procedure; and (5) the fitting of cumulative distributions to represent the probabilities associated with total flood depths. However, steps 2 through 4 were used only to assess the goodness-of-fit of several distributions fitted to the total flood depths.

In order to determine the cumulative probabilities of total flood depths, the steps that were required included: (1) the calculation of riverine inundation depths based on gage measurements; (2) the calculation of tidal inundation depths based on gage measurements; (3) the calculation of pluvial inundation depths; (4) the calculation of total flood depths by summing the corresponding riverine, tidal, and pluvial inundation depths; and (5) the fitting of distributions to the total flood depth data series. Within this procedure, uncertainties were present in the processes used to calculate flood inundation depths based on the gage data. Further uncertainty may have been introduced through the choice of distribution fitted to the total flood depths, as clear guidance for choosing one distribution or copula family over the other possibilities does not exist.

8.3.1.1. Generation of Simulated Gage Measurements and Calculation of Flood Depths

The riverine, tidal, and pluvial flood depths were first calculated. The riverine and tidal flood depths were calculated as described in Section 8.2.1.1. In order to generate the pluvial flood depths, the first step was to generate a series of precipitation gage measurements. For the purpose of comparison to the observed data, a series of 17 precipitation gage measurements were generated. The gamma distribution was used to represent the rainfall depths, with a shape parameter of 0.6962 and a scale parameter of 3.7432.

Runoff depths were then calculated based on the generated rainfall depths. The first step to these calculations involved generating a curve number value. The mean unweighted curve number was taken to be 49, which was based on a fully developed urban area in fair condition with soil from soil group A, containing deep sand. The standard deviation of the unweighted curve number was 6.87, from McCuen (2002). The procedure outlined by McCuen (2002) was followed to determine the parameters necessary to represent the curve number using a gamma distribution. The pervious area curve number was represented by subtracting from 100 the value generated from a gamma distribution with shape and scale parameters calculated as:

$$b = \frac{s^2}{\bar{x}} \tag{8-1}$$

$$c = \left(\frac{\bar{x}}{s}\right)^2 \tag{8-2}$$

where c was the shape parameter and b was the scale parameter. To calculate the weighted curve number also required generating the fraction of impervious land. The

fraction of impervious land was generated from a lognormal distribution (mean of 0.5 from the observed data, or log-mean of -0.6987, and a standard deviation of 0.053, or a log-standard deviation of 0.1057). From Stankowski (1974), for a combination of single-family and multiple-family residential land use categories, an average range of impervious surface area was identified as 24%. Then, from Snedecor and Cochran (1968), the standard deviation range was identified as 0.222. The standard deviation of the fraction of impervious area was then calculated as 0.222*0.24, or 0.053, which was converted to a log-parameter and used in the generation of the log-normal variates. Once the weighted curve number had been calculated, the storage could be calculated. Finally, based on the calculated storage and the generated rainfall depths, the runoff depth could be calculated.

After the runoff depths were calculated the corresponding runoff discharges were calculated. The watershed length was set to 4,232 feet, the observed length, and the watershed slope was generated from a lognormal population with a mean of 0.004 (log-mean of -5.5326) and a standard deviation of 0.006, identified as the expected error of 15% times the mean value, (log-standard deviation of 0.1493). Using the watershed length, the watershed slope, and the weighted curve number, the time of concentration was calculated. The ratio of initial abstraction to precipitation depths was used to identify from tabled values the appropriate values of C_0, C_1, and C_2. These tables were obtained from the TR-55 manual. From this information, the logarithm of q_{um} was calculated as:

$$\log q_{um} = C_0 + C_1 * \log(t_c) + C_2 * (\log(t_c))^2 \quad (8\text{-}3)$$

where t_c was the time of concentration. From this, q_{um} was calculated, and then q_p, the peak discharge from the runoff, was calculated.

Finally, based on the runoff discharges, pluvial flood depths were calculated. The flood depths were calculated by solving Manning's equation for flood depths based on the calculated discharges. Manning's roughness coefficient was generated from a log-normal distribution with a mean of 0.012 (log-mean of -4.4659), the observed value, and a standard deviation of 0.0036 (log-standard deviation of 0.2936), which was calculated as the expected error of 30% of the observed value. Watershed slope was previously generated, and the street dimensions presented in Chapter 7 were used. From this information, the pluvial flood depths were calculated.

8.3.1.2. Fitting Distributions to Total Flood Depths

Once the riverine, tidal, and pluvial flood depths that correspond to the generated gage measurements had been determined, the total flood depths that impacted the location of interest could be calculated. As explained in Chapters 6 and 7, the observed riverine, tidal, and pluvial flood depths that occurred together were summed in order to determine the total flood depths. For each simulation in the uncertainty analysis, the generated riverine, tidal, and pluvial flood depths were summed to calculate the total flood depths.

Once the total flood depths were calculated, a distribution was required to provide the non-exceedance probabilities corresponding to each flood depth. The analysis procedure examined the suitability of the following four distributions: the LP3, the GEV, the gamma, and the normal distributions. The LP3, GEV, and gamma distributions were all typically observed to perform fairly well, while the normal

distribution was not typically found to provide a good fit to the data. Therefore, the uncertainty analysis examined the potential distinctions between the LP3, GEV, and gamma distributions. The three distributions were fitted to each set of total flood depth values calculated in the uncertainty analysis. Then, each distribution was used to calculate the total flood depths corresponding to the 2-year, 10-year, 25-year, 50-year, and 100-year events. This provided a measure of the uncertainty that could be caused in the final flood depth estimates due to uncertainty in the parameters used to calculate the flood inundation depths based on the observed gage data and due to the uncertainty in the choice of the distribution used to model the non-exceedance probabilities of the total flood depths.

8.3.1.3. Analysis of Simulation Results

The objective of this analysis was to demonstrate the range of flood depths that could result from different parameter values and different choices of distributions fitted to the total flood depths. For each of the three distributions evaluated, several flood depths were identified. These flood depths corresponded to the 2-year, 10-year, 25-year, 50-year, and 100-year events. For each flood depth, from the 3,000 simulation runs, the minimum depth, the 5^{th} percentile, the 10^{th} percentile, the 25^{th} percentile, the 50^{th} percentile, the 75^{th} percentile, the 90^{th} percentile, the 95^{th} percentile, and the maximum flood depth were identified. This provided information similar to a box-and-whisker plot about the likely range of flood depths that could occur, depending on the level of uncertainty in the parameters and distributions. This analysis also allowed for comparisons to be made across distributions, to determine whether certain combinations of distributions resulted in more or less uncertainty in

the resulting flood depths. The same range of flood depths was also identified based on the entire set of simulated data as well, in order to identify the full range of uncertainty in flood depths while taking into account the possible variations due to parameters and choice of distributions. This information was provided in tabular form and graphically. The results of the observed data set were plotted against the full set of uncertainty analyses to place the results of the observed data into context.

8.3.2. Results and Discussion

The flood depths that correspond to the 2-year, 10-year, 25-year, 50-year, and 100-year return periods were identified from each simulation run. For each of the three distributions fitted to the total flood depths, a range of flood depths was identified that corresponded to each return period of interest. The range of flood depths is presented for each distribution and then the combined results for all three distributions are compared to the observed flood depths.

8.3.2.1. Log-Pearson Type III Distribution

The flood depths that corresponded to the exceedance probabilities of 0.5, 0.1, 0.04, 0.02, and 0.01 were calculated based on the log-Pearson Type III distribution when three flood sources contributed to the total flood depths. Table 8-9 and Figure 8-5 illustrate the range of flood depths calculated for each return period. A wide range of flood depths could result based on uncertainties in the parameters used to calculate the flood depths for each individual source that then contributed to the total flood depths. The range was less dramatic for the smaller return periods, for instance, an exceedance probability of 0.5 than for the larger return periods such as an exceedance probability of 0.01. For the smaller events, the flood depths could range

from approximately 1 foot to 7 feet. However, for an event that could be expected to occur on average every two years, this is a fairly wide range. The largest event considered, the 100-year flood, could range in depth from 4.88 to 44.78 feet, which is a very wide range. The reasonableness of the larger predicted flood depths will be discussed in more detail shortly. From this information, it should be apparent that the parameters used to calculate the riverine, tidal, and pluvial flood depths can greatly impact the flood depths calculated for each source; thus, these parameters greatly impact the total flood depths determined based on all three flood sources. By impacting the total flood depths, the parameters also impact the distributions fitted to the total flood depths, so they can also greatly impact the flood depths predicted by the fitted distributions. Thus, care should be taken to identify the most appropriate parameter values prior to calculating the flood depths for each source.

Table 8-9: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the LP3 Distribution is Used to Represent the Total Flood Depths

	Exceedance Probabilities				
Flood Depths	0.5	0.1	0.04	0.02	0.01
Minimum	1.00	4.44	4.80	4.86	4.88
5th Percentile	1.74	6.11	7.03	7.33	7.46
10th Percentile	1.98	6.56	7.80	8.53	8.58
25th Percentile	2.27	7.72	10.08	11.23	12.20
50th Percentile	2.72	9.29	12.38	14.41	16.39
75th Percentile	3.31	11.06	15.48	18.95	22.88
90th Percentile	3.98	13.23	19.12	23.65	28.32
95th Percentile	4.47	14.99	22.18	25.73	29.86
Maximum	7.04	21.18	31.61	38.65	44.78

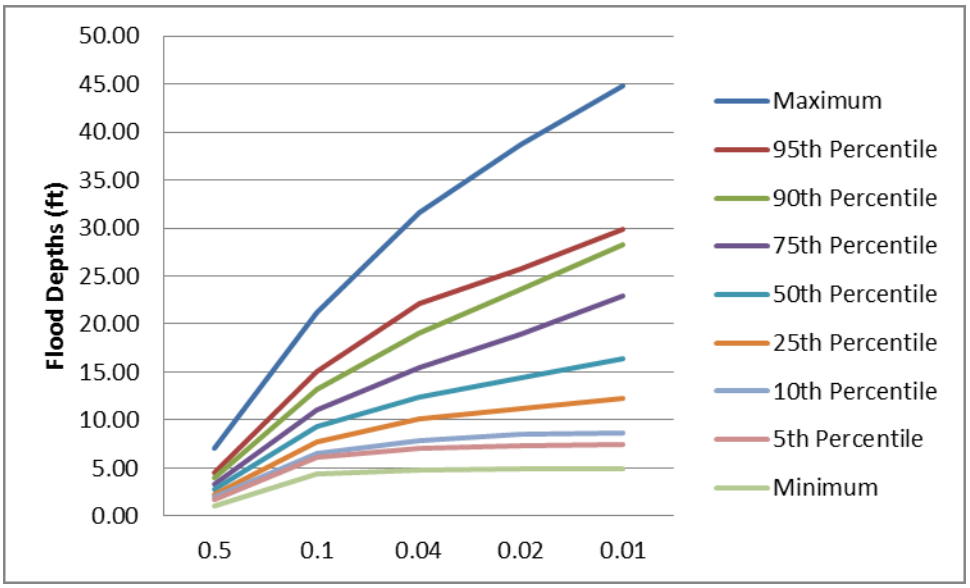

Figure 8-5: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the LP3 Distribution Fitted to Total Flood Depths

The maximum flood depths predicted for the 100-year flood event in the series of simulation runs conducted was nearly 45 feet. This does not appear to be a reasonable flood depth to occur at the location of interest. Unreasonably large predicted flood depths were also found in the process of conducting an uncertainty analysis for two flood sources. These unreasonably large flood depths were found to result from scenarios in which the distribution did not fit the total flood depth sample very well. These samples typically contained numerous very small flood depth values, typically due to zero-flood years in the riverine sample, and one or two comparatively large total flood depth samples. Typically, the maximum non-exceedance probability calculated for these samples was fairly low, on the order of 0.9, meaning the flood depths that corresponded to the larger flood events (for example a non-exceedance probability of 0.99) needed to be predicted based on the fitted distribution. The behavior of the upper tail of the distribution in these cases resulted in predictions of very large, unreasonable flood depths for the larger magnitude events. In the case of three flood sources, the pluvial sample also typically

had numerous zero-flood years, in addition to the riverine sample, which could further contribute to total flood depth samples with numerous very small depth values and just a few comparatively large depth values.

It is of interest to compare the range of predicted flood depths calculated based on consideration of three flood sources to those calculated based on consideration of two flood sources to determine how the additional flood source may have influenced the uncertainty in the analysis. For the smaller return periods (2-years, 10-years, and 25-years) the predicted flood depths based on three flood sources were larger. For the 2-year event the difference in flood depths was fairly small, but for the 10-year and 25-year events the difference in predicted flood depths could be up to several feet. For the two larger return periods (50-year and 100-year events) the differences in predicted flood depths were fairly small. However, the maximum flood depth predicted for each of these return periods based on three flood sources was actually smaller than the predicted flood depth based on two flood sources. Because these return periods were more likely to be predicted based on the upper tails of the fitted distributions when the distributions did not fit the total flood depth samples well, this suggests that the addition of the third flood source may actually make it somewhat easier to fit a distribution to the total flood depths. Even though the pluvial flood depth samples typically contained numerous zero-flood years, the addition of the third flood source may have helped round out the total flood depth sample so that it did not contain so many very low depth values along with a few comparatively large flood depth values, so the fitted distributions resulted in slightly less unreasonable upper tails. Overall, the addition of the third flood source does not

appear to have greatly influenced the level of uncertainty in the analyses, based on the range of predicted flood depths. This was not unexpected given that the pluvial flood depth sample typically did not contain very deep flood depths; thus, it did not greatly alter the magnitude of the total flood depth samples.

8.3.2.2. Generalized Extreme Value Distribution

The second distribution assessed was the GEV distribution. Table 8-10 and Figure 8-6 illustrate the range of flood depths that were calculated based on variations in the parameters used to calculate the riverine, tidal, and pluvial flood depths. A very wide range of flood depths was evident for all return periods of interest. For the 2-year event, the flood depths ranged from approximately 0.9 feet to 6.2 feet, which is a wide range of depths of an event that would be expected to occur approximately every 2 years. As the return period increased (or the exceedance probability decreased) the range of flood depths calculated also increased. For the 100-year event, the flood depths ranged from approximately 5 feet to above 80 feet, which is a very wide range. The reasonableness of the larger predicted flood depths will be discussed in more detail. Based on the ranges of flood depths calculated, it can be surmised that the parameters used to calculate the flood depths for the three flood sources would be very important and have great influence on the flood depths ultimately predicted to impact the location of interest.

Table 8-10: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the GEV Distribution is Used to Represent the Total Flood Depths

	Exceedance Probabilities				
Flood Depths	0.5	0.1	0.04	0.02	0.01
Minimum	0.91	3.86	4.41	4.72	4.97
5th Percentile	1.81	5.55	7.05	8.14	8.98
10th Percentile	1.96	5.94	8.00	9.24	10.91
25th Percentile	2.26	7.03	9.99	12.57	15.77
50th Percentile	2.71	8.32	13.06	17.41	22.61
75th Percentile	3.11	9.93	16.94	25.52	37.71
90th Percentile	3.80	11.31	21.31	34.52	54.65
95th Percentile	4.13	13.61	24.39	38.30	63.27
Maximum	6.21	19.35	32.61	51.26	82.62

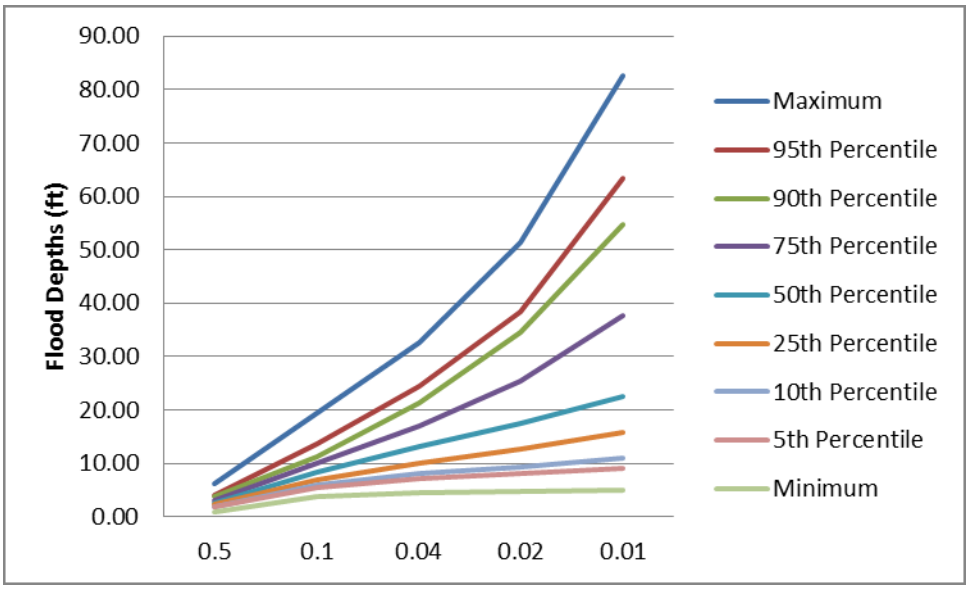

Figure 8-6: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the GEV Distribution Fitted to Total Flood Depths

The maximum flood depths predicted for the 100-year flood event based on the GEV distribution was above 80 feet, which was quite a bit larger than the maximum flood depth predicted by the LP3 distribution. The GEV distribution also resulted in the largest maximum predicted flood depths when only two flood sources were considered. These large flood depths were again attributed to poor fits to the total flood depth samples in these scenarios. When the total flood depth samples consisted of numerous very low depth values and a few comparatively large depth

values, the GEV distribution did not typically fit the total flood depth sample well, and the maximum non-exceedance probability was typically on the order of 0.9. Based on this, the flood depths that corresponded to the larger events had to be predicted outside the range of the total flood depth data. The upper tail of the GEV distribution typically increased sharply outside the range of the total flood depth data; thus, for the larger events the predicted flood depths tended to be unreasonably deep.

The predicted flood depths calculated based on the GEV distribution should be compared to the predicted flood depths calculated based on the LP3 distribution to determine whether one distribution introduced more uncertainty into the analysis than the other. The range of the flood depths predicted using the GEV distribution was quite a bit wider than the range of predicted flood depths based on the LP3 distribution. This suggests that the use of the GEV distribution resulted in greater uncertainty than the LP3 distribution. Differences in the behavior of the upper tails of the two distributions resulted in these differences in predicted flood depths. As this was also observed when two flood sources were considered, this result was expected.

It was also of interest to compare the results of the uncertainty analysis for three flood sources to the results for two flood sources to determine how the additional flood source impacted the level of uncertainty. The range of flood depths calculated for each return period was fairly similar between the analyses conducted for two flood sources and three flood sources. The maximum flood depth calculated for the 100-year flood event was slightly larger when three flood sources were considered than when two flood sources were considered. This suggested that the additional flood source may add some uncertainty to the analysis, which would be

expected given the additional uncertain parameters added to the analysis in the process of calculating the pluvial flood depths. However, the impact of the third flood source on the level of uncertainty seems to be fairly small overall. As was suggested when the results were assessed based on fitting the LP3 distribution to the total flood depths, because the pluvial flood depth samples were typically not very deep, they did not greatly impact the total flood depths calculated and therefore did not noticeably impact the level of uncertainty in the procedure.

8.3.2.3. Gamma Distribution

The final distribution assessed in the uncertainty analysis for three flood sources was the gamma distribution. Table 8-11 and Figure 8-7 demonstrate the range of flood depths calculated for each of the return periods of interest when the gamma distribution was fitted to the total flood depth samples. As has been previously observed, the uncertainty is larger for the larger return periods. For the 2-year event, the calculated flood depths range from slightly above 1 foot to 6 feet deep, while the flood depths calculated for the 100-year event range from approximately 8 feet to 43 feet deep. The range for each return period is quite deep, and for the larger return periods in particular, the maximum flood depths seem to be unreasonable. This suggests that the uncertainty in the parameters used to calculate the riverine, tidal, and pluvial flood depth samples greatly influence the flood depths predicted when the gamma distribution is fitted to the total flood depths.

Table 8-11: Flood Depths (in feet) Calculated for Multiple Return Periods to Assess Uncertainty in the Analysis Procedure When the Gamma Distribution is Used to Represent the Total Flood Depths

	Exceedance Probabilities				
Flood Depths	0.5	0.1	0.04	0.02	0.01
Minimum	1.18	4.49	5.97	7.06	8.14
5th Percentile	1.77	6.01	8.01	9.53	10.97
10th Percentile	1.96	6.39	8.80	10.48	12.00
25th Percentile	2.22	7.35	10.07	11.96	13.95
50th Percentile	2.57	8.64	12.11	14.70	17.34
75th Percentile	3.03	10.05	14.29	17.61	20.93
90th Percentile	3.59	11.93	17.33	21.14	24.95
95th Percentile	4.00	13.94	19.60	24.41	28.77
Maximum	6.02	21.04	29.72	36.31	42.93

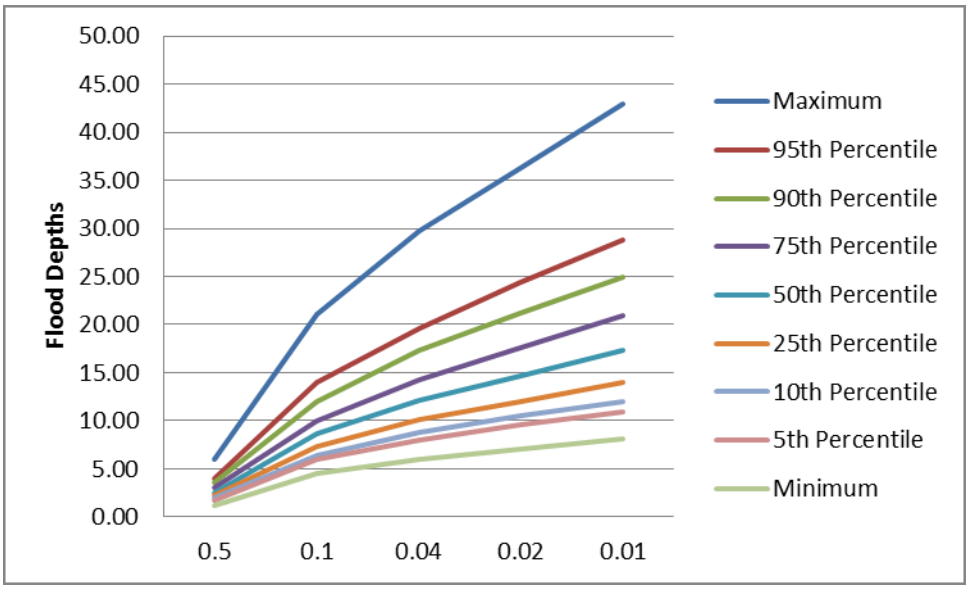

Figure 8-7: Plot Illustrating Variability in Expected Flood Depths for Various Exceedance Probabilities Based on the Gamma Distribution Fitted to Total Flood Depths

As was observed when the LP3 and GEV distributions were fitted to the total flood depths, the maximum predicted flood depths were unreasonably high, especially for the largest return periods. These high flood depths were the result of poor fits of the gamma distribution to the total flood depth sample simulated in these scenarios. In these scenarios, the maximum non-exceedance probability calculated was on the low side, perhaps on the order of 0.9. Because of this, the fitted gamma distribution

had to be used to predict the flood depths for the larger return periods outside the range of the simulated total depth data. The upper tail of the gamma distribution increased fairly steeply outside the range of the total depth data; which resulted in the large flood depths predicted in these scenarios.

The flood depths predicted based on the gamma distribution for three flood sources were compared to the predicted flood depths based on the LP3 and GEV distributions. The maximum flood depths predicted based on the gamma distribution were smaller than those predicted using either of the other distributions. The upper tail of the gamma distribution was not observed to rise as steeply as the upper tail of either the LP3 or GEV distributions, resulting in lower flood predicted flood depths. The gamma distribution was also observed to result in the lowest predicted flood depths when only two flood sources were considered as well.

Finally, the predicted flood depths calculated using the gamma distribution when three flood sources were considered were compared to the predicted flood depths based on the gamma distributions when only two flood sources were considered to determine whether or not the addition of the third flood source increases the level of uncertainty in the analysis. The maximum flood depths for each return period increased when the third flood source was added into consideration. This suggests that the addition of the third flood source does increase the level of uncertainty within the analysis. However, overall, the difference is fairly small between two flood sources and three flood sources. Because the pluvial flood depths are typically very small, they did not tend to greatly influence the total flood depths calculated, so they did not tend to influence the flood depths predicted by the

distribution fitted to the total flood depths. Thus, the additional uncertainty in the parameters used to calculate the pluvial flood depths did not greatly influence the final results of the analysis.

8.3.2.4. Consolidated Results of Uncertainty Analyses

In order to understand the combined effect of the parameters used to calculate the riverine, tidal, and pluvial flood depths and the choice of distribution to represent the total flood depths the results of all simulations were examined. Table 8-12 and Figure 8-8 demonstrate the range of flood depths calculated for each exceedance probability of interest for all three distributions. Table 8-8 and Figure 8-8 also compare the results of the simulations to the results of the observed data set. As expected, the range of flood depths for each exceedance probability was quite wide. The range of flood depths also increased for the larger return periods, as was previously observed. This suggests that uncertainties due to the parameters used to calculate the flood depths and the distribution used to represent the total flood depths have more impact on the larger return period events than on the smaller return period events.

Table 8-12: Range of Flood Depths (in feet) Calculated for Various Exceedance Probabilities for All Distributions

	Exceedance Probabilities				
Flood depths	0.5	0.1	0.04	0.02	0.01
Minimum	0.91	3.86	4.41	4.72	4.88
5th Percentile	1.75	5.78	7.31	8.11	8.87
10th Percentile	1.97	6.37	8.18	9.43	10.61
25th Percentile	2.26	7.35	10.05	11.96	13.77
50th Percentile	2.65	8.72	12.41	15.18	18.07
75th Percentile	3.15	10.31	15.50	19.87	24.55
90th Percentile	3.80	12.48	19.60	26.60	36.87
95th Percentile	4.22	14.17	22.26	31.98	47.26
Maximum	7.04	21.18	32.61	51.26	82.62
Observed	2.71	5.37	7.18	8.77	10.58

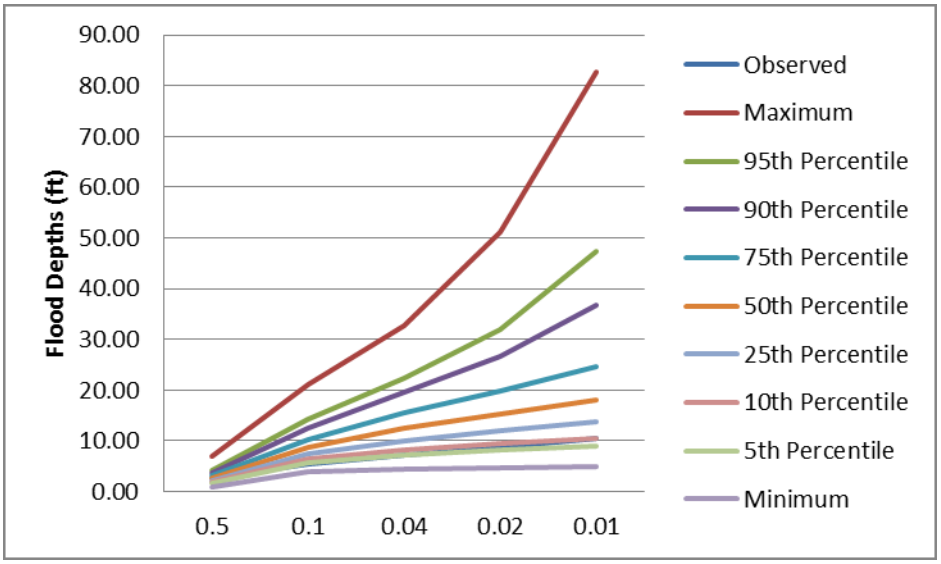

Figure 8-8: Plot Illustrating Range of Flood Depths for Various Exceedance Probabilities Calculated for All Distributions

As discussed for each individual distribution, the maximum flood depths predicted, especially for the larger return periods, seem quite unreasonable. For instance, a 100-year flood depth of 82 feet is clearly not possible. However, as was discussed for each of the three distributions individually, these high predicted flood depths were due to a poor fit between the distribution chosen to represent the total flood depths and the total flood depth sample in these scenarios. The maximum non-exceedance probability calculated in these scenarios was low, which meant that the

flood events that corresponded to the larger return periods had to be predicted based on the fitted distributions. When the flood depths had to be predicted outside the range of the total flood depths, the behavior of the upper tails of the distribution determined predicted flood depths. When the upper tails increased steeply outside the range of the data, the predicted flood depths tended to be unreasonably deep. This suggests that under certain circumstances, the method of fitting a distribution to the total flood depths may not produce a reasonable understanding of the total flood hazard at the location of interest. From the uncertainty analysis, it would appear that the when numerous zero-flood years in the riverine and pluvial flood depth samples resulted in a total flood depth sample with numerous very low values and several comparatively high flood depth values, the method developed to determine the combined flood hazard for the location of interest may be too sensitive to these factors, which makes extrapolation in the tails quite inaccurate.

Finally, the comprehensive results of the uncertainty analysis for three flood sources should be compared to the comprehensive results of the uncertainty analysis for two flood sources to determine the impact of the additional flood source. The range of flood depths predicted when three flood sources contribute to the total flood depths was quite similar to the range of flood depths predicted based on two flood sources. Though the addition of the third flood source introduces several additional uncertain parameters, the pluvial flood depths are typically so low that they do not greatly impact the total flood depths. Thus, the addition of the third flood source does not greatly impact the distribution fitted to the total flood depths or the flood depths predicted based on those fitted distributions.

8.4. Conclusions

An uncertainty analysis was conducted in order to understand the impact of various parameters and distribution choices on the final probabilities of occurrences of total flood depths. Similar analyses were conducted when the total flood depths were contributed to by both two and three flood sources. Based on these analyses, the choice of parameters and distribution could introduce significant uncertainty into the probabilities that correspond to total flood depths. In particular, the parameters used to calculate the riverine, tidal, and pluvial flood depths based on the gage measurements were found to result in wide variations in calculated flood depths. These variations then influenced the total flood depths calculated, as well as the distributions fitted to the total flood depths. Thus, it was determined that it was very important to obtain as accurate information as possible about the watershed and flow characteristics, in order to determine the parameters used to calculate the flood depths as accurately as possible.

The distributions fitted to the total flood depths in order to represent the non-exceedance probabilities corresponding to the total flood depths were also found to introduce uncertainty into the analyses. Clear guidance does not exist for selecting one distribution over another distribution to represent a set of flood depths. Of the three distributions evaluated in this uncertainty analysis, the gamma distribution resulted in the narrowest range of predicted flood depths for the exceedance probabilities of interest, while the GEV distribution resulted in the widest range of flood depths. Similar trends were observed for consideration of both two and three flood sources. The differences in distributions were especially evident for larger

return periods, for example the 100-year flood event. The different behaviors of the three distributions in the tails resulted in these differences. This fact must be considered carefully, as the larger return periods correspond to the larger flood events that are typically of the most interest and concern. Given the possible variation in predicted flood depths based on the choice of distribution used to represent the total flood depths, it would be very important to have the results of the copula analysis developed in Chapters 4 and 5 or the results of Equation 1-3, if the flood sources were determined to be independent, in order to confirm the accuracy of the fitted distribution. Further, care must be taken in using this methodology when the riverine or pluvial sample consists of a large number of low flood depths and a few high flood depth values, as the distributions were not observed to fit the total flood depth samples well in these cases. This resulted in the unreasonably high predicted flood depths for the larger return periods observed in some cases in the uncertainty analyses. In these cases, the flood depths predicted by the fitted distribution for each return period will not provide a reasonable understanding of the flood hazard at the location of interest.

CHAPTER 9

CONCLUSIONS AND RECOMMENDATIONS

9.1. Introduction

The research conducted and described in the previous chapters focused on assessing the flood hazard for a location of interest influenced by more than one flood source. This recognizes that two or three flood sources could individually or in some combination impact the location of interest. Simulation studies were first conducted to understand the impact of sample size and correlation between the samples. Then observed data for a location in Florida was used to develop a comprehensive flood hazard assessment and to illustrate the comprehensive flood risk for the location. Conclusions that resulted from this research and recommendations for further research are presented herein.

9.1.1. Need for the Research

Prior to conducting this research, a review of the literature was conducted to determine the current state of the art. Several methods of developing joint distributions for multiple flood sources were identified and considered. In evaluating these methods, the strengths and weaknesses of each were identified. Considering these strengths and weaknesses along with the goals of this research assisted in identifying the methodology that would be used and the improvements that would be needed in order to fulfill the objectives of the research.

9.1.1.1. State of the Art

Currently, flood hazard and flood risk assessments are undertaken for individual flood sources. Even for locations which may be impacted by multiple flood sources, the probability and risk of flooding is typically considered for each of those flood sources individually. For locations that could be impacted by more than one flood source, considering these sources individually does not necessarily provide a comprehensive understanding of the likelihood or risk of flooding.

The literature provides a few examples where flooding from multiple sources was considered. Morris (1982) outlined a method for combined-population flood hazard assessment. This method was applied to locations that were impacted by multiple populations of flooding, such as rainfall and snowmelt-caused flooding. This method could also be applied to locations that were impacted by multiple sources of flooding, such as riverine and tidal flooding. However, this method requires the assumption that the flood sources are independent of each other. This may or may not be a realistic assumption, depending on the specific location of interest. Loganathan *et al.* (1987) developed a method of determining joint probabilities for streamflows and tides in a tidal estuary. The riverine and tidal measurements were transformed to normal distributions using the Box-Cox transformation. A joint normal distribution was then used to represent the occurrence of various combinations of streamflow and tidal values. This method did not require an assumption of independence between the riverine and tidal flood sources. However, this method only provided the probabilities of combinations of riverine and

tidal events. It could not be used to provide the probability of a given flood depth that could be caused by many possible combinations of riverine and tidal flood depths.

Other methods are available for developing a joint distribution for multiple flood sources or flood characteristics. The method used by Loganathan *et al.* (1987) was limited in that it required the riverine flows and tidal heights to be represented by normal distributions, though the normal distribution is not typically used in hydrologic applications. Another method of developing joint distributions that has emerged in the hydrologic field in the last several decades makes use of copulas to develop joint distributions. Copulas have been used in numerous studies to develop joint distributions for various characteristics of flood events, such as peak flow, duration, and intensity of riverine flood events. Several studies have also evaluated the ability of copulas to model the joint distributions of riverine flows in locations near the confluence of two tributaries. Using copulas to develop these joint distributions has several advantages. First, they do not require an assumption that the flood sources are independent of each other. Copulas are, in fact, able to model the dependence between the flood sources. Also, copulas do not require the two flood sources to be represented by the same marginal distributions, so each flood source can be represented by the most appropriate marginal distribution. Based on these advantages, the copula method appears to have several strengths over the method proposed by Loganathan *et al.* (1987). However, the copulas still only provide a joint distribution representing the various combinations of riverine and tidal flood depths. They do not provide the probability of occurrence of a given flood depth of interest that may occur through many combinations of riverine and tidal flood depths.

9.1.1.2. Improvements Due to Research

The research conducted in this study expanded on the work described in Section 9.1.1.1 in several ways. The methodology developed in this research allowed a comprehensive flood hazard assessment to be developed for a location of interest that could be impacted by multiple flood sources. By using a copula equation to represent the joint distribution, each flood source could be represented by a different marginal distribution. Thus, the most appropriate distribution could be used to represent each flood source rather than requiring the normal distribution to represent each flood source. Also, in a review of the literature, examples where the copula was used to develop joint distributions corresponding to different flood sources were not found. The copula appears to have been primarily used to date in developing joint distributions based on different characteristics of riverine flood events.

Further, the literature has primarily focused on developing joint distributions. The work by Loganathan *et al.* (1987) developed the joint distribution of riverine flows and tidal heights, but did not take a step further to illustrate the impact of those combined events on a location of interest. The research conducted in this study used the joint distributions in order to determine the probability of a given flood depth that occurred at a location of interest while taking into account the many possible combinations of riverine and tidal flood depths that could have caused the flood depth of interest. While the equation presented in Morris (1982) could be used to calculate a probability of a certain flood depth occurring at the location of interest, taking into account that a flood depth could occur in multiple ways, this approach requires an assumption that the flood sources are independent of each other. By using the copula

approach, the methodology developed in this research can account for dependence between the flood sources.

9.2. Conclusions

A summary of conclusions drawn based on results of assessments conducted for both two flood sources and three flood sources are presented herein. The implications of these conclusions will be further discussed.

9.2.1. Two Flood Sources

Both riverine and tidal flood data were used to develop a comprehensive flood frequency assessment for a given location. The procedures developed were tested using both simulated and observed data. From the marginal distributions, which were the cumulative distributions fitted to the riverine and tidal simulated and observed samples, a copula equation was used to develop joint probability distributions and cumulative distributions, which provided the probability of various combinations of riverine and tidal flood depths that occurred jointly. From the joint distribution, a set of non-exceedance probabilities were calculated by taking the double integral under the joint distribution for various total flood depths. Probability distributions were fitted to these probabilities to represent the population of the total flood depths. The probabilities that correspond to total flood depths provided a comprehensive understanding of the likelihood of flooding when both riverine and tidal flood sources could impact the location of interest.

9.2.1.1. Assessment of the Flood Hazard

The primary objective of this research was to assess the flood hazard for a location of interest that could be impacted by more than one source of flooding. The

method developed in this research relied on the joint distributions developed for the two flood sources of interest. Based on the joint distributions, non-exceedance probabilities were calculated for total flood depths. Conclusions drawn from the development and evaluation of this procedure using both simulated and observed data are presented.

9.2.1.1.1. Simulation Studies

Because only short records of riverine and tidal measurements were available for the location of interest, simulation studies were first used to understand the effect of sample size. Since the observed data set had a sample size of only 17, it was important to understand how well the marginal distributions could fit a small sample. Further, because assessing the impact of correlation between the riverine and tidal samples would require investigations at more than one location of interest, which would introduce additional variables into the analyses, simulation studies also assessed the impact of correlation between the samples. After samples of varying size and level of correlation had been developed, marginal distributions were fitted to the samples. Trends in marginal distribution parameters fitted to the sample data were not observed as either sample size or level of correlation varied. Though a trend was not evident, sample size would be expected to impact the fitting of the marginal distributions. As the sample size increased, the sample would be expected to better represent the population, such that the fitted distribution would be more accurate. Though the level of correlation between the samples was not observed to result in trends in the marginal distribution parameters, the level of correlation would impact

the marginal distributions because the generated samples would need to be varied in order to meet the desired level of correlation.

The next step in the procedure was to use the copula equation to develop joint distributions for each generated sample. Sample size was not observed to impact the copula parameters calculated or the joint pdfs or cdfs. Because the copula parameter is calculated as a function of the ranks of the generated samples, it would not be expected that sample size would have an impact on the development of the copula parameters or the joint distributions. However, as with the marginal distributions, increasing sample size would be expected to lead to increasing accuracy in the joint distributions. The sample correlation was observed to impact the development of a joint distribution. As the correlation between the riverine and tidal generated samples increased the calculated value of Kendall's τ and the copula parameters were also observed to increase. Distinct differences were observed in the joint pdfs and cdfs as the level of correlation between the generated riverine and tidal flood depth samples increased. In plots of the joint pdfs, the peaks were observed to become steeper and narrower as the level of correlation increased. As the level of correlation was increased, the range of tidal flood depth values that could be generated to correspond to a given riverine flood depth value was constrained in order to meet the desired level of correlation between the samples. Thus, the narrower peaks in the pdfs indicate that for higher levels of correlation there is a smaller range of tidal flood depth values that could be expected to occur jointly with given riverine flood depth values. Further, as the level of correlation increased, the choice of copula family was also observed to have more significant impact on the resulting joint pdfs and cdfs

calculated. For low values of correlation between the samples, all three copula families could be used to represent the samples with similar results, while the joint distributions resulting from each copula family were observed to differ more for higher levels of correlation.

These results suggest that the level of correlation between the riverine and tidal samples has serious implications on the joint distributions developed for a set of flood depth data. High levels of correlation suggest stronger relationships between the riverine and tidal samples; thus, the variation in levels of correlation was expected to impact the joint distributions. It should be noted, though, that these results were based on only one scenario, and different trends and relationships may be evident in studies based on different scenarios. The joint pdf was used to calculate the non-exceedance probabilities that correspond to total flood depths; therefore, the level of correlation will ultimately influence the flood frequency predictions that correspond to the total flood depths. It was also important to understand the way that the level of correlation between the riverine and tidal samples influences the performance of each of the three copula families. As is typically the case in fitting a distribution to a set of data, numerous copula families could be used to represent the joint distribution, but none of the copula families can necessarily be considered the "correct" family to represent a given set of data. Several goodness-of-fit tests, such as the Akaike's Information Criterion or Q-Q plots can be used to determine the most appropriate copula family to represent a given set of data. For highly correlated data, the results obtained from the simulation study suggested that greater care must be taken in calculating these goodness-of-fit measures and in interpreting their results.

The final step of the analysis was to calculate non-exceedance probabilities that correspond to the total flood depths. These were calculated by taking the double integral under the joint pdfs. These non-exceedance probabilities were then compared to distributions fitted to the total flood depths using the method of moments and maximum likelihood estimation methods. The distribution that agreed most closely with the non-exceedance probabilities calculated using the double integral was used to represent the probabilities of the total flood depths for the location of interest. Sample size, which had not been observed to result in any trends in the fitting of marginal distributions or the calculation of the joint distributions, was not observed to result in any trends in the fitting of distributions to the total flood depths. Sample size would be expected to impact the accuracy of the fitted distributions, as discussed in the fitting of the marginal distributions and the joint distributions. The level of correlation between the riverine and tidal flood sources was observed to result in trends in the distributions fitted to the total flood depths. These trends were observed in the fitted parameters for each of the distributions evaluated to represent the total flood depths. Despite these trends in the parameters, calculation of the flood depths that corresponded to several common exceedance probabilities using each distribution did not illustrate significant differences in the calculated flood depths as the correlation varied. This suggests that the trends in the distribution parameters did not have a significant impact on the probabilities calculated for the total flood depths.

These results suggest that the degree of correlation between the samples of riverine and tidal flood depths could influence the exceedance probabilities of the total flood depths for a location of interest. Since flood frequency analyses are used

to design flood protection systems, the correlation between the riverine and tidal flood sources would then influence the flood protection system designed for a location of interest. Prior to fitting a distribution to the total flood depths to assess the flood hazard for a location it would be important to understand the relationship between the riverine and tidal flood depth samples. It should be noted that these studies were exploratory in nature, and based on only one location. The trends, or lack thereof, based on either sample size or level of correlation would likely differ in studies based on a different location. In particular, different relationships between the flood sources for a different location may influence the size, shape, and location of the joint pdf peaks, leading to different results than those discussed in this study. These results suggest that there would be benefit to evaluating these methods further; though strong conclusions about the performance of the methodology cannot be drawn based on the studies conducted in this research.

9.2.1.1.2. Observed Data

The flood hazard was assessed for a location of interest in Florida based on observed riverine and tidal flood depth data. Flood depths were calculated for each source based on observed gage measurements. In calculating the flood depths for both sources, flood events were determined not to have impacted the location of interest in numerous years of the record. Because it would not be expected that this location of interest would flood every year these zero-flood years were determined to be rational from a hydrologic perspective. However, the zero-flood years did make fitting marginal distributions challenging. Because the log-Pearson Type III and lognormal distributions were based on logarithms of the computed flood depths, and

could not be used to represent the zero-flood years, a small value of approximately 1% of the maximum flood depth was added to the riverine flood record (Jennings and Benson, 1969), so that all four of the distributions considered in the simulation studies could also be considered to represent the observed data. Even with this adjustment, it was very difficult to obtain adequate fits to the observed data. Approximately one-half of the flood depths transposed from the riverine record were zero-flood years, while approximately one-quarter of the flood depths transposed from the tidal record were zero-flood years. Attempts to fit a distribution to a set of data that consists of a large number of identical values do not result in as accurate a fit. This is a challenge that is expected to occur at any location for which this method was applied, as there is almost no location outside of the floodplain that could be expected to flood on an annual basis.

Once marginal distributions had been fitted to the two flood sources, the joint distributions were calculated using the copula equations. Given the sample size and the level of correlation between the observed riverine and tidal flood depth samples, the copula parameter calculated for the observed data agreed well with the results and trends observed in the simulation studies. However, the copula parameter calculation would also have been impacted by the number of zero-flood years in the observed data set. The copula parameter is dependent on the ranks of the riverine and tidal flood depth samples, but the zero-flood years would all receive the same rank. Thus, the large number of zero-flood years may also make it more difficult to fit a joint distribution to the observed data in addition to making it difficult to fit the marginal distributions. The observed riverine and tidal flood sources did not have a

strong relationship. In fact, high flood depths from one of the two flood sources typically occurred with low, or even zero flood depths from the other source. Comparing the joint distributions calculated in the simulation studies to the observed data it was determined that the joint distributions were better determined when a stronger relationship existed between the flood sources. This result would be expected because a stronger relationship between the flood sources suggests that they are more likely to occur jointly, making it easier to define a joint distribution for the flood sources.

Based on the joint pdf calculated for the observed data, the non-exceedance probabilities that correspond to the total flood depths were calculated by taking the double integral under the joint pdf. Several probability distributions were fitted to the total flood depths and compared to these non-exceedance probabilities in order to determine the distribution that best represented the total flood depths. In the simulation studies, excellent agreements were typically observed between the fitted distributions and the non-exceedance probabilities calculated using the double integral under the joint pdf. However, for the observed data, the agreement between the fitted distributions and the non-exceedance probabilities calculated using the double integral method was not observed to be as good. The poor agreement was most evident in the upper tail of the distributions, where the distributions predicted a larger non-exceedance probability than did the double integral method for a given flood depth. The influence that the zero-flood years had on the joint distribution, as well as the lack of a strong relationship between the riverine and tidal flood sources are believed to have influenced the non-exceedance probabilities calculated using the

double integral under the joint pdf. This suggests that prior to undertaking this method, an examination of the relationship between the flood sources should be conducted. If the relationship between the flood sources is found to be weak, the copula method may not be the most appropriate approach for determining the probabilities that correspond to the total flood depths for the location of interest. As will be discussed in Section 9.2.2.2., a weak relationship between the flood sources may suggest that they can be considered to be independent of each other, in which case Equation 1-3 can be used instead of the copula approach to determine the probabilities that correspond to the total flood depths.

9.2.1.2. Assumption of Independence or Dependence Between Flood Sources

For the observed data, flood hazard assessments were made based on both the individual annual maximum riverine and tidal flood depths as well based on consideration of both contributing flood sources. Two methods of calculating flood hazard based on both flood sources were considered. The first method, which used the copula, was based on the assumption that the flood sources were dependent, while the second method was based on the assumption that the flood sources were independent of each other. These analyses illustrated that if only one of the potential flood sources was considered, a much different understanding of the flood hazard would occur. The probability of riverine flooding was found to differ widely from the probability of tidal flooding. For lower flood depths, the exceedance probabilities differed significantly between the two flood sources, though they converged toward very low exceedance probabilities as the flood depths increase. The general curvature of the exceedance probabilities also differed widely between the flood sources. The

exceedance probabilities calculated based on both flood sources differed from both independent samples, but appeared to have a mixture of the characteristics of both independent sources. This suggests that considering both flood sources does, in fact, provide a more comprehensive understanding of the probability of flooding for a given location, as desired.

The assumption of independence or dependence between the two flood sources, when considering them jointly, was not found to significantly impact the calculated exceedance probabilities. Slight differences in exceedance probabilities were observed in the range of three to five feet of total flood depth, over which range the difference in slope of the exceedance probabilities of the independent riverine and tidal flood depths was most apparent. As previously discussed, a strong relationship did not exist between the observed riverine and tidal flood depth samples. Thus, the observed data for these two flood sources could be considered to be independent of each other. For cases in which the flood sources do not have a strong relationship, the more straightforward and less complex method of determining the probability of total flood depths, using Equation 1-3, can be used to assess the flood hazard for the location of interest. This method does not require a measure of the dependence between the flood sources, such as Kendall's τ, nor does it require the calculation of the joint distributions based on the copula equations. However, it is expected that in many locations a strong relationship would exist between the two flood sources. It is expected that in such situations a more significant difference would be evident in the exceedance probabilities calculated based on the assumptions of independence and dependence between the flood sources. Thus, in order to determine the most

appropriate method of calculating the probabilities that correspond to the total flood depths, and to determine whether or not an assumption of independence or dependence should be made, the relationship between the flood sources should be evaluated by calculating the correlation between the riverine and tidal annual maximum flood depth samples. If a low correlation is calculated, the assumption of independence is likely acceptable, and the more simple method can be used, while if a moderate correlation is calculated, the assumption of dependence should likely be made and the method based on the copula joint distributions should be used.

9.2.1.3. Flood Risk Calculations

After the assessment of the flood hazard was completed, the flood risk was calculated for the simulated and observed data sets. For the observed data set, the flood risk was calculated for each of four scenarios (each flood source individually and both methods of evaluating the probability of joint flood events). The effects of sample size and sample correlation on the flood risk calculations were investigated with the simulation study. The sample size was not observed to have an impact on the flood risk calculations, as sample size was not observed to impact the assessment of the flood hazard. However, as was discussed for the assessment of the flood hazard, the larger sample sizes can be expected to be more accurate representations of the flood risk, due to an expected higher accuracy of the calculated probabilities of flooding. The degree of correlation between the riverine and tidal samples was observed to impact the flood risk calculations. The flood risk was observed to increase slightly as the level of correlation between the samples increased. This impact related to the impact of varying sample correlation on the exceedance

probabilities. The flood risk calculations based on the observed data were found to agree well with those calculated for the simulation studies, given the trends that were identified through the simulation studies.

As expected based on the comparisons made of the flood hazard, considering only one of the flood sources in risk calculations was found to produce quite different understandings of the risk than considering both of the flood sources. These trends were calculated based on only one location of interest, and it is expected that if the developed methodology were used for a different location, different trends in the results would be observed. Either method of considering both flood sources jointly would result in a more complete understanding of the flood risk for the location of interest; however, the assumption of independence or dependence was not observed to influence the resulting risk calculations. The flood risks calculated based on the exceedance probabilities calculated using both assumptions were observed to be very similar. Because the assumption of dependence or independence was not observed to significantly influence the exceedance probabilities calculated, it was not expected that the assumption would influence the flood risk calculations. Thus, the riverine and tidal flood depth data should be examined prior to use in calculation of the flood hazard or flood risk. This assessment should evaluate whether or not a relationship exists between the two flood sources. Based on this assessment, it should be possible to determine whether or not the assumption of independence, which results in simpler and more straightforward calculations, can be justified. If a relationship is observed between the two flood sources, the method developed in this research, based on the

copula, should be used to calculate the likelihood of flooding and the resulting risk of flooding.

9.2.1.4. Uncertainty Analysis

An uncertainty analysis was conducted to determine the impact of the choice of various parameters and distributions on the results of the flood hazard assessment. In order to assess the flood hazard based on multiple flood sources for the location of interest, the flood depths from each source had to be calculated, which required estimating several parameters. Though these parameters were estimated with the best information possible, a level of uncertainty in them existed, which could lead to uncertainty in the calculated probabilities that correspond to total flood depths. Probability distributions also had to be fitted to the total flood depths in order to estimate the non-exceedance probabilities that correspond to the total flood depths. Numerous probability distributions could be used in this case and there is not clear guidance as to which distribution is the most appropriate distribution.

The uncertainty analysis conducted for two flood sources evaluated the impact of the uncertainty in these parameters and the choice of distributions to represent the total flood depths by calculating the 2-year, 10-year, 25-year, 50-year, and 100-year flood depths based on each of three distribution choices for numerous simulations in which the flood depths were calculated based on different sets of parameters. The uncertainty analyses demonstrated the importance of choosing these parameter values and distributions. The parameters used to calculate the flood depths were observed to result in wide variations in the calculated flood depths, as would be expected. The calculated values of riverine and tidal flood depths then influenced the

total flood depths, which ultimately impacted the distributions fitted to the total flood depths. Thus, care should be taken to obtain as accurate values as possible for these parameters needed to determine the riverine and tidal flood depths.

The choice of distribution fitted to the total flood depth was also observed to have a significant effect on the flood depths calculated for the return periods of interest. Of the three distributions evaluated in this uncertainty analysis, the gamma distribution resulted in the narrowest range of predicted flood depths for the exceedance probabilities of interested, while the GEV distribution resulted in the widest range of flood depths. The differences in distributions were especially evident for larger return periods, for example the 100-year flood event. The different behaviors of the three distributions in the tails resulted in these differences. This is a significant factor to consider as it is the tails of the distributions that contain the larger flood events typically of most concern. Clear guidance does not exist as to which distribution should be chosen to represent a given set of data, but numerous goodness-of-fit tests can be used to assist in identifying the distribution most able to represent the data. The results of these goodness-of-fit tests should be carefully evaluated and considered when choosing a distribution to represent the total flood depths to ensure that the distribution represents the total flood depths as accurately as possible. The results of the copula procedure or Equation 1-3 are also critical for comparing the results of the distribution fitting to the estimated non-exceedance probabilities for the total flood depths. This analysis was carried out based on only one location of interest. If similar analyses were carried out for other locations, it is likely that each

of these factors would have had a different influence on the overall uncertainty in the analysis.

The uncertainty analyses also illustrated that, under certain conditions, the developed method does not result in reasonable predicted flood depths that correspond to return periods of interest. For all three distributions considered, the maximum flood depths identified for each return period were excessively large and clearly unreasonable for the watershed. In evaluating the conditions under which these large flood depths were predicted, it was determined that particularly poor fits between the distributions and the generated total flood depth samples were responsible. In some cases, the maximum non-exceedance probability calculated when the distributions were fitted to the sample were quite low, for instance approximately 0.9, which meant that the predicted flood depths for the larger return events had to be predicted outside the range of the total flood depth sample. Thus, the behavior of the upper tail of the distribution was responsible for the predicted flood depths for these return periods. The poor fit between the distributions and the total flood depths seemed most likely to occur when the total flood depth sample consisted of numerous very low values and just a few relatively large values. The zero-flood years, particularly in the riverine sample, seemed to be responsible for creating the total flood depth samples that could not be easily fit by the desired distributions. This suggests that the procedure developed may not produce reliable predicted flood depths when the riverine and tidal samples have a particularly large number of zero-flood years.

9.2.2. Three Flood Sources

Once the flood hazard based on two flood sources was understood the flood hazard based on three flood sources was next investigated. A comprehensive flood frequency assessment was developed for the location of interest based on a riverine, tidal, and pluvial flood source. The procedure was tested using both simulated and observed data. Cumulative distributions were fitted to the riverine, tidal, and pluvial flood depth samples to provide the marginal distributions. These distributions were used as input to the copula equations to develop joint distributions that provided the probability of various combinations of riverine, tidal, and pluvial flood depths that occurred jointly. Non-exceedance probabilities that corresponded to total flood depths were calculated by using the triple integral under the joint distribution, and these probabilities were compared to distributions fitted to the total flood depths. These probabilities provided a comprehensive understanding of the likelihood of flooding when riverine, tidal, and pluvial flood sources could all impact the location of interest.

9.2.2.1. Assessment of the Flood Hazard

The aim of this research was to provide a comprehensive flood hazard assessment for a location of interest based on three possible flood sources. This comprehensive flood hazard assessment was based on the joint distributions developed for each of the three flood sources. Conclusions drawn from the development and evaluation of this procedure using simulated and observed data are presented.

9.2.2.1.1. Simulation Studies

The first step to conducting this research was to test the developed procedure using simulated data sets. The simulation studies evaluated two factors that could not be easily assessed using observed data; the sample size and the level of correlation between the three flood depth samples. The observed data set obtained for the location of interest in Florida consisted of only 17 years of data, which is a small sample. Thus, the simulation studies assisted in understanding the impact that this small sample could have on the results of the assessment. Evaluating the impact of the level of correlation between the flood depth samples would require investigations at more than one location of interest, which would introduce additional variables into the analyses. Thus, the simulation studies assisted in understanding the impact that the level of correlation that existed between the observed data sets had on the results of the assessment. After samples of varying size and level or correlation had been generated, the first step to the procedure was to fit marginal distributions to each generated flood depth sample. Trends in the marginal distribution parameters fitted were not observed based on either sample size or level of correlation. Though a trend in the distribution parameters was not observed, sample size was observed to have some influence on the distributions fitted to the samples. As the sample size increased, the fitted distributions appeared to more accurate represent the generated data, as would be expected given that larger sample sizes provide a better representation of the population.

For each set of generated riverine, tidal, and pluvial flood depth samples, the next step in the procedure was to use copula equations to develop joint distributions.

A general trend was observed in the calculation of the copula parameters, that as the sample size increased, the copula parameter value generally decreased. However, as was observed when only two flood sources were considered, the standard error of the copula parameters was quite wide, such that the variation in the copula parameters for these scenarios did not result in significant variations in the developed joint distributions. For the scenarios in which the level of correlation was varied, it was observed that as the level of correlation between the flood depth samples decreased, the copula parameter also decreased. However, the differences in copula parameters for these scenarios were within the standard error of the copula parameters, such that the variations in copula parameters had minimal impact on the developed joint distributions.

The choice of copula family used to represent the samples was also observed to have minimal impact on the joint distributions developed. Both of these findings differed from the results observed when only two flood sources were considered. In that case, as the level of correlation between the samples increased, the copula parameters were observed to vary significantly, such that the developed joint distributions did vary. Further, for two flood sources, as the level of correlation increased, the choice of copula family was observed to impact the joint distribution developed. However, a wider range of levels of correlation were examined for two flood sources than for three flood sources. Because a slight trend was observed in the analysis of three flood sources as the level of correlation varied, it is expected that a more significant difference in the copula parameters and the joint distributions developed would be evident for higher levels of correlation.

The final step to the analysis was to calculate non-exceedance probabilities that corresponded to the total flood depths. The triple integral under the joint pdf was used to calculate these probabilities, which were then compared to the cdfs fitted to the total flood depths. Based on this comparison, a distribution was selected to represent the total flood depths. Sample size was observed to impact the results of this step to some degree. It was somewhat difficult to fit distributions to the sample of 10 total flood depths, due to the small sample size. Further, the four distributions considered to represent the total flood depths were observed to result in quite different estimates of the probabilities that corresponded to given total flood depth values. However, once the sample size increased to at least 25, the distributions were observed to fit the total flood depth samples very well. Further, the LP3, GEV, and gamma distributions were observed to fit the samples almost identically, such that any of the three distributions would be adequate to represent the sample. This suggests that once the sample size was large enough to be considered accurate, sample size had minimal impact on the results of the analysis. A slight trend in the fitted distributions was also observed as the level of correlation between the samples varied. As the level of correlation decreased, the gamma distribution shape parameter increased and the scale parameter decreased. These differences were minimal and not observed to significantly impact the calculated non-exceedance probabilities. However, a more significant impact on the distribution fitted to the total flood depths might be observed if a wider range of correlations were evaluated.

These results have assisted in understanding the impact that the sample size and the level of correlation between the samples would have on the flood hazard

assessment developed based on three flood sources. These findings will greatly clarify the results obtained using observed data for the location of interest in Florida. These findings suggest that the results obtained from a small sample should be examined very carefully, as they may not be accurate enough to be used in designing flood protection systems. They also suggest that the level of correlation between the samples may impact the understanding of the flood hazard, especially for higher levels of correlation, so the relationship between the three flood depth samples should be understood prior to using the developed procedure. It should be noted that the trends, or lack of trends, based on sample size or level of correlation identified by this research were based on conditions at a single location. This research was exploratory in nature, and thus considered only one location. Different conditions and different relationships between the flood sources at other locations would likely result in different trends in the results if similar studies were carried out for other locations.

9.2.2.1.2. Observed Data

Once the simulation studies had been completed to better understand the impact of the sample size and level of correlation between the samples, the flood hazard was assessed for a location in Florida based on observed riverine, tidal, and pluvial flood depth samples. In calculating the flood depths for each source, several zero-flood years were identified for each source. From a hydrological perspective, these zero-flood years appeared rational, as few locations would be expected to flood on an annual basis. As was discussed when two flood sources were considered, the log-Pearson Type III and lognormal distributions were not able to represent data which included zero-values, so a small value, of approximately 1% of the maximum

observed flood depth (Jennings and Benson, 1969), was added to the riverine and pluvial samples so that all four distributions could be considered to represent these samples. The distributions considered to represent the tidal flood depth sample were able to represent zero-values, so an adjustment was not made to this sample. Even with these adjustments, the zero-flood years made it quite difficult to obtain adequate fits to the observed data. Approximately one-half of the flood depths transposed from the riverine and pluvial records were zero-flood years and approximately one-quarter of the flood depths transposed from the tidal record were zero-flood years. Attempts to fit distributions to samples with such a high percentage of identical values do not result in as accurate a fit. This challenge is expected to be encountered in any location for which this procedure is used, as there are few locations that would be expected to flood on an annual basis.

Once marginal distributions had been fitted to the observed flood depth samples, the joint distributions were calculated using the copula equations. The copula parameters calculated were reasonable, but smaller than those typically calculated for the generated samples. They were most similar to the copula parameters calculated for the scenarios in which low correlation existed between the flood depth samples (scenarios RLA and RLHL), which most resembled the observed data. The calculated values of Akaike's Information Criteria, which were used to assess the fit of the three copula families, were much higher than those calculated in the simulation studies, which suggests that greater variation resulted when fitting the copulas to the observed data than it was with the generated data. Due to the inclusion of the zero-flood years in the observed data sets, the copula parameter was calculated

for the observed data based on numerous identical values. This likely contributed to the greater variation in fitting the joint distributions to the observed data. The joint distributions developed based on the observed data appeared reasonable, given the results previously obtained in the simulation studies. The choice of copula family was observed to have minimal influence on the developed joint distributions.

The probabilities that correspond to the total flood depths were determined based on the joint distributions obtained using the copulas. Non-exceedance probabilities were then calculated by taking the triple integral under the joint pdf developed based on the copula. These probabilities were compared to the cdfs of distributions fitted to the total flood depths, and a distribution was then selected to represent the total flood depths. In the simulation studies, the non-exceedance probabilities calculated by the triple integral agreed very well with the cdfs fitted to the total flood depths. This was not the case of the observed flood depths. The maximum non-exceedance probabilities calculated using the triple integral were quite low, such that the upper tail of the fitted distributions did not agree well at all. This was also observed when only two flood sources were considered, and it was attributed primarily to the large number of zero-flood years in the samples. Because the zero-flood years impacted the calculation of the copula parameter, they also impacted the development of the joint pdf and cdf. The less accurate joint pdf resulted in less accurate calculations of the non-exceedance probability through the triple integral procedure. This suggests that the number of zero-flood years in the data could have a very significant impact on the results of an analysis, and this factor must be carefully

considered prior to using the developed method, as it is unlikely that there would be a location of interest for which zero-flood years did not exist.

9.2.2.2. Assumption of Independence or Dependence Between Flood Sources

Based on the observed data, flood hazard assessments were made based on the assumptions that the flood sources were independent of and dependent on each other. The purpose of this assessment was to determine the importance of this assumption on the resulting flood hazard assessment. When the flood sources were assumed to be dependent on each other, the copula procedure developed and demonstrated in Chapters 6 and 7 was used in the assessment of the flood hazard. When the flood sources were assumed to be independent of each other, the flood hazard assessment was based on Equation 7-19. These analyses indicated the importance of considering all possible flood sources that could impact the location of interest. If only one of the flood sources was considered, a much different understanding of the flood hazard was developed. This conclusion was drawn based on studies of only one location, and further studies for different locations would be beneficial to demonstrate the overall validity of the conclusion. Different conditions would exist for different locations, which might result in different trends in the results.

Regardless of which assumption was made, the consideration of all three flood sources was observed to result in a more complete understanding of the flood hazard for the location of interest than only considering the individual flood sources. The non-exceedance probabilities calculated using both methods were compared and a wide difference was observed for some of the mid-range flood depths in particular. For low and high flood depths, the non-exceedance probabilities calculated using both

methods were very similar. This suggests that the assumption of dependence or independence does have the potential to significantly impact the resulting understanding of the flood hazard. The difference in non-exceedance probabilities based on these assumptions was much greater than was observed when only two flood sources were considered. The riverine and tidal flood sources, which were considered in the two source analysis, were not observed to have a strong relationship between them, which was believed to be responsible for the minimal difference in results for each assumption. However, when the pluvial flood source was added in for consideration, a strong relationship was observed to exist between the riverine and pluvial flood sources. This relationship likely explains why the assumption of independence or dependence was observed to have a more significant impact for three flood sources. This suggests that the relationship between the three flood sources should be understood prior to using the developed method. If a strong relationship does not exist between the three flood sources, it may be more appropriate to use Equation 7-19, which is a simpler method than the developed copula method, to develop an understanding of the flood hazard based on the three flood sources. However, these results and conclusions are based on analysis of only a single location and they need to be confirmed through studies at other locations.

9.2.2.3. Flood Risk Calculations

After the flood hazard had been assessed, the flood risk was calculated for both the simulated and observed data sets. The purpose of this was to demonstrate the impact that using a comprehensive flood hazard assessment could have on the understanding of the flood risk at the location of interest. The impacts of sample size

and level of correlation between the samples were investigated using the simulation studies. The flood risk calculations for the smallest sample of 10 events were observed to differ widely from the flood risk calculations for the larger samples. This was expected given the differences observed in the non-exceedance probabilities calculated for a sample of 10 as compared to the larger samples. However, once the sample size reached or exceeded 25, the flood risk calculations were virtually identical. This suggests that once the sample size is large enough to be considered accurate, the impact of sample size on the calculated flood risk was negligible. A slight trend in the flood risk calculations was also attributed to the level of correlation between the samples. As the level of correlation between the samples decreased the flood risk calculations were also observed to decrease. However, the impact on flood risk calculations for the range of correlations investigated was minimal. A more significant trend might emerge if a wider range of correlations were examined.

Based on the observed data, flood risk was calculated for each flood source independently and for three flood sources jointly based on the assumptions of independence and dependence between them. The three flood sources considered independently resulted in much different understandings of the flood risk for the location of interest, as would be expected given the differences in probabilities calculated for each. Considering all three flood sources jointly, whether independently or dependently, did result in a more comprehensive understanding of the flood risk for the location of interest. However, it was clear that the assumption made could greatly impact the resulting understanding of the flood risk. The differences in probabilities calculated using these assumptions, discussed in Section

9.2.2.2. were magnified when combined with the vulnerability and consequence data. This reinforces the previous conclusion that the relationships between the three flood sources should be understood prior to choosing the method to use in developing the flood hazard assessment. If strong relationships are not evident between the flood sources, using equation 7-19 to develop the flood hazard assessment and resulting flood risk calculations would likely be adequate; however, if strong relationships are evident, the copula method developed in this research would likely be necessary. It should be noted, though, that these conclusions are based on analysis of a single location, and further studies at different locations under different conditions would assist in confirming the validity of these conclusions.

9.2.2.4. Uncertainty Analyses

Uncertainty analyses were also conducted for scenarios in which three flood sources could jointly influence a location of interest. These uncertainty analyses assessed the impact of the parameters used to calculate the riverine, tidal, and pluvial flood depths based on simulated gage measurements for each flood source. Though the best possible information was used in determining the values to use for each of these parameters they all included some level of uncertainty that could impact the calculation of the flood depths and the final results of the analyses. The uncertainty analyses also assessed the impact of the choice of probability distribution chosen to represent the total flood depths, as numerous distributions could be chosen without clear evidence that any one distribution is better than the other possible distributions. To assess the level of uncertainty in the analyses, a total of 3,000 riverine, tidal, and pluvial flood depth samples were simulated based on differing parameter values and

three probability distributions were fitted to the total flood depths. Based on these three distributions, a range of the flood depths calculated for exceedance probabilities of 0.5, 0.1, 0.04, 0.02, and 0.01 (return periods of 2-years, 10-years, 25-years, 50-years, and 100-years) were identified.

A wide range of flood depths were predicted for each return period of interest for all three of the distributions considered. This suggests that the parameters chosen for use in calculating the three flood depth samples could introduce uncertainty into the analyses. Thus, it is very important that the parameters selected for calculation of the flood depths be as accurate as possible. The choice of distribution was also determined to introduce a significant amount of uncertainty into the analyses. The GEV distribution was observed to result in the widest range of predicted flood depths, indicating that it introduced more uncertainty than either of the other distributions, while the gamma distribution was observed to result in the narrowest range of flood depths. Clear guidance does not exist as to which distribution is most appropriate, especially considering that the total flood depths being fitted were influenced by three different flood sources that are typically represented by different distributions. The results of the uncertainty analyses indicate that the choice of distribution to represent the total flood depths is very important. The results and conclusions drawn from this uncertainty analysis were based on the conditions existing at a single location. Different factors may need to be considered for different locations, and each factor may have different impact on the uncertainty in the procedure for a different location.

The maximum flood depths predicted for each return period, especially for the higher return periods, were unreasonably high. This was also observed when only

two flood sources were considered. The high flood depths were determined to occur in scenarios in which the distributions were unable to fit to the total flood depth sample well. This was particularly common when the total flood depth sample consisted of numerous very low depth values, primarily due to zero-flood years in the riverine and pluvial depth samples, with one or two comparatively large flood depth values as well. The distributions were not observed to fit these samples very well, resulting in maximum non-exceedance probabilities of approximately 0.9. Thus, the predicted flood depths that corresponded to the larger return periods had to be predicted based on the fitted distributions outside the range of the data. The upper tails of the distributions tended to rise fairly steeply outside the range of the total depth samples, which resulted in the very large predicted flood depths. This suggests that in scenarios in which the riverine and pluvial flood depth samples consist of numerous zero-flood years and one or two large flood depths, the method developed in this research may not be appropriate. In these scenarios, it is expected that it would not be possible to fit an accurate distribution to represent the total flood depths based on all three flood sources.

The results of the uncertainty analyses conducted based on three flood sources were also compared to those conducted based on two flood sources to determine what impact the addition of the third flood source may have on the level of uncertainty in the analyses. The range of flood depths predicted for two flood sources versus three flood sources were quite similar overall. This suggests that the addition of the third flood source did not significantly increase the level of uncertainty in the analyses. Though the addition of the pluvial flood source introduced several more uncertain

parameters to the analysis procedure, the pluvial flood source did not significantly impact the results because the pluvial flood depths tended to be very low. Thus, the pluvial flood depth sample did not greatly influence the total flood depths or the distributions fitted to the total flood depths. Thus, the flood depths predicted based on the fitted distributions did not vary significantly when three flood sources were considered as compared to when two flood sources were considered.

9.3. Recommendations

The research that has been conducted is only the starting point in developing a comprehensive understanding of the likelihood of flooding influenced by multiple flood sources. Many questions bear further investigation to better understand the ability of this method to provide a more reasonable understanding of flood hazard and flood risk. Recommendations for further research are provided herein.

The first step of this research was to evaluate the likelihood of flooding at a location that could be influenced by two or three different flood sources. A procedure was developed to assess the flood hazard at this location; however, the methodology was only applied to and evaluated at this one location. Thus, the first general recommendation would be to apply the method developed at other locations to assess its performance under different conditions. In particular, it would be interesting to assess the method for varying levels of correlation between observed data sets, which is obviously not possible when evaluating only one location. It would also be of interest to assess this methodology for varying relationships between the flood sources. At the location of interest for this research, the riverine and tidal and the pluvial and tidal flood sources were not strongly related; thus, the flood sources could

essentially be considered as independent of each other. Further studies of this nature would provide confirmation that the method developed is generally applicable across locations which are influenced by more than one potential source of flooding.

9.3.1. Improved Calculation of Flood Depths

The first step of the procedure was to calculate flood inundation depths for a location of interest based on a series of gage measurements. In this research, very general methods were used to calculate flood depths. These methods were certainly not the most accurate methods possible; however, they were chosen because they were fairly straightforward to use. The focus of the research was on the copula analysis, not on the calculation of flood depths, so some potential inaccuracy in flood depth calculations was considered to be acceptable. As long as reasonably accurate flood depths could be calculated, the copula procedure could be tested and demonstrated.

If this method were to be used to develop flood hazard calculations for a specific location with the intent to use the results of the analysis for design work or land use planning, it is highly recommended that more detailed and accurate flood inundation depth calculations be made, to ensure that the likelihood of flooding is calculated as accurately as possible. In particular, obtaining accurate information about channel geometry would be very important to calculating accurate riverine flood depths. The method used to calculate tidal flood depths was very simple, and likely neglected a number of important, but complex, factors that would influence the magnitude of tidal flooding. More detailed information about the factors influencing runoff caused by rainfall, as well as perhaps information about the spatial variation of

the rainfall, would likely result in a more accurate understanding of the pluvial flood depths possible. More detailed methods that are able to take into account these additional factors would be beneficial in obtaining more accurate flood depth calculations.

Additionally, different methods of calculating flood inundation depth, such as the HAZUS or HEC-RAS computer programs, would provide flood inundation depths for a larger area than just one location. This would allow the procedure developed to be applied to a larger geographical area. This would make the results of the analysis more generally useful and applicable if they could be provided for an entire town, for instance, rather than just for one piece of property within the town.

9.3.2. Calculation of Total Flood Depths

In addition to the methods of calculating the riverine, tidal, and pluvial flood depths, the method of calculating the total flood depths impacting the location of interest also bears further investigation. In this research, the corresponding riverine, tidal, and pluvial flood depths were summed to obtain total flood depths. This assumption was supported by the fact that in determining total flows downstream of a confluence the addition of the flows in the two tributaries is used. However, many additional factors may influence the relationship between the riverine, tidal, and pluvial flood sources. Summing the flood depths from each source to determine the total flood depths at the location of interest is likely a simplistic approach. With more specific information about the relationship between the flood sources at the location of interest a more accurate approach to determining total flood depths may be

identified. Such an approach would likely be heavily dependent on the characteristics of the location of interest.

9.3.3. Improved Method of Accounting for Zero-Flood Years

The method developed to assess the flood hazard for the location of interest required fitting marginal distributions to each flood source. However, fitting these marginal distributions was observed to be a challenge because of the number of zero-flood years, especially in the riverine and pluvial flood records. In the riverine case, the stream was not dry but the zero values reflect that the flow in any one year did not reach the point of interest at a higher elevation on the floodplain. This is rational from a hydrologic standpoint, as a specific location would not be expected to flood every year.

For a riverine flood source, the bankfull flow, which is the largest flow that would not exceed the banks of the river, is typically considered to be approximately a 2-year event. Thus, flows large enough to cause a flood at the location of interest should be expected to occur less frequently than every 2 years. When the flood record of interest is the flood depth at the location of interest caused by the riverine flow, this results in a large number of zero-flood years in the record. Though the hydrology of the tidal and pluvial floods differ somewhat from the riverine flood, it would be reasonable to expect zero-flood years to occur for similar reasons in the tidal and pluvial records of interest.

Though the number of zeros observed in the riverine, tidal, and pluvial flood inundation depth samples makes sense from a hydrologic standpoint, they do make fitting marginal distributions difficult. For riverine and pluvial samples, the log-

Pearson Type III distribution is typically used to represent the flood frequencies and this distribution is unable to model flood records that include zero-flood years. The gamma distribution, which can model zero values, was used in this research; however, it remained difficult to fit the distribution because close to half of the riverine flood depth record when transposed to the location of interest consisted of zero values. The Generalized Extreme Value distribution, which can model zero values, is typically used to represent tidal flooding. As with the gamma distribution, though the zero values could be modeled, the number of zeros in the sample led to difficulty in fitting the distribution. Methods have been developed to fit the LP3 distribution in cases in which zero-flood years are part of the record, such as the Jennings and Benson (1969) or adding small values to the discharge record to avoid the zero values; however, these methods do not assist in fitting a distribution to a set of data with a large number of zeros. Thus, a method of fitting marginal distributions to data sets in which up to 50% of the record consist of zero-flood years, would improve the ability to develop joint distributions and assess the flood hazard for a location of interest impacted by two flood sources.

9.3.4. Improved Flood Risk Calculations

The primary focus of the research conducted was on developing a comprehensive flood hazard analysis. The impact of the comprehensive flood frequency analysis on flood risk calculations was then briefly investigated. In these risk calculations, only the probability of flooding depended on the flood source, the vulnerability and consequence curves did not depend on the flood source. However, it is likely that these curves should also vary with flood source.

It is not necessarily logical that either vulnerability or consequences would be the same for equal depths of flooding from different sources. The vulnerability may differ depending on the source of the flooding because the systems protecting against each source are likely different. For instance, a seawall may protect against tidal flooding and a levee may protect against riverine flooding. Thus, the vulnerability to each source of flooding would likely differ because the different protective systems offer different levels of protection.

Further, consequences of equal depths of flooding from different sources may differ. A riverine or pluvial flood of 3 feet may not cause the same level of damage as a tidal flood of 3 feet. Other factors beyond the depth of flooding may also impact the level of consequences caused by each flood source. For instance, damage from tidal floods may come both for the depth of inundation and from wave action, whereas damage from riverine floods may be caused by flow velocity or flow duration in addition to inundation depth. These differing factors between flood sources may result in differing levels of damage for the same inundation depth. The appropriateness of varying vulnerability and consequence curves for different flood sources should be investigated. Nadal *et al.* (2010) provided an excellent analysis of building damage due to both riverine and coastal flooding. This source may be considered as a starting point in the process of developing a more complete method of evaluating consequences of multiple flood sources.

9.3.5. Use of Methods to Account for Multiple Populations

An additional item of interest to this research was flooding from multiple populations, which is in concept fairly similar to flooding from multiple sources.

Multiple populations might include riverine flooding caused by either rainfall or snowmelt events, or flooding caused by convective storms versus hurricanes, for example. Though specific tests were not conducted, it is believed that the methods developed and demonstrated to determine the flood hazard for a location of interest based on multiple flood sources could also be used to determine the flood hazard for a location impacted by multiple flood populations.

One significant difference that may need to be addressed for multiple populations is the number of gages available to provide information about the flood events. Depending on the populations of interest in these situations, there may only be one gage available from which to obtain direct information about the magnitude of flooding. In the situation in which multiple flood sources impact a location, both flood sources should have a gage available to provide information about the flood magnitudes from that source. However, if riverine flooding were caused by either rainfall or snowmelt events, the only gage available would be the riverine gage, which does not indicate the cause of the flooding. In these situations; however, it is believed that there would typically be enough information available from other sources to assist in separating that one flood record into multiple records based on the causal population. The most significant challenge in this approach would come when two or more populations contribute to a single flood event. It is believed that a method of proportioning the measured flood event based on other knowledge about the contributing populations could be developed for these situations. The copula method described in this research should be further tested using scenarios in which

flooding is caused by multiple populations in order to confirm the applicability of the method.

9.3.6. Accounting for Climate Change

Traditional methods of flood frequency analysis are known to be limited in their ability to model or predict the effects of climate change. Though non-stationarity is known to exist in the data, methods of accounting for or adjusting for this non-stationarity have not been settled upon and incorporated. Because the marginal distributions used in the process of developing the joint distributions using the copula were determined by using traditional flood frequency analyses, these challenges will also apply to the results of the procedure developed and demonstrated in this research. However, this method is very flexible and the marginal distributions could be developed using other approaches as well. Should a different approach for conducting flood frequency analyses be proposed that is better able to account for non-stationarity, this method could be used to develop the marginal distributions used as input to the copulas.

Bibliography

Apel, H., Thieken, A.H., Merz, B., and Bloschl, G. (2004). "A probabilistic modeling concept for the quantification of flood risks and associated uncertainties." International Environmental Modeling and Software Society International Conference, p. 977.

Apipattanavis, S., Rajagopalan, B., and Lall, U. (2010). "Local Polynomial-Based Flood Frequency Estimator for Mixed Population." *Journal of Hydrologic Engineering*, 15(9): 680-691.

American Society of Civil Engineers (ASCE). (1996). Hydrology Handbook, 2nd ed. New York, NY.

Balakrishnan, N., and Lai, C-D. (2009). Continuous Bivariate Distributions, 2nd ed. Springer, New York, NY.

Bales, J.D., and Wagner, C.R. (2009). "Sources of uncertainty in flood inundation maps." *Journal of Flood Risk Management,* 2(2): 139-147.

Beard, L.R. (1962). Statistical methods in hydrology, USACE, Sacramento CA.

Beltaos, ,S. (2003). "Numerical modeling of ice-jam flooding on the Peace-Athabasca delta." *Hydrological Processes*, 17(18): 3685-3702.

Brooks, E.S., and Boll, J. (2005). "A Simple GIS-Based Snow Accumulation and Melt Model." Proceedings of the Western Snow Conference, 73: 123-129.

Buchele, B., Kreibich, H., Kron, A., Thieken, A., Ihringer, J., Oberle, P., Merz, B., and Nestmann, F. (2006). "Flood-risk mapping: contributions towards an

enhanced assessment of extreme events and associated risks." *Natural Hazards and Earth System Sciences*, 6(4): 485-503.

Carlson, R.F., and Fox, P. (1976). "A Northern Snowmelt-Flood Frequency Model." *Water Resources Research,* 12(4): 786-794.

Carter, N.T. (2005). Flood Risk Management: Federal Role in Infrastructure. Congressional Research Service, Library of Congress, Washington, D.C.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). Copula Methods in Finance. John Wiley & Sons, Ltd., Hoboken, NJ.

Cobby, D., Morris, S., Parkes, A., and Robinson, V. (2009). "Groundwater flood risk management: advances toward meeting the requirements of the EU floods directive." *Journal of Flood Risk Management,* 2(2): 111-119.

Collier, C.G. (2009). "Pluvial Flooding: A Perspective Prepared for the Flood Risk from Extreme Events (FREE) Steering Committee." *Flood Risk From Extreme Events: A Strategic Program*, 1: 1-7.

Cullmann, J., Krausse, T., and Philipp, A. (2009). "Communicating flood forecast uncertainty under operational circumstances." *Journal of Flood Risk Management,* 2(4): 306-314.

De Michele, C., and Salvadori, G. (2003). "A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas." *Journal of Geophysical Research,*108(D2): 15-1-15-11.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R. (2005). "Bivariate Statistical Approach to Check Adequacy of Dam Spillway." *Journal of Hydrologic Engineering,*10(1): 50-57.

Dijkman, J.P.M., and Heynert, K.V. (1993). "Changing approach in flood management along the Rhine River in the Netherlands." In Marchand, M., Heynert, K.V., van der Most, H., and Penning, W.E. (ed.), Dealing with flood risk, Delft Hydrualics Select Serires I, Delft University Press, The Netherlands, pp. 23-34.

Doornkamp, J.C. (1998). "Coastal flooding, global warming, and environmental management." *Journal of Environmental Management,* 52(4): 327-333.

Dunne, T., and Leopold, L.B. (1978). Water in Environmental Planning. W.H. Freeman and Company, New York.

European Commission. (2009). "Early warning system for pluvial flooding." *Science for Environment Policy: DG Environment News Alert Service,* 166:1.

Exponent. (2010). "Deterministic and Probabilistic Assessments." July 15, 2010. http://www.exponent.com/deterministic_assessments/

Falconer, R. (2007). "Pluvial and Groundwater Flooding New Approaches in Flood Warning and Risk Management." Presented at CIWEM Global Environment Conference, 10/11/2007.

Falconer, R.H., Cobby, D., Smyth, P., Astle, G., Dent, J., and Golding, B. (2009). "Pluvial flooding: new approaches in flood warning, mapping, and risk management." *Journal of Flood Risk Management,* 2(3): 198-208.

Favre, A-C., Adlouni, S.E., Perreault, L., Thiemonge, N., and Bobee, B. (2004). "Multivariate hydrological frequency analysis using copulas." *Water Resources Research*, 40(1).

Federal Emergency Management Agency (FEMA). (2003). "Appendix D: Guidance for Coastal Flooding Analyses and Mapping." Guidelines and Specifications for Flood Hazard Mapping Partners.

Federal Emergency Management Agency (FEMA). (2010). "Hurricane Katrina Flood Recovery (Louisiana)." December 12, 2010. http://www.fema.gov/hazard/flood/recoverydata/katrina/katrina_la_faqs.shtm

Finch, J.W., Bradford, R.B., and Hudson, J.A. (2004). "The spatial distribution of groundwater flooding in a chalk catchment in southern England." *Hydrological Processes,* 18(5): 959-971.

Frederick, R.H., Myers, V.A., and Auciello, E.P. (1977). "Five-To-Sixty-Minute Precipitation Frequency for the Eastern and Central U.S." NOAA Technical Memorandum NWS Hydro-35.

Frees, E.W., and Valdez, E.A. (1998). "Understanding Relationships Using Copulas." *North American Actuarial Journal*, 2(1): 1-25.

Garrity, N.J., Battalio, R., Hawkes, P.J., Roupe, D. (2007). "Evaluation of Event and Response Approaches to Estimate the 100-Year Coastal Flood for Pacific Coast Sheltered Waters." Coastal Engineering 2006, Volume 2: Proceedings of the 30th International Conference: 1651-1664.

Genest, C., and MacKay, J. (1986). "The Joy of Copulas: Bivariate Distributions with Uniform Marginals." *The American Statistician*, 40(4): 280-283.

Genest, C., and Rivest, L-P. (1993). "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association,* 88(423): 1034-1043.

Genest, C., and 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.

Gerard, R., and Calkins, D.J. (1984). "Ice related flood frequency analysis: application of analytical estimates." *Proc., Cold Regions Engrg. Spec. Conf.,* Can. Soc. For Civ. Engrg., Motreal, Quebec, Canada.

Golding, B.W., (2009). "Uncertainty propagation in a London flood simulation." *Journal of Flood Risk Management,* 2(1): 2-15.

Gottfried, G.J., Neary, D.G., and Folliott, P.F. (2003). "Snowpack-Runoff Relationships for Forested Mid-Elevation Watersheds and a High-Elevation Watershed in Arizona." Proceedings of the Western Snow Conference, 71: 49-58.

Griffis, V.W., and Stedinger, J.R. (2007a). "Evolution of Flood Frequency Analysis with Bulletin 17." *Journal of Hydrologic Engineering,* 12(3): 283-297.

Griffis, V.W., and Stedinger, J.R. (2007b). "Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. I: Distribution Characteristics." *Journal of Hydrologic Engineering,* 12(5): 482-491.

Grimaldi, S., and Serinalidi, F. (2006a). "Asymmetric copula in multivariate flood frequency analysis." *Advances in Water Resources,* 29(8): 1155-1167.

Grimalidi, S., and Serinaldi, F. (2006b). "Design hyetograph analysis with 3-copula function." *Hydrological Sciences Journal,* 51(2): 223-238.

Hawkes, P.J. (2005). "Use of Joint Probability Methods in Flood Management: A Guide to Best Practice." Defra/Environment Agency Flood and Coastal Defence R&D Programme.

Hawkes, P.J. (2008). "Joint probability analysis for estimation of extremes." *Journal of Hydraulic Research*, 46(2):246-256.

Hawkes, P.J., Gouldby, B.P., Tawn, J.A., and Owen, M.W. (2002). "The joint probability of waves and water levels in coastal engineering design." *Journal of Hydraulic Research,* 40(3): 241-251.

Hayse, J.W. (2000). "Using Monte Carlo Analysis in Ecological Risk Assessments." September 5, 2010. http://web.ead.anl.gov/ecorisk/issue/pdf/montecarlo.pdf

Hoozemans, F.M.J., and Hulsbergen, C.H. (1995). "Sea-level rise: a world-wide assessment of risk and protection costs." In Eisma, D. (ed.), Climate Change: Impact on Coastal Habitation, Lewis Publishers, London, pp. 137-163.

Hussain, A. "Thousands more face floods: Half a million more homes – from Chelsea to Salford – will be at risk from next month." The Sunday Times. July 6, 2008. http://www.timesonline.co.uk/tol/money/property_and_mortgages/article4275882.ece

Interagency Advisory Committee on Water Data (IACWD). (1982). "Guidelines for Determining Flood Flow Frequency. Bulletin #17B of the Hydrology Subcommittee. Washington, DC.

International Federation of Red Cross and Red Crescent Societies (IFRCRS). (2001). World Disasters Report: Focus on Recovery.

Jennings, M.E., and Benson, M.A. (1969). "Frequency Curves for Annual Flood Series with Some Zero Events of Incomplete Data." *Water Resources Research,* 5(1): 276-280.

Johnson, W.K. (1991). "Importance of Surface-Ground Water Interactions to Corps Total Water Management: Regional and National Examples." Research Document (RD)-32, Davis California, U.S. Army Corps of Engineers Hydrologic Engineering Center.

Jonkman, S.N., Kok, M., van Ledden, M., and Vrijling, J.K. (2009). "Risk-based design of flood defence systems: a preliminary analysis of the optimal protection level for the New Orleans metropolitan area." *Journal of Flood Risk Management,* 2 (3): 170-181.

Kao, S-C., and Govindaraju, R.S. (2007a). "A bivariate frequency analysis of extreme rainfall with implications for design." *Journal of Geophysical Research,* 112(13): 1-15.

Kao, S-C., and Govindaraju, R.S. (2007b). "Probabilistic structure of storm surface runoff considering the dependence between average intensity and storm duration of rainfall events." *Water Resources Research,* 43(6): 1-15.

Kao, S-C., and Govindaraju, R.S. (2010). "A copula-based joint deficit index for droughts." *Journal of Hydrology,* 380(1-2): 121-134.

Karmakar, S., and Simonovic, S.P. (2008a). "Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques." *Journal of Flood Risk Management,* 1(4): 190-200.

Karmakar, S., and Simonovic, S.P. (2008). "Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions." *Journal of Flood Risk Management,* 2(1): 32-44.

Kattelmann, R. (1990). "Floods in the high Sierra Nevada, California, USA." *Hydrology in Mountainous Regions II – Artificial Reservoirs, Water and Slopes,* Proceedings of the Two Lausanne Symposia.

Kattelmann, R. (1991). "Peak flows from snowmelt runoff in the Sierra Nevada, USA." *Snow, Hydrology and Forests in High Alpine Areas,* Proceedings of the Vienna Symposium: 203-211.

Kattelmann, R., Berg, N., and McGurk, B. (1991). "A History of Rain-on-Snow Floods in the Sierra Nevada." Proceedings of the Western Snow Conference, 59: 138-141.

Kirshen, P., Watson, C., Douglas, E., Gontz, A., Lee, J., and Tian, Y. (2008). "Coastal Flooding in the Northeastern United States due to climate change." *Mitigation and Adaptation Strategies for Global Change,* 13(5-6): 437-451.

Klein, B., Pahlow, M., Hundecha, Y., and Schumann, A. (2010). "Probability Analysis of Hydrological Loads for the Design of Flood Control Systems Using Copulas." *Journal of Hydrologic Engineering,* 15(5): 360-369.

Kundzewicz, Z.W., and Hirabayashi, Y. (2010). "River Floods in the Changing Climate – Observations and Projections." *Water Resources Management,* 24(11): 2633-2646.

Leitao, J.P., Boonya-aroonnet, S., Prodanovic, D., and Maksimovic, C. (2009). "The influence of digital elevation model resolution on overland flow networks for

modeling urban pluvial flooding." *Water Science & Technology,* 60(12): 3137-3149.

Loganathan, G.V., Kuo, C.Y., and Yannaccone, J. (1987). "Joint Probability Distribution of Streamflows and Tides in Estuaries." *Nordic Hydrology,* 18(4): 237-246.

Macdonald, D., Hall, R. Carden, D., Dixon, A., Cheetham, M., Cornick, S., and Clegg, M. (2007). Investigating the Interdependencies Between Surface and Groundwater in the Oxford Area to Help Predict the Timing and Location of Groundwater Flooding and to Optimize Flood Mitigation Measures.

Maksimovic, D., Boonya-Aroonnet, S., Leitao, J.P., Dyordjevic, S., and Allitt, R. (2009). "Overland flow and pathway analysis for modeling of urban pluvial flooding." *Journal of Hydraulic Research,* 47(4): 512-523.

Marks, D., Kimball, J., Tingey, D., and Link, T. (1998). "The sensitivity of snowmelt processes to climate conditions and forest cover during rain-on-snow: a case study of the 1996 Pacific Northwest flood." *Hydrological Processes,* 12(10-11): 1569-1587.

McCuen, R.H. (2002). "Approach to Confidence Interval Estimation for Curve Numbers." *Journal of Hydrologic Engineering*, 7(1): 43-48.

McCuen, R.H., (2005). Hydrologic Analysis and Design, 3rd edition. Pearson Prentice Hall, Upper Saddle River, New Jersey.

McCuen, R.H., (2003). Modeling Hydrologic Change: Statistical Methods. Taylor and Francis Group, Boca Raton, Florida.

McCuen, R.H., and Levy, B.S. (2000). "Evaluation of Peak Discharge Transposition." *Journal of Hydrologic Engineering,* 5(3): 278-289.

McCuen, R.H., and Beighley, R.E. (2003). "Seasonal flow frequency analysis." *Journal of Hydrology*, 279(1-4): 43-56.

McInnes, K.L, Hubbert, G.D., Abbs, D.J., and Oliver, S.E. (2002). "A numerical modeling study of coastal flooding." *Meteorology and Atmospheric Physics,* 80(1-4): 217-233.

Merwade, V., Olivera, F., Arabi, M., and Edleman, S. (2008). "Uncertainty in Flood Inundation Mapping: Current Issues and Future Directions." *Journal of Hydrologic Engineering,* 13(7): 608-620.

Middelmann, M.H. (2010). "Flood damage estimation beyond stage-damage functions: an Australian example." *Journal of Flood Risk Management,* 3(1): 88-96.

MindTools. (2010). "Risk Analysis and Risk Management: Evaluating and Managing the Risks You Face." August 27, 2010. http://www.mindtools.com/pages/article/newTMC_07.htm

Murdersbach, C., and Jensen, J. (2009). "Nonstationarity extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline." *Journal of Flood Risk Management,* 3(1): 52-62.

Morris, E.C. (1982). "Mixed Population Frequency Analysis." Training Document 17, Davis California, U.S. Army Corps of Engineers Hydrologic Engineering Center.

Mutel, C.F., ed. (2010). A Watershed Year: Anatomy of the Iowa Floods of 2008. University of Iowa Press, Iowa City, Iowa.

Nadal, N.C., Zapata, R.E., Pagan, I., Lopez, R., and Agudelo, J. (2010). "Building Damage Due to Riverine and Coastal Floods." *Journal of Water Resources Planning and Management,* 136(3): 327-336.

National Hurricane Center (NHC). (2010). "Storm Surge Overview." November 19, 2010. http://www.nhc.noaa.gov/ssurge/index.html

National Oceanic and Atmospheric Administration (NOAA) Digital Coast: NOAA Coastal Services Center. "Approaches: Coastal Inundation Toolkit." November 1, 2010. http://www.csc.noaa.gov/digitalcoast/inundation/index.html

National Research Council Committee on Risk-Based Analysis for Flood Damage Reduction Studies (NRC). (2000). Risk Analysis and Uncertainty in Flood Damage Reduction Studies. National Academy Press, Washington, D.C.

National Research Council Building Research Advisory Board (NRC). (1977). Methodology for Calculating Wave Action Effects Associated with Storm Surges. National Academy of Sciences, Washington, D.C.

Nelsen, R.B. (1999). An Introduction to Copulas. Springer, New York, NY.

Olsen J.R., Stedinger, J.R., Matalas, N.C., and Stakhiv, E.Z. (1999). "Climate Variability and Flood Frequency Estimation for the Upper Mississippi and Lower Missouri Rivers." *Journal of the American Water Resources Association,* 35(6): 1509-1523.

Prinos, P. (2009). "Review of Flood Hazard Mapping." Integrated Flood Risk Analysis and Management Methodologies.

Payton, E.A., and Brendecke, C.M. (1985). "Rainfall and Snowmelt Frequency in an Alpine Watershed." Proceedings of the Western Snow Conference, 53: 25-36.

Pethick, J. (1993). "Shoreline Adjustments and Coastal Management: Physical and Biological Processes under Accelerated Sea-Level Rise." *The Geographical Journal,* 159(2): 162-168.

Pirazzoli, P.A., Costa, S., Dornbusch, W., and Tomasin, A. (2006). "Recent evolution of surge-related events and assessment of coastal flooding risk on the eastern coasts of the English Channel." *Ocean Dynamics,* 56(5-6): 498-512.

Pitt, M. (2008). Learning Lessons from the 2007 Floods.

Plate, E.J. (2002). "Flood risk and flood management." *Journal of Hydrology,* 267(1-2): 2-11.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., and Stromberg, J.C. (1997). "The Natural Flow Regime." *BioScience*, 47(11): 769-784.

Poulter, B., and Halpin, P.N.(2008). "Raster modeling of coastal flooding from sea-level rise." *Journal of Geographical Information Science,* 22(2): 167-182.

Purnell, R. (2002). "Flood risk – a government perspective." In *Civil Engineering SpecialIssue – Floods – A New Approach,* 150(1): 10-14.

Purvis, M.J., Bates, P.D., and Hayes, C.M. (2008). "A probabilistic methodology to estimate future coastal flood risk due to sea level rise." *Coastal Engineering,* 55(12): 1062-1073.

Reeve, D., Chadwick, A., and Fleming, C. (2004). Coastal Engineering: Process, Theory, and Design Practice. Spon Press, New York, NY.

Renard, B., and Lang, M. (2007). "Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology." *Advances in Water Resources,* 30(4): 897-912.

Salvadori, G., and De Michele, C. (2007). "On the Use of Copulas in Hydrology: Theory and Practice." *Journal of Hydrologic Engineering,* 12(4): 369-380.

Samuels, P. (2002). "A new joint probability appraisal of flood risk." *Proceedings of the Institution of Civil Engineers*, 154(2): 109-115.

Serinaldi, F., and Grimaldi, S. (2007). "Fully Nested 3-Copula: Procedure and Application on Hydrologic Data." *Journal of Hydrologic Engineering,*12(4): 420-430.

Setty and Associates. (2006). "Flood Mitigation and Prevention." General Services Administration, Washington, D.C.

Simoes, N.E., Leitao, J.P., Maksimovic, C., Marques, A.S., and Pina, R. (2010). "Sensitivity analysis of surface runoff generation in urban flood forecasting." *Water Science & Technology*, 61(10): 2595-2601.

Simonovic, S.P. (2009). "Managing flood risk, reliability, and vulnerability." *Journal of Flood Risk Management,* 2(4): 230-231.

Singh, V.P., Jain, S.K., and Tyagi, A.K. (2007). Risk and Reliability Analysis: A Handbook for Civil and Environmental Engineers. ASCE Press, Reston, VA.

Singh, V.P., and Zhang, L. (2007). "IDF Curves Using the Frank Archimedean Copula." *Journal of Hydrologic Engineering*, 12(6): 651-662.

Sklar, A. (1959). "Fonctions de repartition a n dimensions et leurs marges." *Publ. Inst. Stat. Univ. Paris*, 8, 229-231.

Smith, D.I. (1994). "Flood damage estimation – A review of urban stage-damage curves and loss functions." Water S.A., 20(3): 231-238.

Snedecor, G.W., and Cochran, W.G. (1968). Statistical Methods, 6th Edition. The Iowa State University Press, Ames, Iowa.

Solver.com. (2010). "Risk Analysis Overview." September 3, 2010. http://www.solver.com/risk-analysis/index.html

Stamatelatos, M., et al. (2002). "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practioners." http://www.hq.nasa.gov/office/codeq/doctree/praguide.pdf

Stankowski, S.J. (1974). Magnitude and Frequency of Floods in New Jersey with Effects of Urbanization: New Jersey Department of Environmental Protection Special Report 38.

Stedinger, J.R., and Griffis, V.W. (2008). "Flood Frequency Analysis in the United States: Time to Update." *Journal of Hydrologic Engineering,* 13(4): 199-204.

Ten Veldhuis, J.A.E., and Clemens, F.H.L.R. (2009). "Uncertainty in risk analysis of urban pluvial flooding: a case study." *Water Practice & Technology,* 14(1): 1-8.

The GeoInformation Group. (2008). "Surface Water Flood Risk Mapping." *Cities Revealed*.

Troendle, C.A., and Porth, L.S. (2000). "Predicting the 1.5-, 5-, 10-, and 25-Year Event from the Snow Zones of Colorado and Wyoming." Proceedings of the Western Snow Conference, 68: 41-48.

Tuthill, A.M., Weubben, J.L., Daly, S.F., and White, K.D. (1996). "Probability Distributions for Peak Stage on Rivers Affected by Ice Jams." *Journal of Cold Regions Engineering*, 10(1): 36-57.

United States Army Corps of Engineers (USACE). (2000). "Application of Risk-Based Analysis to Planning Reservoir and Levee Flood Damage Reduction Systems." Technical Paper 160, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (2003). "Generic Depth-Damage Relationships for Residential Structures with Basements." Economic Guidance Memorandum (EGM) 04-01, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (1999). "Groundwater Hydrology." Engineering Manual 1110-2-1421, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (1993). "Hydrologic Frequency Analysis." Engineering Manual 1110-2-1415, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (2002). "Ice Engineering." Engineering Manual 1110-2-1612, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (1991). "Ice-Influenced Flood Stage Frequency Analysis." Engineering Technical Letter 1110-2-325, Department of the Army, Washington, DC.

United States Army Corps of Engineers (USACE). (1998). "Runoff from Snowmelt." Engineering Manual 1110-2-1406, Department of the Army, Washington, DC.

United States Geological Survey (USGS). (1997). "Predicting Coastal Flooding and Wetland Loss." USGS Fact Sheet 094-97, Lafayette, Louisiana.

United States Geological Survey. (2000). "Ground-Water Flooding in Glacial Terrain of Southern Puget Sound, Washington." USGS Fact Sheet 111-00, Takoma, Washington.

United States Nuclear Regulatory Commission (USNRC). (2007). "Fact Sheet on Probabilistic Risk Assessment." September 3, 2010. http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/probabilistic-risk-assess.html

Visocky, A.P. (1995). "Determination of 100-Year Ground-Water Flood Danger Zones for the Havana and Bath Areas, Mason County, Illinois." Contract Report 584, Prepared for the Illinois Emergency Management Agency.

Wang, C., Chang, N-B., and Yeh, G-T. (2009). "Copula-based flood frequency (COFF) analysis at the confluences of river systems." *Hydrological Processes,* 23(10): 1471-1486.

Wang, Y., Ma, H., Sheng, D., and Wang, D. (2012). Assessing the Interactions between Chlorophyll a and Environmental Variables Using Copula Method. *Journal of Hydrologic Engineering,* 17(4): 495-506.

Waterworlds. (2008). “Pluvial Flooding.” July 23, 2010. http://waterworlds.wordpress.com/2008/01/03/13-pluvial-flooding/

Waylen, P., and Woo, M-K. (1982). “Prediction of Annual Floods Generated by Mixed Processes.” *Water Resources Research,* 18(4): 1283-1286.

Wheater, H.S. (2002). “Progress in and prospects for fluvial flood modeling.” *Philosophical Transactions of the Royal Society,* 360: 1409-1431.

White, K.D., Pederson, K.J., and Pomerleau, R.T. (2000). “Ice-Affected Flood Frequency.” Ice Engineering, 24:1-4.

Yue, S. (2000). “Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls.” *Hydrological Sciences Journal,* 45(2): 315-326.

Yue, S., Ouarda, T.B.M.J., and Bobee, B. (2001). “A review of bivariate gamma distributions for hydrological application.” *Journal of Hydrology,* 246(1-4): 1-18.

Yue, S., Ouarda, T.B.M.J., Bobee, B., Legendre, P., and Bruneau, P. (1999). “The Gumbel mixed model for flood frequency analysis.” *Journal of Hydrology,* 226(1-2): 88-100.

Zerger, A., and Wealands, S. (2004). “Beyond Modeling: Linking Models with GIS for Flood Risk Management.” *Natural Hazards,* 33(2): 191-208.

Zhang, L., and Singh, V.P. (2006). “Bivariate Flood Frequency Analysis Using the Copula Method.” *Journal of Hydrologic Engineering,* 11(2): 150-164.

Zhang, L., and Singh, V.P. (2007a). "Gumbel-Hougaard Copula for Trivariate Rainfall Frequency Analysis." *Journal of Hydrologic Engineering*, 12(4): 409-419.

Zhang, L., and Singh, V.P. (2007b). "Trivariate Flood Frequency Analysis Using the Gumbel-Hougaard Copula." *Journal of Hydrologic Engineering,* 12(4): 431-439.

Zuzal, J.F., and Greenwalt, R.N. (1985). "Probability Distributions of Rain on Snow for North Central Oregon." Proceedings of the Western Snow Conference, 53:37-43.