# Evaluating the Impact of Climate Change on Rainfall Intensity-Duration-Frequency (IDF) Curves in Alabama Using Dynamically-Downscaled Precipitation Data

by

Golbahar Mirhosseini

A dissertation submitted to the Graduate Faculty of Auburn University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

> Auburn, Alabama August 3, 2013

Keywords: Climate change, Global Climate Models (GCMs), Intensity-Duration-Frequency (IDF) curve, Temporal downscaling

Copyright 2013 by Golbahar Mirhosseini

Approved by

Puneet Srivastava, Co-chair, Associate Professor of Biosystems Engineering Xing Fang, Co-chair, Associate Professor of Civil Engineering Latif Kalin, Associate Professor, School of Forestry and Wildlife Sciences Keith Ingram, Associate Research Scientist of Agricultural and Biological Engineering, University of Florida

#### Abstract

Rainfall intensity-duration-frequency (IDF) curves are extensively used for hydrologic designs. A design based on inaccurate rainfall characteristics can lead to malfunction of infrastructure, excessive design cost, and loss of life. It is expected that the frequency and magnitude of future extreme rainfalls will change due to increases in greenhouse gas (GHG) concentrations. Quantifying potential effects of climate change by reviewing and updating IDF curves for future climate scenarios and adapting to them is one way to reduce vulnerability. This study was undertaken to evaluate expected changes in IDF curves from the current climate to the projected future climate. Six combinations of global and regional climate models were used to develop the IDF curves under future climate scenarios. Three-hourly precipitation data were temporally downscaled using two different disaggregation methods: stochastic and artificial neural network (ANN). In the second chapter, stochastic method was used to downscale three-hourly precipitation into 15-min precipitation amount and IDF curves were developed. The results of all six climate models suggest that the future rainfall intensities are expected to decrease for short duration events (i.e., less than 2 hours). However, for longer duration events, the results are not consistent across the models. In the third chapter, a feed-forward, backpropagation ANN model was developed to estimate maximum 15-, 30-, 45-, 60-, and 120-min precipitations. The results were also compared with the disaggregated rainfalls derived using the stochastic method.

Results indicate that future rainfall intensities for short duration (< 2 hours) events are expected to decrease by 33% to 74%. However, a large uncertainty exists in the projected rainfall intensities of longer duration events. In chapter 4, uncertainty associated with the studied models was quantified using kernel density estimator and probability-based IDF curves were created. The resultant IDF curves incorporating all models were also developed. Although the results derived from different climate projections show large uncertainty associated with climate models, all of them indicate decrease in future rainfall intensity for short duration rainfall events, especially for durations less than 2 hours. The methodology developed can be used for developing IDF curves for other parts of the United States and the world.

#### Acknowledgments

I read somewhere that "The only way to find out how to do a Ph.D. is to do one". Well, I did one, but not alone. Many have helped and supported me along this difficult, long journey and I will be forever thankful to them. I would like to express my sincere gratitude to my major advisor, Dr. Puneet Srivastava for his help, support and encouragement along the way. I would like to express my heartfelt gratitude to my coadvisor, Dr. Xing Fang for his valuable suggestions, help and guidance. I am so grateful to my advisory committee, Dr. Latif Kalin and Dr. Keith Ingram for helpful suggestions, guidance and support. I owe them my heartfelt appreciation. Special thanks to Dr. Lydia Stefanova for her valuable suggestions and assistance during this journey. I will be forever grateful to her. I would also like to thank Dr. Sumit Sen, Dr. Vaishali sharda, Dr. Ajay Sharda and Dr. Suresh Sharma for their friendship and encouragement during the course of this study.

I am indebted to my parents for their unconditional love, continues encouragement, support and for telling me that I could do anything and actually believing it. I am thankful to my sister and brother-in-law for encouraging me to dream big and sending their constant support and love across the world for the past four years. I would not be where I am today if it wasn't for them. I would not be able to finish this journey if not for my husband. I cannot even begin to thank him enough for all he has done for me. His love and support means the world to me. He believed in me when I didn't and never stopped encouraging me. My deepest gratitude and love to him for believing in me so I could believe in myself. I could not have asked for a better partner to share my life with.

Making this dream come true was not able without the love, support and encouragement of all my family and friends. With sincere thanks.

Golbahar Mirhosseini

July 2013

### Table of Contents

Abstrac	t	ii
Acknow	vledgments	iv
List of ]	Гables	. ix
List of I	Figures	X
1. Int	roduction	1
1.1.	Climate Change	1
1.2.	General Circulation Models (GCMs)	2
1.3.	Regional Circulation Models (RCMs)	3
1.4.	Climate Model Scenarios	5
1.5.	Bias Correction	8
1.6.	Downscaling Techniques	8
1.6	5.1. Spatial Downscaling	8
1.6	5.2. Temporal Disaggregation	12
1.7.	Intensity-Duration-Frequency (IDF) Curves	14
1.8.	Problem Statement	19
1.9.	Dissertation Objectives	20
1.10.	Dissertation Structure	20
2. Th	e Impact of Climate Change on Rainfall Intensity-Duration-Frequency (IDF)	
Curves	in Alabama	22
2.1.	Abstract	22
2.2. I	ntroduction	23
2.3. N	Materials and Methods	26
2.3	3.1. Data and Model Used	26
2.3	3.2. Bias correction	34
2.3	3.3. Temporal Downscaling	35

2.3.4	. Performance	37
2.3.5	. Creating IDF curves	38
2.4. Re	sults and Discussion	40
2.4.1	. Bias correction	40
2.4.2	. Performance	40
2.4.3	. IDF curves	45
2.5. Su	nmary and Conclusions	54
3. Deve Under Fu	eloping Rainfall Intensity-Duration-Frequency (IDF) Curves For Alabama ture Climate Scenarios Using Artificial Neural Network (ANN)	56
3.1.	Abstract	56
3.2.	ntroduction	57
3.3.	Methodology	59
3.3.1	. Data and Model Used	59
3.3.2	. Bias Correction and Temporal Downscaling	60
3.3.3	. Artificial Neural Network	61
3.3.4	. ANN Model Input Data	62
3.3.5	. ANN Model Training and Performance Measures	65
3.3.6	. Intensity-Duration-Frequency (IDF) Curves	66
3.4.	Results and Discussion	66
3.4.1	. Performance	66
3.4.2	. IDF Curves	71
3.5.	Summary and Conclusions	80
4. Deve	eloping Probability-based IDF Curves Using Kernel Density Estimator	82
4.1.	Abstract	82
4.2.	ntroduction	83
4.3.	Methodology	86
4.3.1	. Data	86
4.3.2	. Probability Density Function (PDF)	87
4.3.3	. Kernel Density Estimation	88
4.3.4	Probability-based IDF Curves	96
4.4.	Results and Discussion	96

	4.4.1	Optimal bandwidth selection	96
	4.4.2	Probability-based IDF Curves	97
4.	5. S	ummary and Conclusions	113
5.	Conc	lusions	115
5.	1. S	ummary and Conclusions	115
	5.1.1	Objective 1	116
	5.1.2	Objective 2	117
	5.1.3	Objective 3	118
6.	Futur	e research	
Refe	erence	S	121
Арр	endix	A	134
Mor	nthly C	Cumulative Distribution Functions (CDF) To Perform Bias Correction	134
Арр	endix	В	144
Rair	nfall Ir	Itensity-Duration-Frequency Maps	144
Арр	endix	С	
Scat	ter Plo	ots of Observed and Modeled Rainfall Depths Using the ANN Model a	und the
Stoc	chastic	Method	186
App	endix	D	
IDF	Maps	, using ANN model for rainfall disaggregation	206
App	endix	Е	226
Proł	oability	y-based IDF curves for Alabama	226
App	endix	F	325
IDF	curve	s for Alabama	325

# List of Tables

Table 2-1 Performance measures between the disaggregated time series and measured      statistics of 15-min rainfall
Table 2-2 Performance measures between the maximum rainfall of disaggregated time series and measured statistics
Table 2-3 Performance measures between the sample and theoretical cumulative      distributions using Kolmogorov-Smirnov (K-S) test
Table 2-4 Comparisons of IDF curves under the current and future climate scenario for      Auburn, AL
Table 3-1 The Akaike Information Criterion (AIC) and Normalized Mean Square Error (NMSE) calculated for the nine combinations of input variables in order to determine optimal input variables of the ANN model
Table 3-2 Performance of the ANN model for training, validation, and testing datasets for      every month and every rainfall duration
Table 3-3 Comparisons of IDF curves under the current and future climate scenarios for      Auburn, AL using two temporal downscaling methods, the ANN model and the      stochastic method.
Table 4-1 Examples of kernel functions 90
Table 4-2 Optimal bandwidth selection for Auburn, 15-min rainfall duration
Table 4-3 Uncertainty quantification for 15-min rainfall duration, Auburn, AL101
Table 4-4 Uncertainty quantification of rainfall depths (mm) projected from differentclimate models for Auburn, AL and London, ON

# List of Figures

Figure 1.1 Representation of a climate model setup4
Figure 1.2 Predicted changes in winter precipitation over Central/ Southern Europe between the present day and 2080 predicted by the RCM (right) and GCM (left)
Figure 1.3 Total global annual CO <sub>2</sub> emissions [in gigatons of carbon (GtC/year)] for six scenario groups
Figure 1.4 Downscaling techniques allow the use of GCMs outputs as inputs into regional climate and weather models. From http://www.southwestclimatechange.org/climate/modeling/downscaling retrieved on 03/12/2013
Figure 1.5 Rainfall Intensity-Duration-Frequency (IDF) curve15
Figure 2.1 Location of rain gauge stations (red squares) and NARCCAP 50-km resolution grid centers (green dots, HRM3-HadCM3) used to develop IDF curves for Alabama27
Figure 2.2 Monthly Cumulative Distribution Functions (CDFs) to perform bias correction- Model data is from HRM3-HadCM3 model42
Figure 2.3 Rainfall intensity map; 50-year rainfall of 6-hr, 12-hr, 24-hr and 48-hr durations (mm) under future climate using HRM3-HadCM3 projected data
Figure 2.4 IDF curves under current and future climate using HRM3-HadCM3, CRCM-CGCM3 and HRM3-GFDL models for Auburn, AL
Figure 2.5 IDF curves under current and future climate using CRCM-CCSM, RCM3-GFDL and ECP2-GFDL models for Auburn, AL
Figure 2.6 Rainfall intensity vs. return period under current and future climate for a 12-hr rainfall using a) HRM3-HadCM3, b) HRM3-GFDL, c) RCM3-GFDL, d) CRCM-CGCM3, e) CRCM-CCSM and f) ECP2-GFDL model

Figure 3.1 Scatter plots of observed and modeled maximum a) 15-min, b) 30-min, c) 45-min, d) 60-min, and e) 120-min rainfall depths using the ANN model (left panels) and the stochastic method (right panels) for July
Figure 3.2 Rainfall intensity map; 100-year rainfall of 15-, 30-, 45- and 60-min durations (mm) under future climate using HRM3-HadCM3 projected data72
Figure 3.3 IDF curves for 10-yr and 100-yr return periods under current and future climate (using HRM3-HadCM3, CRCM- CGCM3, and HRM3-GFDL CRCM-CCSM and RCM3-GFDL models) for Auburn, AL
Figure 4.1 Comparison of kernel density estimate (right) and histogram (left). The individual kernels are the red dashed curves and the kernel density estimate the blue curves
Figure 4.2 Probability-based IDFs when stochastic method used for disaggregation for 15-, 30-, 45-min,1-,2-, 3-, 6-, 12-, 24- and 48-hr rainfall duration-graphs are presented for Auburn, AL
Figure 4.3 Resultant IDF curves for Auburn, AL using the stochastic method100
Figure 4.4 Projected rainfall depth range from six climate models (bars show minimum and maximum) and the red dot represents the rainfall depth derived from the probability-based IDF
Figure 4.5 Projected rainfall depth range (bars show minimum and maximum and the red dot represents the rainfall depth derived from the probability-based IDF) for Auburn, AL (left side) and London, Ontario, Canada (right side)104
Figure 4.6 Resultant IDF curves under future and current climate for Auburn, AL when the stochastic method was used for rainfall disaggregation108
Figure 4.7 Probability-based IDFs when ANN model used for disaggregation for 15-, 30-, 45-min,1- and 2-hr rainfall duration-graphs are presented for Auburn, AL110
Figure 4.8 Resultant IDF curves for Auburn, AL using ANN model111
Figure 4.9 Resultant IDF curves under future and current climate for Auburn, AL112

#### Chapter 1

#### 1. Introduction

#### 1.1. Climate Change

Climate is a complex system consisting of the oceans, land, atmosphere, snow, and ice. Changes in climate can happen due to natural events such as volcanic eruptions, solar radiation variations, earth internal dynamics, orbital irregularities or human activities, etc. "Climate change" is a term usually used to describe changes caused by human activities (Solaiman 2011). During the last century, the concentration of carbon dioxide (CO<sub>2</sub>) and other greenhouse gases (GHGs) in the earth's atmosphere has risen due to increased industrial activities (Prodanovic et al. 2007). This increase in GHG concentrations is causing large-scale variations in atmospheric processes, which can then lead to changes in precipitation and temperature characteristics.

The global average temperature has increased 0.74°C during the 20th century, and projections indicate that it will continue to increase over the next hundred years (IPCC 2007). Changes in temperature can be associated with changes in precipitation, causing more drought, floods, heat waves, melting ice caps, sea level rise, or more frequent extreme rainfalls (IPCC 2007). Based on the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC 2007), between 75 and 250 million people may experience increased water stress in the next 10 years because of climate change (Muller 2007).

All these changes will increase challenges with respect to water resources as they become more pronounced. Degradation of water quality, property damage, and potential loss of life due to flooding are caused by extreme rainfall events. Damage from erosion can impact areas from farm fields to stream banks adjacent to important infrastructure (Wright et al. 2010). To prepare for future climate changes, it is imperative that we review and update current standards for water management infrastructure design. This would prevent water management infrastructures from performing below the designated guidelines in the future (Prodanovic et al. 2007).

#### **1.2.General Circulation Models (GCMs)**

The first climate model was developed by Norman Phillips (1956). This mathematical model could successfully show monthly and seasonal patterns in the troposphere (Cox 2002). Other scientists started working to develop General Circulation Models (GCMs) in the late 1960s, when the first oceanic and atmospheric model was developed at the National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory (Lynch 2006, NOAA 2007). The National Center for Atmospheric Research (NCAR) developed the Community Atmosphere Model in the early 1980s (Collin 2004), and many more climate models have been created since then. These General Circulation Models (GCMs) are currently the best tools to study the effects of climate change. They are the current state-of-the-art in climate science. The main goal of these models is to identify the climate system's functioning, employing various sciences such as physics and chemistry (Prodanovic et al. 2007). GCMs use complicated equations representing the physical processes of the atmosphere. Significant computing resources are required to produce meaningful forecasts in this complex system, which includes

interactions between the different components of ocean layers, temperature gradients, trade winds, solar radiation, clouds, land masses and ocean currents (Prodanovic et al. 2007). To capture large-scale circulation flows, GCMs operate at medium coarse resolutions (between 250 and 600 km). GCMs picture the climate in three dimensions throughout the globe, with 10 to 20 vertical layers in the atmosphere and as many as 30 layers in the ocean (IPCC 2007). GCMs are the best tools available to obtain projections of different variables, such as precipitation, temperature, humidity, solar radiation, wind speed, etc. Figure 1.1 displays how a climate model is built. The world is broken up into many smaller areas or grid cells depending of the resolution of the model. The location of each grid cell is determined based on its position on earth. There are vertical layers above each grid cell that can be defined at different pressure levels.

The lower left picture shows all the physical processes happening in each grid cell of the model. Different kinds of surfaces (e.g. ocean, land, ice) are defined in the model, and various variables such as temperature, wind, and many others will be calculated in the model (http://www.cmmap.org).

#### **1.3. Regional Circulation Models (RCMs)**

Regional Climate Models (RCMs) perform the same as GCMs but over a restricted domain, which allows higher resolution and leads to more detailed outputs (Barron and Sorooshian 1997). They are downscaling tools that add high-resolution information to large-scale projections of GCMs. By resolving features down to 50 km or less, RCMs provide more accurate depictions of surfaces, such as coastlines, mountain topographies, small islands, etc. RCMs also make a more realistic representation of small islands as compared to global models, where their climate would be considered the same as the surrounding ocean (Barron and Sorooshian 1997, Jones et al. 2004).



**Figure 1.1** Representation of a climate model setup (from http://www.cmmap.org/learn/modeling/whatIs2.html retrieved on 03/12/2013)

Like GCMs, they are full, physically-based models that represent all the interactions and processes between climate system components (Jones et al. 2004). Dynamical, statistical, and hybrid (statistical-dynamical) methods are three techniques used to obtain regional climate projections; RCMs are categorized in a dynamical technique group (Jones et al. 2004). They can be thought of as being composed of three layers. One layer is largely driven by the GCM, another layer builds on some locally specific data, and the third layer uses its own physics-based equations to resolve the model based on data from model based on data from the other two. The results are comparatively local projections that are informed by both local specifics and global models. This process requires significant computational resources because it depends on the use of complex models. Figure 1.2 demonstrates an example of precipitation prediction by RCM and GCM over the Alps and Pyrenees between the present and 2080. It is obvious that more detailed results can be obtained by incorporating RCM that can represent the effects of these mountains on the weather much better than GCM.



Figure 1.2 Predicted changes in winter precipitation over Central/ Southern Europe between the present day and 2080 predicted by the RCM (right) and GCM (left). (from http://climateprediction.net/content/regional-climate-models retrieved on 03/12/2013)

#### **1.4.Climate Model Scenarios**

To predict or estimate the changes that might be expected to happen in future climate, some assumptions about the status of future (e.g.  $CO_2$  concentration, population growth, and economic development) need to be made (IPCC 2007). IPCC has developed a set of future emission scenarios called Special Report on Emissions Scenarios (SRES). Four qualitative storylines yield four sets of scenarios called "families": A1, A2, B1, and B2. The A1 scenario family assumes a future of global population that reaches the highest point at mid-century and drops from that time forward, very rapid economic growth, and rapid introduction of new and more efficient technologies (IPCC 2007). The A1 scenario family includes three groups identified by their technological emphasis: fossil-intensive (A1FI), non-fossil energy sources (A1T), or a balance across all sources (A1B). The A1FI and A1T scenarios specifically analyze the alternative energy technology developments and hold other driving forces constant.

The A2 scenario (Nakicenovic et al. 2000) family assumes a population of 15 billion by 2100. Self-reliant nations, slower changes in technology, and increased incomes are some of the factors that define the scenario. Economic development is regionally oriented, and economic growth and technological changes are more fragmented and slower than in other storylines. A continuous increase in  $CO_2$  emissions related to land use as a result of the increasing population without a proportional increase in agricultural productivity is also demonstrated in this scenario (Nakicenovic et al. 2000).

The B1 scenario family considers a world with the same global population as A1, rapid changes in the economy, and the introduction of clean and resource-efficient technologies. This scenario's emphasis is on global solutions to economic, social, and environmental sustainability without additional climate initiatives.

The B2 scenario family highlights local solutions to economic, social, and environmental sustainability. The world's population is increasing at a lower rate compared to the A2 scenario. There are intermediate levels of economic development and less rapid and more diverse technological change than in the B1 and A1 scenarios.

6

In most scenarios, the global forest area will decrease for some decades due to the population increase and income growth. They will increase by the end of the  $21^{st}$  century, with the greatest increase in the B1 and B2 scenario families. Agricultural land use changes are associated with changing food demands. Many other social, technological, and economic factors affect agricultural, forest, and other types of land change. All of the driving forces mentioned above will affect CO<sub>2</sub> and other GHG emissions. Figure 1.3 shows the range of total CO<sub>2</sub> emissions in gigatons of carbon (GtC/year) from all sources (energy, industry, and land-use change) for different scenario groups from 1990 to 2100.



**Figure 1.3** Total global annual CO<sub>2</sub> emissions [in gigatons of carbon (GtC/year)] for six scenario groups.

#### **1.5. Bias Correction**

General Circulation Models (GCMs) are important tools in evaluating climate change impacts and decision-making (Wood et al. 2004, Li et al. 2010). However, since models are not perfect, a 20<sup>th</sup> century climate projected from a model is not the same as the climate of the 20<sup>th</sup> century based on observations. Hence, GCM precipitation outputs cannot be used in hydrological models or in decision-making without performing some form of bias correction (Sharma et al. 2007, Hansen et al. 2006, Feddersen and Andersen 2005). A realistic presentation of future precipitation from global climate models is extremely important for vulnerability and impact assessment (Wood et al. 2004, Schneider et al. 2007). Therefore, modelers use bias correction techniques to represent more realistic GCM outputs by establishing a relationship between climate model outputs and observations and then using that relationship to transform the simulated 21<sup>st</sup> century climate to a "best guess" 21<sup>st</sup> century climate. These techniques are given a variety of names in the literature, such as statistical downscaling, histogram equalizing, and quantile-based mapping (Piani et al. 2010).

#### **1.6. Downscaling Techniques**

#### 1.6.1. Spatial Downscaling

GCMs are developed to simulate the current climate and predict the future climate change under different climate scenarios. However, the projected climate at a global scale

does not provide a satisfactory output to use in a hydrological model at a watershed or regional scale (Solaiman 2011). The spatial resolution of GCMs usually ranges from 250 km to 600 km, which is a coarse resolution for studies at a watershed scale. The accuracy of GCMs decreases at smaller temporal and spatial scales. Therefore, the projected results from the GCMs are downscaled to a higher resolution level (Solaiman 2011). "Downscaling" is based on the view that the regional climate is conditioned by the climate on a larger scale (Von Storch 1995, 1999a). Information is cascaded "down" from larger to smaller scales. The regional climate is the result of the interplay of the overall atmospheric or oceanic circulation and of regional specifics, such as topography, land-sea distribution, and land use (Von Storch 1995, 1999a). The confidence that may be placed in downscaled climate change information is dependent on the validity of the large-scale fields from the GCM. Since different variables have different characteristic spatial scales, some variables are considered more realistically simulated by GCMs than others. However, there is no consensus in the community about what level of spatial aggregation is required for the GCM to be considered skillful.

Two common downscaling techniques are dynamical downscaling and statistical downscaling. Running a set of regional climate models, dynamical downscaling process uses the relatively coarse resolution output from the GCMs as a continuously updated boundary condition for the high-resolution regional models. The most common technique to perform dynamical downscaling is using RCMs (Brissette et al. 2007, Solaiman 2011). Utilizing dynamical downscaling requires computational efforts that make it impractical when several GCMs and emission scenarios are used (Maurer et al. 2007). Figure 1.3

shows how the downscaling techniques can provide a better representation of GCMs with higher resolution output to use in regional climate change studies.



**Figure 1.4** Downscaling techniques allow the use of GCMs outputs as inputs into regional climate and weather models. From http://www.southwestclimatechange.org/climate/modeling/downscaling retrieved on 03/12/2013

http://www.southwesternhateenange.org/ennhate/modernig/downsearing/fetrieved on 03/12/2013

Statistical downscaling is more popular due to its relative computational ease. Generally, they are categorized as weather generators, transfer functions and weather typing schemes. Weather generators are statistical models that generate random but realistic daily sequences of variables, such as precipitation and temperature. These data are sometimes called synthetic data. Usually, they generate precipitation first, and other data are later derived using the statistical relationships between precipitation and these data (Von Storch 1999b, Katz and Parlange 1996). Transfer functions use a quantitative relationship, such as regression models or linear or non-linear interpolations (Von Storch 1999b). Sailor and Li (1999) modeled local temperature at different stations in the U.S. using multiple regression models, and Dehn and Buma (1999) specified precipitation at a French Alpine site. An alternative to linear regression is piecewise linear or nonlinear interpolation, such as kriging tools (Wackernagel 2003). Biau et al. (1999) used this approach to relate local precipitation to large-scale pressure distributions. Other studies have used cubic splines to specify precipitation (Buishand and Klein Tank 1996, Buishand and Brandsma 1997, 1999). Another non-linear technique is based on Artificial Neural Networks (ANNs), which is more powerful than other methods, but the interpretation of the dynamical character of the relationships is less easy (Von Storch 1999b).

Weather typing schemes are based on a more traditional synoptic climatology concept, including analogs and phase space partitioning. They define weather classes related to local and regional climate variations (Von Storch 1999b). These weather classes are defined synoptically or fitted specifically for downscaling purposes by constructing indices of airflow (Conway et al. 1996). Weighting the local climate states with the relative frequencies of the weather classes and then the frequency, distribution, or mean of the regional climate are derived, and climate change is estimated by determining the frequency of weather classes (Von Storch 1999b). The regional and local climates are usually obtained from observations. Zorita et al (1995) introduced the analog method for downscaling and its applicability for the specification of precipitation was tested by Cubasch et al. (1996), Martin et al. (1996), and Biau et al. (1999). Techniques that partition the large-scale state phase space such as Classification Tree Analysis are other methods which are mathematically more demanding techniques (Zorita et al. 1995, Lettenmaier 1999).

#### **1.6.2.** Temporal Disaggregation

Many hydrologic models or studies related to hydrology and continuous watershed simulations require long-term meteorological data, such as precipitation (Burian et al. 2000). However, most of the available precipitation data are in hourly or daily intervals, which are still not small enough time intervals to estimate the accurate hydrologic response for small watersheds or developing IDF curves. Time intervals on the order of minutes (e.g. 15-min or 30-min) are suggested for accurate characterization in urban settings or many other hydrologic studies (Burian et al. 2000). To obtain such high temporal resolution data, the hourly or daily data need to be disaggregated. Several techniques have been developed and proposed to disaggregate precipitation data. Most of these methods discuss disaggregation of the daily or longer time interval precipitation (e.g. Hershenhorn and Woolhiser 1987, Glasby et al. 1995). Liou (1970) developed a method using four explicit distributions to disaggregate the hourly rainfall to 15-min. rainfall. The selection of distribution to use in a specific storm hour was based on a comparison of precipitation depths in that hour with those of the preceding and successive hours. Ormsbee (1989) developed two disaggregation models to disaggregate hourly precipitation data. One model, called a discrete model, disaggregates the hourly

precipitation to three 20-min intervals precipitation. The other model — a continuous model— disaggregates the hourly rainfall into 1- to 30-min time intervals. The probability distribution of the accumulated rainfalls at different time steps was used by Pandey (1994) to investigate the scaling nature of probability distribution. They used the multi-fractal multiplicative cascades theory to explain the rainfall probability distributions at different time scales and tested the method on an example in Quebec. Results showed that the method can successfully estimate the distribution of short duration rainfall.

Socolofsky et al. (2001) introduced a method to downscale daily precipitation data to hourly data. The method breaks daily precipitation into possible storm intensity patterns by selecting samples of measured event statistics from hourly observed precipitation data.

Rodriguez-Iturbe et al. (1987, 1988) investigated two point-process based models. In the Rectangular Pulses Model, storms arrive according to a Poisson process. In this model, each event has a random duration and constant intensity. The second model simulates a storm arising from a Poisson process when each storm is associated with a cluster of cells and each of the cells has a random depth and duration. Examples of clustered Poisson models are the Bartlett-Lewis and Neyman-Scott models (Rodriguez-Iturbe et al. 1987, 1988, Islam et al. 1990, Onof and Wheater 1993). Other models have been used to capture the variability between and within storms but they usually rely on extrapolating scaling at course resolution to finer scales (Schertzer and Lovejoy 1987, Gupta and Waymire 1993, Olsson 1996, Veneziano et al. 1996). Hershenhorn and Woolhiser (1987) and Econonpouly et al. (1990) developed a model to disaggregate daily rainfall by simulating the number of storms in a day and the amount, duration, and starting time of each event. The number of rainfall events, duration, and start time are obtained from an estimated probability distribution fit to observed hourly data within the same climatological regime.

Several other studies investigated the possibility of using Artificial Neural Networks (ANNs) for rainfall temporal downscaling. Burian et al. (2000) compared the results of rainfall disaggregation using an ANN model with two other methods; linear model and a continuous deterministic rainfall disaggregation model (Ormsbee 1989). Results of their study indicated that the latter two models underpredicted maximum rainfall intensities and were outperformed by the ANN model. Dibikie and Coulibaly (2006) evaluated the ANN model in disaggregating precipitation and compared it with the linear regression model and concluded that ANN outperforms the statistical models. Other studies have also reported better performance in predicting heavy rainfall events using the ANN model compared to linear regression downscaling methods (Weichert and Burger 1998).

#### 1.7. Intensity-Duration-Frequency (IDF) Curves

An IDF curve presents the probability of a given rainfall intensity and duration expected to occur at a particular location. Figure 1.5 shows an IDF graph. Lines on the graph represent probability. For example, the five-year line would represent rainfall events with a probability of occurring once every five years. Another way to explain it is that the probability of a five-year magnitude storm happening in any given year is 1/5 or 20%. The information presented on the IDF graph is based on a statistical analysis of an actual storm prediction (Bedient and Huber 1988).

Each line in the IDF graph shows rainfall events with the same occurring probability for different durations. A 50-year storm, for example, can be a 50-year, 15-min storm, a 50-year, 1-hr storm or a 50-year 12-hr storm. Rainfall intensity is higher for short storms and lower for longer storms (Bedient and Huber 1988).



Figure 1.5 Rainfall Intensity-Duration-Frequency (IDF) curve

An IDF curve is one of the most common hydrologic tools for designing various structures, such as, stormwater management facilities, erosion and sediment control structures, flood protection structures, and drainage structures in urban areas (McCuen 1989, Prodanovic et al. 2007). Typically, standards are set for designs that require the minimum capacity in terms of rainfall return periods. For example, storm sewers are typically designed to carry a minimum of a five-year storm meaning that all the runoff

from a five-year storm from the upstream of the sewer system must fit in the sewer without overflowing. The first step toward designing different water management structures (e.g. dams, channels, detention ponds) is to identify the characteristics of design storm in terms of duration, return period and intensity. Time of concentration of the watershed draining to the hydraulic structure usually dictates the design storm duration. Storm return period is assigned based on economic assessments and risk analysis (probability of damage, loss of life in case of failure). For example, a 5- to 10-yr return period is used for designing roadside channels (Brown et al. 2001), where the cost of failure is negligible. Whereas, a much larger return period is used for designing small dams (100-yr and longer), where there is a great risk of loss of life in case of failure. With knowledge of the storm return period, the design rainfall intensity is then acquired from the developed IDF curves of the region. The rainfall intensity and duration of design storm dictate the cost of the hydraulic structure, and any uncertainty in estimation of these parameters can impose great design uncertainties.

IDF curves are created by analyzing historic rainfall event statistics. More complete and longer data provide a better statistical analysis. Cumulative frequency analysis is used to analyze the rainfall records by determining the event intervals with a specific magnitude for a range of metrics within the rainfall data (e.g. a 30-min storm with the maximum rainfall value that occurs at a 25-yr interval). A Cumulative Distribution Function (CDF) is usually used for fitting the precipitation data. For a given return period ( $T_r$ ), the cumulative frequency (F) is:

$$F = \mathbf{1} - \frac{\mathbf{1}}{T_r} \tag{1}$$

With knowledge of the cumulative frequency, the maximum rainfall intensity can be found using an appropriate theoretical distribution function. The most common distribution functions used for developing IDF curve are Gumbel, Log-Pearson Type III, Weibull, and Generalized Extreme Value (GEV) distributions.

In the U.S., IDFs are presented in the form of isohyetal maps in three reports: Technical Paper No. 40 (TP-40) (Hershfield 1961) presents the IDF curves for durations from 30 minutes to 24 hours and return periods from one to 100 years; Technical Paper No. 49 (Miller 1964) shows the IDFs for durations from two to 10 days for return periods of two to 100 years; and Hydro-35 (Frederick et al. 1977) showed the IDFs for durations from five to 60 minutes for just two return periods; two and 100 years. NOAA Atlas 14 presents updated precipitation-frequency atlas of the United States for durations from five minutes to 60 days (Bonnin et al. 2004). Updating precipitation-frequency is still an ongoing process and is expected to be completed by June 2013.

Many different studies have been undertaken to develop the rainfall IDF curves. Huard et al. (2010) estimated the IDF curves using Bayesian analysis instead of the classical approach. Svensson et al. (2007) compared different methods of IDF curve estimation from fragmented records for Scotland. Even though studies have investigated various methods for developing IDF curves, studies associated with changes to IDF curves due to climate change are limited. Simonovic and Peck (2009) updated the rainfall IDF curves for the period from 1961 to 2002 under two climate scenarios for the city of London, Ontario, Canada. In a continuous study, Solaiman and Simonovic (2011) developed future IDF curves using 27 climate change scenarios for the city of London, Ontario, Canada. IDF curves were created for durations from one to 24 hours and return periods from two to 100 years. The results of their study showed 20-40% changes in different durations for all return periods. The climate change impact on rainfall IDF curves was evaluated by Wang et al. (2013) in the Apalachicola River basin (Florida Panhandle coast) using two climate models: HRM3–HadCM3 and RCM3–GFDL. Their study suggested that, based on HRM3–HadCM3 projections, there will be no significant changes in rainfall intensity at the upstream and middle stream stations, but higher rainfall intensity is expected at the downstream station. Analysis of the RCM3–GFDL projections, on the other hand, showed that rainfall intensity is expected to increase from upstream to downstream. The future IDF curves developed by Nguyen et al. (2006, 2007) were developed using two climate model projections; HadCM3 and CGCM2 showed a large difference in the IDF curves of these two models.

Using the projections from CGCM2 B2 to develop the IDF curves for the Grand River and Kenora Rainy River regions in Ontario indicated an increase from 24% to 35% in the 24-hour rainfall intensity for the 2050s and 2080s (Coulibaly and Shi 2005). Another study investigated the climate change impact on rainfall IDF curves using the outputs from the CRCM climate model for Southern Quebec (Mailhot et al. 2009). An analysis of the results showed a 50% decrease by 2050 for rainfall durations of two and six hours and a 32% decrease for 12- and 24-hour rainfall durations. The IDF curves developed from the projections of RCM A2 for the next 80 years for Denmark showed that an increase of 2 to 15% in extreme rainfalls is expected (Onof and Arnbjeg-Nielsen 2009).

#### **1.8. Problem Statement**

Changes in extreme rainfall events can lead to a revision of standards for designing civil engineering infrastructures. It can also lead to the reconstruction and/or upgrade of existing civil engineering infrastructures. Current design standards are based on historic climate information. For example, a dam that is designed to control a 100-year flood event will provide a significantly lower level of protection if the intensity and duration of the 100-year flood event increases. To prepare for future climate changes, it is imperative that we review and update the current standards for water management infrastructure design. This would prevent water management infrastructures from performing below the designated guidelines in the future (Prodanovic et al. 2007).

Most of the studies related to the impact of climate change on future rainfall intensity do not consider short rainfall durations for developing the IDF curves (e.g. 10or 15-min). The reason is that the future projections derived from GCMs are usually provided at coarse temporal resolutions; as a result, downscaling is required to obtain a high temporal resolution data. However, in most studies (Wang et al. 2013, Svensson et al. 2007), temporal disaggregation has been skipped and IDF curves developed with the available precipitation data from the GCMs.

Another issue with some of the studies is using only one or two GCM projections to develop the future rainfall IDFs (Wang et al. 2013, Nguyen et al. 2006, 2007, Coulibaly and Shi 2005, Mailhot et al. 2009). There are significant uncertainties in the results of different GCMs. The difference in physical parameterizations, especially of radiative and

precipitation-forming processes, amongst different GCMs and RCMs, the simplification of complex processes in the atmosphere, and the difference in initial and boundary conditions for each climate projection are some of the reasons that explain the existing uncertainty. Therefore, using only one climate model may represent one of many possible results and cannot be considered representative of the future. Therefore, it is important to use more than one GCM for a complete and thorough assessment of possible changes in the future.

#### **1.9. Dissertation Objectives**

The objectives of this dissertation are as follows:

- To develop rainfall IDF curves for Alabama under future climate change scenarios; using a stochastic method to disaggregate three-hourly precipitations into 15minute precipitations.
- Use Artificial Neural Network (ANN) to estimate the maximum 15-, 30-, 45-, 60-, and 120-min rainfall depths from three-hourly precipitations for the creation of future IDF curves for Alabama.
- 3. To develop probability-based IDF curves incorporating climate projections from six different climate models.

#### **1.10. Dissertation Structure**

This dissertation includes six chapters. Chapter 1 provides an introductory overview, a review of literature, and the objectives of the study. A brief review of different climate change models, climate model scenarios, bias correction and downscaling techniques,

and a review of studies related to the assessment of climate change on IDF curves is presented. Chapters 2 through 4 present a discussion of the methodology and results of the three objectives outlined above. Chapter 2 discusses the development of rainfall IDF curves for Alabama under future climate change scenario. The details of performing temporal disaggregation using stochastic method are also presented. The results of this chapter have already been published in the Regional Environmental Change (Mirhosseini et al. 2012). Chapter 3 describes the methodology used to develop an ANN model to estimate the maximum rainfall depths (15-, 30-, 45-, 60-, and 120-min) from three-hourly precipitations. It also presents the future IDF curves developed using the ANN disaggregation model and compares the outcomes with the previous chapter's results. This chapter has been submitted for publication to the Hydrological Processes Journal (Mirhosseini et al. 2013). Chapter 4 documents the development of a probability-based IDF curve using a kernel density estimator using all climate models investigated in this research. This chapter will be submitted to Journal of Hydrologic Sciences for publication. Chapter 5 presents the conclusions of this research. Finally, suggestions for future work are presented in Chapter 6.

#### Chapter 2

## 2. The Impact of Climate Change on Rainfall Intensity-Duration-Frequency (IDF) Curves in Alabama

#### 2.1. Abstract

Changes in the hydrologic cycle due to increase in greenhouse gases (GHG) are projected to cause variations in intensity, duration, and frequency of precipitation events. Quantifying the potential effects of climate change and adapting to them is one way to reduce vulnerability. Since rainfall characteristics are often used to design water management infrastructures, reviewing and updating rainfall characteristics (i.e., Intensity-Duration-Frequency (IDF) curves) for future climate scenarios is necessary. This study was undertaken to assess expected changes in IDF curves from the current climate to the projected future climate. To provide future IDF curves, 3-hourly precipitation data simulated by six combinations of global and regional climate models were temporally downscaled using a stochastic method. Performance of the downscaling method was evaluated and IDF curves were developed for the State of Alabama. The results of all six climate models suggest that the future precipitation patterns for Alabama are expected to veer towards less intense rainfalls for short duration events. However, for long duration events (i.e. > 4 hours), the results are not consistent across the models. Given that a large uncertainty existed for projected rainfall intensity of these six climate models, developing an ensemble model as a result of incorporating all six climate

models, performing an uncertainty analysis and creating a probability based IDF curves could be proper solutions to diminish this uncertainty.

#### **2.2. Introduction**

There is a misunderstanding on the subject of climate change that causes great confusion about its potential effects on natural resources and human lives (Miller and Yates 2006). Much of the public is left with two extreme and opposing views about climate change: either climate change is happening at the greatest rate and soon nothing will be left on earth, or climate change is a fiction that can be safely ignored. However, neither of these impressions are useful in making management decisions about natural resources.

An analysis of temperature records shows a significant warming trend around the world during the 20<sup>th</sup> century (Miller and Yates 2006). The global average temperature increased by 0.74°C over the 20th century. The changing temperature has caused changes in the seasonal timing of runoff in mountainous areas (Stewart et al. 2004). Despite the total increase in winter precipitation in the western U.S. and some parts of Canada; spring snowpacks have been melting earlier (Stewart et al. 2004). Along with the change in temperature, there will be changes in the hydrologic cycle, leading to changes in precipitation and runoff. Sea level rise and new stresses on ecological systems (e.g. forests and freshwater aquatic systems) may be additional climate change impacts on biological and physical systems.

There is scientific evidence that human activities play a significant role in the temperature increase. During the last century, the concentration of carbon dioxide ( $CO_2$ ) and other greenhouse gases (GHGs) in the earth's atmosphere have risen due to increased

industrial activities (e.g. burning fossil fuels can release carbon dioxide into the atmosphere) (Prodanovic et al. 2007). This increase in GHG concentrations is causing large-scale variations in atmospheric processes, which can lead to changes in precipitation and temperature characteristics. There is scientific agreement on some important features of the possible hydrologic changes: there will be an increase in the average precipitation around the world (Miller and Yates 2006). However, this does not imply that more precipitation will occur everywhere. Climate change impacts on precipitation are less certain on the regional scale, and climate models are not consistent in predicting the expected precipitation, which may change with global climate change. While it is not possible to make very accurate predictions regarding the overall quantity of precipitation for a specific region, scientific theory suggests that there will be more intense but less frequent periods of precipitation. In other words, there may be longer periods of drought alternating with heavy rainfalls.

Changes in extreme rainfall events often present challenges for water management utilities. Heavy runoff can degrade the water quality, forcing additional costs for treatment and increasing the risk of water supply contamination by pathogens (Miller and Yates 2006). Property damage and potential loss of life due to flooding are also caused by extreme rainfall events (Wright et al. 2010). Heavy precipitation caused by a hurricane in 1999 left a water treatment plant in Greenville, NC, surrounded by floodwaters and damaged the utilities' infrastructure, costing about 11 million dollars (Miller and Yates 2006). Drought can also decrease the capability of water utilities to meet water demands and enforce emergency restrictions. A severe drought in 2002 left many Colorado reservoirs dangerously low (Miller and Yates 2006). Forest fires are becoming common as the temperature increases. Fire can cause serious effects on downstream water quality and reservoir sedimentation. Erosion can damage areas from farm fields to stream banks adjacent to important infrastructure (Wright et al. 2010). For example, in 1996, a flood occurred after a Buffalo Creek fire, which caused severe sediment and debris flow into Denver's Strontia Springs Reservoir. It required cleanup and resulted in long-term water quality effects, imposing heavy costs on the utility (Miller and Yates 2006).

Such extreme events may become common and more difficult to anticipate in the future because of global climate change. They can lead to a revision of standards for designing civil engineering infrastructures. It can also lead to the reconstruction and/or upgrade of existing civil engineering infrastructures. Current design standards are based on historic climate information. For example, a dam that is designed to control a 100-year flood event will provide significantly less protection if the intensity and duration of the 100-year flood event increases. Many industries have already started long-term planning in the context of uncertainty. Accounting for potential changes in water consumption patterns because of socio-economic or demographic changes is an example of water utilities' long-term planning. Climate change, as an extra source of uncertainty, is becoming significantly relevant to water resource managers (Miller and Yates 2006). Like any other source of uncertainty, the best practice is to attempt to understand the possible changes that can happen in the future and the consequences that these changes can have on the management of infrastructure. Therefore, to prepare for future climate
changes, it is imperative that we review and update the current standards for water management infrastructure design. This will prevent water management infrastructure from performing below the designated guidelines in the future (Prodanovic et al. 2007). This study was funded by the National Oceanic and Atmospheric Agency (NOAA) Regional Integrated Sciences and Assessments (RISA) program. Its main objective was to create IDF curves for Alabama using high-resolution projections (for 2038-2070) derived from dynamical downscaling of General Circulation Models (GCMs) by Regional Climate Models (RCMs) and to evaluate the impact of climate change on IDF curves.

#### **2.3. Materials and Methods**

### 2.3.1. Data and Model Used

The stations providing long-term historical precipitation data for Alabama are shown in Figure 2.1. The observed (historical) precipitation data at 15-minute intervals were obtained from the NOAA National Climatic Data Center (NCDC Online Climate Data Directory 2005). Historical simulations of precipitation for 1968 to 2000 and future projections for 2038 to 2070 were obtained from the North American Regional Climate Change Assessment Program (NARCCAP) at three-hour intervals with a spatial resolution of 50 km (Sebastien et al. 2007, Richard et al. 2007). NARCCAP was designed to investigate the uncertainties in future climates at a regional scale (Mearns et al. 2007). To this end, it uses several GCM historical simulations and projections from the Intergovernmental Panel on Climate Change (IPCC) Coupled Model Intercomparison Project and dynamically downscales them using a set of RCMs. The regional downscaling process uses the relatively coarse resolution output from the GCMs as a continuously updated boundary condition for the high-resolution RCMs.



Figure 2.1 Location of rain gauge stations (red squares) and NARCCAP 50-km resolution grid centers (green dots, HRM3-HadCM3) used to develop IDF curves for Alabama.

The regional downscaling domain of NARCCAP covers the U.S. and most of Canada (Mearns et al. 2007, 2009, Sebastien et al. 2007, Richard et al. 2007).

As mentioned above, future projections of precipitation data for this study were obtained from the NARCCAP website. Six different dynamically downscaled datasets were used. In the entries below, the datasets are named with the RCM identifier, followed by the identifier of the GCM providing boundary conditions.

## 1. HRM3-HadCM3:

The Hadley Centre Coupled Model, version 3 (HadCM3) was developed at the Hadley Centre in the United Kingdom. It is a coupled atmosphere-ocean general circulation model (AOGCM) and was one of the major models used in the IPCC Third Assessment Report in 2001 (Gordon et al. 2000, Pope et al. 2000, Collins et al. 2001). Two components of HadCM3 are the atmospheric model (HadAM3) and the ocean model (HadOM3), which includes a sea ice model. The model's resolution is about 300 km (Jana and Majumder 2010). Unlike the earlier models developed at the Hadley Centre and elsewhere, HadCM3 does not need a flux adjustment to produce a good simulation. It has higher ocean resolution than HadCM2, which is one of the major features of this model. It also provides a good match between the oceanic and atmospheric components. The ocean mixing scheme has been improved in this model, as well. The HadOM3 resolution is 1.25 degrees longitude x 1.25 degrees latitude. It has 20 levels and a time step of one hour. The HadAM3 horizontal resolution is 3.75 longitude x 2.5 degrees latitude. There are 19 levels in the vertical.

The Hadley Centre Regional Model, version 3 (HRM3) is a configuration of the HadCM3 model that provides high resolution projections of regional climates for impact studies; the resolution is about 50 km (Gordon et al. 2000; Pope et al. 2000; Johns et al. 2003). The Hadley Centre for the UK climate impact programme, first used the model in 2002, and it has been used for many high-resolution simulation studies since then. HRM3 includes the same atmospheric component as HadCM3, but some modifications were made to the model physics. It provides a higher spatial resolution of a local area by taking the boundary conditions of coarse resolution global model simulations (Gordon et al. 2000, Pope et al. 2000, Johns et al. 2003).

## 2. CRCM-CGCM3:

The Coupled Global Climate Model, version 3 (CGCM 3.1) was developed at the Canadian Centre for Climate Modelling and Analysis (CCCma). CCCma is a division of the Climate Research Branch of the Meteorological Service of Canada of Environment Canada, based at the University of Victoria in Victoria, British Columbia. Three atmosphere and three coupled atmosphere-ocean GCMs have been produced since 2000 at this center.

CGCM3 uses the same component that was used in developing CGCM2, and it also uses the updated atmospheric component of the Third Generation Atmospheric General Circulation Model, AGCM3 (Flato and Boer 2001, Kim et al. 2002, 2003). Like AGCM2, the horizontal structure of the main prognostic variables is presented by the spectral transform method, and the vertical representation is in terms of a rectangular finite element defined for a hybrid vertical coordinate (Laprise and Girard 1990). AGCM3 has a higher horizontal resolution than AGCM2. The vertical domain is deeper, and the vertical resolution is higher. CGCM3 extends from surface to the stratopause region, and this region spans 32 layers.

CRCM 4.2 is a Canadian Regional Climate Model developed at the Université du Québec à Montréal based on the fully elastic, non-hydrostatic Euler equations (Laprise et al. 1998, Caya and Laprise 1999). The physical parameterization is mostly based on CCCma GCM3. Solar radiative transfers were improved with three bands in the near-infrared region and one band in the visible region (Puckrin et al. 2004). The radiative effects of GHGs are considered separately for CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, CFC<sub>11</sub>, and CFC<sub>12</sub>. The land surface scheme is the Canadian LAnd Surface Scheme (CLASS) 2.7 (Verseghy 1991, Verseghy et al. 1993). CLASS uses three soil layers (0.1 m, 0.25 m and 3.75 m thickness). It includes prognostic equations for energy and water conservation for the three soil layers and a thermally and hydrologically distinct snowpack where applicable.

## 3. HRM3-GFDL:

GFDL CM 2.1 is a coupled atmosphere-ocean general circulation model (AOGCM) developed by the Geophysical Fluid Dynamics Laboratory (GFDL) at Princeton University (Anderson et al. 2004). It is one of the major models used in the Fourth Assessment Report, along with HadCM3. The atmospheric component (AM 2.1) covers 180° to 45°W longitude, 10°N to 75°N latitude and has a 24-level atmosphere run at a resolution of two degrees in the east-west and 2.5 degrees in the north-south direction. This resolution is sufficient to resolve the large mid-latitude cyclones responsible for weather variability.

The atmosphere model includes a representation of radiative fluxes, mixing in the atmospheric boundary layer, representations of the effects of stratus and cumulus clouds, a scheme for representing drag on upper-level winds caused by gravity waves, changes in the spatial distribution of ozone, and the ability to represent the impact of multiple greenhouse gasses.

The ocean component has 50 levels running at one degree in the east-west direction, and it varies from one degree to 1/3 degree along the equator in the north-south direction. This resolution is good enough to resolve the equatorial current system. The other parameterization includes a free surface height that changes in response to evaporation, the precipitation and convergence of ocean currents, the absorption of sunlight tied to observed chlorophyll concentrations, and a representation of the oceanic mixed layer.

## *4. CRCM-CCSM:*

The Community Climate System Model (CCSM) is a coupled climate model developed at the National Center for Atmospheric Research (NCAR) in Boulder, Colorado. It simulates past, present, and future climates (Collins et al. 2006a). The model includes ocean, atmosphere, sea ice and land surface. CGCM3 is the third generation of coupled models designed to produce realistic simulations and includes an atmosphere model (CAM3) (Collins et al. 2004, 2006b), the land surface (CLM3) (Oleson et al. 2004; Dickinson et al. 2006), the sea ice (CSIM5) (Briegleb et al. 2004), and an ocean model based on Parallel Ocean Program (POP) version 1.4.3 (Smith and Gent 2002). Each component is designed to conserve total water, energy, mass, and fresh water. The CAM3 is based on the Eulerian spectral dynamical core with triangular spectral truncation

(Collins et al. 2006b). The zonal resolution at the equator ranges from 3.75 degrees to 1.41 degrees. It includes 16 vertical layers (Collins et al. 2006b).

The ocean model has a horizontal resolution of three degrees or one degree. There are 25 levels extending to 4.75 km in the three-degree version and 40 levels extending to 5.37 km in the one-degree version. The sea-ice model is integrated on the same horizontal grid as the ocean model (Collins et al. 2006b).

The CLM3 model is built based on the nested sub-grid hierarchy of scales representing land units, soil or snow columns, and plant functional types (Oleson et al. 2004). CCSM3 considers the effect of competition for water among plant functional types. One of the major goals in developing the land model was to reduce the positive continental temperature biases during boreal winter. The relationship between fractional snow coverage and snow height was modified. The biogeophysics formulation was modified to increase the sensible and latent heat fluxes over sparsely vegetated surfaces. In previous versions of CCSM, a constant value for dense vegetation was considered for the turbulent transfer coefficient between soil and the overlying canopy air. The new formulation considers this coefficient to be dependent on the density of the canopy, characterized by leaf and stem area indices (Oleson et al. 2004, Collins et al. 2006a).

### RCM3-GFDL:

Regional Climate Model (RCM3 or RegCM) version 3, is a three-dimensional, hydrostatic, compressible, primitive equation,  $\sigma$ -coordinate regional climate model originally developed at the National Center for Atmospheric Research (NCAR). It uses the NCAR CCM3 radiation scheme. The solar component which accounts for the effect of H<sub>2</sub>O, O<sub>3</sub>, CO<sub>2</sub>, and O<sub>2</sub>, uses the  $\delta$ -Eddington approximation (Keihl et al. 1996). It includes 18 spectral intervals from 0.2 to 5µm. The surface physics are performed using Biosphere-Atmosphere Transfer Scheme version 1e (BATS1e) (Dickinson et al. 1993). BATS is a surface package designed to describe the role of vegetation and interactive soil moisture in modifying the surface-atmosphere exchanges of momentum, energy, and water vapor. The model has a vegetation layer; a snow layer; a surface soil layer, 10 cm thick, or a root zone layer, 1-2 m thick; and a third deep soil layer, 3 m thick. The model also includes a planetary boundary layer scheme, convective precipitation schemes, a large-scale precipitation scheme, a pressure gradient scheme, a lake model, and an aerosols and dust model.

# 5. *ECP2-GFDL*:

The updated data from the Regional Spectral Model were developed at the Experimental Climate Prediction Center (ECPC) at the Scripps Institute of Oceanography. It was originally called ECPC. The main difference between ECPC and ECP2 is the employment of spectral nudging. Spectral nudging forces the regional model solution to closely follow its driven global model for scales that are usually close to and above the synoptic scale. ECPC combines a spectral damping of horizontal wind tendency and areal average correction of temperature, moisture, and surface pressure, while ECP2 does not use the moisture correction component of the spectral nudging, which results in better simulation of precipitation. Also, a larger domain of ECP2 can more closely match the region modeled by the other RCMs.

### 2.3.2. Bias correction

Models are not perfect; a 20<sup>th</sup> century simulated climate, projected from a model is not the same as the climate of the 20<sup>th</sup> century coming from observations. Hence, GCM precipitation outputs cannot be used in hydrological models or in decision making without performing some form of bias correction (Sharma et al. 2007; Hansen et al. 2006; Feddersen and Andersen 2005). A realistic presentation of future precipitation from global climate models is extremely important for vulnerability and impact assessment (Wood et al. 2004; Schneider et al. 2007). Therefore, modelers use bias correction techniques to represent more realistic GCM outputs by establishing a relationship between climate model outputs and observations, then using that relationship to transform the simulated 21<sup>st</sup> century climate to a "best guess" 21<sup>st</sup> century climate. These techniques are given a variety of names in the literature, such as statistical downscaling, histogram equalizing and quantile-based mapping (Piani et al. 2010).

For this study, a quantile-based mapping method proposed by Li et al. (2010) was used. In this method, monthly rainfall values were used to define the CDF error of historical model runs relative to observations. This error was used to correct the model CDF for the future period by calculating a scaling factor from the monthly totals. The scaling factor is defined as bias-corrected rainfall total for a given month, divided by the non-bias-corrected total. Prior to disaggregation of 3-hourly rainfall events, the 3-hourly totals were multiplied by this scaling factor (Li et al. 2010).

## 2.3.3. Temporal Downscaling

High temporal resolution (e.g., 15-min, 30-min, and 1-hr) data are needed to create the IDF curves. Since NARCCAP provides future climate data at 3-hour intervals, it is necessary to temporally downscale the precipitation data. Different types of downscaling techniques have been developed over the years; they can be categorized as weather generators, transfer functions (e.g. linear regression, stochastic method and artificial neural networks) and weather typing schemes (Von Storch 1999b).

The temporal downscaling method employed in this study is a modified version of a stochastic method introduced by Socolofsky et al. (2001). They used the method to downscale daily precipitation data to hourly data. The method developed for this project breaks 3-hourly precipitation into possible storm intensity patterns by selecting samples of measured event statistics from a 15-min observed precipitation data (Socolofsky et al. 2001). Most stochastic methods are based on two popular models, Neyman-Scott and Bartlett-Lewis (Rodriguez-Iturbe et al. 1987, Islam et al. 1990). One of the main reasons for selecting the Socolofsky method was the reduction of computational effort needed to perform the method compared with the above-mentioned approaches.

## 2.3.3.1 Stochastic method

For the purpose of disaggregation, a month-specific database of observed

precipitation data was created. This database includes several observed rainfall events with high temporal resolution (15 minute interval). An "event" was defined as a continuous sequence of precipitation, separated by 30 minutes of dry weather. In case an event was longer than 3 hours, it was further divided into 3-hr subintervals starting from the beginning of the rainfall. As such, the database was composed of multiple years of observed events, separated by month, where each event was no longer than 3 hours.

For each month in the database, the events were sorted based on their total accumulated rainfall depth and created a CDF using the rainfall depths, such that each point on the CDF corresponds to actual, measured storm event in the database. In other words, each observed event is corresponded with a probability.

To disaggregate a 3-hr GCM forecast for August  $1^{st}$ , 2038 between 9:00 to 12:00 a.m., with a magnitude of 300 mm (call it  $D_T$ ), the downscaling algorithm stochastically selects several (and not <u>one</u>) observed events from the database and distributes them randomly over the 3-hr period. This means, the algorithm follows these steps:

1- Given a 3-hourly rainfall with magnitude of  $D_T$ , the algorithm first searches the historic CDF for the ordinate "a" corresponding to  $D_T$ .

2- A random number is generated from uniform distribution between 0 and "a", call it "u<sub>1</sub>".

3- " $u_i$ " is the probability of the randomly selected historic event. The magnitude of the observed event with probability of " $u_i$ " is read from the historic CDF (name  $D_i$ ) and its distribution is retrieved from the database.

4- New  $D_T$  is calculated by subtracting  $D_i$  from  $D_T$ .

5- This process then repeats for the new  $D_T$ 

6-Stop when  $\sum_{i=1}^{n} D_i = D_{T^*}$  n is the number of observed events selected.

# 2.3.4. Performance

This method's performance was tested to ensure its ability to disaggregate the 3hourly precipitation data to 15-min data. 3-hourly precipitation time series at each 15-min gauge were created by adding the measured 15-min precipitation, and the disaggregation method was tested in its ability to disaggregate the 3-hourly data into a 15-min synthetic time series. To evaluate the performance of the disaggregation method, the statistics of measured and synthetic time series were compared as suggested by Socolofsky et al. (2001) and Choi et al. (2008). The statistical parameters of the maximum rainfall values were also calculated to ensure the model captured the peaks. Because the disaggregation method is stochastic, 30 model runs were performed at each station and the mean value of statistics over all 30 runs was used in the error quantification.

One of the statistics used to evaluate the performance of the disaggregation method was "zero rainfall probability". According to the explained method, the stochastic algorithm selects a number of observed events randomly from a database and distributes them stochastically over a 3-hr period (12 time slots of 15 minutes). Some of the slots might remain empty (no rainfall). The final product is a newly generated distribution which has no similarity to any of the individual historic events. For validation, the algorithm was run 30 times on some selected events with known distribution and reported the average metric. Becasue the method is stochastic, for each run, different historic events are selected from the database and randomly distributed over 12 time slots. So the "zero rainfall probability" from each run is different from the other. The other statistic used for validation of the method was performance measures between the maximum rainfall of disaggregated time series compared to reference rainfall. As explained earlier, several reference rainfall events with known distributions were selected and disaggregated using the explained scheme. The resultant distributions were compared to the reference distribution and the statistical analysis was performed to acquire the goodness of fit measures (e.g.  $R^2$  of zero rainfall probability and maximum rainfall). When a single 3-hr event was disaggregated, the maximum partial rainfall depth for a given period within that event was calculated. These maxima are not compared by taking the maximum of the monthly or annual disaggregated rainfalls but by taking the maximum of disaggregated selected events for each month and rainfall duration. For any of these events, the model was run 30 times and the mean statistics was presented.

## 2.3.5. Creating IDF curves

Generalized Extreme Value (GEV) distribution was selected as the best probability distribution for Alabama based on different tests (e.g. probability plots, goodness of fit and L-moment ratio) in a study by Durrans and Brown (2001). The GEV distribution is a continuous probability distribution that combines Gumbel, Frechet and Weibull distributions and it is based on extreme value theory (Coles, 2001). This distribution was used in this study for creating IDF curves. The GEV distribution has cumulative distribution function:

$$F(x; \mu, \sigma, \xi) = \exp\left[-t(x)\right]$$

$$t(x) = \begin{cases} (1 + (\frac{x - \mu}{\sigma})\xi)^{-1/\xi} & \text{if } \xi \neq \mathbf{0} \\ e^{-(x - \mu)/\sigma} & \text{if } \xi = \mathbf{0} \end{cases}$$

Where  $\begin{cases} \mu \in R \text{, location parameter} \\ \sigma > 0 \text{, scale parameter} \\ \xi \in R, shape \text{ parameter} \end{cases}$ 

GEV parameters have been estimated using the method of moments (MOM) (Hosking et al. 1985, Bhunya et al. 2007). Kolmogorov-Smirnov (K-S) test was used to evaluate the performance of the fit.

The steps below describe the process of creating IDF curves:

- Obtain annual maximum series of precipitation depth for a given duration (15, 30 and 45 min, 1, 2, 3, 6, 12, 24 and 48 hr)
- Use GEV distribution to find precipitation depths for different return periods
   (2, 5, 10, 25, 50 and 100 years)
- 3. Repeat the first two steps for different durations
- 4. Plot depth versus duration for different frequencies

## 2.4. Results and Discussion

## 2.4.1. Bias correction

As mentioned earlier, the quantile-based mapping method proposed by Li et al. (2010) was used in this study. Bias correction was done on a monthly basis for each of the six climate models. CDFs of observed data from all stations and CDFs of climate model data for the same period were compared to each other. The 21<sup>st</sup> century projections were then corrected based on the differences between theses CDFs. In this method, monthly rainfall values were used to define the CDF error of historical model runs relative to observations and this error was used to correct the model CDF for the future period by calculating a scaling factor from the monthly totals (Li et al. 2010). The resulting bias-corrected model projections were used for the remainder of this study. Figure 2.2 shows the observed and modeled CFDs for each month when bias correction was performed for HRM3-HadCM3 model. The CDFs for the remaining models are presented in Appendix A.

# 2.4.2. Performance

The statistical measures used in the error quantification for typical months in winter (February) and summer (August) are presented in Table 2-1. In addition, statistical parameters of maximum rainfall values were calculated to make sure that the disaggregation model was capturing the peaks (Table 2-2).





Figure 2.2 Monthly Cumulative Distribution Functions (CDFs) to perform bias correction- Model data is from HRM3-HadCM3 model

These results show that the method was performing well in disaggregating the 3hour interval precipitation to 15-min data. Performance of GEV parameter estimation was also evaluated using Kolmogorov-Smirnov (K-S) test (Massey 1951). The critical value between sample and theoretical cumulative distributions at 95% level of confidence ( $\alpha$ =0.05) was 0.234. The maximum distance between the sample and theoretical cumulative distributions needs to be less than the critical value.

 Table 2-1 Performance measures between the disaggregated time series and measured statistics of 15-min rainfall

Month	Station	Statistics <b>R</b> <sup>2</sup>		MAE	RRSE	
February	Auburn, AL	$P_0 \over \sigma^2$	0.91 0.82	0.01 0.0003	0.31 0.62	
August	Auburn, AL	$P_0 \over \sigma^2$	0.82 0.78	0.05 0.002	0.69 0.81	

 $P_0$ : probability of zero rainfall,  $\sigma^2$ : the variance of 15-min rainfall,  $R^2$ : linear regression, MAE: mean absolute error and RRSE: root relative squared error

Table 2-3 shows statistical measures used for this evaluation. Based on the K-S test results, on all attempts GEV distribution fit to the sample CDFs with minimal error. The test results were always smaller than the critical value at 95% confidence and had a small standard error.

			$\mathbf{R}^2$					MBE*		
min	15	30	45	60	120	15	30	45	60	120
Jan	0.67	0.69	0.75	0.80	0.81	-14.6	-9.08	-10.9	-9.5	-8.9
Feb	0.75	0.75	0.78	0.73	0.75	-8.57	-10.8	-10.4	-11.5	-7.4
Mar	0.77	0.78	0.81	0.82	0.85	17.6	-3.9	-10.6	-10.03	-10.3
Apr	0.65	0.68	0.80	0.85	0.85	7.42	-9.7	-8.6	-8.4	-8.21
May	0.68	0.70	0.73	0.75	0.77	-5.9	-7.5	-8.3	-9.2	-10.4
Jun	0.72	0.75	0.78	0.78	0.82	-6.7	-11.1	-9.3	-9.8	-12.9
Jul	0.67	0.69	0.72	0.73	0.81	-5.01	-9.8	-10.7	-9.2	-8.97
Aug	0.69	0.73	0.75	0.77	0.79	-8.40	-8.18	-9.2	-10.4	-9.6
Sep	0.75	0.78	0.79	0.80	0.81	-3.2	-9.2	-8.6	-8.1	-8.9
Oct	0.78	0.80	0.80	0.81	0.82	-9.02	-9.8	-9.5	-11.5	-8.4
Nov	0.65	0.70	0.76	0.78	0.83	-24.6	-6.4	-9.4	-8.9	-4.9
Dec	0.73	0.75	0.77	0.81	0.84	-12.9	-6.8	-8.3	-10.8	-9.7
Ave.	0.71	0.73	0.77	0.78	0.81	-6.15	-8.52	-9.48	-9.78	-9.04

**Table 2-2** Performance measures between the maximum rainfall of disaggregated time series and measured statistics

\* MBE: Mass Balance Error (%)

Table 2-3 Performance 1	neasures between the	sample and theoretical	cumulative distributions
	using Kolmogorov	-Smirnov (K-S) test	

	Kolmogorov-Smirnov				<b>R</b> <sup>2</sup> **			
Models	Average	Max	Min	Std.*	Average	Max	Min	Std.*
HRM3-HadCM3	0.092	0.201	0.043	0.026	0.980	0.996	0.926	0.011
CRCM-CGCM3	0.086	0.195	0.039	0.024	0.982	0.996	0.937	0.010
HRM3-GFDL	0.089	0.175	0.038	0.024	0.981	0.996	0.922	0.011
CRCM-CCSM	0.087	0.178	0.035	0.023	0.982	0.997	0.933	0.01
RCM3-GFDL	0.104	0.218	0.042	0.034	0.971	0.997	0.910	0.01
ECP2-GFDL	0.088	0.178	0.031	0.024	0.982	0.997	0.933	0.01

*The critical value at*  $\alpha$ =0.05 *is* 0.234

\* Standard deviation

\*\* Coefficient of determination

### 2.4.3. IDF curves

IDF curves for Alabama were created as a series of 60 maps for each of the 6 NARCCAP regional climate projections (360 maps total) for 10 different rainfall durations and 6 different return periods.

An example of the type of maps that can be generated is illustrated in Figure 2.3 using HRM3-HadCM3 projections. Maps for the remaining models are presented in Appendix B. Comparing these maps with NOAA Technical Paper 40 (TP-40) (Hershfield 1961) shows that changes in future IDF curves are expected with the future climate (TP-40 is not shown in the figure).

For example, for a 50-yr return period with 6-hr duration, 191 mm of precipitation on average is expected to fall in the southwestern part of the state. Based on TP-40, this amount is 152 mm— about 25 % less than what is predicted for the future. The largest projected 12-hr rainfall value (about 329 mm) is expected to happen in southwest Alabama. For the same region and duration, TP-40 (Hershfield 1961) shows about 203 mm of rainfall, 62% less than what is expected in future. Changes in future rainfall intensity are expected to continue for other rainfall durations and return periods, but it is not possible to discuss the results of all 360 maps in details. Therefore, City of Auburn in Alabama was selected as an example from which to discuss the results in more detail. Figure 2.4 and figure 2.5 show the future and current IDF curves using all six NARCCAP regional climate projections for Auburn, AL for the two different return periods of 10 and 100 years.



Figure 2.3 Rainfall map; 50-year rainfall of 6-hr, 12-hr, 24-hr and 48-hr durations (mm) under future climate using HRM3-HadCM3 projected data

Figure 2.4 (a-b) shows IDF curves under both the future and current climates for Auburn when HRM3-HadCM3 projections were used to develop future IDF curves. Figure 2.4a demonstrates that the projected rainfall intensity for a 10-yr return period tends to decrease by 20% when the rainfall duration is less than 4 hours, and is expected to increase by 42% for rainfall durations of more than 4 hours. Also, rainfall intensity tends to increase by 38% if the rainfall duration exceeds 6 hours and is expected to decrease by 20% for durations less than 6 hours when the return period is 100 years (Figure 2.4b).

Figure 2.4 (c-d) depicts the changes in IDF curves when CRCM-CGCM data were used. For a 10-yr return period, rainfall intensity tends to be reduced by 51%. Results also show a 49% decrease for a 100-yr return period. Using HRM3-GFDL projections presents a 26% decline in rainfall intensity for durations less than 4 hours and a 65% increase for durations more than 4 hours (Figure 2.4e). Likewise, a 22% decrease and a 129% increase are expected to be observed for durations of less than and more than 4 hours, respectively (Figure 2.4f).

Figure 2.5 (a-b) displays future and current IDF curves developed using the CRCM-CCSM model. Forty three percent and 49% declines are expected to be noticed for all durations when the return period is 10-yr and 100-yr, respectively. Results of utilizing the RCM3-GFDL model for developing IDF curves are presented in Figure 2 .5 (c-d). A 34% rainfall intensity reduction was observed for durations less than 5 hours, while there was a 27% increase for longer durations (Figure 2.5c). A 14% decline for durations of less than 3 hours and a 55 % increase in rainfall intensities for longer durations was also observed (Figure 2.5d). Figure 2.5 (e-f) demonstrates changes in rainfall intensity employing the ECP2-GFDL model. Both the 10 and 100-yr return periods saw 33% and 37% reductions, respectively for all rainfall durations.



Figure 2.4 IDF curves under current and future climate using HRM3-HadCM3, CRCM- CGCM3 and HRM3-GFDL models for Auburn, AL



Figure 2.5 IDF curves under current and future climate using CRCM-CCSM, RCM3-GFDL and ECP2-GFDL models for Auburn, AL

Table 2-4 summarizes the discussed results for different return periods using six climate models for Auburn, AL.

Model	Return period	iod Average Percentage difference in rainfall intensity			
	2-yr and 5-yr	duration(t) >3 hrs: 49% increase t <3 hrs: 22 % decrease			
HRM3-HadCM3	10-yr	t >4 hrs: 42% increase t <4 hrs: 20 % decrease			
	50-yr	t >5 hrs: 33% increase t <5 hrs: 20 % decrease			
	100-yr	t >6 hrs: 38% increase t <6 hrs: 20 % decrease			
CRCM-CGCM3	all return periods	50% decrease for all durations			
HRM3-GFDL	2-5 and 10 years	t >4 hrs: 50% increase t <4hrs: 29 % decrease			
	50- and 100-yr	t >4 hrs: 100% increase t <4hrs: 22 % decrease			
CRCM-CCSM	all return periods	44% decrease for all durations			
	2-yr	t >12 hrs: 13% increase t <12 hrs: 38 % decrease			
	5 years	t >6 hrs: 23% increase t <6 hrs: 34 % decrease			
RCM3-GFDL	10 years	t >5 hrs: 27% increase t <5 hrs: 34 % decrease			
	50 years	t >4 hrs: 44% increase t <4 hrs: 20 % decrease			
	100 years	t >3 hrs: 55% increase t <3 hrs: 14 % decrease			
ECP2-GFDL	2-yr	t >12 hrs: 15% increase t <12 hrs: 46 % decrease			
	5-yr	t >17 hrs: 3% increase t <17 hrs:41 % decrease			
	10, 50 and 100-yr	36% decrease for all durations			

 Table 2-4 Comparisons of IDF curves under the current and future climate scenario for Auburn,

 AL

As the results above clearly demonstrate, the six different NARCCAP-based projections are not identical. Analyzing all maps for the state of Alabama shows the same disparity as Auburn. Two of the models (CRCM-CGCM3 and CRCM-CCSM) show a decrease in future rainfall intensity for all return periods and all rainfall durations for Alabama. The other four suggest that, depending on the return period, future rainfall intensities could decrease below and increase above a specific rainfall duration. The disparity in results could be due to many factors. Dai (2006) performed a study in which precipitation characteristics in eighteen climate models were analyzed and compared with historical data. The study pointed out that some of the climate models' deficiencies in measuring tropical rainfall were correlated with biases in the Sea Surface Temperature (SST) (Dai, 2006). The SST biases in the CGCM3 model are in accordance with dry biases in the Caribbean Sea and the Gulf of Mexico, so it may underestimate variables such as precipitation (Dai, 2006). It was also noted that the HadCM3 model simulates a realistic precipitation pattern, but that the results of climate models vary for different regions in the world (Dai, 2006). Therefore, it should be noted that these models are different in nature, and many variables could be involved in creating discrepancies. Differing results could be attributed to different types of GCMs and RCMs or to initial conditions and boundary conditions for each climate projection— but they all agree that for short durations of rainfall (usually less than 4 hours), rainfall intensity is expected to decrease or remain close to the current values. Results suggests that the current standard and guidelines, which use short rainfall durations for designing water management infrastructures (e.g. a roadside channel, a detention pond for a small drainage area), can serve their purpose in the future well.

As mentioned earlier the results of six different NARCCAP-based projections are not consistent with respect to larger events. To further explore the results of larger events, graphs presented in Figure 2.6 were prepared. In this figure rainfall intensity for a 12-hr rainfall under future and current climate was plotted for different return periods. Figure 2.6a presents the results when HRM3-HadCM3 projections were used to develop future IDF curves. It shows that if a given rainfall intensity under current climate occurs once every 20 years (the probability of that given rainfall happening in any year; p=5%), the same rainfall intensity is expected to happen once every 2 years (p=50%), under future climate. Figure 2.6 (b-c) also show increase in rainfall intensity under future climate using HRM3-GFDL and RCM3-GFDL projections. On the other hand, Figure 2.6d that presents the results of using CRCM-CGCM3 projections, suggests that if a given rainfall intensity under current climate occurs once every 2 years (p=50%), the same rainfall intensity is expected to happen once every 10 years (p=10%) under future climate. Likewise, Figure 2.6e and 2.6f show reduction in future rainfall intensity. How these results will affect designing different structures will be discussed below.

The first step towards designing different water management structures (e.g. dams, channels, detention ponds) is to identify the characteristics of design storm in terms of duration, return period and intensity. Time of concentration of the watershed draining to the hydraulic structure usually dictates the design storm duration. Storm return period is assigned based on economic assessments and risk analysis (probability of damage, loss of life, etc. in case of failure). For example, a 5- to 10-yr return period is used for designing roadside channels (Brown et al. 2001) where the cost of failure is negligible whereas a much larger return period is used for designing small dams where



**Figure 2.6** Rainfall intensity vs. return period under current and future climate for a 12-hr rainfall using a) HRM3-HadCM3, b) HRM3-GFDL, c) RCM3-GFDL, d) CRCM-CGCM3, e) CRCM-CCSM and f) ECP2-GFDL model

there is a great risk of life in case of failure. Knowing the storm return period, the design rainfall intensity is then acquired from developed IDF curves of the region. Rainfall intensity and duration of design storm dictate the cost of the hydraulic structure, and any uncertainty bound to estimation of these parameters can greatly impose design uncertainties. For example, figure 2.6 a, b and c suggest design rainfall intensities of 16, 28 and 23 mm/hr, respectively for a 100-yr return period and 12-hr duration while figure 2.6d and 2.6f recommend 9 mm/hr and figure 2.6e proposes 8 mm/hr as a design rainfall intensity (design rainfall intensity based on current IDF curves for this specific example is 12 mm/hr). As can be clearly seen, there is a large uncertainty existed on projected rainfall intensity of these six climate models for long durations. Developing probability-based IDF curves as a result of incorporating all six climate models could be a better way to present the results.

# **2.5. Summary and Conclusions**

This study developed IDF curves under future climate scenarios for Alabama, which were then compared to the IDF curves under the current climate. Six highresolution projections (for 2038-2070) derived from dynamical downscaling of GCMs by RCMs were used in this study. Results of the four climate model projections— HRM3-HadCM3, HRM3-GFDL, RCM3-GFDL and ECP2-GFDL— suggest that future rainfall intensity could decreases or increases depending on the return period. Results of the remaining two model projections— CRCM-CGCM3 and CRCM-CCSM— indicate a reduction in future rainfall intensity for all return periods and all rainfall durations for Alabama. A large uncertainty of projected rainfall intensity of these six climate models for long durations makes it difficult to obtain strong conclusions about the expected changes on future rainfall intensity in Alabama. A variety of factors cause the differing results; a likely reason is the difference in physical parameterizations, especially of radiative and precipitation-forming processes, amongst different GCMs and RCMs, as well as the difference in initial and boundary conditions for each climate projection-but the result they all have in common is that the precipitation pattern for Alabama veers toward less intense rainfalls for short rainfall durations. From results discussed above, we can conclude that the current standards and guidelines for designing municipal management infrastructures based on short rainfall durations can continue to serve well in the future. This conclusion is solely based on the results of the six climate model projections used in this study and not all existing climate models and scenarios. Using additional climate model projections in the future will help to make a stronger conclusion in this regard. Also, given the large uncertainty in the output from GCMs and disparity in the results of these models, creating a probability based IDF curves could be a better way to presents the IDF curves.

## Chapter 3

# 3. Developing Rainfall Intensity-Duration-Frequency (IDF) Curves For Alabama Under Future Climate Scenarios Using Artificial Neural Network (ANN)

# 3.1. Abstract

Hydrologic design of water management infrastructures is based on specific design storms derived from historic rainfall events available in the form of intensity-durationfrequency (IDF) curves. However, it is expected that the frequency and magnitude of future extreme rainfalls will change due to increase in greenhouse gas concentrations in earth's atmosphere. This study evaluated potential changes in current IDF curves for Alabama under projected future climate scenarios. Three-hourly precipitation data simulated by five combinations of global and regional climate models were temporally downscaled using Artificial Neural Networks (ANNs). A feed-forward, back-propagation model was developed to estimate maximum 15-, 30-, 45-, 60-, and 120-min precipitations. The results were also compared to disaggregated rainfalls derived using the stochastic method. Comparison of these two methods indicated that the ANN model provided superior performance in estimating maximum rainfall depths, while the stochastic method tended to under-predict maximum rainfall depths. Developed IDF curves indicate that future rainfall intensities for the rainfall events less than 2 hours are expected to decrease by 33% to 74% from current ones when ANN model is used, while large uncertainty exists in the projected rainfall intensities of longer duration events. This finding was independent of the temporal downscaling method used.

## **3.2. Introduction**

Hydrologic design of storm water management facilities, culverts, flood protection structures, erosion and sediment control structures, and many other hydraulic structures are based on specific design storms derived from historic rainfall events as available in the form of intensity-duration-frequency (IDF) curves (McCuen 1998, Prodanovic et al. 2007). An IDF curve is a tabular or graphical illustration of the probability that a given rainfall intensity will happen (Wolcott et al. 2009). A design based on an inaccurate design storm can result in malfunction of infrastructure, loss of life in the case of failure, or excessive cost if the design storm is overestimated.

An increase in greenhouse gas concentrations, such as carbon dioxide ( $CO_2$ ), in the earth's atmosphere during the last century can cause large-scale variations in atmospheric processes and, as a result, changes in precipitation and temperature characteristics (Prodanovic et al. 2007). One of the expected impacts of the changes in precipitation characteristics is a change in the frequency and magnitude of extreme rainfalls, which can lead to revision of existing standards for designing civil engineering infrastructures. It can also require the rebuilding and/or upgrading of existing infrastructures.

One way to be prepared for possible changes, and to decrease vulnerability of hydraulic infrastructures to climate change, is to predict potential effects (as manifested by IDF curves) and adapt to them (Prodanovic et al. 2007).

A recent study by Mirhosseini et al. (2012) (Chapter 2) developed future IDF curves for Alabama using high-resolution projections (for 2038–2070) derived from dynamical downscaling of General Circulation Models (GCMs) by Regional Climate Models (RCMs). A stochastic method was used in order to disaggregate 3-hourly precipitation to 15-min precipitation. Following the previous study, the main objective of the present study was to create IDF curves for Alabama using an Artificial Neural Network (ANN) to estimate the maximum 15-, 30-, 45-, 60-, and 120-min precipitation depths. An additional objective was to compare the performance of ANN model downscaling with the performance of the stochastic model used by Mirhosseini et al. (2012).

The artificial neuron concept was first introduced by McCulloch and Pitts (1943) and ANNs have existed since then (Dwason and Wilby 2001). ANN models have been used for both spatial and temporal downscaling of temperature and precipitation. They also have been used for rainfall–runoff modelling and flood forecasting (Dwason and Wilby 2001). There are studies investigating performance of ANN models to temporally disaggregate rainfall. They compared results of ANN disaggregation with other methods as well. Burian et al. (2000) developed an ANN model and evaluated its performance by comparing it with two other methods; a linear model and a continuous deterministic rainfall disaggregation model (Ormsbee 1989). Viability of an ANN model for rainfall disaggregation was evaluated in a study by Dibikie and Coulibaly (2006). Evaluation of ANN model was done by comparing the results with linear regression model.

In this chapter developing an ANN model to estimate the maximum 15-, 30-, 45-, 60-, and 120-min rainfall depths from 3-hourly precipitations will be discussed. Results were used to create future IDF curves and then were compared to the results of previous chapter, the impact of climate change on the IDF curves was evaluated afterwards.

### **3.3. Methodology**

## 3.3.1. Data and Model Used

Historical (observed) rainfall data at 15-minute intervals were obtained from the NOAA National Climatic Data Center Online Climate Data Directory. Locations of 34 rain gauge stations in Alabama are presented in Figure 1. Simulated historical precipitations for the period 1968–2000 and projected precipitations for the period 2038–2070 at a temporal resolution of 3 hours and a spatial resolution of 50 km were obtained from the North American Regional Climate Change Assessment Program (NARCCAP) (Sebastien et al., 2007, Richard et al., 2007). Maximum and minimum temperatures for the same periods were obtained at daily intervals.

Five different dynamically downscaled datasets were used for the study and derived from five NARCCAP projections: HRM3-HadCM3 (Hadley Centre Regional Model and Hadley Centre Coupled Model, version 3), CRCM-CGCM3 (Canadian Regional Climate Model and Coupled Global Climate Model), HRM3-GFDL (Hadley Centre Regional Model and Geophysical Fluid Dynamics Laboratory model), CRCM-CCSM (Canadian Regional Climate Model and The Community Climate System Model), and RCM3-GFDL (Regional Climate Model and Geophysical Fluid Dynamics Laboratory model). A complete description of these models can be found in chapter 2, section 2.3.1. It should be noted that ECP2-GFDL model used in the previous study was not used in the current study because of the lack of availability of the future temperature variables needed to develop an ANN model.

# 3.3.2. Bias Correction and Temporal Downscaling

As mentioned in chapter 2, a quantile-based mapping method proposed by Li et al. (2010) was used to perform bias correction. For this method, monthly rainfall values were used to define the cumulative distribution function (CDF) of errors of historical model runs relative to the observations. This error was used to correct the model CDF for the future period by calculating a scaling factor from the monthly totals for each month (January–December). This error was used to correct the model CDF for the future period by calculating factor from the monthly totals. The scaling factor is defined as the bias-corrected rainfall total for a given month divided by the non-bias-corrected total. Prior to the disaggregation of 3-hourly rainfall events, 3-hourly totals were multiplied by their corresponding monthly scaling factors (Li et al., 2010). The same method was used for bias correction of temperature data.

Creating IDF curves requires high temporal resolution rainfall data. NARCCAP future climate data are provided at 3-hour intervals, which make temporal downscaling a necessary task. Different types of downscaling techniques have been discussed in many studies (Von Storch 1999b, Rodriguez-Iturbe et al. 1987, Islam et al. 1990, Socolofsky et al. 2001). The temporal downscaling method employed in a recent study by Mirhosseini et al. (2012) was a modified version of the stochastic method introduced by Socolofsky et al. (2001). Results of that study showed that stochastic method tends to underpredict

precipitation. A study by Burain et al. (2000) compared the results of rainfall disaggregation using a feed-forward ANN model and two other methods; linear model and a continuous deterministic rainfall disaggregation model introduced by Ormsbee (1989). Based on their study both of these models underpredicted maximum rainfall intensities as compared to ANN model. Several other studies also showed the same results (Dibikie and Coulibaly 2006, Weichert and Burger 1998). Therefore, in the current study, ANN model was used to find the maximum 15-, 30-, 45-, 60-, and 120-min rainfall depths from 3-hourly precipitations.

## 3.3.3. Artificial Neural Network

ANN models are not deterministic models. They learn from examples, train themselves using the input and target outputs presented to them, and adjust parameters until they are able to provide meaningful outputs (Burian et al., 2001). Artificial Neural Networks have been used to solve different hydrological and water resources problems (Halff et al. 1993, Smith and Eli 1995, Minns and Hall, 1996), and have proven to be powerful tools for solving difficult problems. ANN models have also been successfully used for temporal downscaling of precipitation. For example, Burian et al. (2000, 2001) used an ANN model to disaggregate hourly rainfall into sub-hourly and found it to be a viable method.

The ANN disaggregation model developed for this study was a feed-forward, backpropagation model. The model includes three layers: input, hidden, and output layers. The connections between neurons in feed-forward networks are in the forward direction. Output is calculated from each neuron (processing unit) of the ANN model, starting from
the input layer and moving forward through the hidden layers to the outputs (Burian et al., 2001). The function that determines the behavior between the input and output layers is called transfer function. The hyperbolic tangent sigmoid transfer function was used in this study.

Training and learning functions are used to adjust the weights and biases of the network. The Levenberg-Marquardt back propagation training function and the gradient descent with momentum weight/bias learning function were used to develop the ANN model. Back-propagation explains the direction of propagation for calculating the errors. Making adjustments to the network weight (learning process) is a separate step which depends on the results of the training step (Burian et al., 2001). Determination of the learning rate has a significant impact on the performance of the model. Having too large values, the ANN model would remember only the training data due to the weight oscillation. On the other hand, too small values will significantly increase training time (Burian et al. 2000, 2001, Isik et al. 2012). Learning rate used to develop the ANN model in this study was determined to be 0.16, as compared with a range of 0.01 to 0.95 in other studies (Maier and Dandy 2000).

#### **3.3.4.** ANN Model Input Data

Finding appropriate input variables is an important step when developing an ANN model (Maier and Dandy 2000, Dawson and Wilby 2001, Bowden et al. 2005). Many methods of selecting significant input variables have been developed, such as, using prior knowledge of the system, linear cross-correlation methods, and using partial mutual information algorithms (Bowden et al. 2005, May et al. 2008). The Akaike Information Criterion (AIC) and Normalized Mean Square Error (NMSE) were two criteria or

performance parameters used in this study to find the best input datasets. The AIC has been widely used as a model input selection method (Qi and Zhang 2001, Ren and Zhao 2002, Zhao et al. 2008, Kalin et al. 2010, Isik et al. 2012), and can be calculated using the equations below (Qi and Zhang 2001):

$$AIC = \log(\hat{\sigma}_{MLE}^2) + \frac{2m}{T}$$

where *m* is the number of model parameters (input variables), *T* is the number of observations and  $\hat{\sigma}_{MLE}^2$  is the maximum likelihood estimate of variance of the residual term or the mean square error (MSE) between the observed and simulated data.

$$\hat{\sigma}_{MLE}^2 = \frac{\sum (y_{obs} - y_{sim})^2}{T}$$

where  $y_{obs}$  is the observed data and  $y_{sim}$  is the output of the ANN model.

More precisely,  $y_{obs}$  used to train, validate, and test ANN models are the maximum 15-, 30-, 45-, 60-, and 120-min precipitations derived from observed rainfall data at stations in Alabama. The model with the minimum AIC value was selected as the best one.

The NMSE estimates the overall deviations between the observed and simulated values, and it ranges from 0 to  $+\infty$ . A model with a zero NMSE is a perfect model. If NMSE is 1, the model is as good as the observed mean values, and greater than 1 shows a poor model (Weigend and Gershenfeld, 1993; Singh et al., 2007).

NMSE = 
$$\frac{\sum_{i=1}^{n} (y_{obs} - y_{sim})^{2}}{\sum_{i=1}^{n} (y_{obs} - \bar{y}_{obs})^{2}}$$

where  $\overline{y}_{obs}$  is the mean of the observed data.

There are 60 (12 months by 5 rainfall durations) ANN models developed in this study. For each model, nine different combinations of possible input variables were

selected for each month and rainfall duration (15-, 30-, 45-, 60-, and 120-min). The above two criteria were then calculated for each combination in order to identify the optimum input variables. As an example, Table 3-1 presents the results of selecting input variables for January. The results in Table 3-1 shows that 3-hourly precipitation ( $P_3$ ), daily precipitation ( $P_{daily}$ ), monthly precipitation ( $P_{monthly}$ ), and maximum and minimum temperatures ( $T_{max}$  and  $T_{min}$ ) were selected as optimum input data or variables for the ANN model. The results for the remaining months also suggest that the above inputs are the optimum data for developing the ANN model.

**Table 3-1** The Akaike Information Criterion (AIC) and Normalized Mean Square Error (NMSE) calculated for the nine combinations of input variables in order to determine optimal input variables of the ANN model.

		AIC				NMSE						
No.	Input Variables	15-	30-	45-	60-	120-		15-	30-	45-	60-	120-
		min	min	min	min	min		min	min	min	min	min
1	P <sub>t-3</sub> , P <sub>3</sub> , P <sub>t+3</sub>	1.80	1.95	1.95	1.94	1.53	(	).48	0.32	0.24	0.19	0.05
2	$P_{t-3}$ , $P_3$ , $P_{t+3}$ , $P_{daily}$ , $P_{monthly}$	1.76	1.91	1.92	1.89	1.49	(	).44	0.30	0.22	0.17	0.05
3	$\begin{array}{c} P_{t\text{-}3},P_3,P_{t\text{+}3},P_{daily},\\ P_{monthly},T_{max},T_{min} \end{array}$	1.74	1.88	1.89	1.87	1.50	(	).42	0.28	0.20	0.16	0.05
4	P <sub>3</sub> , P <sub>daily</sub> , P <sub>monthly</sub>	1.76	1.92	1.92	1.89	1.52	(	).44	0.30	0.22	0.17	0.05
5	P <sub>3</sub> , P <sub>daily</sub> , P <sub>monthly</sub> , P <sub>annual</sub>	1.76	1.91	1.91	1.90	1.49	(	).44	0.30	0.22	0.18	0.05
6	$P_3$ , $P_{daily}$ , $P_{monthly}$ , $T_{max}$ , $T_{min}$	1.71	1.88	1.88	1.87	1.48	(	).32	0.22	0.18	0.14	0.04
7	$P_3, P_{daily}, P_{monthly}, P_{annual}, T_{max}, T_{min}$	1.74	1.89	1.91	1.87	1.53	(	).42	0.28	0.21	0.17	0.05
8	$P_3$ , $P_{daily}$ , $P_{monthly}$ , $P_{annual}$ , $T_{max}$ , $T_{min}$ , elev.	1.74	1.89	1.89	1.87	1.48	(	).42	0.28	0.20	0.16	0.04
9	$P_{t-3}$ , $P_3$ , $P_{t+3}$ , $T_{max}$ , $T_{min}$	1.77	1.92	1.93	1.89	1.54	(	).45	0.30	0.22	0.17	0.05

 $P_3$  is 3-hourly precipitation;  $P_{t-3}$ ,  $P_{t+3}$  are precipitations 3 hr before and 3 hr after  $P_3$ ;  $P_{daily}$ ,  $P_{monthly}$ , and  $P_{annual}$  are 24-hr, monthly, and annual sums of precipitation, respectively;  $T_{max}$ ,  $T_{min}$  are maximum and minimum temperatures; and elev. is the elevation of the station.

In this study, target data of the ANN model are the maximum 15-, 30-, 45-, 60-, and 120-min precipitation. The main advantage of this study is that the output of the ANN model are the maximum precipitation of different durations that are needed later to create the IDF curves. This method is computationally more efficient as compared to finding the disaggregated rainfall for each interval and then finding the maximum rainfalls for five different durations.

### 3.3.5. ANN Model Training and Performance Measures

The ANN model was trained with input and target datasets using the training function discussed in section 3.3.3. Model accuracy usually increases with an increase in training, but there is a point at which more training does not improve model performance, and sometimes even gets worse (Burian et al., 2000; 2001). Identifying this point can be done with trial and error. For the ANN model evaluated in this study, 60% of observed data were used for training, 20% for validation, and the last 20% for testing. Burian et al. (2001) investigated a number of training iterations to develop an ANN model for hourly rainfall disaggregation and found that the performance of the model improves when the iteration numbers increase from 100 to 1000–1500, and that it worsens for higher values. For this study, the iteration number of 1000 was selected. The determination of the number of hidden neurons in an ANN model is a matter of experimentation (Burian et al. 2000, 2001, Isik et al. 2012). In our investigation, performance of the ANN model for 1-30 hidden neurons was evaluated, and eventually 6 hidden neurons were selected for developing the model.

The performance measures for evaluation of the ANN model were Nash-Sutcliffe Efficiency coefficient (NSE) and correlation coefficient (R). Again, Normalized Mean Square Error (NMSE) and Akaike Information Criterion (AIC) were used to find the optimum input data. The results of the ANN model were also compared to the results of the rainfall disaggregation using stochastic method from the previous chapter.

#### 3.3.6. Intensity-Duration-Frequency (IDF) Curves

Future IDF curves were created for Alabama using Generalized Extreme Value (GEV) distribution. As mentioned in Chapter 2, this method was selected as the best probability distribution for Alabama in a study by Durrans and Brown (2001). Method of moments (MOM) was used to estimate the GEV parameters (Hosking et al., 1985; Bhunya et al., 2007) and performance of the fit was evaluated by Kolmogorov-Smirnov (K-S) test (Massey 1951). Based on the K-S test results, GEV distribution fitted to the sample CDFs with minimal error. Statistical measures used for this evaluation are presented in Table 2-3 (Chapter 2). IDF curves where created by obtaining the annual maximum precipitation depth for different rainfall durations (e.g. 15-, 30- and 45- min, 1- and 2-hr.). GEV distribution was then used to find the precipitation depths for different return periods (e.g. 2, 5, 10 years) and it was repeated for different durations. Finally depth versus duration was plotted for different frequencies to develop the IDF curves.

## **3.4.Results and Discussion**

## 3.4.1. Performance

Temporal disaggregation was performed for each month and rainfall duration separately. The statistical measures used in the error quantification are presented in Table 3-2. Both R and NSE for all months are greater than or equal to 0.6, and higher R and

NSE values were obtained for longer durations. These results show that the developed ANN model performed very well in disaggregating the 3-hour interval precipitations and finding the maximum of 15-, 30-, 45-, 60-, and 120-min rainfall depths.

	-	Training		Valid	Validation		oct			
Performance	Month	$\frac{11a}{R}$	NSE	$\frac{Vand}{R}$	NSE	$\frac{1}{R}$	NSE	R	NSE	
	Jan	0.73	0.60	0.80	0.63	0.80	0.63	0.60	0.80	
	Feb	0.80	0.61	0.83	0.70	0.81	0.70	0.80	0.65	
	Mar	0.75	0.60	0.82	0.70	0.82	0.70	0.80	0.65	
	Apr	0.80	0.62	0.80	0.63	0.82	0.70	0.80	0.64	
	May	0.83	0.70	0.84	0.70	0.86	0.73	0.84	0.70	
15 min	Jun	0.90	0.74	0.87	0.75	0.90	0.80	0.90	0.75	
15-11111	Jul	0.90	0.74	0.86	0.74	0.90	0.80	0.90	0.75	
	Aug	0.86	0.73	0.90	0.80	0.88	0.80	0.86	0.75	
	Sep	0.80	0.65	0.94	0.84	0.92	0.82	0.90	0.77	
	Oct	0.81	0.65	0.86	0.74	0.88	0.77	0.84	0.71	
	Nov	0.80	0.61	0.81	0.66	0.84	0.70	0.81	0.65	
	Dec	0.71	0.60	0.80	0.60	0.84	0.70	0.80	0.61	
	Jan	0.82	0.70	0.86	0.75	0.90	0.80	0.85	0.72	
	Feb	0.85	0.72	0.90	0.80	0.87	0.75	0.86	0.75	
	Mar	0.84	0.70	0.88	0.80	0.88	0.77	0.90	0.74	
30-min	Apr	0.85	0.73	0.86	0.73	0.90	0.80	0.86	0.75	
	May	0.90	0.80	0.90	0.78	0.91	0.82	0.90	0.80	
	Jun	0.91	0.84	0.92	0.85	0.93	0.87	0.92	0.85	
	Jul	0.92	0.84	0.94	0.86	0.93	0.86	0.92	0.85	
	Aug	0.92	0.84	0.93	0.86	0.94	0.88	0.92	0.85	

 Table 3-2 Performance of the ANN model for training, validation, and testing datasets for every month and every rainfall duration.

	Sep	0.88	0.75	0.97	0.92	0.97	0.91	0.94	0.90
	Oct	0.86	0.74	0.90	0.80	0.91	0.81	0.90	0.80
	Nov	0.86	0.74	0.87	0.76	0.90	0.80	0.87	0.80
	Dec	0.81	0.70	0.85	0.72	0.90	0.80	0.85	0.73
	Jan	0.90	0.80	0.90	0.81	0.92	0.83	0.90	0.80
	Feb	0.88	0.78	0.91	0.82	0.91	0.81	0.90	0.80
	Mar	0.90	0.80	0.92	0.84	0.92	0.85	0.91	0.82
	Apr	0.90	0.80	0.90	0.81	0.92	0.84	0.90	0.82
	May	0.93	0.86	0.93	0.87	0.94	0.90	0.93	0.90
15-min	Jun	0.95	0.90	0.95	0.91	0.96	0.92	0.95	0.91
45-11111	Jul	0.95	0.90	0.96	0.92	0.96	0.92	0.95	0.91
	Aug	0.95	0.84	0.95	0.86	0.96	0.88	0.95	0.86
	Sep	0.91	0.84	0.98	0.94	0.97	0.93	0.96	0.91
	Oct	0.90	0.81	0.93	0.86	0.94	0.85	0.92	0.84
	Nov	0.90	0.81	0.91	0.82	0.93	0.84	0.91	0.82
	Dec	0.86	0.74	0.90	0.80	0.93	0.84	0.90	0.80
	Jan	0.90	0.81	0.92	0.85	0.93	0.90	0.92	0.84
	Feb	0.91	0.83	0.93	0.86	0.93	0.85	0.92	0.84
	Mar	0.92	0.84	0.94	0.89	0.94	0.89	0.93	0.87
	Apr	0.92	0.86	0.93	0.86	0.94	0.88	0.93	0.87
60-min	May	0.95	0.89	0.95	0.90	0.96	0.92	0.95	0.90
	Jun	0.96	0.93	0.97	0.94	0.97	0.94	0.97	0.94
	Jul	0.97	0.93	0.97	0.94	0.97	0.94	0.97	0.94
	Aug	0.96	0.93	0.97	0.94	0.98	0.95	0.97	0.94
	Sep	0.94	0.88	0.99	0.97	0.98	0.96	0.98	0.94
	Oct	0.93	0.86	0.95	0.90	0.96	0.90	0.94	0.90
	Nov	0.93	0.86	0.93	0.87	0.95	0.90	0.94	0.87

	Dec	0.90	0.80	0.92	0.83	0.94	0.86	0.92	0.83
	Jan	0.97	0.95	0.98	0.96	0.98	0.96	0.98	0.96
	Feb	0.98	0.95	0.98	0.96	0.98	0.96	0.98	0.96
	Mar	0.98	0.96	0.99	0.97	0.99	0.97	0.98	0.97
	Apr	0.99	0.97	0.98	0.97	0.99	0.98	0.99	0.97
	May	0.99	0.98	0.99	0.98	0.99	0.98	0.99	0.98
120 min	Jun	0.99	0.98	0.99	0.99	0.99	0.98	0.99	0.98
120-11111	Jul	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	Aug	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	Sep	0.99	0.95	0.99	0.97	0.99	0.97	0.99	0.97
	Oct	0.98	0.97	0.99	0.98	0.99	0.97	0.99	0.97
	Nov	0.98	0.96	0.98	0.96	0.99	0.97	0.98	0.96
	Dec	0.97	0.95	0.98	0.96	0.99	0.97	0.98	0.96

R: Correlation coefficient

ENS: Nash-Sutcliffe model efficiency coefficient

ANN model results were also compared with the results of the stochastic method used by Mirhosseini et al. (2012). Figures 3.1(a-e) show scatter plots of observed and predicted maximum 15-, 30-, 45-, 60-, and 120-min rainfalls in a typical summer month (July) from the ANN model and the stochastic method. Scatter plots for the remaining months are presented in Appendix C. Comparison of the results from these two methods indicates that the ANN model performed better in estimating maximum rainfall. The perfect prediction line (1:1 line) shown on each plot divides the under-predictions from over-predictions. These plots suggest that ANN model predictions closely distribute along the perfect prediction line, and that the stochastic method tends to under-predict maximum rainfall intensities.





**Figure 3.1** Scatter plots of observed and modeled maximum a) 15-min, b) 30-min, c) 45-min, d) 60-min, and e) 120-min rainfall depths using the ANN model (left panels) and the stochastic method (right panels) for July.

#### 3.4.2. IDF Curves

The future IDF curves for Alabama were developed for five GCM models and for six return periods (2-, 5-, 10-, 25-, 50-, and 100-yr). The results of future IDF curves were also presented as a series of maps (total 150 maps = 5 durations  $\times$  6 return periods  $\times$  5 GCM models) for the five NARCCAP regional climate projections. An example of the generated maps is illustrated in figure 3.2. Figure 3.2 shows the maps created for a 100-yr return period and durations of 15-, 30-, 45-, and 60-min.



Figure 3.2 Rainfall; 100-year rainfall of 15-, 30-, 45- and 60-min durations (mm) under future climate using HRM3-HadCM3 projected data

The remaining maps for other durations and other models are shown in Appendix D. Comparing these maps with NOAA Technical Paper 40 (TP-40) (Hershfield 1961) and NWS HYDRO-35 (Frederick et al. 1977) shows that changes in future IDF curves are expected with the future climate. The largest projected rainfall for all the durations shown in figure 3.2 is expected to happen in the southwest Alabama. For example, for a 100-yr return period with 30-min duration, between 51 and 72 mm of rainfall is expected to fall in the southwestern part of the state. Tp-40 reports amounts from 76 and 89 mm (Hershfield 1961). This amount is expected to be between 66 and 86 mm for 1-hr rainfall duration. Based on TP-40 (Hershfield 1961), this amount is reported between 102 and 114 mm— about 35% and 25% more than what is predicted for the future. The smallest projected rainfall amounts are expected to occur in the north and northwest Alabama. For a 100-yr return period and a 15-min rainfall duration, between 20 and 23 mm of rainfall is predicted. HYDRO-35 (Frederick et al. 1977) reported this amount to be in a range of 44 and 46 mm which is about 52% on average more than the predicted rainfall for future. Since it is not possible to discuss the results of all 150 maps in detail, City of Auburn in Alabama was selected as an example to discuss the results. Figure 3.3 presents the future IDF curves at Auburn using five NARCCAP regional climate projections for the two different return periods of 10 and 100 years and rainfall durations of 15-, 30-, 45-, 60and 120-min (Results for durations longer than 2-hr are presented in Mirhosseini et al. (2012) so they will not be discussed here). Future results include the IDFs created with disaggregated rainfalls from both the ANN models and the stochastic method.

Figures 3.3a and b depict the IDF curves under the future and current climates using HRM3-HadCM3 projections for developing the future IDFs. Figure 3.3a shows that if the

ANN model is used for rainfall disaggregation, projected rainfall intensity for a 10-yr return period tends to be reduced by 34%. However, the stochastic method suggested a decrease of 20%. Figure 3.3b presents the changes in rainfall intensity for a 100-yr return period. It shows that the rainfall intensity tends to decrease by 20% when using stochastic method and 41% when ANN model was used for rainfall disaggregation.

The CRCM-CGCM3 projections show that the rainfall intensity for a 10-yr return period was reduced by 71% when the ANN model was used for rainfall disaggregation, compared to a 68% decrease using the stochastic method (Figure 3.3c). Figure 3.3d shows the results for a 100-yr return period. Figures 3.3e and f display the IDF curves developed using the HRM3-GFDL model. Future rainfall intensity is expected to decline by 40% when the ANN model is used for disaggregation, while the stochastic method suggests a 26% decline (Figure 3.3e). For a 100-yr return period, future rainfall intensity tends to decrease by 22% and 41% when the stochastic method and ANN model are used for rainfall disaggregation, respectively (Figure 3.3f).

Figures 3.3g-j show IDF curves developed using CRCM-CCSM and RCM3-GFDL projections. Graphs show the future IDF curves developed using two different rainfall disaggregation methods (the ANN model and the stochastic method) and compare them to the current IDF curves. Sixty percent (stochastic method) and 63% (ANN model) declines in rainfall intensities are expected for a 10-yr return period (Figure 3.3g). For the 100-yr return period, 59% (stochastic method) and 66% (ANN model) decreases in future rainfall intensities are expected.





**Figure 3.3** IDF curves for 10-yr and 100-yr return periods under current and future climate (using HRM3-HadCM3, CRCM-CGCM3, and HRM3-GFDL CRCM-CCSM and RCM3-GFDL models) for Auburn, AL.

Figure 3.3i displays the results when RCM3-GFDL data were used for developing the IDF curves. Utilizing the ANN model for disaggregation indicates that rainfall intensities are expected to decline by 44% for a 10-yr return period, while the stochastic method suggests a 34% decrease. A 41% decline in rainfall intensities was also observed for a 100-yr return period while utilizing the ANN model for temporal downscaling. This

value showed an approximate 14% decrease when the stochastic method was used (Figure 3.3j).

Table 3-3 summarizes the results for different return periods using five climate models for Auburn, AL. The ANN model and stochastic method agree that future rainfall intensities for durations less than 2 hours are expected to decrease. However, they don't provide identical results. When HRM3-HadCM3 projections were used to develop the IDF curves, the ANN model suggested a decrease by 37% on average for Auburn, while the stochastic method indicated a 21% decline on average. The ANN model and stochastic method indicate almost the same results when CRCM-CGCM3 and CRCM-CCSM models were used to create IDF curves under the future climate. Utilizing HRM3-GFDL data shows a 41% decrease on average when the ANN model was used as a disaggregation method, while the stochastic method indicates a 26% decrease on average. Using RCM3-GFDL data for developing future IDF curves, ANN model suggests that future rainfall intensity is expected to decrease by 44% on average while the stochastic method shows a 31% decline. Several studies have investigated performance of ANN models for rainfall disaggregation and have compared their result with other methods. Burian et al. (2000) compared the results of rainfall disaggregation using an ANN model with two other methods; linear model and a continuous deterministic rainfall disaggregation model (Ormsbee 1989). Their study indicated that the latter two models underpredicted maximum rainfall intensities and were outperformed by the ANN model ANN model.

	• • •	Average Percentage difference in rainfall intensity					
Model R	eturn period	ANN model	Stochastic method				
	2-yr	33% decrease	24% decrease				
	5-yr	36% decrease	21% decrease				
	10-yr	34% decrease	20% decrease				
HKIVIS-HauCIVIS	50-yr	40% decrease	20% decrease				
	100-yr	41% decrease	20% decrease				
	Average	37% decrease	21% decrease				
	2-yr	66% decrease	68% decrease				
	5-yr	70% decrease	68% decrease				
CDCM CCCM2	10-yr	71% decrease	68% decrease				
CKUM-CGUM3	50-yr	73% decrease	66% decrease				
	100-yr	74% decrease	65% decrease				
	Average	71% decrease	68% decrease				
	2-yr	38% decrease	32% decrease				
	5-yr	40% decrease	28% decrease				
HDM2 CEDI	10-yr	41% decrease	26% decrease				
HKM3-GFDL	50-yr	41% decrease	23% decrease				
	100-yr	41% decrease	22% decrease				
	Average	41% decrease	26% decrease				
	2-yr	58% decrease	60% decrease				
	5-yr	62% decrease	60% decrease				
CDCM CCSM	10-yr	63% decrease	60% decrease				
	50-yr	66% decrease	60% decrease				
	100-yr	66% decrease	59% decrease				
	Average	63% decrease	60% decrease				
	2-yr	45% decrease	48% decrease				
	5 years	47% decrease	40% decrease				
DCM2 CEDI	10 years	44% decrease	34% decrease				
NUMJ-GFDL	50 years	42% decrease	20% decrease				
	100 years	42% decrease	14% decrease				
	Average	44% decrease	31% decrease				

**Table 3-3** Comparisons of IDF curves under the current and future climate scenarios for Auburn,

 AL using two temporal downscaling methods, the ANN model and the stochastic method.

Dibikie and Coulibaly (2006) evaluated the viability of an ANN model to disaggregate precipitation and compared it with a liner regression model and concluded that an ANN outperforms the statistical models. Other studies also reported better performance in predicting heavy rainfall events when using an ANN model as compared to linear regression downscaling methods (Weichert & Burger 1998). One reason for better performance of ANNs could be that a neural network can approximate complex, highly non-linear relationships without *a priori* assumptions (Dibikie and Coulibaly 2006). Also, ANNs allow the data to define the functional form while regression models assume a functional form. Therefore, it is believed that neural networks are more powerful tools than other downscaling methods (Von Storch et al. 2000, Dibikie & Coulibaly 2006).

Results show that no matter which temporal downscaling method was used, the five different climate model projections were not identical. Analyzing all the developed IDF curves for the state of Alabama (data not presented) also shows a similar disparity as observed for Auburn. The CRCM-CGCM3 and CRCM-CCSM models demonstrate a decrease in future rainfall intensity for Alabama, while the other three models suggest that a decrease below and increase above a specific rainfall duration can occur, depending on the return period. As was mentioned by Mirhosseini et al. (2012), the difference in the climate model results could occur due to different initial and boundary conditions for each of the climate projections, or the differences in physical parameterizations (such as precipitation-forming processes) amongst different GCMs and RCMs. Despite the disparity in future projections from the investigated models, they all suggest that future rainfall intensities for short durations are expected to decrease. The results of larger events and their effects on designing different water management structures were explored in further detail in Mirhosseini et al. (2012).

#### **3.5. Summary and Conclusions**

This study investigated the viability of using ANNs to estimate maximum 15, 30, 45, 60, and 120-min precipitations in order to develop IDF curves under the future climate scenarios. The results were also compared to disaggregated rainfall using the stochastic method. Comparison of these two methods indicates that the ANN model provides superior performance in estimating the maximum 15-, 30-, 45-, 60-, and 120-min precipitations.

Scatter plots created to evaluate the performance of the ANN model suggest that the predictions of the ANN model more closely approximated the perfect prediction line, and that the stochastic method under-predicts the maximum rainfall amounts. Using the ANN model for rainfall disaggregation, IDF curves for Alabama under the future climate scenario were developed from dynamically downscaled NARCCAP projections. Results were further compared to the current IDF curves. Comparison of the developed ANN model with the stochastic method shows that both methods agree that future rainfall intensities (for duration less than 2 hours) are expected to decrease. Both methods estimate almost similar decreases in rainfall intensities when CRCM-CGCM3 and CRCM-CCSM models were used to create IDF curves under the future climate. However, the ANN model suggests greater declines in rainfall intensities compared to the stochastic method when HRM3-HadCM3, HRM3-GFDL, and RCM3-GFDL projections were used.

Analyzing the results for longer durations presented in details in Mirhosseini et al. (2012), revealed that there is a large uncertainty in the projected rainfall intensity of the investigated climate models for long durations (> 4 hours). This uncertainty prevents us from giving strong conclusions about the expected changes in future rainfall intensity for long duration rainfall events in Alabama. However, projections from all of the climate models (disaggregated with both temporal downscaling methods) show a decrease for shorter rainfall durations (i.e., less than 2 hours). Therefore, it can be concluded that the existing IDF curves for designing water management infrastructures based on short rainfall durations will continue to be useful in the future. This conclusion is based solely on the results of the selected climate model projections used in this and the previous study by Mirhosseini et al. (2012), and not on all existing climate models and scenarios. Perhaps using additional model projections will help researchers to better understand the possible changes in future rainfall intensities. Also, given the large uncertainty in the output from the GCMs, performing an uncertainty analysis and creating probability based IDF curves, which can better present the results, especially for longer durations, is being undertaken in order to complete the current and previous studies.

#### Chapter 4

## 4. Developing Probability-based IDF Curves Using Kernel Density Estimator

#### 4.1. Abstract

Many hydrologic structures are designed based on Intensity-Duration-Frequency (IDF) curves. A design based on an inaccurate design storm can cause problems, such as, malfunction of the infrastructure, excessive cost, or loss of life. In Chapter 2, IDF curves under future climate scenarios were created for Alabama using six different NARCCAPbased projections. The results demonstrated that these models do not project identical results, and there is uncertainty regarding future rainfall intensities projected by these six climate models. Understanding the uncertainties associated with climate model outputs can help decision-makers to explain the impacts of climate change with more confidence. Therefore, the main objective of this chapter is to develop probability-based IDF curves incorporating climate projections from six different climate models using a kernel density estimator. The resulting IDF curves help understand uncertainties associated with projected rainfall intensities. IDF curves were previously created using two different temporal disaggregation methods: a stochastic method and an ANN model. A kernel density estimator was applied to the resulting estimated rainfall intensities from both methods and probability-based IDFs were developed. In addition to the probability-based IDFs, typical IDF curves that incorporated all models were also developed. Using the median of the distribution as the "most likely" outcome the results are presented in the form of IDF graphs. A comparison of the results with the current IDF curves for Auburn indicated that, when the stochastic method was used for rainfall disaggregation, future rainfall intensity could decreases or increases depending on the return period. The resultant IDF curves for Auburn indicate that future rainfall intensities are expected to decrease by 29% to 39% for durations less than six to eight hours and to increase by 14% to 19% for longer durations. Results of the ANN model for durations less than two hours indicates that the precipitation pattern for Alabama veers toward less intense rainfalls for the investigated durations and for all the return periods. This decrease is expected to be between 48% and 52% for Auburn.

## 4.2.Introduction

As mentioned in previous chapters, many hydrologic structures are designed based on IDF curves (McCuen 1998, Prodanovic et al. 2007), and a design based on an inaccurate design storm can result in malfunction of infrastructure, loss of life in the case of failure, or excessive cost if the design storm is overestimated. In the previous chapters, the development of future IDF curves was discussed, and the results were compared with current IDF curves. The results demonstrated that the six different NARCCAP-based projections are not identical and that there is uncertainty regarding the projected rainfall intensities of these six climate models. Decision-makers can benefit from understanding the uncertainties associated with climate model outputs to explain the impacts of climate change with more confidence (Colglazier 1991, Solaiman 2011). Uncertainties in the outputs of climate models arise from imperfect knowledge of physical processes, inadequate information, or analytical resources. For example, complex atmospheric and

oceanographic processes are being simplified, and imprecise assumptions about different climatic processes have been made.

Uncertainties also exist in estimating the future GHG concentrations, and, as a result, changes in the carbon cycle can be another source of uncertainty of rainfall projections. Despite the fact that the fundamentals of the GHG physics effect are well-founded and an increase in  $CO_2$  concentrations has been recorded since the industrial revolution, there is still discussion and debate about this issue (Solaiman 2011).

In addition, the data are available at a coarse spatial and temporal resolution, which can cause disparities between GCMs when at a regional scale. The earth system's inborn complexity and the inability to predict exact and accurate human behavior are other sources of climate model uncertainty (Wilby and Harris 2006, Stainforth et al. 2007, Buytaert et al. 2009, Solaiman 2011).

The spatial resolution of GCMs usually ranges from 200 to 600 km, but there is an increased need for high-resolution data for climate change assessment and impact studies, and it is accepted that the accuracy of GCMs decrease at finer scales. Variables such as temperature and precipitation from GCMs can be misrepresented because of the coarse resolution. Details of land surfaces, water surfaces and topography are not represented well in some climate models, so the models are unable to predict the high variability in clouds and precipitation and to provide accurate projections (Widmann et al., 2003, Brissette et al. 2007).

Different initial boundary conditions are another source of uncertainty (Stainforth et al. 2007). Another source of uncertainty could be forcing uncertainty. For example, using model simulations based on different scenarios of GHG concentrations in the future

completely depends on the activities undertaken to control these emissions (Cubasch et al. 2001).

Quantifying uncertainties from climate models has been the subject of many studies related to impact assessment (Solaiman 2011). Selecting the most suitable models is one of the key factors in performing reliable climate change assessment research (Tebaldi and Smith 2009). Many studies have used just one climate model to evaluate the impact of climate change in the future. The problem with using one GCM is that, with the presence of significant uncertainties, one model is only one of many possible realizations and may not be representative of the future. Therefore, it is very important to use as many models as possible to present the results in a probabilistic way. Unfortunately, few studies have used a collection of climate models. New and Hulme (2000) used Bayesian Monte Carlo simulations to introduce a probabilistic framework to quantify the climate model's uncertainty. Another study used a CMIP2 multi-model ensemble with equally probable realizations to estimate the probabilities of climate change (Raisanen and Palmer 2001).

Recently, Solaiman (2011) used a kernel density estimator to quantify the uncertainty associated with 11 GCMs for the city of London, Canada. Another study by Giorgi and Mearns (2003) introduced the Reliability Ensemble Averaging (REA) method to estimate the probability of regional precipitation and temperature change. Nine GCMs with two emission scenarios —the A2 and B2 emission scenarios—were used in this study. Bayesian statistics was applied to compute a future climate distribution by using observed data and corresponding simulations from GCMs (Tebaldi et al. 2004, 2005). It was assumed that GCM ensembles illustrate an example of a full potential climate model space consistent with the observed climate using Probability Distribution Functions

(PDFs) at a regional scale. Their study was extended by Smith et al. (2009). They introduced a univariate approach to take into account one region at a time.

A probabilistic framework was developed by Wilby and Harris (2006) to combine the results of four GCMs and two emission scenarios. A Monte Carlo approach was used to find the elements of uncertainty. Cumulative Distribution Functions seemed to be the most sensitive in selecting the climate change scenarios and downscaling different models. Another study by Ghosh and Mujumdar (2007) developed a method to evaluate the uncertainty of climate models to investigate future drought scenarios in a non-parametric manner. More promising results were obtained compared to other parametric methods.

Since all of the investigated climate models in this study provided different results, developing probability-based IDF curves seemed to be an appropriate solution to quantify uncertainties in projected rainfall intensities. Therefore, this chapter presents a methodology to develop probability-based IDF curves using non-parametric kernel density estimation.

#### 4.3. Methodology

#### 4.3.1. Data

As mentioned in the previous two chapters, simulated precipitations for 2038 to 2070 at a temporal resolution of three hours and a spatial resolution of 50 km were obtained from the North American Regional Climate Change Assessment Program (NARCCAP) (Sebastien et al., 2007, Richard et al., 2007). Simulated temperature projections at daily intervals were also obtained from NARCCAP. Using these datasets, future rainfall intensities (for each duration and frequency) were estimated for each NARCCAP regional climate projections. The estimated rainfall intensities were then used to develop the probability-based IDFs.

Probability-based IDF curves were developed from the results of rainfall intensities estimated by the stochastic method (for 15-, 30-, 45-min, 1-, 2-, 3-, 6-, 12-, 24- and 48-hr as well as the ANN model (for 15-, 30-, 45-min, 1-hr and 2-hr).

#### **4.3.2.** Probability Density Function (PDF)

To deal with the uncertainty associated with climate models, a practical method that is robust and flexible enough to handle data variety should be used. It should also be statistically consistent in application across regional, local, and global scales. It needs to provide consistent results to be able to get as much information as possible from the data (Solaiman 2011). Most parametric approaches are unable to meet all mentioned requirements (Solaiman 2011, Zucchini 2003).

The probability distribution of a continuous random variable (x) is described in terms of a Probability Distribution Function (PDF). A PDF is usually used to explain the nature of data (Zucchini 2003). Probabilities associated with (x) can be estimated using the following relationship:

$$P(a \le x \le b) = \int_{a}^{b} f(x) dx$$
 (4.1)

The goal is usually to estimate f(x) from a sample of data  $x_1, x_2, \dots, x_n$ . It can be estimated using a parametric approach that assumes f(x) is a member of some parametric family of distributions and then estimates the assumed distribution's parameters from the

data. This approach is easy to apply but is not useful if the distribution assumption is incorrect (Miller et al. 1965, Zucchini 2003). Lack of flexibility is one of the major drawbacks of parametric approaches (Lall et al. 1996). Therefore, using non-parametric approaches prevents us from making assumptions about the f(x) and estimates the PDF directly from the data (Zucchini 2003). A histogram is widely known as a non-parametric estimator of the PDF. It is very simple to use, but its disadvantage is discontinuity. Moreover, there are other non-parametric methods that are better than histograms (Adamowski 1985, Solaiman 2011).

A non-parametric approach to estimate the PDF can be used for uncertainty quantification of climate models. Methods such as K-nearest neighbor methods, kernel methods, Bayesian-spline methods, maximum likelihood methods, and orthogonal series methods are examples of non-parametric approaches that have been used (Adamowski 1985, Solaiman 2011).

#### 4.3.3. Kernel Density Estimation

Among the methods mentioned above is a method called the kernel density estimation method. This method has been used widely in hydrology, flood frequency analysis, and rainfall resampling as a more reliable and flexible approach than parametric approaches (Sharma et al. 1997, Lall 1995, Adamowski 1985).

Weight functions or kernels convolution centered on the empirical frequency distribution of the data, form a kernel density estimator. The kernel density function can be estimated using the following equation:

$$\int_{-\infty}^{+\infty} K(x) dx = 1$$
 (4.2)

where K(x) is the kernel function. Thus, a PDF can be a kernel function.

The kernel density estimator is sometimes called Parzen-Rosenbalt after Emanuel Parzen and Murray Rosenbalt, who are assumed to have created the kernel density estimator in its current form (Rosenbalt 1956, Parzen 1962).

The kernel density estimator of a sample  $x_1, x_2, \dots, x_n$  from a distribution with an unknown density function f can be determined from:

$$\hat{f}_{h}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{h}(x - x_{i}) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - x_{i}}{h})$$
(4.3)

where  $K\left(\frac{x-x_i}{h}\right)$ , also called K(t), is the kernel function or weight function, and h is a smoothing parameter or bandwidth.

Selecting the appropriate bandwidth and kernel function are the important factors for applying kernel density estimation successfully (Wang et al. 2007; Sharma et al. 1997). Figure 4.1 displays a kernel density estimate and a histogram created from the same data.

Some of the common kernel functions are triangular, Epanechnikov, rectangular, biweight, and normal. Table 4-1 shows the kernel function for these examples (Bergman 2009, Zucchini 2003). The normal kernel is usually used as the kernel function, and it was applied to this study as well.



**Figure 4.1** Comparison of kernel density estimate (right) and histogram (left). The individual kernels are the red dashed curves and the kernel density estimate the blue curves. (From http://en.wikipedia.org/wiki/File:Comparison\_of\_1D\_histogram\_and\_KDE.png retrieved on 03/12/2013

Kernel function	K(t)						
Triangular	$\begin{cases} 1 -  t  &  t  < 1 \\ 0 & otherwise \end{cases}$						
Epanechnikov	$\begin{cases} \frac{3}{4} \left( 1 - \frac{1}{5} t^2 \right) &  t  < \sqrt{5} \\ 0 & otherwise \end{cases}$						
Rectangular	$\begin{cases} \frac{1}{2} &  t  < 1 \\ 0 & otherwise \end{cases}$						
Biweight	$\begin{cases} \frac{15}{16} (1 - t^2)^2 &  t  < 1 \\ 0 & otherwise \end{cases}$						
Normal	$\frac{1}{\sqrt{2\pi}}e^{-(1/2)t^2}$						

Table 4-1 Examples of kernel functions

Kernel bandwidth has a strong effect on the kernel estimate. Many studies have investigated the selection of a smoothing parameter or bandwidth and found that it is even more important than the selection of the kernel function (Wang et al. 2007, Scott 2009, Sheather et al. 1991, Marron et al. 1987). A change in bandwidth can greatly change the shape of the kernel estimate (Sheather et al. 1991, Marron et al. 1987). Various methods have been used to quantify the estimator. One method is the Mean Squared Error (MSE) and its two components: bias and standard error (or variance) (Scott 2009, Zucchini 2003, Sheather et al. 1991, Marron et al. 1987). MSE can be measured using:

$$MSE\left(\hat{f}(\mathbf{x})\right) = E\left[\hat{f}(\mathbf{x}) - f(\mathbf{x})\right]^{2}$$
$$= \left[E\hat{f}(\mathbf{x}) - f(\mathbf{x})\right]^{2} + E\left(\hat{f}(\mathbf{x}) - E\hat{f}(\mathbf{x})\right)^{2}$$
$$= bias^{2}\left(\hat{f}(\mathbf{x})\right) + var(\hat{f}(\mathbf{x})) \qquad (4.4)$$

There is usually a trade-off between the variance and bias of the estimator in equation 4.4. Bias can be reduced by increasing variance, and vice versa, by changing the smoothing parameter.

The most common global measure of  $\hat{f}$  accuracy is the Mean Integrated Squared Error (MISE) (Rosenbalt 1956, Scott 2009, Zucchini 2003, Sheather et al. 1991, Adamowski 1985), given by:

$$MISE(\hat{f}) = E \int_{-\infty}^{+\infty} \left(\hat{f}(x) - f(x)\right)^2 dx$$

$$= \int_{-\infty}^{+\infty} MISE(\hat{f}(x)) dx$$
$$\int_{-\infty}^{+\infty} Bias^{2}(\hat{f}(x)) dx + \int_{-\infty}^{+\infty} Var(\hat{f}(x)) dx \qquad (4.5)$$

# 4.3.3.1. Bias, Variance and MSE

=

The bias is given by (Scott 2009, Zucchini 2003):

$$E\hat{f}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} EK(\frac{x - x_i}{h})$$
$$= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} \int_{-\infty}^{+\infty} K\left(\frac{x - t}{h}\right) f(\mathbf{t}) dt$$
$$= \frac{1}{h} \int_{-\infty}^{+\infty} K\left(\frac{x - t}{h}\right) f(\mathbf{t}) dt \qquad (4.6)$$
If  $z = \frac{x - t}{h}$ , then  $t = x - zh$ ,  $\left|\frac{dz}{dt}\right| = \frac{1}{h}$ 
$$E\hat{f}(\mathbf{x}) = \int_{-\infty}^{+\infty} K(\mathbf{z}) f(\mathbf{x} - zh) dz$$

Using the Taylor series to expand the f(x - zh) yields:

$$f(x - zh) = f(x) - hzf'(x) + \frac{1}{2}(hz)^2 f''(x) + o(h^2)$$

where  $o(h^2)$  represents terms that converge to zero faster than  $h^2$  as h approaches zero. Thus:

$$E\hat{f}(x) = \int_{-\infty}^{+\infty} K(z)f(x)dz - \int_{-\infty}^{+\infty} K(z)hzf'(x)dz + \int_{-\infty}^{+\infty} K(z)\frac{(hz)^2}{2}f''(x)dz + o(h^2)$$
$$= f(x) + \frac{h^2}{2}k_2f''(x) + o(h^2) \qquad (4.7)$$
$$Bias(\hat{f}(x)) \cong \frac{h^2}{2}k_2f''(x) \qquad (4.8)$$

where  $k_2$  is the variance of the kernel,  $h^2$  is the smoothing parameter or bandwidth, and f''(x) is density curvature at x.

The variance is given by (Scott 2009, Zucchini 2003):

$$Var\left(\hat{f}(x)\right) = Var\left(\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-x_{i}}{h}\right)\right)$$

$$= \frac{1}{n^2 h^2} \sum_{i=1}^{n} Var\left(K\left(\frac{x-x_i}{h}\right)\right)$$

Therfore, since  $x_1, x_2, \dots x_n$  are distributed independently:

$$Var\left(K\left(\frac{x-x_{i}}{h}\right)\right) = E\left(K\left(\frac{x-x_{i}}{h}\right)\right)^{2} - \left(EK\left(\frac{x-x_{i}}{h}\right)\right)^{2}$$
$$= \int K\left(\frac{x-t}{h}\right)^{2} f(\mathbf{t}) dt - \left(\int K\left(\frac{x-t}{h}\right) f(\mathbf{t}) dt\right)^{2}$$

$$Var\left(\hat{f}(\mathbf{x})\right) = \frac{1}{n} \int \frac{1}{h^2} K\left(\frac{x-t}{h}\right)^2 f(\mathbf{t}) dt - \frac{1}{n} \left(\frac{1}{h} \int K\left(\frac{x-t}{h}\right) f(\mathbf{t}) dt\right)^2$$
$$= \frac{1}{n} \int \frac{1}{h^2} K\left(\frac{x-t}{h}\right)^2 f(\mathbf{t}) dt - \frac{1}{n} \left(f(\mathbf{x}) + Bias\left(\hat{f}(\mathbf{x})\right)\right)^2$$

Using Taylor approximation and substituting  $z = \frac{x-t}{h}$ , when **n** becomes large and **h** becomes small, the approximation of above equation will be:

$$Var\left(\hat{f}(x)\right) \approx \frac{1}{nh} f(x) \int K^2(z) dz$$
 (4.9)

It is clear that h increases as variance decreases. Combining equations 4.8 and 4.9 leads to:

$$MSE\left(\hat{f}(\mathbf{x})\right) = Bias^{2}\left(\hat{f}(\mathbf{x})\right) + Var\left(\hat{f}(\mathbf{x})\right)$$
$$\approx \frac{1}{4}h^{4}k_{2}^{2}f''(\mathbf{x})^{2} + \frac{1}{nh}f(\mathbf{x})j_{2} \qquad (4.10a)$$

where  $k_2^2 = \int z^2 K(z) dz$  and  $j_2 = \int K(z)^2 dz$ 

Integrating equation 4.10a leads to:

$$MISE(\hat{f}) \approx \frac{1}{4} h^4 k_2^2 \beta(f) + \frac{1}{nh} j_2$$
 (4.10b)

where  $\beta(f) = \int f''(x)^2 dx$ .

As shown in equation 4.10b, MISE is a function of the bandwidth. The first term in equation 4.10 becomes large when h values are large, and the second term becomes large when h decreases. There is a value for h that minimizes the MISE, and that is the optimal bandwidth (Rosenbalt 1956, Scott 2009, Zucchini 2003, Sheather et al. 1991, Adamowski 1985, Jones et al. 1996). Equation 4.10 is the measure of the estimator's performance. By minimizing equation 4.10, the optimal bandwidth can be calculated using the following equation:

$$\frac{dMISE(\hat{f})}{dh} = h^3 k_2^2 \beta(f) - \frac{1}{nh^2} j_2$$

Setting above equation equal to zero will give the optimal bandwidth,  $h_{opt}$ .

$$h_{opt} = \left(\frac{1}{n} \frac{j_2 k_2^{-2}}{\beta(f)}\right)^{1/5}$$
 (4.11)

Substituting equation 4.11 in 4.10, the minimum MISE for the given PDF and kernel can be calculated using:

$$MISE_{opt}(\hat{f}) = \frac{5}{4} \left(\frac{\beta(f)j_2^4k_2^2}{n^4}\right)^{1/5}$$
(4.12)

Finding the optimal bandwidth depends on the sample size, kernel function, and an unknown PDF. A variety of methods have been developed to simplify equation 4.11 based on the selected kernel function, methods such as rule of thumb, cross-validation methods, the plug-in approach, and the smoothed bootstrap approach (Silverman 1986, Rudemo 1982, Bowman 1984, Hall 1992, Sheather 1986, 1983, Sheather et al. 1991, Polansky and Baker 2000). Considering a standard normal density as a kernel function, Polansky and Baker (2000) proposed an equation for optimal bandwidth selection. The optimal bandwidth is given by the following equation and was used in this study (Polansky and Baker 2000):

$$h_{opt} = 1.587 \hat{\sigma} n^{-1/3}$$
 (4.13)

where *n* is the sample size and  $\hat{\sigma} = \min\left\{S_{I}\frac{IQR}{1.349}\right\}$  (Silverman 1986), where *S* is the sample standard deviation and *IQR* is the interquartile range. *IQR* is a measure of statistical dispersion. It is defined as the difference between the upper and lower quartiles  $(IQR = Q3 - Q\mathbf{1})$ .

#### 4.3.4. Probability-based IDF Curves

The results of the created IDFs from the previous chapters were used to develop the probability-based IDF curves. The kernel density estimator was applied to the projected rainfall depths for each return period, and each rainfall duration and probability versus rainfall depth was plotted. To develop the IDF curves, the median of the distribution (probability: 0.5) was selected as the most likely outcome, and the results are presented in the form of IDF graphs.

Probability-based IDF curves were developed from the results of rainfall intensities estimated by the stochastic method (for 15-, 30-, 45-,min, 1-hr and 2-hr) in combination with NARCCAP projections (for 3-,6-,12-,24- and 48-hr) and estimated rainfall intensities using the ANN model (for 15-, 30-, 45-min, 1-hr and 2-hr).

#### **4.4.Results and Discussion**

#### 4.4.1. Optimal bandwidth selection

The equation proposed by Polansky and Baker (2000) was used to calculate the optimal bandwidth. Rainfall intensities for 10 rainfall durations (e.g. 15-min, 30-min) and 6 return periods (e.g. 2-yr, 5-yr) were estimated using six different climate model projections (Chapter 2). Selecting 15-min rainfall duration for Auburn as an example, Table 4-2 shows calculated bandwidth when stochastic method was used as a disaggregation method.

Climato Models	15-min rainfall (mm)									
	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr				
CRCM-CCSM	10.19	13.73	16.17	19.39	21.89	24.49				
CRCM-CGCM3	8.50	11.16	12.77	14.63	15.91	17.09				
HRM3-GFDL	14.88	20.39	23.86	28.06	31.05	33.91				
RCM3-GFDL	16.58	22.28	25.55	29.17	31.54	33.65				
HRM3-HadCM3	12.53	18.40	22.46	27.81	31.94	36.17				
GFDL-ECP2	11.76	15.46	17.85	20.79	22.91	24.98				
			sample si	ize (n)= 6						
Standard deviation(S)	2.97	4.20	4.96	5.90	6.63	7.39				
IQR/1.349	3.48	4.93	5.70	6.43	7.15	6.98				
Min{S,IQR/1.349}	2.97	4.20	4.96	5.90	6.63	6.98				
h <sub>opt</sub> (Equ. 4.13)	2.59	3.67	4.33	5.16	5.79	6.10				

Table 4-2 Optimal bandwidth selection for Auburn, 15-min rainfall duration

The same procedure was applied to calculate the bandwidth for the remaining rainfall durations for all 34 stations in Alabama. Optimal bandwidths were also calculated to develop the probability-based IDF curves when ANN model was used for rainfall disaggregation.

#### 4.4.2. Probability-based IDF Curves

## 4.4.2.1 Stochastic Method

Probability-based IDF curves were developed for 34 stations in Alabama (Figure 2.1). The results of the developed probability-based IDFs when the stochastic method was
used for disaggregation are shown in Figure 4.2 for Auburn. The remaining graphs for other stations in Alabama are presented in Appendix E.





Figure 4.2 Probability-based IDFs when stochastic method used for disaggregation for 15-, 30-, 45min,1-,2-, 3-, 6-, 12-, 24- and 48-hr rainfall duration-graphs are presented for Auburn, AL.

Figure 4.2 provides additional probability information compared to a typical IDF graph, which allows designers and decision-makers to use the IDF curve with more confidence. For example, for a 10-yr return with one-hour rainfall duration, there is a probability of 0.6 that rainfall amount will be less than or equal to 35 mm.

Considering the median of distribution as the most likely outcome, typical IDF curves were developed for all stations in Alabama. Figure 4.3 shows how the resultant IDF was created for Auburn. A probability-based IDF for 15-min rainfall duration is presented as an example to explain the process.



Figure 4.3 Resultant IDF curves for Auburn, AL using the stochastic method

The dashed line presents the rainfall depths (mm) with a probability of 0.5 for 15-min rainfall durations and for all the return periods. Rainfall depths for the remaining rainfall durations (Figure 4.2) were also derived from the probability-based IDF curves. These values were then used to create the resultant IDF curves. IDF curves for the remaining stations (33 stations) are presented in Appendix F.

To discuss the magnitude of uncertainty using a particular GCM projection on deriving IDF curves, Table 4-3 presents rainfall depths for 15-min rainfall durations. The upper section of the table is similar to Table 4-2. The lower section summarize the rainfall depths with a most likely probability of 0.5 derived from probability-based IDF curve (Figure 4.3), minimum and maximum of the rainfall depths that presents the rainfall depth range.

Climata Madala -	15-min rainfall (mm)								
	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr			
CRCM-CCSM	10.19	13.73	16.17	19.39	21.89	24.49			
CRCM-CGCM3	8.50	11.16	12.77	14.63	15.91	17.09			
HRM3-GFDL	14.88	20.39	23.86	28.06	31.05	33.91			
RCM3-GFDL	16.58	22.28	25.55	29.17	31.54	33.65			
HRM3-HadCM3	12.53	18.40	22.46	27.81	31.94	36.17			
GFDL-ECP2	11.76	15.46	17.85	20.79	22.91	24.98			
Rainfall depth from probability based-IDF	12.34	16.94	19.93	23.64	26.30	28.93			
Min	8.50	11.16	12.77	14.63	15.91	17.09			
Max	16.58	22.28	25.55	29.17	31.94	36.17			
Percentage	-31.10	-34.12	-35.93	-38.12	-39.52	-40.93			
difference (%)	34.40	31.52	28.19	23.38	21.42	25.02			

Table 4-3 Uncertainty quantification for 15-min rainfall duration, Auburn, AL

Table 4-3 shows that rainfall depths for 2-yr return period range from 8.5 mm (projection from CRCM-CGCM3) to 16.58 mm (projection from RCM3-GFDL). Using 50% as most likely probability, 2-yr, 15-min rainfall is projected as 12.34 mm. Therefore, 8.5 mm projection from CRCM-CGCM3 is 31% underestimated but 16.58 mm projection from RCM-GFDL is about 34% overestimated. Results for the remaining return periods (5-, 10-, 25-, 50- and 100-yr) are presented in Table 4-3. Figure 4.4 also displays the rainfall depth range and the resulting rainfall depth derived from probability-based IDF curve.



**Figure 4.4** Projected rainfall depth range from six climate models (bars show minimum and maximum) and the red dot represents the rainfall depth derived from the probability-based IDF.

Figure 4.4 shows that rainfall depth range (difference between max and min) increases for longer return periods. For example, for 100-yr return period, rainfall depth derived from the probability-based IDF projects this amount to be about 29 mm.

Projected rainfall depth from HRM3-HadCM is 36.17 mm which is 25% overestimated compared with rainfall depth derived from the probability-based IDF. CRCM-CGCM3 projection for the same return period and same duration is 17.09 mm that shows about 41% underestimation.

Solaiman and Simonovic (2011) developed probability-based IDF curves using 27 climate change scenarios for the city of London, Ontario, Canada. To compare the results of their study with this one, Figure 4.5 shows projected rainfall depth range (bars) and rainfall depth derived from probability-based IDF (red dots) for Auburn and London. Results are presented for 1-, 2-, 6-, 12- and 24-hr rainfall durations.





Figure 4.5 Projected rainfall depth range (bars show minimum and maximum and the red dot represents the rainfall depth derived from the probability-based IDF) for Auburn, AL (left side) and London, Ontario, Canada (right side).

Table 4-4 shows the rainfall depths derived from probability-based IDF, ranges of the rainfall depths projected by different future climate models for Auburn, AL and London, ON. Results are presented for 1-, 2-, 6-, 12- and 24-hr rainfall duration. Percentage differences (Table 4-4) were calculated between minimum or maximum rainfall depths projected by different future climate models and rainfall depth derived from probability-based IDF for each duration. The percent differences between minimum rainfall depths and the rainfall depths derived from probability-based IDF for five durations from only six future climate models from Auburn range from -33.5% to -43.5% with an average of -37.2%. These percent differences are comparable to ones obtained from 27 future climate models for London, Ontario, i.e. ranging from -18.3% to -22.1% with average of -20.6%. The percent differences of maximum rainfall depths are also comparable to ones obtained from Solaiman and Simonovic (2011).

Figure 4.5 and data in Table 4-5 show that rainfall depth ranges for Auburn using six climate models are comparable to the ones for London while using 27 climate change scenarios. Using more GCMs to develop the probability-based IDFs will lead to obtain more accurate results but Figure 4.5 shows that even though only six GCM projections were used, the IDF curves developed using the probability method are more representative than only one model to what it would happen in the future. Analyzing the results from Table 4-3 and the remaining rainfall durations (not shown in the table) for Auburn shows that projections from CRCM-CGCM3 are likely to be underestimated but RCM3-GFDL and in some cases HRM3-GFDL most likely overestimates rainfall in Alabama, and projections from HRM3-HadCM3 are the most representative.

City	Auburn, AL					London, ON						
Duration	1-hr											
Return period	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr
50%*	18.0	25.6	30.9	37.8	43.1	48.5	33.0	43.0	52.0	62.0	67.0	77.0
Min	10.4	14.1	17.0	21.0	24.5	28.3	24.1	33.1	38.8	46.0	51.4	56.7
Max	29.1	40.4	47.6	56.3	62.4	71.1	42.1	59.6	71.2	85.9	96.8	107.6
Range	18.7	26.3	30.6	35.2	38.0	42.8	18.0	26.6	32.4	39.9	45.4	50.9
Percentage	-42.0	-44.9	-45.2	-44.3	-43.2	-41.7	-26.9	-23.1	-25.4	-25.8	-23.3	-26.4
diff (%)**	61.7	57.7	53.9	48.8	44.8	46.6	27.5	38.7	37.0	38.5	44.4	39.7
Duration	2-hr											
50%	26.0	35.2	41.3	49.2	55.2	61.2	40.0	55.0	67.0	81.0	91.0	101.0
Min	14.9	18.3	20.8	24.3	27.2	30.4	30.4	40.3	46.9	55.2	61.3	67.5
Max	41.4	53.6	60.7	71.1	79.5	87.6	55.9	78.7	93.8	113.6	129.1	144.5
Range	26.5	35.3	39.9	46.8	52.2	57.2	25.5	38.4	46.9	58.4	67.8	77.0
Percentage	-42.9	-48.1	-49.7	-50.6	-50.7	-50.3	-24.1	-26.7	-30.0	-31.9	-32.6	-33.2
diff (%)	59.1	52.4	46.9	44.6	43.9	43.2	39.7	43.1	40.0	40.3	41.9	43.0
Duration	6-hr											
50%	48.9	68.3	81.5	98.7	111.9	124.4	52.0	72.0	84.0	101.0	113.0	125.0
Min	27.9	35.6	41.4	49.7	56.6	64.2	40.3	53.5	62.1	73.0	81.0	89.0
Max	81.8	103.4	117.7	146.3	169.0	204.5	71.9	101.4	120.9	145.6	163.8	182.0
Range	53.9	67.8	76.3	96.6	112.4	140.3	31.6	47.9	58.8	72.6	82.8	93.0
Percentage	-42.9	-47.8	-49.2	-49.7	-49.4	-48.4	-22.5	-25.8	-26.1	-27.8	-28.3	-28.8
diff (%)	67.2	51.5	44.5	48.1	51.0	64.4	38.3	40.8	43.9	44.1	45.0	45.6
Duration		12-hr										
50%	67.8	95.7	114.5	138.4	155.1	167.5	62.0	83.0	97.0	114.0	128.0	141.0
Min	41.5	54.9	65.0	79.6	91.9	99.6	49.4	66.6	77.4	90.2	99.7	109.2
Max	110.0	141.2	178.6	233.6	280.8	333.9	85.8	117.2	138.0	164.3	183.8	203.2
Range	68.5	86.4	113.6	154.0	188.9	234.2	36.4	50.6	60.6	74.0	84.1	94.0
Percentage	-38.8	-42.7	-43.2	-42.5	-40.7	-40.5	-20.3	-19.8	-20.2	-20.8	-22.1	-22.6
diff (%)	62.2	47.6	56.0	68.8	81.1	99.3	38.4	41.3	42.3	44.1	43.6	44.1
Duration						24-]	hr					
50%	87.1	121.0	144.7	177.2	202.4	232.1	71.0	94.0	110.0	129.0	144.0	158.0
Min	57.9	77.7	92.7	114.2	123.7	131.1	56.4	76.8	87.8	102.0	112.6	123.1
Max	124.9	184.6	241.4	328.1	405.2	494.7	98.2	133.0	156.0	185.0	206.6	228.2
Range	67.0	106.9	148.7	213.9	281.5	363.6	41.8	56.2	68.2	83.0	94.1	105.1
Percentage	-33.5	-35.8	-36.0	-35.6	-38.9	-43.5	-20.6	-18.3	-20.1	-20.9	-21.8	-22.1
diff (%)	43.4	52.6	66.8	85.1	100.2	113.1	38.3	41.4	41.8	43.4	43.5	44.5

 Table 4-4 Uncertainty quantification of rainfall depths (mm) projected from different climate models for Auburn, AL and London, ON

\* 50%: rainfall depth derived from the probability-based IDF, considering 50% probability as a most likely outcome.

\*\* Percentage diff (%): percentage difference of min or max from 50%.

A comparison of the resultant IDF curves with the current IDFs is presented in Figure 4.6. Figure 4.6a shows that, for a 2-yr return period, rainfall intensity is reduced by 39% for durations of less than 12 hours. A 14% increase is expected for durations of more than 12 hours. Figure 4.6b displays a 36% decrease for a 5-yr return period for durations of less than eight hours and a 17% increase for durations longer than 8 hours. For a 10-yr return period, a 34% decline for durations of less than 8 hours and an 18% increase in rainfall intensities are expected (Figure 4.6c).

Figure 4.6d depicts the results for a 50-yr return period. It shows that the projected rainfall intensity tends to decrease by 31% for durations of less than 6 hours and is expected to increase by 19% for rainfall durations of more than six hours.



Figure 4.6 Resultant IDF curves under future and current climate for Auburn, AL when the stochastic method was used for rainfall disaggregation

#### 4.4.2.2 ANN Model

Figure 4.7 displays probability-based IDFs when the ANN model was used to disaggregate rainfall. For example, if the rainfall depth of a one-hour storm for a 10-yr return period is 30 mm, the maximum probability of this specific rainfall is about 0.6 (Figure 4.7). The resultant IDF curves created using the probability of 0.5 as the most likely probability are presented in Figure 4.8. The results were also compared to current IDF curves. Figure 4.9 shows the IDF curves under the future and current climate for different return periods and compares the result of the ANN model with the stochastic method for durations of less than two hours.

Figure 4.9a shows a 48% decrease for a 2-yr return period for durations less than two hours. The stochastic method shows almost the similar results. There was a 50% decrease for a 5-yr return period. The stochastic method estimates a 46% decrease for the same return period (Figure 4.9b). Figure 4.9c displays the IDF curves under the future and current climate for a 10-yr storm. It shows that the rainfall intensity is expected to decrease by 51%. A 44% decrease is projected when the stochastic method is used. Figure 4.9d shows that the rainfall intensity in expected to decrease by 52% for a 50-yr return period. The stochastic method shows a decrease of 40%. A 52% decrease in rainfall intensities is expected for a 100-yr return period (Figure 4.9e). A 38% decrease is estimated by the stochastic method.



**Figure 4.7** Probability-based IDFs when ANN model used for disaggregation for 15-, 30-, 45-min,1- and 2-hr rainfall duration-graphs are presented for Auburn, AL.



Figure 4.8 Resultant IDF curves for Auburn, AL using ANN model

The results of both the ANN model and the stochastic method suggest that future rainfall intensities for rainfall durations of less than two hours are expected to decrease. They estimate almost the same decrease for two-, five-, and ten-year return periods.



Figure 4.9 Resultant IDF curves under future and current climate for Auburn, AL

#### **4.5.Summary and Conclusions**

Developing the IDF curves under the future climate for the state of Alabama was investigated in the previous chapters, and the results showed that, due to the uncertainty associated with climate models, these models are unable to project identical results. Because all of the investigated climate models in this study provided different results, developing probability-based IDF curves offered a more reliable approach. Therefore, this chapter presented a methodology to develop probability-based IDF curves using nonparametric kernel density estimation. A probability-based IDF curve provides additional probability information that allows designers and decision-makers to use IDF curves with more confidence.

The resultant IDF curves were also created by selecting the median of the distribution as the most likely outcome. Analyzing the results of the ANN model and the stochastic method for durations of less than two hours indicated that they both agree that the precipitation pattern for Alabama veers toward less intense rainfalls for the investigated durations and for all the return periods. The results indicated that, for small duration events, future rainfall intensity is expected to decrease by 54% on average for the State. The result is not exactly the same when stochastic method was use but still indicates a decline in future rainfall intensities. A 47% decrease on average is expected for durations of less than 2 hours.

A comparison of the results with the current IDF curves for Alabama indicated that, when the stochastic method was used for rainfall disaggregation, future rainfall intensity could decreases or increases depending on the return period. Results displays a 35% decrease on an average for Alabama. This decline mostly is expected for durations of less than six hours and 12 hours, depending on the location and return period. For longer durations, an 11% increase is projected.

The results of this study indicated that incorporating all six NARCCAP projections to develop the resultant IDF curves project that future rainfall intensities mostly for rainfall durations of less than six hours is expected to decrease, and this finding is independent of the temporal downscaling method used.

# **Chapter 5**

## 5. Conclusions

#### **5.1. Summary and Conclusions**

Many hydrologic designs of water management infrastructures are based on specific design storms derived from historic rainfall events available in the form of intensityduration-frequency (IDF) curves. However, it is expected that the frequency and magnitude of future extreme rainfalls will change due to increase in greenhouse gas concentrations in earth's atmosphere (Prodanovic et al. 2007). There is scientific agreement that there will be an increase in the average precipitation around the world. However, this does not imply that more precipitation will occur everywhere. Extreme events can present great challenges for water management utilities. Degradation of water quality, forcing additional costs for treatment, property damage and potential loss of life due to flooding are some of the things that are caused by extreme events (Miller and Yates 2006). On the other hand, drought can decrease the capability of water utilities to meet water demands and enforce emergency restrictions (Miller and Yates 2006). Changes in precipitation characteristics may require revision of existing standards for designing civil engineering infrastructures. It can also require the rebuilding and/or upgrading of existing infrastructures. One way to be prepared for possible changes, and to decrease vulnerability of hydraulic infrastructures to climate change, is to predict potential effects (as manifested by IDF curves) and adapt to them. Therefore, the overall

objective of this study was to evaluate the impact of climate change on rainfall IDF curves for Alabama. Three objectives were presented in the beginning of this dissertation. For each of the objectives, major findings are summarized below.

## **5.1.1.** Objective 1

To develop rainfall IDF curves for Alabama under future climate change scenario; using a stochastic method to disaggregate three-hourly precipitations into 15-minute precipitations.

Six high-resolution projections (for 2038-2070) derived from dynamical downscaling of General Circulation Models (GCMs) by Regional Climate Models (RCMs) were used in this study. Three-hourly projected precipitations were disaggregated using a stochastic method and rainfall IDF curves under future climate scenario were created. Major conclusions were:

- 1. The rainfall pattern for Alabama will change in future due to climate change.
- 2. Results of the six NARCCAP projections were not identical. A variety of factors can be responsible for differing results; a likely reason is the difference in physical parameterizations, especially of radiative and precipitation-forming processes, amongst different GCMs and RCMs, as well as the difference in initial and boundary conditions for each climate projection.
- Four of these models— HRM3-HadCM3, HRM3-GFDL, RCM3-GFDL and ECP2-GFDL— showed both increase and decrease in predicted rainfall intensity. This increase/decrease depends on the rainfall duration and return period of the storm.

- 4. The remaining two model projections— CRCM-CGCM3 and CRCM-CCSM indicated a reduction in future rainfall intensity for all return periods and all rainfall durations for Alabama.
- 5. There is a large uncertainty on the projected rainfall intensity of these six climate models for long durations (i.e., larger than 4 hours).
- 6. Precipitation pattern for Alabama veers toward less intense rainfalls for short rainfall durations.

## **5.1.2.** Objective 2

Use Artificial Neural Network (ANN) to estimate the maximum 15-, 30-, 45-, 60-, and 120-min rainfall depths from three-hourly precipitations for the creation of future IDF curves for Alabama.

A feed-forward, back-propagation model was developed to estimate maximum precipitations for different rainfall durations. The performance measures for evaluation of the ANN model were Nash-Sutcliffe Efficiency coefficient (NSE) and correlation coefficient (R). Normalized Mean Square Error (NMSE) and Akaike Information Criterion (AIC) were used to find the optimum input data. The results of the ANN model were also compared to the results of the rainfall disaggregation using stochastic method from the previous chapter and it was found that:

1. Scatter plots created to evaluate the performance of the ANN model suggest that the predictions of the ANN model more closely approximated the perfect prediction line.

- 2. The ANN model provided a superior performance in estimating maximum rainfall depths, while the stochastic method under-predicted maximum rainfall depths.
- Analyzing the developed IDF curves for Auburn indicated that future rainfall intensities for the studied duration (< 2 hours) events are expected to decrease by 33% to 74% from current ones.
- 4. Comparison of the developed ANN model and the stochastic method showed that both methods are in agreement that future rainfall intensities (for duration less than 2 hours) are expected to decrease.
- ANN model and the stochastic method estimated almost similar decreases in rainfall intensities when CRCM-CGCM3 and CRCM-CCSM models were used to create IDF curves under the future climate.
- The ANN model displayed more declines in rainfall intensities compared to the stochastic method when HRM3-HadCM3, HRM3-GFDL, and RCM3-GFDL projections were used.

# 5.1.3. Objective 3

To develop probability-based IDF curves incorporating climate projections from six different climate models.

Results of the developed IDF curves clearly indicated that due to the uncertainty associated with climate models, these models are unable to project identical results. Therefore, developing probability-based IDF curves seemed to offer a more reliable approach to incorporate all the models. Non-parametric kernel density estimator was used to develop probability-based IDF curves that provides additional probability information and allows designers and decision-makers to use IDF curves with more confidence.

Major findings of this study were:

- 1. With the presence of significant uncertainties associated with climate models, using one single GCM provides one of many possible realizations and may not be representative of the future. Utilizing a collection of models can provide more realistic information about the possible changes in future.
- 2. Using the ANN model to disaggregate rainfall, the resultant IDF curve for Alabama indicated that future rainfall intensity is expected to decrease by 54% on average for durations less than two hour.
- A decline by 47% on average for durations less than 2 hours was estimated for Alabama when stochastic method was used.
- 4. For durations more than 2 hours, changes in rainfall intensity depends on the rainfall duration and return period. It could decrease or increase for a given duration and return period.
- 5. Probability-based IDF curves developed using the kernel density estimator provide very useful information that can help decision-makers to use IDF curves with more confidence.
- 6. Although the results derived from different climate projections shows large uncertainty associated with climate models, all of them indicate decrease in future rainfall intensity for short durations, especially durations of less than 2 hours.

# Chapter 6

## 6. Future research

This study offers two temporal downscaling methods to disaggregate precipitation and develop the IDF curves. It also presented a framework for uncertainty quantification associated with six NARCCAP climate models. This chapter provides some recommendations for future research.

- Using a collection of models can provide more realistic information about the possible changes in future than only one model. Six climate models were used in this study, but incorporating more models will improve the results and allows us to use the information of IDF curves with higher level of confidence.
- The NARCCAP modelers used only one emission scenario for simulation. In addition to using more climate models, considering more scenarios can also lead to improvements in IDF curves.
- 3. Spatial and temporal resolution of the projections used in this study was 50 km and 3 hour, respectively. Incorporating models with higher spatial and temporal resolution could significantly improve the results.
- 4. Uncertainties resulting from different downscaling methods were not considered in this study. Therefore, future research may include uncertainty investigation of the disaggregation method.

### References

- Adamowski, Kaz. "Nonparametric Kernel Estimation of Flood Frequencies." *Water Resources Research* 21, no. 11 (1985): 1585-90.
- Anderson, JL, V Balaji, Anthony J Broccoli, William F Cooke, Thomas L Delworth, Keith W Dixon, Leo J Donner, *et al.* "The New Gfdl Global Atmosphere and Land Model Am2-Lm2: Evaluation with Prescribed Sst Simulations." *J. Climate* 17, no. 24 (2004): 4641-73.
- Barron, E., and S. Sorooshian. "Assessing the Impacts of Climate on Regionalwater Resources." In 12th Mission to Planet Earth Observing System Investigators Working Group Meeting, 1997.
- Bedient, Philip B, and Wayne Charles Huber. "Hydrology and Floodplain Analysis." (1988).
- Bergman, Stefan, and Menahem Schiffer. *Kernel Functions and Differential Equations*. Vol. 4: Academic Press, 2009.
- Bhunya, PK, SK Jain, CS Ojha, and A Agarwal. "Simple Parameter Estimation Technique for Three-Parameter Generalized Extreme Value Distribution." *Journal of Hydrologic Engineering* 12, no. 6 (2007): 682-89.
- Biau, Gérard, Eduardo Zorita, Hans von Storch, and Hans Wackernagel. "Estimation of Precipitation by Kriging in the Eof Space of Thesea Level Pressure Field." *Journal of Climate* 12, no. 4 (1999): 1070-85.
- Bonnin, Geoffrey M, Deborah Martin, Bingzhang Lin, Tye Parzybok, Michael Yekta, and David Riley. "Precipitation-Frequency Atlas of the United States." *NOAA atlas 14*, no. 2 (2004).
- Bowden, Gavin J, Graeme C Dandy, and Holger R Maier. "Input Determination for Neural Network Models in Water Resources Applications. Part 1—Background and Methodology." *Journal of Hydrology* 301, no. 1 (2005): 75-92.
- Bowman, Adrian W. "An Alternative Method of Cross-Validation for the Smoothing of Density Estimates." *Biometrika* 71, no. 2 (1984): 353-60.
- Briegleb, BP, CM Bitz, EC Hunke, WH Lipscomb, MM Holland, JL Schramm, and RE Moritz. "Scientific Description of the Sea Ice Component in the Community Climate System Model, Version Three." Technical Note TN-463STR, NTIS# PB2004-106574, National Center for Atmospheric Research, Boulder, CO, 2004.

- Brissette, FP, M Khalili, and R Leconte. "Efficient Stochastic Generation of Multi-Site Synthetic Precipitation Data." *Journal of Hydrology* 345, no. 3 (2007): 121-33.
- Brown, SA, SM Stein, and JC Warner. "Urban Drainage Design Manual, Hydraulic Engineering Circular 22, Federal Highway Administration (Fhwa)." Report No. FHWA-NHI-01-021 HEC-22, 2001.
- Buishand, T Adri, and Theo Brandsma. "Dependence of Precipitation on Temperature at Florence and Livorno (Italy)." *Climate Research* 12 (1999): 53-63.
- Buishand, T, and Theo Brandsma. "Comparison of Circulation Classification Schemes for Predicting Temperature and Precipitation in the Netherlands." *International Journal of Climatology* 17, no. 8 (1997): 875-89.
- Buishand, TA, and AMG Klein Tank. "Regression Model for Generating Time Series of Daily Precipitation Amounts for Climate Change Impact Studies." *Stochastic Hydrology and Hydraulics* 10, no. 2 (1996): 87-106.
- Burian, Steven J, S Rocky Durrans, Stephan J Nix, and Robert E Pitt. "Training Artificial Neural Networks to Perform Rainfall Disaggregation." *Journal of Hydrologic Engineering* 6, no. 1 (2001): 43-51.
- Burian, Steven J, S Rocky Durrans, Saša Tomic, Russell L Pimmel, and Chung Ngai Wai. "Rainfall Disaggregation Using Artificial Neural Networks." *Journal of Hydrologic Engineering* 5, no. 3 (2000): 299-307.
- Buytaert, Wouter, Rolando Célleri, and Luis Timbe. "Predicting Climate Change Impacts on Water Resources in the Tropical Andes: Effects of Gcm Uncertainty." *Geophysical Research Letters* 36, no. 7 (2009).
- Caya, Daniel, and Rene Laprise. "A Semi-Implicit Semi-Lagrangian Regional Climate Model: The Canadian Rcm." *Monthly Weather Review* 127, no. 3 (1999): 341-62.
- Choi, Janghwan, Scott A Socolofsky, and Francisco Olivera. "Hourly Disaggregation of Daily Rainfall in Texas Using Measured Hourly Precipitation at Other Locations." *Journal of Hydrologic Engineering* 13, no. 6 (2008): 476-87.
- Coles, Stuart. An Introduction to Statistical Modeling of Extreme Values. London ; New York: Springer, 2001.
- Colglazier, E William. "Scientific Uncertainties, Public Policy, and Global Warming: How Sure Is Sure Enough?". *Policy studies journal* 19, no. 2 (1991): 61-72.
- Collins, M, SFB Tett, and C Cooper. "The Internal Climate Variability of Hadcm3, a Version of the Hadley Centre Coupled Model without Flux Adjustments." *Climate Dynamics* 17, no. 1 (2001): 61-81.
- Collins, William D, Cecilia M Bitz, Maurice L Blackmon, Gordon B Bonan, Christopher S Bretherton, James A Carton, Ping Chang, *et al.* "The Community Climate System Model Version 3 (Ccsm3)." *Journal of Climate* 19, no. 11 (2006a): 2122-43.
- Collins, William D, Philip J Rasch, Byron A Boville, James J Hack, James R McCaa, David L Williamson, Bruce P Briegleb, *et al.* "The Formulation and Atmospheric

Simulation of the Community Atmosphere Model Version 3 (Cam3)." *Journal of Climate* 19, no. 11 (2006b): 2144-61.

- Collins, William D, Philip J Rasch, Byron A Boville, James J Hack, James R McCaa, David L Williamson, Jeffrey T Kiehl, *et al.* "Description of the Ncar Community Atmosphere Model (Cam 3.0)." *NCAR Tech. Note NCAR/TN-464+ STR* (2004).
- Conway, D, RL Wilby, and PD Jones. "Precipitation and Air Flow Indices over the British Isles." *Climate Research* 7 (1996): 169-83.
- Coulibaly, Paulin, and Xiaogang Shi. "Identification of the Effect of Climate Change on Future Design Standards of Drainage Infrastructure in Ontario." *McMaster University, Hamilton, Ontario: Department of Civil Engineering* (2005).
- Cox, John D. Storm Watchers: The Turbulent History of Weather Prediction from Franklin's Kite to El Niño. Wiley, 2002.
- Cubasch, U, and Coauthors. "Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC." Cambridge University Press, Cambridge, UK, 2001.
- Cubasch, Ulrich, H Von Storch, J Waszkewitz, and E Zorita. *Estimates of Climate Change in Southern Europe Using Different Downscaling Techniques*. Max-Planck-Institut für Meteorologie, 1996.
- Dai, Aiguo. "Precipitation Characteristics in Eighteen Coupled Climate Models." *Journal* of Climate 19, no. 18 (2006): 4605-30.
- Dawson, CW, and RL Wilby. "Hydrological Modelling Using Artificial Neural Networks." *Progress in physical Geography* 25, no. 1 (2001): 80-108.
- Dehn, Martin, and Jelle Buma. "Modelling Future Landslide Activity Based on General Circulation Models." *Geomorphology* 30, no. 1 (1999): 175-87.
- Dibike, Yonas B, and Paulin Coulibaly. "Temporal Neural Networks for Downscaling Climate Variability and Extremes." *Neural Networks* 19, no. 2 (2006): 135-44.
- Dickinson, Robert Earl, PJ Kennedy, and Ann Henderson-Sellers. *Biosphere-Atmosphere Transfer Scheme (Bats) Version 1e as Coupled to the Ncar Community Climate Model.* National Center for Atmospheric Research, Climate and Global Dynamics Division, 1993.
- Dickinson, Robert E, Keith W Oleson, Gordon Bonan, Forrest Hoffman, Peter Thornton, Mariana Vertenstein, Zong-Liang Yang, and Xubin Zeng. "The Community Land Model and Its Climate Statistics as a Component of the Community Climate System Model." *Journal of Climate* 19, no. 11 (2006): 2302-24.
- Durrans, S Rocky, and Pauline A Brown. "Estimation and Internet-Based Dissemination of Extreme Rainfall Information." *Transportation Research Record: Journal of the Transportation Research Board* 1743, no. 1 (2001): 41-48.
- Econopouly, TW, DR Davis, and DA Woolhiser. "Parameter Transferability for a Daily Rainfall Disaggregation Model." *Journal of Hydrology* 118, no. 1 (1990): 209-28.

- Feddersen, Henrik, and Uffe Andersen. "A Method for Statistical Downscaling of Seasonal Ensemble Predictions." *Tellus A* 57, no. 3 (2005): 398-408.
- Flato, GM, and GJ Boer. "Warming Asymmetry in Climate Change Simulations." *Geophysical Research Letters* 28, no. 1 (2001): 195-98.
- Frederick, Ralph H, Vance A Myers, Eugene P Auciello, United States. Soil Conservation Service, and United States. National Weather Service. *Five-to 60-Minute Precipitation Frequency for the Eastern and Central United States*. US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, 1977.
- Ghosh, Subimal, and PP Mujumdar. "Nonparametric Methods for Modeling Gcm and Scenario Uncertainty in Drought Assessment." Water Resources Research 43, no. 7 (2007).
- Giorgi, F, and LO Mearns. "Probability of Regional Climate Change Based on the Reliability Ensemble Averaging (Rea) Method." *Geophysical Research Letters* 30, no. 12 (2003).
- Glasbey, CA, G Cooper, and MB McGechan. "Disaggregation of Daily Rainfall by Conditional Simulation from a Point-Process Model." *Journal of Hydrology* 165, no. 1 (1995): 1-9.
- Gordon, Chris, Claire Cooper, Catherine A Senior, Helene Banks, Jonathan M Gregory, Timothy C Johns, John FB Mitchell, and Richard A Wood. "The Simulation of Sst, Sea Ice Extents and Ocean Heat Transports in a Version of the Hadley Centre Coupled Model without Flux Adjustments." *Climate Dynamics* 16, no. 2-3 (2000): 147-68.
- Gupta, Vijay K, and Edward C Waymire. "A Statistical Analysis of Mesoscale Rainfall as a Random Cascade." *Journal of Applied Meteorology* 32 (1993): 251-51.
- Halff, Albert H, Henry M Halff, and Masoud Azmoodeh. "Predicting Runoff from Rainfall Using Neural Networks." Paper presented at the Engineering Hydrology (1993), 1993.
- Hall, Peter, JS Marron, and Byeong U Park. "Smoothed Cross-Validation." *Probability Theory and Related Fields* 92, no. 1 (1992): 1-20.
- Hansen, James W, Andrew Challinor, Amor Valeriano M Ines, Tim Wheeler, and Vincent Moron. "Translating Climate Forecasts into Agricultural Terms: Advances and Challenges." *Climate Research* 33, no. 1 (2006): 27-41.
- Hershfield, DM. "Rainfall Frequency Atlas of the United States." US Weather Bureau *Technical Paper*, no. 40 (1961): 115.
- Hershenhorn, J, and DA Woolhiser. "Disaggregation of Daily Rainfall." *Journal of Hydrology* 95, no. 3 (1987): 299-322.
- Hershfield, David Michael. *Rainfall Frequency Atlas of the United States: For Durations* from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years. US Department of Commerce, Weather Bureau, 1963.

- Hosking, JRM, James R Wallis, and Eric F Wood. "Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments." *Technometrics* 27, no. 3 (1985): 251-61.
- Huard, David, Alain Mailhot, and Sophie Duchesne. "Bayesian Estimation of Intensity– Duration–Frequency Curves and of the Return Period Associated to a Given Rainfall Event." *Stochastic Environmental Research and Risk Assessment* 24, no. 3 (2010): 337-47.
- IPCC. Impacts, Adaptation, and Vulnerability. Contribution of Working Group Ii to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press Cambridge, UK, 2007.
- Isik, Sabahattin, Latif Kalin, Jon E Schoonover, Puneet Srivastava, and B Graeme Lockaby. "Modeling Effects of Changing Land Use/Cover on Daily Streamflow: An Artificial Neural Network and Curve Number Based Hybrid Approach." *Journal of Hydrology* (2012).
- Islam, S, D Entekhabi, RL Bras, and I Rodriguez-Iturbe. "Parameter Estimation and Sensitivity Analysis for the Modified Bartlett-Lewis Rectangular Pulses Model of Rainfall." *Journal of Geophysical Research* 95, no. D3 (1990): 2093-100.
- Jana, Bipal K, and Mrinmoy Majumder. *Impact of Climate Change on Natural Resource Management*. Springer, 2010.
- Johns, TC, JM Gregory, WJ Ingram, CE Johnson, A Jones, JA Lowe, JFB Mitchell, et al. "Anthropogenic Climate Change for 1860 to 2100 Simulated with the Hadcm3 Model under Updated Emissions Scenarios." *Climate Dynamics* 20, no. 6 (2003): 583-612.
- Jones, M Chris, James S Marron, and Simon J Sheather. "A Brief Survey of Bandwidth Selection for Density Estimation." *Journal of the American Statistical Association* 91, no. 433 (1996): 401-07.
- Jones, Richard. *Generating High Resolution Climate Change Scenarios Using Precis*. Met office, 2004.
- Kalin, Latif, Sabahattin Isik, Jon E Schoonover, and B Graeme Lockaby. "Predicting Water Quality in Unmonitored Watersheds Using Artificial Neural Networks." *Journal of environmental quality* 39, no. 4 (2010): 1429-40.
- Katz, RW, and MB Parlange. "Mixtures of Stochastic Processes: Application to Statistical Downscaling." *Climate Research* 7 (1996): 185-93.
- Kiehl, Jeffrey T, James J Hack, Gordon B Bonan, and B Boville. "Description of the Ncar Community Climate Model(Ccm 3)." *NASA*, no. 19980147981 (1996).
- Kim, S-J, G Flato, and G Boer. "A Coupled Climate Model Simulation of the Last Glacial Maximum, Part 2: Approach to Equilibrium." *Climate Dynamics* 20, no. 6 (2003): 635-61.

- Kim, S-J, G Flato, G Boer, and N McFarlane. "A Coupled Climate Model Simulation of the Last Glacial Maximum, Part 1: Transient Multi-Decadal Response." *Climate Dynamics* 19, no. 5-6 (2002): 515-37.
- Lall, U. "Nonparametric Function Estimation: Recent Hydrologic Contributions." *Reviews of Geophysics* (1995): 1093-102.
- Lall, Upmanu, Balaji Rajagopalan, and David G Tarboton. "A Nonparametric Wet/Dry Spell Model for Resampling Daily Precipitation." *Water Resources Research* 32, no. 9 (1996): 2803-23.
- Laprise, Rene, Daniel Caya, Michel Giguere, Guy Bergeron, Hélène Côté, Jean-Pierre Blanchet, George J Boer, and Norman A McFarlane. "Climate and Climate Change in Western Canada as Simulated by the Canadian Regional Climate Model." *Atmosphere-Ocean* 36, no. 2 (1998): 119-67.
- Laprise, René, and Claude Girard. "A Spectral General Circulation Model Using a Piecewise-Constant Finite-Element Representation on a Hybrid Vertical Coordinate System." *Journal of Climate* 3 (1990): 32-52.
- Lettenmaier, Dennis. "Stochastic Modeling of Precipitation with Applications to Climate Model Downscaling." In *Analysis of Climate Variability*, 199-214: Springer, 1999.
- Li, Haibin, Justin Sheffield, and Eric F Wood. "Bias Correction of Monthly Precipitation and Temperature Fields from Intergovernmental Panel on Climate Change Ar4 Models Using Equidistant Quantile Matching." *Journal of Geophysical Research: Atmospheres (1984–2012)* 115, no. D10 (2010).
- Liou, Earnest Y. "Opset: Program for Computerized Selection of Watershed Parameter Values for the Stanford Watershed Model." (1970).
- Lynch, Peter. *The Emergence of Numerical Weather Prediction: Richardson's Dream*. Cambridge University Press, 2006.
- Maier, Holger R, and Graeme C Dandy. "Neural Networks for the Prediction and Forecasting of Water Resources Variables: A Review of Modelling Issues and Applications." *Environmental modelling & software* 15, no. 1 (2000): 101-24.
- Mailhot, Alain, and Sophie Duchesne. "Design Criteria of Urban Drainage Infrastructures under Climate Change." *Journal of Water Resources Planning and Management* 136, no. 2 (2009): 201-08.
- Marron, JS, and WJ Padgett. "Asymptotically Optimal Bandwidth Selection for Kernel Density Estimators from Randomly Right-Censored Samples." *The Annals of Statistics* 15, no. 4 (1987): 1520-35.
- Martin, E, B Timbal, and E Brun. "Downscaling of General Circulation Model Outputs: Simulation of the Snow Climatology of the French Alps and Sensitivity to Climate Change." *Climate Dynamics* 13, no. 1 (1996): 45-56.
- Massey Jr, Frank J. "The Kolmogorov-Smirnov Test for Goodness of Fit." *Journal of the American statistical Association* 46, no. 253 (1951): 68-78.

- Maurer, Edwin P, Levi Brekke, Tom Pruitt, and Philip B Duffy. "Fine-Resolution Climate Projections Enhance Regional Climate Change Impact Studies." *Eos, Transactions American Geophysical Union* 88, no. 47 (2007): 504-04.
- May, Robert J, Graeme C Dandy, Holger R Maier, and John B Nixon. "Application of Partial Mutual Information Variable Selection to Ann Forecasting of Water Quality in Water Distribution Systems." *Environmental Modelling & Software* 23, no. 10 (2008): 1289-99.
- McCuen, Richard H. *Hydrologic Analysis and Design*. Prentice-Hall Englewood Cliffs, NJ, 1989.
- McCulloch, Warren S, and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *The Bulletin of Mathematical Biophysics* 5, no. 4 (1943): 115-33.
- Mearns, Linda O, William Gutowski, Richard Jones, Ruby Leung, Seth McGinnis, Ana Nunes, and Yun Qian. "A Regional Climate Change Assessment Program for North America." *Eos, Transactions American Geophysical Union* 90, no. 36 (2009): 311.
- Mearns LO, and et al. "The North American Regional Climate Change Assessment Program Dataset." National Center for Atmospheric Research Earth System Grid data portal, Bolder, CO, 2007, updated 2011.
- Miller, Irwin, John E Freund, and Richard Arnold Johnson. *Probability and Statistics for Engineers.* Vol. 4: Prentice-Hall Englewood Cliffs, NJ, 1965.
- Miller, John Frederick. *Two-to Ten-Day Precipitation for Return Periods of 2 to 100 Years in the Contiguous United States.* US Department of Commerce, Weather Bureau, 1964.
- Miller, Kathleen A, and David Noble Yates. *Climate Change and Water Resources: A Primer for Municipal Water Providers*. American Water Works Association, 2006.
- Minns, AW, and MJ Hall. "Artificial Neural Networks as Rainfall-Runoff Models." *Hydrological Sciences Journal* 41, no. 3 (1996): 399-417.
- Mirhosseini, Golbahar, Puneet Srivastava, and Lydia Stefanova. "The Impact of Climate Change on Rainfall Intensity–Duration–Frequency (Idf) Curves in Alabama." *Regional Environmental Change* (2012): 1-9.
- Muller, M. "Climate Change Adaptation and Integrated Water Resource Management." *Global. Water Partnership, policy brief*, no. 5 (2007).
- Nakicenovic, Nebojsa, Joseph Alcamo, Gerald Davis, Bert de Vries, Joergen Fenhann, Stuart Gaffin, Kenneth Gregory, *et al.* "Special Report on Emissions Scenarios: A Special Report of Working Group Iii of the Intergovernmental Panel on Climate Change." Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US), 2000.

- NCDC. "Online Climate Data Directory." NOAA National Climatic Data Center (NCDC 2005).
- New, Mark, and Mike Hulme. "Representing uncertainty in climate change scenarios: a Monte-Carlo approach." *Integrated Assessment 1*, no. 3 (2000): 203-213.
- Nguyen, Tan-Danh, Van-Thanh-Van Nguyen, and Philippe Gachon. "A Spatial-Temporal Downscaling Approach for Construction of Intensity-Duration Curves in Consideration of Gcm-Based Climate Change Scenarios." *Advances in Geosciences, Volume 6: Hydrological Science (HS)* 6 (2006): 11-21.
- Nguyen, V-T-V, T-D Nguyen, and A Cung. "A Statistical Approach to Downscaling of Sub-Daily Extreme Rainfall Processes for Climate-Related Impact Studies in Urban Areas." *Water science and technology: water supply* 7, no. 2 (2007): 183-92.
- NOAA. "The First Climate Model." http://celebrating200years.noaa.gov/.
- Oleson, Keith W, Yongjiu Dai, Gordon Bonan, Mike Bosilovich, Robert Dickinson, Paul Dirmeyer, Forrest Hoffman, *et al.* "Technical Description of the Community Land Model (Clm)." *Tech. Note NCAR/TN-461+ STR* (2004).
- Olsson, J. "Validity and Applicability of a Scale-Independent, Multifractal Relationship for Rainfall." *Atmospheric research* 42, no. 1 (1996): 53-65.
- Onof, C, and Karsten Arnbjerg-Nielsen. "Quantification of Anticipated Future Changes in High Resolution Design Rainfall for Urban Areas." *Atmospheric research* 92, no. 3 (2009): 350-63.
- Onof, Christian, and Howard S Wheater. "Modelling of British Rainfall Using a Random Parameter Bartlett-Lewis Rectangular Pulse Model." *Journal of Hydrology* 149, no. 1 (1993): 67-95.
- Ormsbee, Lindell E. "Rainfall Disaggregation Model for Continuous Hydrologic Modeling." *Journal of Hydraulic Engineering* 115, no. 4 (1989): 507-25.
- Pandey, Ganesh Raj. "Estimation of Short-Duration Rainfall Distribution Using Data Measured at Longer Time Scales." *Water Science & Technology* 29, no. 1-2 (1994): 39-45.
- Parzen, Emanuel. "On Estimation of a Probability Density Function and Mode." *The annals of mathematical statistics* 33, no. 3 (1962): 1065-76.
- Phillips, Norman A. "The General Circulation of the Atmosphere: A Numerical Experiment." *Quarterly Journal of the Royal Meteorological Society* 82, no. 352 (1956): 123-64.
- Piani, C, JO Haerter, and E Coppola. "Statistical Bias Correction for Daily Precipitation in Regional Climate Models over Europe." *Theoretical and Applied Climatology* 99, no. 1-2 (2010): 187-92.
- Polansky, Alan M, and Edsel R Baker. "Multistage Plug—in Bandwidth Selection for Kernel Distribution Function Estimates." *Journal of Statistical Computation and Simulation* 65, no. 1-4 (2000): 63-80.

- Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton. "The Impact of New Physical Parametrizations in the Hadley Centre Climate Model: Hadam3." [In English]. *Climate Dynamics* 16, no. 2-3 (2000/02/01 2000): 123-46.
- Prodanovic, Predrag, and Slobodan P Simonovic. Development of Rainfall Intensity Duration Frequency Curves for the City of London under the Changing Climate.
   Department of Civil and Environmental Engineering, The University of Western Ontario, 2007.
- Puckrin, E, W FJ Evans, J Li, and H Lavoie. "Comparison of Clear-Sky Surface Radiative Fluxes Simulated with Radiative Transfer Models." *Canadian journal* of remote sensing 30, no. 6 (2004): 903-12.
- Qi, Min, and Guoqiang Peter Zhang. "An investigation of model selection criteria for neural network time series forecasting." *European Journal of Operational Research* 132, no. 3 (2001): 666-680.
- Raisanen, Jouni, and T. N. Palmer. "A probability and decision-model analysis of a multimodel ensemble of climate change simulations." *Journal of Climate* 14, no. 15 (2001): 3212-3226.
- Ren, Liqun, and Zhiye Zhao. "An Optimal Neural Network and Concrete Strength Modeling." *Advances in Engineering Software* 33, no. 3 (2002): 117-30.
- Richard, J., MO. Wilfran, and T. Simon. "The North American Regional Climate Change Assessment Program Dataset." National Center for Atmospheric Research Earth System Grid data portal, Bolder, CO, 2007, updated 2011.
- Rodriguez-Iturbe, I, DR Cox, and Valerie Isham. "Some Models for Rainfall Based on Stochastic Point Processes." *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 410, no. 1839 (1987): 269-88.
- Rodriguez-Iturbe, Ignacio, DR Cox, and Valerie Isham. "A Point Process Model for Rainfall: Further Developments." *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 417, no. 1853 (1988): 283-98.
- Rosenblatt, Murray. "Remarks on Some Nonparametric Estimates of a Density Function." *The Annals of Mathematical Statistics* (1956): 832-37.
- Rudemo, Mats. "Empirical Choice of Histograms and Kernel Density Estimators." *Scandinavian Journal of Statistics* (1982): 65-78.
- Sailor, David J, and Xiangshang Li. "A Semiempirical Downscaling Approach for Predicting Regional Temperature Impacts Associated with Climatic Change." *Journal of Climate* 12, no. 1 (1999): 103-14.
- Schertzer, Daniel, and Shaun Lovejoy. "Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Processes." *Journal of Geophysical Research: Atmospheres (1984–2012)* 92, no. D8 (1987): 9693-714.
- Schneider, SH, and S Semenov. "Assessing Key Vulnerabilities and the Risk from Climate Change. Contribution of Working Group Ii to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Ml Parry, of Canziani,

Jp Palutikof, Pj Van Der Linden and Ce Hanson." Cambridge University Press, Cambridge, UK, 2007.

- Scott, David W. *Multivariate Density Estimation: Theory, Practice, and Visualization*. Vol. 383: Wiley, 2009.
- Sebastien, B., C Daniel, and L. René. "The North American Regional Climate Change Assessment Program Dataset." National Center for Atmospheric Research Earth System Grid data portal, Bolder, CO, 2007, updated 2011.
- Sharma, Ashish, David G Tarboton, and Upmanu Lall. "Streamflow Simulation: A Nonparametric Approach." *Water Resources Research* 33, no. 2 (1997): 291-308.
- Sharma, D, A Das Gupta, and MS Babel. "Spatial Disaggregation of Bias-Corrected Gcm Precipitation for Improved Hydrologic Simulation: Ping River Basin, Thailand." *Hydrology and Earth System Sciences Discussions* 11, no. 4 (2007): 1373-90.
- Sheather, Simon. "A Data-Based Algorithm for Choosing the Window Width When Estimating the Density at a Point." *Computational Statistics & Data Analysis* 1 (1983): 229-38.
- Sheather, Simon J. "An Improved Data-Based Algorithm for Choosing the WindowWidth When Estimating the Density at a Point." *Computational Statistics & Data Analysis* 4, no. 1 (1986): 61-65.
- Sheather, Simon J, and Michael C Jones. "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation." *Journal of the Royal Statistical Society*. *Series B (Methodological)* (1991): 683-90.
- Silverman, Bernard W. *Density Estimation for Statistics and Data Analysis*. Vol. 26: Chapman & Hall/CRC, 1986.
- Simonovic, Slobodan P, and Angela Peck. Updated Rainfall Intensity Duration Frequency Curves for the City of London under the Changing Climate. Department of Civil and Environmental Engineering, The University of Western Ontario, 2009.
- Smith, Jason, and Robert N Eli. "Neural-Network Models of Rainfall-Runoff Process." *Journal of water resources planning and management* 121, no. 6 (1995): 499-508.
- Smith, RD, and PR Gent. "Reference Manual for the Parallel Ocean Program (Pop), Ocean Component of the Community Climate System Model (Ccsm2. 0 and 3.0)." Technical Report LA-UR-02-2484, Los Alamos National Laboratory, Los Alamos, NM, http://www. ccsm. ucar. edu/models/ccsm3. 0/pop, 2002.
- Smith, Richard L, Claudia Tebaldi, Doug Nychka, and Linda O Mearns. "Bayesian Modeling of Uncertainty in Ensembles of Climate Models." *Journal of the American Statistical Association* 104, no. 485 (2009).
- Socolofsky, Scott, E Eric Adams, and Dara Entekhabi. "Disaggregation of Daily Rainfall for Continuous Watershed Modeling." *Journal of Hydrologic Engineering* 6, no. 4 (2001): 300-09.

- Solaiman, Tarana A, and Slobodan P Simonovic. "Development of Probability Based Intensity-Duration-Frequency Curves under Climate Change." London, Ontario, Canada, 2011.
- Solaiman, Tarana Aftab. "Uncertainty Estimation of Extreme Precipitations under Climate Change: A Non-Parametric Approach." The University of Western Ontario, 2011.
- Stainforth, David A, Thomas E Downing, Richard Washington, Ana Lopez, and Mark New. "Issues in the Interpretation of Climate Model Ensembles to Inform Decisions." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365, no. 1857 (2007): 2163-77.
- Stewart, Iris T, Daniel R Cayan, and Michael D Dettinger. "Changes in Snowmelt Runoff Timing in Western North America under Abusiness as Usual'climate Change Scenario." *Climatic Change* 62, no. 1-3 (2004): 217-32.
- Svensson, Cecilia, Robin T Clarke, and David A Jones. "An Experimental Comparison of Methods for Estimating Rainfall Intensity-Duration-Frequency Relations from Fragmentary Records." *Journal of Hydrology* 341, no. 1 (2007): 79-89.
- Tebaldi, Claudia, Linda O Mearns, Doug Nychka, and Richard L Smith. "Regional Probabilities of Precipitation Change: A Bayesian Analysis of Multimodel Simulations." *Geophysical Research Letters* 31, no. 24 (2004).
- Tebaldi, Claudia, and Richard L Smith. "Characterizing the Uncertainty of Climate Change Projections Using Hierarchical Models." In *The Handbook of Applied Bayesian Analysis*: Oxford University Press, 2009.
- Tebaldi, Claudia, Richard L Smith, Doug Nychka, and Linda O Mearns. "Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles." *Journal of Climate* 18, no. 10 (2005): 1524-40.
- Veneziano, Daniele, Rafael L Bras, and Jeffrey D Niemann. "Nonlinearity and Self-Similarity of Rainfall in Time and a Stochastic Model." *Journal of Geophysical Research: Atmospheres (1984–2012)* 101, no. D21 (1996): 26371-92.
- Verseghy, Diana L. "Class—a Canadian Land Surface Scheme for Gcms. I. Soil Model." International Journal of Climatology 11, no. 2 (1991): 111-33.
- Verseghy, DL, NA McFarlane, and M Lazare. "Class—a Canadian Land Surface Scheme for Gcms, Ii. Vegetation Model and Coupled Runs." *International Journal of Climatology* 13, no. 4 (1993): 347-70.
- Singh, Vijay P., and Sharad Kumar Jain. Risk and reliability analysis: *a handbook for civil and environmental engineers*. ASCE Publications, 2007.
- Von Storch, H. "Inconsistencies at the Interface of Climate Impact Studies and Global Climate Research." *Meteorologische Zeitschrift* 4, no. 2 (1995): 72-80.

- Von Storch, H, B Hewitson, and L Mearns. "Review of Empirical Downscaling Techniques." *Regional climate development under global warming. General Technical Report* 4 (2000).
- Von Storch, H. "Representation of Conditional Random Distributions as a Problem of "Spatial" Interpolation." In *Geostatistics for Environmental Applications*. Dordrecht, Boston: Kluwer Academic Publishers, 1999.
- Von Storch, Hans. "The Global and Regional Climate System." In *Anthropogenic Climate Change*, 3-36: Springer, 1999.
- Wackernagel, Hans. *Multivariate Geostatistics. An Introduction with Applications. 3rd Edition.* Berlin: Springer, 2003.
- Wang, Bin, and Xiaofeng Wang. "Bandwidth Selection for Weighted Kernel Density Estimation." *arXiv preprint arXiv:0709.1616* (2007).
- Wang, Dingbao, Scott C. Hagen, and Karim Alizad. "Climate Change Impact and Uncertainty Analysis of Extreme Rainfall Events in the Apalachicola River Basin, Florida." *Journal of Hydrology* 480, no. 0 (2/14/2013): 125-35.
- Weichert, Andreas, and Gerd Bürger. "Linear Versus Nonlinear Techniques in Downscaling." *Climate Research* 10 (1998): 83-93.
- Weigend, Andreas S, and Neil A Gershenfeld. "Results of the Time Series Prediction Competition at the Santa Fe Institute." Paper presented at the Neural Networks, 1993., IEEE International Conference on, 1993.
- Widmann, Martin, Christopher S Bretherton, and Eric P Salathé Jr. "Statistical Precipitation Downscaling over the Northwestern United States Using Numerically Simulated Precipitation as a Predictor\*." *Journal of Climate* 16, no. 5 (2003): 799-816.
- Wilby, Robert L, and I Harris. "A Framework for Assessing Uncertainties in Climate Change Impacts: Low-Flow Scenarios for the River Thames, Uk." *Water Resources Research* 42, no. 2 (2006): W02419.
- Wolcott, SB, M Mroz, and J Basile. "Application of Northeast Regional Climate Center Research Results for the Purpose of Evaluating and Updating Intensity-Duration-Frequency (Idf) Curves. Case Study: Rochester, New York." Paper presented at the Proceedings of world environmental and water resources congress, 2009.
- Wood, Andrew W, Lai R Leung, V Sridhar, and DP Lettenmaier. "Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs." *Climatic change* 62, no. 1-3 (2004): 189-216.
- Wright, P., A. DeGeatano, W. Merkel, L. Metcalf, Q. D. Quan, and D. Zarrow.
  "Updating Rainfall Intensity Duration Curves in the Northeast for Runoff Prediction." Paper presented at the ASABE Annual International Meeting, Pittsburgh, Pennsylvania., 2010.

- Zhao, Zhiye, Yun Zhang, and Hongjian Liao. "Design of Ensemble Neural Network Using the Akaike Information Criterion." *Engineering Applications of Artificial Intelligence* 21, no. 8 (2008): 1182-88.
- Zorita, Eduardo, James P Hughes, Dennis P Lettemaier, and Hans von Storch. "Stochastic Characterization of Regional Circulation Patterns for Climate Model Diagnosis and Estimation of Local Precipitation." *Journal of Climate* 8, no. 5 (1995): 1023-42.
- Zucchini, Walter. "Applied Smoothing Techniques, Part 1 Kernel Density Estimation." (2003).
## Appendix A

## Monthly Cumulative Distribution Functions (CDF) To Perform Bias Correction

# A.1. HRM3-GFDL





## A.2.RCM3-GFDL





## A.3. ECP2-GFDL





# A.4.CRCM-CGCM3





# A.5. CRCM-CCSM





# Appendix B

# **Rainfall Intensity-Duration-Frequency Maps**

# B.1. HRM3-HadCM3















### **B.2. HRM3-GFDL**















### **B.3. RCM3-GFDL**















### **B.4. ECP2-GFDL**














### **B.5. CRCM-CGCM3**















### **B.6. CRCM-CCSM**















# Appendix C

## Scatter plots of observed and modeled rainfall depths using the ANN model and the stochastic method

### C.1. 15-min, Jan-Dec

















C.3. 45-min, Jan-Dec
























# Appendix D

## IDF Maps, using ANN model for rainfall disaggregation

## D.1. HRM3-HadCM3









#### D.2. HRM3-GFDL









#### D.3. RCM3-GFDL









## D.4. CRCM-CGCM3









### **D.5. CRCM-CCSM**









# Appendix E Probability-based IDF curves for Alabama

#### E1. Stochastic Method

#### 1. Athens





### 2. Bridgeport





## 3. Boaz





#### 4. Hanceville







## 5. Addison





## 6. Haleyville





## 7. Hamilton





## 8. Vernon




# 9. Berry





#### 10. Warrior





#### 11. Jacksonville





# 12. Birmingham





# 13. Ashland





#### 14. Tuscaloosa





# 15. Thorsby





### 16. Dadeville





#### 17. Marion



258





### 18. Warrior L&D





### 19. Alberta





# 20. Montgomery





# 21. Midway





22. Troy





### 23. Greenville





# 24. Peterman





#### 25. Abbeville





# 26. Enterprise




#### 27. Dothan





#### **28.** Atmore 1





### **29. Atmore 2**







### **30. River Falls**





#### 31. Andalusia





### 32. Thomasville





#### 33. Jackson





#### E.2. ANN model

#### 1. Athens





#### 2. Bridgeport



# 3. Boaz



# 4. Hanceville



Rainfall depth (mm)

# 5. Addison



# 6. Haleyville





# 7. Hamilton



20

40

Rainfall depth (mm)

0.0

0

100-yr

80

60

# 8. Vernon



# 9. Berry



#### 10. Warrior



# 11. Jacksonville



# 12. Birmingham



Rainfall depth (mm)

# 13. Ashland



### 14. Tuscaloosa



# 15. Thorsby



#### 16. Dadeville



Rainfall depth (mm)

# 17. Marion



#### 18. Warrior L&D



40

60

Rainfall depth (mm)

80

0.2

0.0 0

20

10-yr 25-yr

50-yr 100-yr

100

120

# 19. Alberta



#### 20. Montgomery



#### 21. Midway



22. Troy


# 23. Greenville



314

Rainfall depth (mm)

# 24. Peterman



315

Rainfall depth (mm)

#### 25. Abbeville



#### 26. Enterprise



Rainfall depth (mm)

# 27. Dothan



# 28. Atmore 1





# **29. Atmore 2**



#### **30. River Falls**



#### 31. Andalusia



Rainfall depth (mm)

# 32. Thomasville



# 33. Jackson





#### **IDF curves for Alabama**



#### F.1. Stochastic Method











































#### F.2. ANN model

















































Duration (hr)







