Quantifying the Impacts of Future Uncertainties on the Apalachicola-Chattahoochee-Flint Basin

Katherine E. Lownsbery
QUANTIFYING THE IMPACTS OF FUTURE UNCERTAINTIES ON THE APALACHICOLA-CHATTAAHOOCHEE-FLINT BASIN

A Masters Project Presented

by

Katherine Elizabeth Lownsbery

Submitted to the
Department of Civil and Environmental Engineering
University of Massachusetts – Amherst

in partial fulfillment of the requirements for the degree of
Master of Science in Civil Engineering

April 2014
QUANTIFYING THE IMPACTS OF FUTURE UNCERTAINTIES ON THE APALACHICOLA-CHATTahoochee-FLINT BASIN

A Masters Project Presented

by

KATHERINE ELIZABETH LOWNSBERY

Approved as to style and content by:

[Signatures]

Dr. Casey Brown, Chairperson

Dr. Richard N. Palmer, Member

Sanjay Arwade

Civil and Environmental Engineering Department
1 Contents
1 Contents ........................................................................................................... 1
2 Acknowledgements .......................................................................................... 2
3 Abstract ........................................................................................................... 3
4 Introduction ...................................................................................................... 4
5 Background ...................................................................................................... 10
6 Methodology ................................................................................................... 14
6.1 Water Resources Systems Model .................................................................. 16
6.2 System Stress Test ......................................................................................... 17
  6.2.1 Stress Test Data ....................................................................................... 18
  6.2.2 Climate Scenarios .................................................................................... 18
    6.2.2.1 Stochastic Climate Model ................................................................. 19
    6.2.2.2 Hydrologic Models ........................................................................... 20
  6.2.3 Demand Scenarios .................................................................................... 21
6.3 Climate Projections ...................................................................................... 23
6.4 Demand Projections ..................................................................................... 24
6.5 Basin Vulnerability Metrics .......................................................................... 27
7 System Response Results ................................................................................ 30
  7.1 System Marginal Sensitivity ...................................................................... 30
  7.2 System Response Surfaces ........................................................................ 34
  7.3 Likelihood of Future System Performance ................................................. 39
  7.4 Spatial Variability of System Impacts ......................................................... 45
8 System Uncertainty Results ............................................................................ 49
  8.1 ANOVA Procedure ..................................................................................... 53
  8.2 Basin-wide ANOVA Results ....................................................................... 56
  8.3 Spatial Variation in ANOVA Results ........................................................... 58
9 Discussion and Conclusion ............................................................................. 64
10 References ...................................................................................................... 68
Appendix A. Likelihood of Future System Performance ...................................... 76
Appendix B. Testing ANOVA Assumptions ....................................................... 78
Appendix C. Spatial Variation in ANOVA Results .............................................. 80
2 Acknowledgements
I want to gratefully acknowledge funding support from the National Science Foundation (NSF) under grant NSF DGE-0907995 and from the Strategic Environmental Research and Development Program issued through the Department of Defense (DoD). The views expressed in this work are mine and do not necessarily reflect the views of the NSF or the DoD.

I want to thank my advisor and committee chairperson, Dr. Casey Brown, for his valuable guidance to this work. I really appreciate his continual encouragement of my academic growth. I also want to thank Dr. Richard N. Palmer, my committee member, for allowing me to use the basin systems model he originally developed and for his valuable contributions and insights to this work.

I want to thank and recognize Scott Steinschneider for always answering questions and for developing the stochastic weather generator and the hydrologic models used in this work.

And finally, I want to thank everyone who has shaped me and helped me reach this point: my friends, my former teachers and professors, my family – especially my grandfather, Dr. Daniel J. Collins, and my heavenly Father. Your commitment to me has made this work possible, and I am deeply grateful.
3 Abstract

Water resources systems are increasingly stressed by both climate and changing water demands. The uncertainty associated with these stressors greatly complicates risk assessments, especially because the relative and combined impacts of each stressor on water resources systems is often unknown. This study applies a bottom-up *ex post* scenario analysis approach (termed decision-scaling) to explore the spatiotemporal distribution of impacts from multiple stressors on a multi-objective transboundary river basin – the Apalachicola-Chattahoochee-Flint basin. The response of this large water resources system to variability and change in climate (specifically precipitation and temperature), as well as change in water demand (specifically municipal, industrial, and agriculture water demand), is examined using a novel “stress test” approach that simultaneously explores the relative and joint impacts of each stressor. The resulting system response is used to frame available projections of climate and water demand change in terms of risk to system performance. Additionally, an analysis of variance was conducted on the resulting *ex post* scenarios to attribute uncertainty in system response to uncertainty in projections of the various stressors. The findings of this analysis indicate that projected changes in mean precipitation are the dominant source of projected stress in the basin, while the impacts from temperature and demand vary in importance spatially. Early in the 55 year planning period, internal climate variability is the greatest source of uncertainty in system response, while by the end of the planning period uncertainty in projections of trends in precipitation dominates. Uncertainty in projections for other stressors also contributes to uncertainty in system response depending on spatial location. This study demonstrates the application of the decision-scaling methodology to a large, complicated, multi-objective basin with multiple stressors and yields important insights into water resources risk assessments for planning.
4 Introduction
Water resources around the world are increasingly stressed by variability and long-term change in the climate system as well as non-climate factors that result in growing competition for water resources, transboundary water disputes, and increasing environmental flow requirements (Frederick & Major, 1997; Lins & Stakhiv, 1998; UNESCO et al., 2012). For regional and local water resources systems, quantifying the relative impacts of these stressors is essential to developing targeted water management and planning strategies that can effectively mitigate potential risks. However, risk analyses are severely complicated by the uncertainty inherent in projections of future stressor states. This study applies the decision-scaling methodology (Brown et al., 2012) to address these issues in an analysis of the Apalachicola-Chattahoochee-Flint (ACF) basin in the southeastern U.S. The ACF basin is a large, complicated, transboundary basin that exhibits the traits of a stressed water resources system. The essential contribution of this study is an extension of decision-scaling that explores the spatiotemporal distribution and relative importance of impacts from climate and demand stressors on stakeholder-centric concerns in a large, multi-objective, transboundary river basin.

The common approach to quantifying the impacts of climate change on water resources systems is to propagate a set of climate change projections (e.g. spatially downscaled and bias corrected general circulation model (GCM) projections) through hydrology and water resources systems models and then to evaluate the performance of the system (some examples include Alcamo et al., 2007; Arnell, 1999; Arnell, 2004; Lettenmaier et al., 1999; Sun et al., 2008; Vörösmarty et al., 2000). However, there is increasing recognition of the shortcomings inherent in projection-led approaches to climate change risk assessment. Most prominently, the climate change projections do not fully delimit the future climate uncertainties that affect a water resources system. One reason for this shortcoming is that ensembles of climate projections represent only a minimal portion of the range of uncertainty due to: 1) climate model inadequacy and errors in representing the real climate system, 2) the uncertainty of model parameter
values and initial conditions, and 3) the uncertain effects of bias correction and downscaling (Hall, 2007; Stainforth et al., 2007a; Stainforth et al., 2007b). As a result, ensembles of projections form an inadequate sample of climate change, limiting the range of potential performance of the water resources system that can be simulated and understood. Another reason for this shortcoming is that water resources systems are particularly sensitive to internal climate variability in addition to changes in mean climate; yet climate projections provide limited and often biased exploration of the effects of internal climate variability, especially in regards to precipitation (Rocheta et al., 2013). To account for internal climate variability to some extent, an analysis can use multiple simulations from a GCM that are perturbed with different initial conditions (Stainforth et al., 2007b); however, the resulting realizations are still limited and likely still exhibit biases. An additional shortcoming of projections-led approaches is the ambiguity of probabilistic climate information from GCMs (Hall, 2007). Because risk is a function of both probability and impact (Dessai & Hulme, 2004), ambiguous probabilities necessarily imply that the resulting risk assessment is also ambiguous. Other shortcomings include the increasing uncertainty that propagates through each step of the modeling chain (Wilby & Dessai, 2010), and the difficulty of generating and conveying useful information to decision-makers (Brown et al., 2012; Hall, 2007; Kundzewicz & Stakhiv, 2010).

In response to these challenges, alternative approaches to climate risk assessment have been proposed. These approaches typically fall under a general category of “bottom-up” analysis, which identifies system vulnerabilities for the purpose of selecting adaptation strategies that are robust over a wide range of plausible future conditions. Some specific decision making under uncertainty methods include robust decision making (RDM) (Lempert et al., 2006) and info-gap analysis (Ben-Haim, 2006; Korteling et al., 2013). RDM tests vulnerabilities over a wide range of GCM climate scenarios, system parameters, and potential decisions. RDM subsequently uses a sophisticated clustering algorithm to identify key problematic scenarios and corresponding robust adaptations. Info-gap analysis, a methodology applied in a variety of engineering disciplines, tests systems over a wide range of key system parameters. Based on
the tests, info-gap analysis defines “robustness” and “opportuness” measures for each decision, based on
the uncertainty surrounding the best estimate of the future states of the key system parameters. The
bottom-up approach improves upon the common approach to risk analysis by focusing on system
vulnerabilities; however, some bottom-up studies use climate projections to drive the climate analysis
(e.g. Jones, 2001; Lempert et al., 2006; Groves & Lempert, 2007), and are thus still limited by the
shortcomings of climate projections from GCMs as previously discussed.

More recently, analysts have begun to use methods that are bottom-up, but that explore climate change
risk without being led by GCM projections. Prominent methodologies include scenario-neutral planning
(Prudhomme et al., 2010; Wilby & Dessai, 2010) and decision-scaling (Brown et al., 2012). Scenario-
neutral planning is very similar to decision-scaling, but its application has currently been limited to
hydrologic models. Decision-scaling, the methodology applied in this study, emphasizes the
characterization of climate variability that influences system performance and the interpretation of
available climate information such as GCM projections, paleodata, expert opinion, and historic trends.
This approach identifies the vulnerabilities of a water resources system by subjecting the system to a wide
range of mean changes in key system stressors (in this study not only climate change but also changing
water demand). Termed a “stress test”, this essential analysis step is unique to decision-scaling in
comparison to other bottom-up approaches. The range of the stress test extends beyond available
projections of system stressors to ensure that critical system vulnerabilities are identified. In addition to a
range of mean changes in key system stressors, the stress test also samples multiple realizations of
variability for each mean change in climate stressors. The stress test results are then used to define
scenarios based on the system performance, following the \textit{ex post} scenario analysis that is recommended
for cases of deep uncertainty (Herrick & Sarewitz, 2000; Groves & Lempert, 2007), which is defined as
situations where the models to describe the system, the probability distributions on key parameters, and/or
the desirability of outcomes is unknown or highly contested (Lempert et al., 2006). Because scenarios are
based on the stress test results, information on climate and demand changes (such as GCM projections and water demand forecasts) can be categorized based on the scenarios and used to infer the projections-based likelihoods of the scenarios. The decision-scaling methodology has been used in a limited number of studies to investigate the impact of streamflow changes on investment plans for the Niger River basin in West Africa (Brown, 2010), the impact of climate change on regulation plans for the Upper Great Lakes (Brown et al., 2011; Moody & Brown, 2012), the impact of net basin supply changes on a municipal surface water system (Brown et al., 2012), and the impact of climate change on floodplain vulnerability and robust adaptation plans at the Coralville Reservoir in the state of Iowa (Brown, 2013).

Previous applications of decision-scaling have only been used in water planning studies that consider the effects of influential yet uncertain hydroclimatic conditions. While climate impacts on water resources are of enormous interest to many in government entities and in the research community (Kundzewicz et al., 2007), the influence of non-climate factors is significant and potentially even greater than that of climate variability and change (Frederick & Major, 1997; Lins & Stakhiv, 1998; Vörösmarty et al., 2000). For example, Vörösmarty et al. (2000) used population and climate change projections to indicate that growth in population, as compared to climate change, is the more significant source of future water stress. Already, many water resources systems are stressed by resource over-allocation, unstable political conflicts, and economic growth (UNESCO et al., 2012). Analogous to climate change projections, where GCM ensembles represent only the “lower bound on the maximum range of uncertainty” (Stainforth et al., 2007a), projections of non-climate factors are also highly uncertain. Water use scenarios are “notoriously difficult to make” (Arnell, 1999), and human population growth, one driver of water demand, is subject to a myriad of volatile and poorly understood factors (Cohen, 2003). In this study, decision-scaling is extended to examine not only the impacts of climate variability and change, but also the impacts of growth in water demand.
Broad global studies have been performed to investigate the relative impacts of climate change and non-climate factors (often population growth that results in changing water demand) on water resources systems (e.g. Alcamo et al., 2007; Arnell, 1999; Arnell, 2004; Vörösmarty et al., 2000). These global scale studies reveal important long-term and wide-reaching impacts but are also limited due to the large scale of the investigation. Some of the limitations include generalized water resources models that lack regional infrastructure, use generic performance indicators, and use aggregated climate data that lacks high-frequency climate variability. Finer resolution studies at the country, regional, and basin-wide level range from general models and metrics to detailed systems models (e.g. Sun et al., 2008; Lettenmaier et al., 1999) and may or may not examine sub-basin impacts. Previous studies using the decision-scaling methodology have similarly been limited in their consideration of spatially disaggregated measures of system performance, usually examining only one or two performance measures for a single location of interest (with the exception of two locations in the Upper Great Lakes studies). This study extends decision-scaling to a complex water resources system with multiple objectives and control points that encompass a wide geographic area.

In addition to investigating impacts over a wide geographic area, this study extends decision-scaling to investigate the temporal evolution of system impacts from a wide array of transient stressors. On the time scale of a typical planning horizon for a water resources system, often a few decades, many water managers consider climate change to be less important and more uncertain than non-climate factors (Lins & Stakhiv, 1998). Additionally, the most significant source of climate uncertainty, such as internal climate variability, climate scenarios, or model uncertainty, changes through time (Hawkins & Sutton, 2009; Yip et al., 2011). Several decision-scaling studies (Brown, 2010; Brown et al., 2012; Brown et al., 2011) have briefly addressed the transient nature of climate impacts on water resources systems, but are
limited to either a discussion of adaptive management strategies or a summary of changes in system performance for every 20 or 25 years. This study investigates the spatiotemporal variation of uncertainty in system response and how uncertainty in projections contributes to uncertainty in system response.

This study extends the decision-scaling methodology to examine both the relative and combined impacts of climate and non-climate factors on a multi-objective, transboundary, water resources system. Two major analyses are conducted in this study. First, decision-scaling is applied to a complicated river basin for the purpose of analyzing the relative impacts of various stressors on stakeholder-centric metrics, and how those impacts vary spatially throughout the basin. Second, the decision-scaling results are used to analyze the spatiotemporal variation in uncertainty of system response and to attribute that uncertainty to the uncertainty associated with the range of projections for various stressors. As a case study, this study uses the ACF basin, a complex, large-scale, and highly contested water resource. The remainder of this paper develops these analyses and case study in five sections. The first section provides a general background on the ACF basin, a short history of the litigation over its water, and a summary of previous ACF study findings. The second section describes the study methodology, including a discussion of the models and metrics used in the analysis. The third section presents the results of the system response, while the fourth section presents uncertainty in system response due to uncertainty in stressor projections using an analysis of variance procedure. Finally, the fifth section contains the discussion and conclusion.
5 Background
The ACF basin drains 50,764 km² of Alabama, Georgia, and Florida (Figure 1) and is located in the semi-humid southeastern U.S., a region of high population growth (Sun et al., 2008; GA State, 2010). The Chattahoochee River forms the western upstream segment of the basin and is highly regulated. The five major dams along the Chattahoochee River are managed by the U.S. Army Corps of Engineers (USACE). Three of the dams form large reservoirs with significant storage (Lake Lanier with 3.15 x 10⁹ m³, West Point Lake with 9.56 x 10⁸ m³, and Walter F. George Lake with 1.15 x 10⁹ m³), while the other two operate as run-of-river projects with a limited amount of usable storage (George Andrews Lake with 2.24 x 10⁷ m³ and Lake Seminole with 4.91 x 10⁸ m³) (U.S. Army Corps of Engineers, 2012). The storage of Lake Lanier is significantly larger than that of the other reservoirs, despite its location in the headwaters of the basin. The Flint River, which forms the eastern upstream segment of the basin, is largely unregulated and joins the Chattahoochee at Lake Seminole, forming the Apalachicola River. The Apalachicola, surrounded by land that is predominately designated as conservation status, is a richly biodiverse area and contributes 35% of the freshwater that enters the Gulf of Mexico (Ruhl, 2005).

There are many competing stakeholder interests within the basin, including growing water demands for the metropolitan areas of Atlanta and Columbus in Georgia, a variety of water uses in Alabama, substantial agriculture water use in the Flint basin, an oyster and fishing industry in the Apalachicola Bay, and environmental concerns about habitat for commercial fish nurseries and other threatened or endangered species (Carter et al., 2008). The major reservoirs are operated to meet these competing water needs in addition to management for flood damage reduction, peaking hydropower generation, navigation, and recreation throughout the basin (U.S. Army Corps of Engineers, 2012). The Atlanta metropolitan area, located near the headwaters of the Chattahoochee River just below Lake Lanier, is growing rapidly (GA State, 2010) and represents 75% of Georgia’s economic activity (Carter et al., 2008). Metropolitan Atlanta receives nearly all of its water from surface sources, 73% of which is from
the Chattahoochee basin including Lake Lanier. Of the water withdrawn from the Chattahoochee basin, approximately 33% of it is consumed, and the remainder is returned to the system. In total, the Atlanta metropolitan region has an average demand of 600 mgd for municipal, industrial, and thermoelectric cooling uses (AECOM et al., 2009).

The long and complicated water controversy in the ACF basin, termed the “Tri-State Water Wars,” has led to a number of studies on its evolution (e.g. Abrams, 2008; DuMars & Seeley, 2004; Feldman, 2008; Grant, 2004; Ruhl, 2005; Stephenson, 2000). In brief, Atlanta requested that the USACE reallocate storage in Lake Lanier to water supply in the late 1980’s (DuMars & Seeley, 2004), prompting subsequent lawsuits by Alabama, Florida, and other interested stakeholders. In the early 1990’s, the litigation was suspended pending the creation of an interstate compact based on a comprehensive planning study (Palmer et al., 2013). However, the incomplete compact expired in 2003 and the conflict returned to the courts. In 2011, the U.S. Court of Appeals gave the USACE the legal right to decide water supply allocations within predetermined limits (Robbins, 2012; U.S. Court of Appeals, 2011). The ruling was challenged to the U.S. Supreme Court, but was declined a hearing in 2012, effectively ending the litigation. However, the underlying issue of a mutually agreed upon and adaptive basin-wide management plan still remains unresolved (Atkins, 2013; Jackson, 2013; Munson, 2013; Turner, 2013).

National studies of water resources vulnerability to climate change, which treat the ACF as a single hydrologic unit, show low to moderate levels of risk for the basin, especially compared to regions in the West. These studies have investigated water supply, distribution, and consumptive use (Hurd et al., 1999), water quality (Cruise et al., 1999; Hurd et al., 1999), and drought risk (Strzepek et al., 2010) under current and future climate scenarios. A regional study of the southeastern U.S. at the hydrologic unit code 8 level by Sun et al. (2008) performed a projections-led analysis to explore the relative impacts of climate,
population, and land use scenarios on an annual water supply index. The study found that the greatest and most wide-spread impacts in the southeastern U.S. are from climate change, followed by population growth in local regions (such as in the Atlanta region), followed by minor impacts from land use. The primary limitations of the study are: 1) the projections-led analysis that relies on two GCM based climate scenarios and 2) an annual definition of water supply that is limited to each hydrologic unit and does not account for current water infrastructure.

Multiple studies have investigated the ACF basin or sub-portions of the basin. Several studies have created models of the ACF to determine reservoir critical yield or best reservoir management operations under current climate (McMahon & Stevens, 1995; U.S. Army Corps of Engineers, 2010; U.S. Army Corps of Engineers, 2012). Hurd et al. (2004) use a seasonal economic optimization model of water resources in the ACF basin (as well as three other U.S. river basins) to assess the impact of climate change. The climate change scenarios, created by combinations of changes to historic temperature and precipitation mean values, indicate that for the ACF, higher temperatures lead to increased agricultural welfare while greater precipitation leads to economic losses from flooding. The primary limitations of this study are that: 1) the only stressor is climate change, and 2) the simplified systems model is on a seasonal time step and incorporates assumptions of competitive water markets and perfect foresight. Jenicek et al. (2011) examine the water sustainability of Fort Benning, which is centrally located in the basin on the Chattahoochee River, as part of a nation-wide assessment of ten army installations. The study examines the impacts of five scenarios that represent various levels of water demand and drought occurrence. The study concludes that the region is relatively water rich, with risks for short-term periods of water scarcity that are caused by droughts rather than changes in water demand. The primary limitations of this study are that: 1) climate change is only considered as reoccurring droughts, and 2) the regional scope lacks significant impacts of the upstream water demands at Atlanta and reservoir operations at Lake Lanier. The first comprehensive study of the ACF basin was performed by Lettenmaier et al. (1999) as part of a larger
study on several key U.S. river basins. The study investigates the impact of changes in climate, water demand, and operations on total system performance in the ACF, using a systems model developed by Palmer et al. (1995). The study found that “there is relatively little consistency among climate scenarios as to the implications of climate change”. The study also finds that, over a variety of system metrics, climate change has impacts of similar magnitude as those of operational changes and water demand growth. The primary limitation of the study is the climate projection-led, scenario-based approach to water resources systems risk assessment. This study expands upon the work of Lettenmaier et al. (1999) by using the decision-scaling methodology to investigate similar issues, specifically, the impacts of uncertain regional climate and changing water demand.
6 Methodology

The decision-scaling methodology has been described in previous work (Brown et al., 2012) and is briefly summarized here. The approach enables the identification and consideration of water resources risk in situations where future climate and other factors are highly uncertain. The initial step in decision-scaling is the identification of system objectives, performance metrics, and the specification of the models used to represent the system. Subsequently, a stress test is performed that identifies the climate or societal conditions that lead to unacceptable system performance. The framework then determines the relative likelihood of those problematic futures using various sources of evidence such as climate projections, historic climate, paleodata, and population projections. The combination of these two elements leads to estimates of system-specific risk that is useful for planning purposes. However, because the stress test is separated from the analysis of possible future conditions, the approach ensures that the system is stressed over a sufficiently wide range of possible futures to identify important system vulnerabilities. The approach also ensures that future scenarios are not limited to climate model projections.

In this study, the stress test considers both climate and water demand stressors, including the variability and change of precipitation and temperature as well as increased municipal and industrial (M&I) and agriculture water demand due to population growth and economic development. It is assumed that the climate and demand stressors are independent of each other, though physical connections do exist; for example, there are connections between agriculture water demand and temperature and precipitation. Furthermore, it is outside the scope of this study to model operational changes in the system that could mitigate some of the impacts of changes in the stressors. Metrics of system performance are based on reported stakeholder priorities and are evaluated at a variety of locations in the basin over a 55 year planning horizon. Climate and water demand projections inform the likelihood of future system performance. The following sections describe each component of the stress test in detail, which is summarized in Figure 1.
Figure 1: Summary of the components of the stress test and the study methodology. Blue triangles indicate reservoirs with substantial storage and numbers indicate water withdrawal sites.
6.1 Water Resources Systems Model

The water resources systems model is a modified version of the model developed by Palmer et al. (1995) for the comprehensive planning study as a part of the interstate compact process. This model was also used by Lettenmaier et al. (1999). The model, built in STELLA, runs on a monthly time step and was originally calibrated to the 1939 – 1993 time period. For this study, the model begins simulation in 2005, and runs for 55 years. The model is driven by monthly inflows at 22 locations in the basin and by M&I, agriculture, and thermoelectric cooling water demands. As a “control” case, against which to compare other tests of the system, the model was run using M&I and agriculture water demands from 2005 and using the last 55 years of historic climate (1956 – 2010) from 1/8" gridded, daily observed meteorological data from Jan. 1, 1949 to Dec. 31, 2010 (Maurer et al., 2002). The last 55 years of historic climate were chosen because they are the most recent and also contain the drought of record for the system. The control case represents the performance of the system under historic climate and current demand conditions.

Several modifications to the original model were made. The monthly inflows were originally an unimpaired historic database developed as part of the Alabama-Coosa-Tallapoosa (ACT) and ACF River Basins Comprehensive Water Resources Study (U.S. Army Corps of Engineers, 1997), but were updated with hydrologic model simulations (discussed later). Water demands were updated from 1990 to 2005 levels based on county level demand data in the ACF from Marella & Fanning (2011). In that time period, despite increases in M&I and agriculture demand, overall water demand in the basin decreased by nearly 110 mgd due to dramatic decreases in thermoelectric cooling demand. Fixed reservoir evaporation rates were replaced with derived values conditional on future climate scenarios using the Hargreaves method (Hargreaves & Samani, 1982), where monthly extraterrestrial solar radiation was estimated using the method given in Allen et al. (1998). Surface water-groundwater interactions were also updated for certain demand sites in the basin. The originally modeled interactions were retained at demand sites 1-4 (Figure 1) where groundwater is generally unconfined and exhibits high connectivity (Clarke & Pierce, 1985), so
the modeled groundwater is connected to surface water in a 1-to-1 relationship. The originally modeled interactions were also retained at demand sites 5 and 6, where the groundwater is in confined aquifers (Clarke & Pierce, 1985), so there is no modeled connection to surface water. However, at demand sites 7-11, the original interactions were modeled as prescribed monthly volumes of water regardless of water demand. To reflect changing water demand, and based on the connectivity characteristics of the Floridian aquifer (Torak & McDowell, 1996), the originally modeled connections at sites 7-11 were changed such that groundwater is connected to surface water in a 1-to-0.6 ratio. It should be noted that these interactions are only approximate; for example, Environmental Protection Division (1997a) state that groundwater is equivalent to surface water for the portion of the basin that includes demand sites 1-6 and Environmental Protection Division (1997b) state that in the Lower Flint basin (probably including demand sites 7-9) 10 gallons of groundwater are equivalent to 3 gallons of surface water.

6.2 System Stress Test
The system stress test is designed to expose the vulnerabilities of the system by subjecting it to a wide range of changes for each stressor and combination of stressors. The stressors examined in this study are climate variability and change and changing water demand. Climate variability is represented by 25 stochastic realizations of historical climate for the variables of interest, precipitation and temperature, created by a weather generator. Climate change is represented by transient linear changes to the mean of the stochastic realizations, with 5 precipitation changes and 6 temperature changes over the simulation period of 55 years. Each transient change is denoted by the total mean change achieved by the end of the 55 year period; for example, if the transient change results in 200% of historic precipitation by the end of the 55 year period, that change is termed 200%. The modified climate time series are then run through hydrologic models to generate streamflow, which is the input to the systems model. Water demand changes are represented by transient linear increases in M&I demand (6 changes) and in agriculture demand (3 changes), which are direct inputs into the systems model. The full stress test is every combination of both climate and water demand changes, leading to a total of 13,500 model runs. The
“baseline” is defined as model runs where the stressor levels are not changed from the mean of the historic levels. Hence, there are 25 baseline model runs, because there are 25 stochastic realizations of climate. The baseline cases have the same mean stressor values as the control case (discussed previously), but the control model run is based on historic climate, rather than on stochastic realizations of historic climate. Each component of the stress test, including the data, the climate scenarios (created by a weather generator and hydrologic models), and the demand scenarios, are discussed in detail in the following subsections.

6.2.1 Stress Test Data
Data for the stress test was gathered from a variety of sources. Historic climate data, used in the stochastic weather generator and to calibrate hydrologic models, was gathered from historic 1/8° gridded, daily observed meteorological data from Jan. 1, 1949 to Dec. 31, 2010 (Maurer et al., 2002). This historic climate data was aggregated to monthly values and to the catchment for each inflow point in the basin. Unimpaired streamflow data, used to calibrate the hydrologic models, was gathered from a database developed as part of the Alabama-Coosa-Tallapoosa (ACT) and ACF River Basins Comprehensive Water Resources Study (U.S. Army Corps of Engineers, 1997). Flows in the database were reconstructed to remove the effects of anthropogenic influences on the hydrologic data, such as municipal, industrial, and agriculture withdrawals and return flows, as well as evaporation, precipitation, and leakage associated with man-made reservoirs. This unimpaired streamflow data was also the original input to the basin systems model.

6.2.2 Climate Scenarios
Climate scenarios for the stress test are developed by generating stochastic climate time series, which form the inputs of hydrologic models. The outputs of the hydrologic models are streamflow time series,
which form the inputs to the water systems model. The following descriptions for both the stochastic climate model and the hydrologic models are credited to Scott Steinschneider.

6.2.2.1 Stochastic Climate Model
The stochastic weather generator is based on historic climate data and creates monthly time series of transient sequences of spatially disaggregated precipitation and temperature for the study region with prescribed linear mean changes. This weather generator is loosely based on the weather generator described in Steinschneider & Brown (2013) but is adjusted to create monthly time series. The regional climate exhibits a dual peak in precipitation; a summer peak between May and October (M-O), and a winter peak between November and April (N-A). The generator simulates monthly climate variables from these two seasons separately and then combines them to form a continuous time series of climate data. Precipitation and temperature during the N-A season exhibit a moderate correlation with the wintertime (December - February) El Nino – Southern Oscillation (ENSO). During the M-O season, precipitation and temperature are correlated with summertime (July - August) sea surface temperature anomalies (SSTAs) in the North Atlantic Ocean. These large-scale climate drivers exhibit quasi-oscillatory low-frequency structure that propagates into the local climate of the ACF region.

To reproduce this structure in the stochastic weather generator, a wavelet auto-regressive modeling (WARM) approach is adopted (Kwon et al., 2007). The WARM approach extracts low frequency signals in the ENSO and SSTA data using wavelet decomposition and then stochastically simulates each signal using autoregressive time series models. A series of regressions between ENSO (North Atlantic SSTAs) and wintertime (summertime) average precipitation and temperature is used to translate the simulated large-scale climate drivers into seasonal climate across all locations in the ACF. Random samples from a multivariate-normal distribution fit to the residuals of each regression for all sites are then added to the regression predictions. All seasonal climate variables are then disaggregated to a monthly time step using the method of fragments (Srikanthan & McMahon, 2001).
To impose transient linear mean changes, linear trends (multiplicative for precipitation and additive for temperature) are used to adjust the 25 simulated time series from the stochastic weather generator. The ranges of the temperature and precipitation changes are chosen to completely encompass GCM projections. By the end of the planning period, precipitation changes range from 60% to 140% of the historic mean value in increments of 20%, and temperature increases range from 0 °C to 5 °C greater than the historic mean in increments of 1 °C (where in this case historic refers to the 1949 to 2010 time period). A total of 25 stochastic climate trials with each combination of 5 precipitation changes and 6 temperature changes results in 750 different climate simulations. Changes in climate variability outside of the observed statistics of historic variability are not explored.

6.2.2.2 Hydrologic Models
To generate inflows for the STELLA model, each climate simulation is run through calibrated $abcd$ hydrologic models for 22 different inflow points. The $abcd$ model (Thomas, 1981) was chosen for its parsimony and because it has been shown to have a suitable structure for the ACF region, exhibiting good or acceptable performance over both calibration and evaluation time periods (Martinez & Gupta, 2010). The $abcd$ model, including the four $abcd$ parameters and the initial groundwater and soil moisture level, is calibrated to 20 of the 22 inflow locations using historic climate and unimpaired monthly streamflow. For the remaining two locations, the recorded “natural” inflow exhibits highly unnatural behavior (e.g. many negative values, altered seasonal cycle), so calibration based on historic climate data is deemed impractical. Instead, the method of spatial proximity is used to transfer $abcd$ parameters to these two locations from nearby gaged sites (Steinschneider et al., 2014). Calibration is conducted over the period of January 1950 to December 1980 using a shuffled complex evolution optimization algorithm to minimize the root mean squared error between simulated and reconstructed observed flows (Duan et al., 1992). The period from January 1981 to December 1993 is used for model validation. For the evaluation period, the model mean bias is +2.5%. Two outliers have biases of +24% and -20%, but the remaining 20
models have biases within +8.5% and -2% of the observed flows. Furthermore, 15 of the models have Nash-Sutcliffe Efficiency values over 0.6. While 7 models have NSE values between 0.1 and 0.6, which suggests poor performance, this can be largely attributed to errors in the reconstructed “observed” flows used for calibration, many of which have negative values throughout the record.

6.2.3 Demand Scenarios
Demand scenarios for the stress test are developed by adjusting the baseline 2005 M&I and agriculture water demands (Figure 2) by transient linear multiplicative factors. By the end of the planning period, a prescribed level of demand is reached; M&I demand changes range from 100% to 300% of the baseline levels in increments of 40%, and agriculture demand changes range from 100% to 300% of the baseline levels in increments of 100%. The range of M&I demand changes is chosen to completely encompass water demand projections based on population growth in the Metropolitan North Georgia Water Planning District (MNGWPD). For non-MNGWPD municipalities, this is a conservative estimate of demand growth, but the quantity of M&I water demand is also much less than for the MNGWPD. The agriculture water demand range is chosen so that the one projection for the whole basin lies in the middle of the range. For simplicity, thermoelectric cooling demand changes are not considered. Some justification for this simplification is that a large percentage (88%) of thermoelectric cooling demands is returned to the river system and thus poses a water quality issue, which is not addressed in either the model or this study. Additionally, from 1970 to 2005, the percent of total water demand used for thermoelectric cooling has decreased from nearly 80% to 40% (Marella & Fanning, 2011). Furthermore, while a growing population will require more energy, there are factors in the region that will limit increased water demand for thermoelectric cooling. For example, recent applications for nuclear plants in the southeast U.S. have not been in the ACF basin, and new cooling towers such as those installed at the Southern Company Plant McDonough in the Atlanta region will dramatically decrease water withdrawals even with increased water consumption (Feldman et al., 2008).
One complicating factor to determining future demand is the interbasin transfers in the headwaters of the Chattahoochee and the Flint with the Ocmulgee, Oconee, and Coosa basins. Most of the transfers are relatively small scale and can be assumed to have a net negligible impact on water withdrawal. The exception is a large interbasin transfer in the Atlanta region, where water from the Chattahoochee basin is delivered to surrounding Georgia counties that service areas outside the ACF and that also have local water sources; in 2006 the interbasin transfer was 127 mgd and the Georgia counties total demand was approximately 195 mgd (AECOM et al., 2009). Therefore, when projecting demands into the future, it is assumed that the local water sources can supply the difference between the transfer and the demand, 68 mgd, and that any further unsatisfied demand is satisfied by the Chattahoochee basin.
6.3 Climate Projections
Climate projections for the ACF region (Figure 3) are taken from bias corrected and spatially downscaled GCM projections from the two most recent Coupled Model Intercomparison Projects (Maurer et al., 2007; Reclamation, 2013). From Phase 3 (CMIP3), 95 ensemble members are used, based on the emissions scenarios A1B, A2, and B1. The A1B scenario represents rapid and successful economic growth globally with a balanced use of energy sources, while the A2 scenario represents regionalized and differentiated economic growth, and the B1 scenario represents balanced economic development with global environmental and social consciousness (Nakicenovic & Swart, 2000). From Phase 5 (CMIP5), 150 ensemble members are used, based on all four representative concentration pathways (RCPs) as determined by the Intergovernmental Panel on Climate Change (IPCC). The RCPs are termed 2.6, 4.5, 6 and 8.5, where the numbers represent the amount of radiative forcing (W/m$^2$) that can be expected by the end of the twenty-first century (Van Vuuren et al., 2011). No preference is attached to any scenario or representative pathway for both CMIP3 and CMIP5 (Nakicenovic & Swart, 2000; Van Vuuren et al., 2011); therefore, ensembles members based on the different scenarios are used non-discriminately in this study and are considered equally likely to occur.

Trends in climate are calculated by spatially aggregating the climate projections and finding the difference between the temporally aggregated periods of 2045-2074 and 1990-2019. These periods are 30 year windows centered around the ending and starting years of the study – 2060 and 2005. In CMIP3, temperature and precipitation tend to increase co-linearly, whereas in CMIP5 precipitation change is not coupled (or decreases slightly) relative to temperature increases. CMIP5 spans the widest temperature range and has the lowest precipitation outlier, whereas CMIP3 has the highest precipitation outliers. The mean of both CMIP3 and CMIP5 projections indicate an increase in temperature of 1.6 °C and a decline in precipitation to about 96% of the historic average. However, there is significant spread in the projections for both variables.
Figure 3: a) Bias corrected and spatially downscaled CMIP3 and CMIP5 climate projections, where the projections are the difference between the time periods 2045-2074 and 1990-2019. b) The region used for the climate projections, encompassing the ACF basin.

6.4 Demand Projections
M&I water demand projections are based on population and per capita water use projections for the MNGWPD; this district consists of the 15 counties that encompass the Atlanta metropolitan region. Historic population data is from U.S. Census data for the MNGWPD (Forstall, 1995; U.S. Census Bureau, 2002; U.S. Census Bureau, 2010). The population projections are from the state of Georgia in 2010 (GA State, 2010), a water supply and conservation management report prepared for the MNGWPD in 2009 (AECOM et al., 2009), and the Atlanta Regional Commission in 2009 and 2011 (ARC, 2009; ARC, 2011), and are linearly extrapolated to 2060 as necessary (Figure 4). The projections indicate that population will increase from 4.5 million people in 2006 to between 8 and 10 million by 2050. Per capita water use is based on projected values for the MNGWPD for 2035 for a “without conservation” plan of 154.7 gallons per capita per day (gpcd), a “baseline” plan of 146.7 gpcd, and a “recommended
conservation” plan of 134.9 gpcd (AECOM et al., 2009). These per capita water use projections do not account for the impacts of climate change. AECOM et al. (2009) provide observed 2006 municipal water demands and projections for 2035 and 2050. To create M&I water demand projections for 2060, the various per capita water uses are multiplied by the different population projections. The projected M&I water demand (Figure 4) ranges from 185% to 267% of the 2006 M&I water demand, with a mean of 217%. Because evaluating the skill of projections is outside the scope of this study, the projections of water demand are all considered equally likely to occur. However, it should be noted that previous projections for the MNGWPD from the 2003 water supply and conservation management report have been criticized for overly high population projections and incomplete conservation efforts, which both result in overestimations of water demand (Pacific Institute, 2006). Additionally, water demands for the whole ACF basin do not exhibit a clearly increasing trend from 1970 to 2005 (in five-year increments) even with large increases in population, and high water demand observations may be due to dry conditions (Marella & Fanning, 2011). It should also be noted that water demand in major cities, specifically Boston, MA and New York, NY, has shown a decreasing trend (Laskey, 2012; Strickland, 2012), though whether or not these trends can be generalized to other regions or Atlanta specifically is another question. Therefore, these projections of population and water demand should be considered as conservative estimates.
Figure 4: Historic and projected a) population and b) M&I water demand in the MNGWPD.
The agriculture water demand projections are from estimates for the ACF basin made by the Natural Resources Conservation Service (NRCS) (NRCS, 1996) for the ACF compact study. The historical data is created by aggregating county level data to the basin (Marella & Fanning, 2011). The NRCS projections assume normal rainfall conditions and a medium level of water demand without water conservation. The projections are linearly extrapolated out to 2060, which is 210% of the 2005 demands.

![Figure 5: Historical and projected agriculture water demands for the ACF basin.](image)

### 6.5 Basin Vulnerability Metrics

The ACF basin is composed of a wide range of stakeholder interests that vary along the length of the basin. Some of the most prominent stakeholders, defined broadly, are North Georgia, which includes the MNGWPD and residents around Lake Lanier, Georgia residents downstream of the MNGWPD such as the City of Columbus (and the military training installation at Fort Benning), the eastern portion of Alabama, the Apalachicola region in the panhandle of Florida, the navigation (especially barge) industry,
environmental interests, and federal environmental agencies (Feldman, 2008). These numerous stakeholders can be placed in three broad categories with interests that can be represented using sets of metrics (Table 1).

Table 1: Stakeholders and interests in the ACF and the corresponding metrics.

<table>
<thead>
<tr>
<th>Stakeholders</th>
<th>Interests</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNGWPD Residents around Lake Lanier</td>
<td>Adequate water supply</td>
<td>Frequency and magnitude of water supply shortfalls</td>
</tr>
<tr>
<td></td>
<td>Recreation on Lake Lanier</td>
<td>Lake Lanier storage</td>
</tr>
<tr>
<td>Georgia residents downstream of MNGWPD (i.e. City of Columbus and Fort Benning) Eastern Alabama</td>
<td>Sufficient water quality and quantity to support agricultural, municipal, and navigation interests</td>
<td>Frequency and magnitude of water supply shortfalls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lake Lanier and West Point Lake storage</td>
</tr>
<tr>
<td>Apalachicola Florida Navigation (barge) industry</td>
<td>Water quantity for navigation</td>
<td>Frequency and magnitude of navigation flow shortfalls</td>
</tr>
<tr>
<td>Environmental interests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federal environmental agencies</td>
<td>Environmental flows of quality water for oyster and fishing industry</td>
<td></td>
</tr>
</tbody>
</table>

The frequency and magnitude of shortfalls are given by reliability and vulnerability metrics respectively, which were formalized by Hashimoto et al. (1982). Here, reliability, which ranges from 0 to 100% where 100% indicates perfect system performance, is calculated as

$$Reliability = 100 \left( 1 - \frac{\text{# of shortfalls}}{\text{# of timesteps}} \right).$$ \hspace{1cm} \text{Eq. 1}

Here, vulnerability, which ranges from 0 to 100% where 0% indicates perfect system performance, is calculated as

$$Vulnerability = 100 \text{mean} \left( \frac{\text{shortfall}}{\text{demand}} \right).$$ \hspace{1cm} \text{Eq. 2}

Shortfalls of zero are not included in the vulnerability calculation; therefore, vulnerability as defined here indicates the mean magnitude of shortfall (normalized to the demand at a given site) if a shortfall occurs.
Note that the reliability metric is cumulative; that is, reliability is evaluated at the end of the model run. Vulnerability is also cumulative in that the value is based on all the observed shortages in the simulation record. Both the reliability and vulnerability metrics can be evaluated at any demand site in the basin. Reliability and vulnerability metrics can be further processed to determine the marginal sensitivity of the system to each stressor. The marginal sensitivity shows the change in the metric due to a change in one stressor relative to the metric value observed for the baseline case, which changes for every stochastic realization. The marginal sensitivity is calculated individually for each stressor, assuming baseline levels for the other stressors. Thus, for every stochastic realization, marginal sensitivity is calculated as

$$\text{Marginal Sensitivity}_{\text{metric}} = 100 \left( \frac{\text{metric}}{\text{metric}_{\text{baseline}}} - 1 \right)$$

Eq. 3

The performance of the reservoirs in the basin is characterized by metrics of minimum and maximum annual storage for the whole basin and for individual reservoirs. The whole basin storage (whether minimum or maximum) is calculated in each year as

$$\text{Whole Basin Storage Metric}_{\text{year}} = 100 \, \text{function}_{\text{year}} \left( \frac{\text{LL Storage} + \text{WPL Storage} + \text{WFG Storage} + \text{LS Storage}}{\text{capacity(\text{LL Storage} + \text{WPL Storage} + \text{WFG Storage} + \text{LS Storage})}} \right)$$

Eq. 4

where \text{function} is either minimum or maximum, and LL is Lake Lanier, WPL is West Point Lake, WFG is Walter F. George Lake, and LS is Lake Seminole. For individual reservoirs, the equation is modified to

$$\text{Individual Reservoir Storage Metric}_{\text{year}} = 100 \, \text{function}_{\text{year}} \left( \frac{\text{Reservoir Storage}}{\text{capacity(Reservoir Storage)}} \right)$$

Eq. 5

In contrast to the reliability and vulnerability metrics, which represent a cumulative value over the entire time period, the storage metric is evaluated for each year at the month of minimum or maximum storage. The minimum storage metric is primarily related to drought and water supply, while the maximum storage metric is primarily related to flooding.
7 System Response Results
The system response results begin with an analysis of the marginal sensitivity of metrics to the individual stressors. Subsequently, the system response to the individual and combined impacts of the stressors is analyzed using system response surfaces. From the full system response, ex-post scenarios of risk or no risk are defined and the model-projections-based likelihood of future system states is determined. Finally, the spatial variation of stressor impacts throughout the basin is examined.

7.1 System Marginal Sensitivity
The relative impact of each individual stressor is revealed in the marginal sensitivity of system performance, as shown through reliability and vulnerability metrics at key system locations; water supply at Atlanta, Columbus, and Blountstown, and navigation flow at Blountstown (Figure 6).
Figure 6: Marginal sensitivity of a) reliability and b) vulnerability to changes in precipitation, temperature, M&I and agriculture demand calculated relative to the baseline of each stochastic realization at three key locations in the basin. In a) multiple traces for each metric correspond to the 25 stochastic realizations, while in b) the single trace for each metric corresponds to the mean across the stochastic realizations.

In Figure 6a, for all stressors except M&I demand, only the marginal sensitivity of water supply reliability at Blountstown is observed, because the marginal sensitivity of water supply reliability at all three locations is nearly identical. In contrast, for M&I demand, the marginal sensitivity of both Atlanta and Blountstown water supply reliability is observed, because Atlanta is slightly more sensitive to M&I changes than Columbus and Blountstown, which respond identically. This indicates that every stressor
has the same impact on the frequency of water supply shortfalls regardless of location in the basin, except that M&I demand is slightly more impactful at Atlanta, the location of maximum water demand. While this result may seem surprising, one reason this occurs is likely due to the configuration of the systems model. In the systems model, 10% curtailments are imposed at every demand site when total system storage in the conservation zone of the reservoirs falls below 30%. Consequently, shortfalls tend to occur simultaneously at every demand site in the basin except at Atlanta. However, even though shortfalls occur simultaneously, the magnitude of the shortfalls is different at each location (Figure 6b). This behavior is also observed in subsequent water supply reliability and vulnerability results.

At each key location in the basin, the reliability metrics (Figure 6a) respond linearly to changes in temperature and demand, but respond non-linearly to precipitation changes. Across the system, reliability always improves slightly with increased precipitation but shows a much stronger decline for decreased precipitation. Note that flooding was not examined in this study. Precipitation has by far the greatest impact on reliability for the whole system; for example, the percent change in each metric at 80% of precipitation is at least as large as the maximum percent change caused by the other stressors. For other stressors besides precipitation, the impact on marginal sensitivity depends on spatial location and the chosen reliability metric. For water supply reliability at all locations, precipitation change impacts are followed by M&I demand changes, then temperature changes, with negligible effects from agriculture demand changes. In contrast, for navigation flow reliability at Blountstown, precipitation impacts are followed by temperature changes, then agriculture demand changes, and finally M&I demand changes. Blountstown navigation reliability is unique because Lake Lanier, West Point Lake, and Walter F. George Lake all make releases to meet the flow requirement. For Blountstown navigation reliability, the temperature is relatively more important than demand likely because temperature will affect evaporation, which affects the total runoff in the basin and the water stored in the reservoirs. Likewise, agriculture
demand is likely more impactful than M&I demand because the majority of agriculture demand occurs in the Flint basin and at Blountstown (Figure 2).

The marginal sensitivity of the reliability metrics (Figure 6a) is calculated for each of the 25 stochastic realizations, indicating the relative influence of internal climate variability as compared to longer-term trends. The range of sensitivity across the stochastic realizations grows larger as the change in each stressor becomes more severe; for example, the percent change of Atlanta water supply reliability at the M&I demand level of 150% ranges from 0% to -5%, while at the M&I demand level of 300% ranges from -5% to greater than -15%. Therefore, internal climate variability exerts more influence on system reliability when the system is already stressed by unfavorable long-term trends.

The marginal sensitivity of the vulnerability metrics (Figure 6b), calculated for the mean of the 25 stochastic realizations, reveals different responses to the stressors compared to reliability metrics. One key difference is that the marginal sensitivity of water supply vulnerability at the three locations is not identical as discussed above; additionally, precipitation change in one direction (increase or decrease) no longer clearly dominates the response, but is nearly equaled or exceeded by sensitivity to M&I demand and by sensitivity to temperature at locations further downstream in the basin. Therefore, while precipitation clearly affects the frequency of shortfalls, both precipitation and M&I demand affect the magnitude of shortfalls. A second key difference is that agriculture demand has almost no effect on any vulnerability metric relative to any other stressor. A third key difference is that the marginal sensitivity of navigation flow vulnerability is nearly always smaller than that of water supply, regardless the location. These significantly different results for marginal sensitivity exhibited between the reliability and vulnerability metrics underscore the importance of examining several metrics in order to fully characterize the response of the system.
The importance of spatial location within the basin is clearly observed in the marginal sensitivity of water supply vulnerability. Atlanta water supply vulnerability is most affected by precipitation and M&I demand with some affect from temperature. This result likely occurs because Atlanta has the largest water demand, of which a large portion is M&I demand, and because even though Atlanta is downstream of Lake Lanier (a large reservoir), its headwaters location in the basin limits water availability relative to downstream locations. Columbus water supply vulnerability is affected nearly identically by precipitation, M&I demand, and temperature. This result likely occurs because Columbus is downstream of Atlanta (significant water demand) and Lake Lanier and West Point (large reservoirs with evaporation), but also has greater water availability due to a more centralized basin location compared to Atlanta. Finally, Blountstown water supply vulnerability responds to the stressors in nearly the same manner as Columbus, except M&I demand is more impactful and precipitation is less so. This result likely occurs because Blountstown has the second largest M&I demand after Atlanta and is downstream of not only the Chattahoochee and Flint rivers, but also all the major reservoirs in the system and nearly all other M&I demand sites.

7.2 System Response Surfaces
The system response to multiple stressors, a hyper-dimensional space, is represented by 2D response surfaces, which show the joint sensitivity of the system to the combined impacts of changes in two stressors. To determine the response surfaces, the mean system response is calculated across all 25 stochastic realizations of internal climate variability, essentially averaging out the effects of variability to focus the analysis on the impacts of long-term trends in mean precipitation, temperature, and M&I and agriculture demand. Here, selected response surfaces for Atlanta water supply reliability illustrate the insights gained from joint sensitivities and stressor projections (Figure 7).
Figure 7: Response surfaces showing selected joint sensitivities for the Atlanta water supply reliability metric (in units of %), generated by setting the remaining stressors to the baseline level. The black line indicates the contour that corresponds to system performance of the control case. The orange, yellow, and green points indicate projections for the stressors (CMIP3, CMIP5, and demand respectively), and the black point indicates the mean of the projections. Note that superimposing the projections on the c) response surface assumes that temperature has not changed from the baseline level, even though the climate projections are joint precipitation and temperature change values.

Assuming that only climate changes (Figure 7a), and that both M&I and agriculture demand remain at baseline levels, the angle of the contours indicates that precipitation impacts dominate the system while temperature impacts are nearly negligible unless the increase is very large (3° C or more). The projection mean is only slightly lower than the contour that corresponds to the control case, but the individual projections are distributed on either side, indicating that the future performance of the system relative to the control is uncertain; however, a substantial number of projections indicate lower reliability in the future. For climate change only, the lowest system reliability values, the “worst case”, occur at very low precipitation and very high temperature, where the control reliability of 98.9% drops to 55.7%. If the opposite is assumed, that only demand changes and there is no change in climate (Figure 7b), then the angle of the contours indicates that M&I demand changes have the largest impact, with a small but noticeable impact from agriculture demand changes. Here, all projections lie below the historical reliability of the control case, but the severity of the “worst case” is less; reliability only drops to 89.2% at the highest demand levels. When the two most influential stressors for climate (precipitation) and demand (M&I demand) are compared (Figure 7c), an interesting result emerges. Precipitation is still the dominant stressor, but the impact of M&I demand increases at levels greater than 150% of the historic M&I demand. The majority of the projections indicate a lower than historical reliability, but note that
temperature is assumed to be at the baseline level, even though the climate projections are joint precipitation and temperature values. This assumption should not dramatically alter the results, given the relatively small impact from temperature compared to precipitation and M&I demand for Atlanta water supply reliability as shown in Section 7.1.

As discussed in Section 7.1, the system response as expressed by water supply reliability is nearly identical at all demand sites in the basin. However, variations in system response are observed in the water supply vulnerability response surfaces for each key location in the basin (Figure 8).
Figure 8: Response surfaces due to changes in climate only or demand only for water supply vulnerability in units of % at a-b) Atlanta, c-d) Columbus, and e-f) Blountstown. See Figure 7 for an explanation of the image elements.
At the three key locations, if only demands change (Figure 8b, d, and f), M&I demand is the dominate stressor, which corroborates the marginal sensitivity results (Section 7.1). If only climate changes (Figure 8a, c, and e), precipitation is the dominate stressor, which corroborates the Atlanta marginal sensitivity results, but contradicts the Columbus and Blountstown marginal sensitivity results indicating that temperature is at least as influential as precipitation. This discrepancy is explained by recognizing that marginal sensitivity is calculated for one stressor at the baseline levels of the other stressors. For example, for climate change only at Columbus (Figure 8c), a vertical line at 100% precipitation has nearly as much variation as a horizontal line at 0° temperature change. Thus, if only one stressor changes, precipitation and temperature have relatively equivalent influence; however, over the whole response surface, precipitation is the dominant stressor. This indicates the importance of examining not only the impacts from changes in a single stressor, but also the additional impacts caused by combined stressor changes.

The magnitude of vulnerability decreases across the three key locations. At severe changes in climate the vulnerability at Atlanta is nearly 35%, while at severe decreases in precipitation or severe increases in M&I demand the vulnerability at Columbus is not more than 10%, and at severe increases in M&I demand the vulnerability at Blountstown is not more than 7%. It is premature to generalize that this decreasing trend is only caused by the relative upstream to downstream location of the demand site given that only 3 of the 11 sites are considered; further analysis is presented in Section 7.4. At Atlanta (Figure 8a-b), climate projections are distributed on either side of the historical vulnerability, while demand projections indicate higher vulnerability than historically; this is similar to the Atlanta water supply reliability results. In contrast, at Columbus (Figure 8c-d) vulnerability is never greater than historically across the range of stressor changes tested. Therefore, while Columbus might have decreased reliability – more frequent shortfalls, the vulnerability will not increase – the magnitude of shortfalls will not be more than historically. Finally, at Blountstown (Figure 8e-f), vulnerability is never greater than historically.
across the range of climate changes tested; however, M&I demand changes above 150% cause vulnerability to be slightly higher than historically.

7.3 Likelihood of Future System Performance
The likelihood of future system performance can be determined by delimiting the system response into two scenarios that represent either risk (the system performance is worse than current conditions) or no risk (the system performance improves over current conditions). The threshold used to delimit these ex-post scenarios is the current system performance as observed in the control case. These ex-post scenarios are then used to categorize the projections of climate and water demand mean changes, and thus assign model-based probabilities to future system performance. However, it should be noted that these system performance probabilities are to some degree subjective, based on the probabilities assigned to the projections. Here, because every projection is considered equally likely to occur, the probability of the risk versus no risk scenario is simply the percentage of total projections falling below the threshold versus above the threshold. Another potential option for determining probability is to equally weight the space delimited by the projections and assume that the delimited space is the lower limit of the maximum bound of uncertainty as suggested by Stainforth et al. (2007a). Alternatively, further information about the trustworthiness of different projections (if a given GCM has higher skill at simulating climate in the region or if a water demand projection is much more recent than another) or the independence of different projections could lead to different weights assigned to each projection. Assigning weights to climate projections has been investigated in detail by multiple studies, such as Manning et al. (2009) and Greene et al. (2006); for a review of methods, see Tebaldi & Knutti (2007). However, a detailed analysis of projection skill is beyond the scope of this study. This process of assigning ex-post scenarios and using stressor projections to determine the likelihood of future system performance is shown graphically for Atlanta water supply reliability in Figure 9.
Figure 9: A visual representation of the likelihood of future system performance for Atlanta water supply reliability. The control case reliability (the black line) defines the threshold between the risk and no risk scenarios. The other lines are the empirical density functions of the projections. For the independent groups of projections (CMIP3, CMIP5, M&I demand, and agriculture demand) the densities are calculated at baseline levels of the other groups as necessary. For example, CMIP3 climate was evaluated at the baseline levels of M&I and agriculture demands. Note that the empirical density function for agriculture demand is infinity because there is only one projection.

The reliability associated with each projection (Figure 9) was calculated by linearly interpolating the system response, which is represented by discrete points. For simplicity, agriculture demand was evaluated at 200% (one of the simulated levels) rather than interpolated to 210% (the projection). Due to the minimal effect of agriculture demand on the system and the relative closeness of the projection to the simulated level (compared to the next closest level of 300%), this simplification is assumed to have negligible impact on the results. When considering every combination of every projection, 99.4% of the projections are in the risk scenario. This model-projection-based risk has an associated distribution of
potential water supply reliabilities as estimated by the density function. From the distribution, a statistic such as the median can be used to estimate the magnitude of change from historical performance to future performance. For Atlanta, based on the empirical density function of all the projections, there is equal likelihood that future water supply reliability will be above or below the median value of 89.5%. Furthermore, the distribution is not normal, but has greater spread on the lower range of reliability. In the extreme, a small set of projections indicates that reliability could decrease to as low as approximately 70%. This large spread in reliability is due to the combined impacts of changes in all the stressors. If only climate or demand change, then the distributions shift to higher reliabilities; the percent of projections in the risk scenario and median reliability values for each set of projections based on the empirical density function is given in Table 2. It is noteworthy that conclusions about risk and potential future performance are highly dependent on the chosen projection set(s) and thresholds and could vary even more with the addition of new projections. For example, CMIP5 has a greater spread towards lower reliability values than CMIP3. The percent of projections in the risk scenario and the median of the reliability distribution for each set of projections are tabulated in Table 2.

Table 2: Based on the empirical density function, the percent of projections in the risk scenario and the median of the reliability distribution for each set of projections for Atlanta water supply reliability. Note that for agriculture demand there is no distribution. Units are %.

<table>
<thead>
<tr>
<th>Control Reliability = 98.9</th>
<th>Projections in Risk Scenario</th>
<th>Median of Reliability Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Projections</td>
<td>99.4</td>
<td>89.5</td>
</tr>
<tr>
<td>CMIP3</td>
<td>83.3</td>
<td>96.8</td>
</tr>
<tr>
<td>CMIP5</td>
<td>89.7</td>
<td>95.6</td>
</tr>
<tr>
<td>M&amp;I Demand</td>
<td>99.1</td>
<td>96.5</td>
</tr>
<tr>
<td>Agriculture Demand</td>
<td>0</td>
<td>99.3</td>
</tr>
</tbody>
</table>

As discussed in Section 7.1, the system response as expressed by water supply reliability is nearly identical at all demand sites in the basin; figures for Columbus and Blountstown similar to Figure 9 are included in Appendix A. However, there is a slight difference between the results at each location for the M&I demand projections, and consequently for all projections combined, as shown in Figure 10.
Figure 10: Empirical density functions at the three key locations in the basin for a) all projections and b) M&I demand projections. The vertical black line is the control reliability. The empirical density functions are identical for Columbus and Blountstown. Note that the x-axes span dramatically different ranges.

The process of understanding model-projection-based risk at each of the three locations in the basin can also be applied to water supply vulnerability. The marginal sensitivity results and system response surface results in Sections 7.1 and 7.2 indicate that there should be significant differences in vulnerability across the three locations, which is shown in Figure 11. Here, the projections are represented as histograms rather than empirical density functions to highlight the relative number of projections for each set and the discrete nature of the projections that is obscured by the empirical density functions.
Figure 11: Water supply vulnerability histograms at each of the three key locations in the basin based on the sets of projections, where a) is all the projections combined, b) is the climate projections only, and c) is the demand projections only. The thresholds of historical vulnerability from the control case are shown by vertical lines. See Figure 10 for more information on how these vulnerabilities are calculated.

As expected based on the system response surfaces discussed in Section 7.2, within each set of projections, Atlanta vulnerability is higher than Columbus, which is higher than Blountstown (Figure 11). Relative to the climate projections, there are a relatively small number of projections for demand changes. Creating an empirical density function for the demand projections as done for Atlanta water supply reliability is a broad generalization, given the discrete and limited nature of the data. However,
empirical density functions for all the projections combined and for the climate projections seem more appropriate. Some commonalities between the results at all three demand sites are: 1) CMIP5 climate projections indicate higher vulnerabilities than CMIP3 climate projections (Figure 11b), which corresponds to the lower reliabilities from CMIP5 compared to CMIP3 (Figure 9), and 2) agriculture demand projection always correspond to lower than historic vulnerability (Figure 11c).

In addition to the commonalities between the demand sites, other results are specific to each site. At Atlanta, nearly all the combined projections (Figure 11a) fall in the risk scenario, where future water supply vulnerability is greater than historical. When only climate changes, the majority of the projections indicate lower than historic vulnerability (Figure 11b), while when only M&I demand changes, all the projections lead to greater vulnerability (Figure 11c). At Columbus, all the projections for both climate change and demand change fall into the no risk scenario (Figure 11b-c), at vulnerabilities below the threshold which seems to correspond to an upper bound on vulnerability (corresponding to the results from the system response surfaces in Section 7.2). Further investigation is required to determine whether this threshold is truly an upper bound and what are the reasons for this behavior. The results at Blountstown are neither as concerning as at Atlanta, nor as reassuring as at Columbus. The combined projections and the M&I demand projections fall in both scenarios (Figure 11a, c), while all climate projections result in lower than historic vulnerabilities (Figure 11b). Overall, Atlanta is much more vulnerable than the other two locations. When thinking about the projections however, the vulnerabilities dramatically change depending on which stressors change (i.e. vulnerability decreases if only climate changes compared to if both climate and demand change) and change noticeably depending on the chosen iteration of climate projections (CMIP3 or CMIP5).
7.4 Spatial Variability of System Impacts
The spatial variability of impacts throughout the basin has important implications for local water resources planning. Spatial variability in impacts is explicitly examined by comparing the water supply reliability and vulnerability at every demand site in the basin for representative combinations of stressor values (Figure 12). For simplicity, three representative combinations were chosen; the control case that represents historic conditions, the baseline cases that represent stochastic realizations of historical climate, and the projection mean cases that represent the future system performance if the mean of the projections occurs.
Figure 12: Spatial variation in water supply a) reliability and b) vulnerability from upstream to downstream in the western and eastern portions of the basin.

Reliability at every demand site (Figure 12a) is identical for the control case, the baseline cases, and the projection mean cases, except at Atlanta (demand site number 2), confirming previous results in Sections 7.1 – 7.3. For the projection mean cases, Atlanta’s water supply reliability is slightly lower, indicating
that Atlanta is adversely affected by changes in the stressors to a slightly greater degree than any other demand site. Even with the slight discrepancy at Atlanta, several basin-wide conclusions can be made. First, reliability is much lower for the projection mean cases (median around 90%) than for the baseline cases (median just under 100%), and indicates problematic conditions across the whole basin. Second, the variability of reliability across the stochastic realizations is much larger for the projection mean cases in comparison to the baseline cases. This can be interpreted two ways; either there is minimal variation for the baseline cases because the system operates nearly perfectly regardless of climate variability, or climate variability has a more significant effect on water supply reliability as the stressors deviate from historic conditions. The latter explanation corroborates the results found for the system marginal sensitivity in Section 7.1. Finally, the control case has a relatively low reliability value compared to the majority of the baseline cases, indicating that the historic climate contained considerable challenges to water resources management relative to most of the stochastic realizations.

In contrast to reliability, water supply vulnerability (Figure 12b) varies across the basin and between the baseline and projection mean cases, indicating that the magnitude of shortfall as a percent of total demand is not spatially uniform. The spatial variation is not clearly linked to position upstream to downstream in the basin; if so, vulnerability would uniformly increase or decrease across demand sites 1 – 5, 6 – 8, and 9 – 11. Vulnerability at each demand site is likely due to a variety of factors, including the upstream to downstream position in the basin, as well as the location relative to a reservoir, and the magnitude of demands; however, determining the exact role of each factor requires further detailed investigation. The vulnerability associated with the control case generally matches the mean of the baseline cases, except at Blountstown (demand site 10), indicating that the system response to the stochastic realizations of climate aggregates to the system response to the historic climate when considering vulnerability. The range of vulnerability across the stochastic realizations is larger for the baseline cases, with the exception of Atlanta, which could indicate that internal climate variability is less important to the magnitude of
shortfalls as the stressor change is more severe. This is opposite to what was observed for reliability. Overall, vulnerabilities are higher for the projection mean cases compared to the baseline cases, though there seems to be an upper limit to the vulnerability at certain locations, perhaps due to the model configuration as noted in Section 7.3. The vulnerability at Atlanta for the projection mean cases is dramatically higher in comparison to the other demand sites, a cause for concern because the magnitude of shortfalls at Atlanta is much more affected by stressor changes than the rest of the basin. The unique results at Atlanta likely occur because the M&I demand alone at Atlanta is greater than the demand for all sectors at every other site in the basin except at Jim Woodruff (demand site number 9) and thermoelectric cooling demand at Atlanta is even greater than M&I demand. Additionally, Atlanta is located in the headwaters of the basin, which limits the availability of water, despite storage in Lake Lanier. The spatial variation in impacts has important implications for both Atlanta and the rest of the basin; while the frequency of shortfalls (reliability) is the same, the magnitude of shortfalls (vulnerability) varies dramatically. This difference also highlights the importance of selecting multiple metrics when examining spatial distribution of performance across a complex system.
8 System Uncertainty Results
In addition to understanding the system response as discussed in Section 7, an important aspect of characterizing the system is analyzing change over time and analyzing the sources of uncertainty that affect system response over the planning horizon. Sources of uncertainty can be generalized to four broad categories: forcing uncertainty, initial condition uncertainty, model uncertainty, and model inadequacy (Stainforth et al., 2007b). A method for examining the effects of various sources of uncertainty is to simulate the system under a variety of plausible options for each source of uncertainty. The variety could be explored with Monte Carlo experiments, such as varying parameters to quantify model uncertainty, or with ensemble experiments, such as varying scenarios to quantify forcing uncertainty. The resulting range of system response or performance is then attributed to the various sources of uncertainty through methods such as an analysis of variance (ANOVA) or similar techniques. Attribution of uncertainty studies have been performed for climate model predictions (Deser et al., 2012; Hawkins & Sutton, 2009; Hawkins & Sutton, 2011; Stainforth et al., 2005; Yip et al., 2011), hydrologic model predictions (Bosshard et al., 2013; Dobler et al., 2012; Tang et al., 2006), hydropower systems management (Finger et al., 2012; Schaefl et al., 2007), and water quality modeling (Radwan et al., 2004). This study contributes to the body of literature on sources of uncertainty with an in-depth examination of the importance of diverse types of forcing uncertainty on a water resources system. Forcing uncertainty is caused by uncertainty in the stressors due to the ranges of projections, which includes climate variability, mean temperature and precipitation change, and mean M&I and agriculture demand change. Thus, the stress test samples the system response to uncertain stressors, where the plausible range of changes in uncertainty sources is defined by the span of the stressor projections. An ANOVA is used to attribute system response to the various types of forcing uncertainty.

Before determining the effects of sources of uncertainty on systems, it is useful to define several general categories that encapsulate the myriad sources of uncertainty that affect any modeled system. Stainforth et
Stainforth et al. (2007b) define the following types of uncertainty: forcing uncertainty associated with the external factors that affect the system, initial condition uncertainty associated with the starting state of the system, and model imperfections associated with the inability of the model to recreate earth’s physical processes. Model imperfections are further specified as model uncertainty associated with model parameter values and model inadequacy associated with the simplicity of the model compared to physical processes. While Stainforth et al. (2007b) define these categories for climate models, here it is maintained that these categories are also useful when considering sources of uncertainty for a wide range of models, including hydrologic models and water systems models. For example, for river water quality modeling, Radwan et al. (2004) define the following categories of sources of uncertainty that cause uncertainty in results: model-input uncertainty (analogous to forcing uncertainty from Stainforth et al. (2007b)), model-structure uncertainty (analogous to model inadequacy), and model-parameter uncertainty (analogous to model uncertainty). Here, the categories given by Stainforth et al. (2007b) are used to coherently compare previous literature on attributions of the effects of sources of uncertainty.

Studies of the attribution of uncertainty in climate models have focused on forcing uncertainty, initial condition uncertainty, and model uncertainty. Stainforth et al. (2005) simulate the temperature increase of one GCM in response to multiple sets of parameter values and initial conditions, thus exploring model and initial condition uncertainty. As would be expected, the study found that model uncertainty alone causes a large range of simulated temperatures, but that the range grows larger with the addition of initial condition uncertainty. Deser et al. (2012) examine the specific climate processes in GCMs that lead to internal climate variability, which corresponds to initial condition uncertainty. In two pivotal papers, Hawkins & Sutton (2009, 2011) examine the uncertainty in predictions of global and regional precipitation and temperature due to multiple GCMs, multiple emissions scenarios, and internal climate variability, thus exploring all the types of uncertainty defined previously. Using a fitted polynomial method to partition uncertainty, Hawkins & Sutton (2009, 2011) show that in general the relative
importance of internal climate variability (or initial condition uncertainty) and model uncertainty decreases with increasing lead time, while the relative importance of scenario (or forcing) uncertainty increases with increasing lead time, though the exact nature of these trends varies by region. Yip et al. (2011), examining temperature only, furthered these studies by using an ANOVA to improve upon the polynomial fitting method, showing that both individual and joint effects from each source of uncertainty significantly affect GCM projections.

Studies of the attribution of uncertainty in water-related models (including water quality, hydrologic, and water resources models) focus on a variety of sources of uncertainty. Radwan et al. (2004) investigate multiple sources of uncertainty in a water quality model by partitioning variance in the results, using a method similar to an ANOVA, and find that forcing uncertainty is most important, followed by model uncertainty, and finally model inadequacy. In hydrologic models, model uncertainty is often investigated through a sensitivity analysis to determine which model parameters are most influential. For example, Tang et al. (2006) compare the effectiveness of different sensitivity techniques as applied to a lumped hydrologic model and find that ANOVA is one of the more skillful methods partially due to its computational efficiency. Dobler et al. (2012) investigate not only hydrologic model uncertainty, but also forcing uncertainty (GCM choice, regional climate model (RCM) choice, and bias correction method). By quantifying the spread of the results in response to each uncertainty source, the study finds that forcing uncertainty (specifically GCM and RCM choice) is most important. Similarly, Bosshard et al. (2013) investigate hydrologic model inadequacy and forcing uncertainty (GCM choice, statistical post-processing method) using an ANOVA combined with a sub-sampling routine to partition uncertainty. The study finds that forcing uncertainty (specifically GCM choice) dominates, especially in the summer and fall, but that the other sources of uncertainty and the interactions are also important, especially in the winter and spring.
When hydrologic models are combined with water systems models, the sources of uncertainty become further complicated. For hydropower production from glacial runoff, Schaeffli et al. (2007) investigate forcing uncertainty (mean trends in temperature and regional scaling relationships) and model uncertainty/inadequacy (hydrologic, glacier, and management model parameters and assumptions) in a Monte Carlo experiment that iteratively adds uncertainties together and evaluates how the probability density functions of hydropower performance metrics change in response. The study finds that model uncertainties are of the same order as forcing uncertainty (specifically mean trends in temperature). Also for hydropower production from glacial runoff, Finger et al. (2012) investigate model uncertainty (hydrologic model parameters) and forcing uncertainty (climate scenarios and glacier extents) using an ANOVA. The study finds that forcing uncertainty (specifically climate scenarios) is most important in spring and fall, while in the summer, forcing uncertainty (specifically glacial extent) is most important at the beginning of the study period, and model uncertainty is most important at the end of the study period.

A commonality among the studies that investigate more than one source of uncertainty, whether for climate, hydrologic, or water systems models, is the relative importance of forcing uncertainty compared to every other source of uncertainty. Given the importance of forcing uncertainty, and that the previous literature focuses solely on forcing uncertainty due to climate or glaciers, this study contributes to and expands the existing literature by examining the importance of diverse types of forcing uncertainty on a complex, multi-purpose water resources system – the ACF. The goal is not to characterize the full range of uncertainty in system response, which would require testing all the sources of uncertainty defined above, but rather to characterize the uncertainty in system response due to uncertainty in the ranges of the stressor projections. The stressor projections include projections of climate variability (represented by the stochastic realizations), mean temperature and precipitation change, and mean M&I and agriculture
demand change. The effects of forcing uncertainty on system response are sampled by the individual model runs of the stress test. Because the stress test explores system response beyond the range of the projections, the uncertainty analysis is limited to the portion of the stress test that encompasses the space defined by the stressor projections. An ANOVA is used to attribute system response to the various types of forcing uncertainty. An ANOVA has previously been used in several studies that attribute the effects of predictions or system response to various sources of uncertainty (Bosshard et al., 2013; Finger et al., 2012; Yip et al., 2011). The advantages of an ANOVA are that the method accounts for the full variance in system response without resorting to fitting polynomials (e.g. Hawkins & Sutton, 2009; Hawkins & Sutton, 2010), and the method quantifies both the individual and joint effects of various sources of uncertainty unlike other methods discussed above (e.g. Dobler et al., 2012; Radwan et al., 2004; Schaeffli et al., 2007). The following sections describe the use of an ANOVA to partition uncertainty, the basin-wide ANOVA results, and the spatial variation of the ANOVA results within the basin.

8.1 ANOVA Procedure
The goal of the ANOVA analysis is to determine the uncertainty in future system response caused by uncertainty in the ranges of the stressor projections, which are the expected future of the basin. Therefore, the ANOVA is performed on the portion of the stress test that represents system response to the stressor projections; this portion of the stress test does not cover the full range of mean changes tested, but rather includes the mean changes that encompass the space defined by the stressor projections (Figure 13).
Figure 13: The projections space used for the ANOVA analysis is only a subset of the full range of stressor changes, but it encompasses the M&I demand projections coupled with the CMIP3 and CMIP5 climate projections (orange and yellow respectively).

Because only one projection for agriculture demand exists, variability across this projection cannot be assessed. Consequently, the ANOVA is performed at the level of agriculture demand (200% of historic) that is closest to the projection of agriculture demand (210%). With agriculture demand set at a given level, a three-factor fixed effects ANOVA is used to determine the uncertainty in system response attributable to uncertainty in M&I demand, precipitation, and temperature (denoted as factors A, B and C) and their interactions as well as the uncertainty in system response attributable to internal climate variability represented by the stochastic realizations. The ANOVA is performed on a metric of annual reservoir storage for the basin as a whole (Eq. 4) and for individual reservoirs (Eq. 5). In each case, the metric is aggregated to 5 year periods, such that there are 11 periods out of the 55 year planning horizon. Note that the storage metric is normalized to the capacity of each reservoir, or the combined capacity of all reservoirs for the basin as a whole, to facilitate comparisons of storage across different reservoirs. To understand spatial variation in the results, both maximum and minimum annual reservoir storage is examined. In contrast to the water supply reliability and vulnerability metrics used in the previous section,
reservoir storage is more suited to this analysis because it is a metric that more fully combines the impacts caused by all the stressors. Additionally, reservoir storage is a more meaningful metric to examine at different points in time as opposed to reliability and vulnerability, because the reservoir storage in a given year is not calculated from all the previous years of reservoir storage, though physically there is some interannual relationship. Finally, by examining reservoir storage, the discussion shifts from water supply to reservoir management, which affects not only water supply, but also navigation, recreation, flooding, and the ecological health of the system.

The general model for a three-factor fixed effects ANOVA is given by

\[ X_{ijkl} = \mu + \alpha_i + \beta_j + \delta_k + \gamma^{AB}_{ij} + \gamma^{AC}_{ik} + \gamma^{BC}_{jk} + \gamma_{ijk} + \epsilon_{ijkl} \]  

where \( X_{ijkl} \) is the metric for a given level \( ijk \) of the factors (corresponding to factors A through C, respectively) and a given observation (out of 25 stochastic realizations of climate) \( l \), where \( \mu \) is the mean over all \( ijk \) for \( l \), where \( \alpha_i, \beta_j, \) and \( \delta_k \) denote the main effects of each factor A, B, and C respectively, where \( \gamma^{AB}_{ij}, \gamma^{AC}_{ik}, \) and \( \gamma^{BC}_{jk} \) denote the interactions of two factors, where \( \gamma_{ijk} \) denotes the interaction of all three factors, and where \( \epsilon_{ijkl} \) are the residuals that are assumed independent and identically distributed as \( N(0, \sigma^2) \). The residuals calculate the variance across the \( l \) observations and thus represent the internal climate variability, following Yip et al. (2011). For each five-year period over the planning horizon, the key ANOVA result is a partitioning of the total sum of squared deviations from the mean, \( SS \), into separate sums of squared deviations from the mean due to individual and joint effects of the factors. This relationship is given mathematically by

\[ SS_{total} = SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{ABC} + SS_{error} \]  

where the \( SS_{error} \) term represents uncertainty caused by internal climate variability due to the variance across the stochastic realizations. The magnitude of ANOVA results, given in Eq. 7, can also be
normalized by $SS_{total}$ to find the relative fraction of uncertainty contributed by the individual and joint effects of the factors. An ANOVA assumes that the residuals are independent and identically distributed as $N(0, \sigma^2)$. These assumptions were tested in two ways. First, the Shapiro-Wilk normality test (Shapiro & Wilk, 1965) for the residuals was computed in each time period; however, the residuals fail this test, meaning that the distribution of the residuals is not $N(0, \sigma^2)$. Second, for each time period the residuals were plotted against the fitted ANOVA values; in early time periods the residuals generally exhibit constant variance, while in later time periods, the residuals exhibit heteroskedastic behavior. Plots illustrating the non-normality and heteroskedasticity of the residuals in each time step for the basin-wide ANOVA results are shown in Appendix B. While these assumption violations affect the significance and in later time periods the variance analysis of the ANOVA results, they do not change the overall conclusions that can be drawn and would likely be present in any variance analysis of a water resources system metric.

8.2 Basin-wide ANOVA Results
The basin-wide ANOVA results, calculated with the maximum storage metric for the whole basin, are shown in Figure 14. The magnitude of variance in system response attributed to the factors is hereafter termed the uncertainty in system response, even though it is recognized that the uncertainty delimited here is only due to uncertainty in the ranges of the stressor projections and thus forcing uncertainty. Exploring the full extent of uncertainty in system response due to forcing uncertainty and model uncertainty/inadequacy is beyond the scope of this paper. Furthermore, the forcing uncertainty quantified here is a limited portion of the true uncertainty in the system performance, because it explores the uncertainty due to the projection ranges and the range of internal climate variability simulated by the stochastic climate realizations, which is only a partial sample of the full uncertainty associated with future forcing to the system. Despite these limitations, the results shown below demonstrate that an ANOVA is a
useful procedure for partitioning uncertainty and for determining the relative importance of sources of uncertainty in space and time.

Figure 14: For the basin maximum storage metric, the ANOVA results in both a) magnitude (units of squared %) and b) fraction (unitless).

The total magnitude of uncertainty increases dramatically over time, with non-linear growth (Figure 14a). This result is consistent with intuitive expectations that uncertainty is greater further in the future. The growth in variance is mainly caused by uncertainty in the range of precipitation mean change projections; by the end of the time period nearly 80% of the uncertainty in system response can be attributed to the uncertainty associated with future precipitation (Figure 14b). Uncertainty associated with internal climate variability is a significant portion of the uncertainty in system response throughout the time period, and the dominant portion of uncertainty in the first 20 years. However, because the magnitude of internal climate variability remains relatively constant, its influence decreases with increasing time. A small fraction of uncertainty in system response is also associated with M&I demand projections after 20 to 25 years, while temperature and all joint effects have negligible impact. The slight discontinuities in the results, especially 35 to 45 years in the future, are due to the variation between different stochastic realizations of internal climate variability and would likely be smoothed out with an increased number of stochastic realizations. Computational limitations prevented additional runs in this analysis. The results
observed here for basin maximum storage are very similar to the results observed for the basin minimum storage.

The basin-wide ANOVA results lead to several key insights. First, the dramatic growth of uncertainty through time implies that uncertainty in system response is increasingly large further in the future due to uncertainty in projections. Consequently, treatment of uncertainty becomes increasingly important in the planning process as the planning horizon increases. For time periods furthest in the future, the vast majority of system uncertainty is associated with uncertainty in precipitation projections with small but potentially non-negligible contributions of uncertainty from internal climate variability and M&I demand projections. At typical water resources planning horizons, on the order of 25 years in the future, uncertainty in system response is equally influenced by uncertainty in internal climate variability and precipitation projections. In contrast, short-term planning horizons should be mainly concerned with internal climate variability, which explains the majority of system response uncertainty within that time frame. The importance of internal climate variability underscores the necessity of characterizing extreme events for short-term planning. Furthermore, because internal climate variability is an irreducible result of the chaotic nature of the earth’s climate system, these results imply that water planning decisions are subject to a large degree of irreducible uncertainty. These insights apply to both the maximum and minimum (not shown) storage in the basin, indicating that every aspect of reservoir management is affected.

8.3 Spatial Variation in ANOVA Results
The basin-wide ANOVA results lead to important general insights, but reveal nothing about the spatial variation of uncertainty within the basin or the variation in individual reservoir response. To explore the relative sources of uncertainty at more local scales and to assess potential differences that arise based on
location within the basin, an ANOVA was performed for each main reservoir in the basin. Note that the same caveats regarding the limited exploration of uncertainty due to uncertainty in the range of projections as discussed for the basin-wide ANOVA results is also applicable here. Regardless, the ANOVA results still reveal important information about the system. The total uncertainty in the response of each reservoir using both the maximum and minimum storage metric is shown in Figure 15. Note that both the maximum and minimum storage metrics are normalized to the capacity of each reservoir, to facilitate comparisons across all reservoirs.

Figure 15: Total magnitude of the ANOVA results for the individual lakes in the system (units of squared %). Solid lines indicate the maximum storage metric and dashed lines indicate the minimum storage metric. The order of the reservoirs in the legend is upstream to downstream.

The ANOVA results for individual reservoirs reveal that the total magnitude of the uncertainty in reservoir response decreases the further downstream the reservoir. As observed in the basin-wide results, the magnitude of uncertainty increases with time for each reservoir. At Lake Lanier, the uncertainty is at
least an order of magnitude greater than at other reservoirs after approximately 30 years. Thus, future reservoir management for Lake Lanier must consider a much wider range of possible reservoir response to uncertainties in stressors than any other reservoir in the basin, with the second greatest variation observed at West Point Lake. The results also reveal that for Walter F. George Lake and Lake Seminole the uncertainty in minimum storage is always higher than the uncertainty in maximum storage. Similar behavior is observed at Lake Lanier and West Point Lake until late in the time period, when the uncertainty associated with the maximum storage becomes greater than that of the minimum storage.

The spatial variation in uncertainty for both the maximum and minimum storage metrics could be caused by several factors, such as the relative location of the reservoir in the basin (upstream vs. downstream), the size of the reservoir, and the type of reservoir operations. Both upstream reservoirs, Lake Lanier and West Point Lake, have dedicated flood storage, while both downstream reservoirs, Walter F. George Lake and Lake Seminole, do not. Furthermore, Lake Seminole is operated as a run-of-river reservoir with some pondage. Lake Lanier has by far the largest storage volume, Walter F. George and West Point Lake have similar volumes but are second and third largest respectively, and Lake Seminole is the smallest (U.S. Army Corps of Engineers, 2012). Thus, size is not necessarily related to the uncertainty in reservoir response in the ANOVA results, because West Point Lake is smaller than Walter F. George Lake but exhibits greater uncertainty. Rather, upstream reservoirs with flood control exhibit high levels of uncertainty relative to downstream reservoirs with no flood control and consequently more limited operations.

To further illustrate the differences between results at the level of individual reservoirs, the full ANOVA results for the maximum storage metric at a representative upstream (Lake Lanier) and downstream
reservoir (Lake Seminole) are shown in Figure 16. The full ANOVA results for the maximum storage metric at West Point Lake and Walter F. George Lake are shown in Appendix C.

Figure 16: The ANOVA results based on the maximum storage metric for a representative upper basin reservoir (Lake Lanier) in a) magnitude (units of squared %) and b) fraction (unitless), and the ANOVA results for a representative lower basin reservoir (Lake Seminole) in c) magnitude (units of squared %) and d) fraction (unitless). Note that in a) and c) the scale of the y-axis is different by several orders of magnitude.

The discontinuities observed in the basin-wide results (Figure 14), caused by the internal climate variability across the stochastic realizations, are more pronounced in the individual reservoir results (Figure 16); the discontinuities are especially pronounced for Walter F. George Lake, shown in Appendix
The shape of the magnitude (Figure 16a & c) of the ANOVA results is very different between the two reservoirs, which typifies the shape for the upstream and downstream reservoirs respectively. At the end of the time period, the steep growth in magnitude for Lake Lanier begins to level, while the growth in magnitude for Lake Seminole resembles the first 30 years of Lake Lanier. The ANOVA results for Lake Lanier (Figure 16a & b), specifically the shape of the magnitude curve and the relative fraction contributed by various effects, is very similar to that of the basin-wide results (Figure 14). This similarity is especially noticeable in the uncertainty associated with M&I demand projections, which is observed at Lake Lanier and for the basin-wide results, but is not observed at any other reservoir.

The difference in magnitude of variation has already been discussed; however, the fractional variance plots (Figure 16b & d) indicate that the importance of internal climate variability through time changes for each reservoir. For Lake Lanier, effects that do not include internal climate variability contribute 40% of the uncertainty after 20 years, whereas for Lake Seminole, the same 40% of uncertainty from non-internal climate variability effects is barely obtained by the end of the 55 year time period. This trend, where internal climate variability is increasingly important for the whole time period the further downstream, is observed for all the reservoirs when examining the maximum storage metric. In contrast, when the minimum storage metric is examined (not shown), the relative fraction of uncertainty contributed by each factor is very similar at each reservoir. Thus the magnitude of uncertainty varies spatially for both the maximum and minimum storage metrics; however, the fraction of uncertainty only varies spatially for the maximum storage metric.

These observations based on the individual reservoir ANOVA results lead to several key insights related to spatial scale that augment the insights gained from the basin-wide results. First, the relative importance of the different sources of uncertainty changes depending on the scale of the analysis. At the basin scale,
uncertainty associated with projections of mean changes in climate is the dominant source of uncertainty, although internal climate variability played a significant role, especially earlier in the time period. At the scale of individual reservoirs, internal climate variability dominates for reservoirs in the lower basin, whereas mean precipitation changes are still significant in the upper basin. This result implies that for planning in the lower basin far into the future, it is essential to develop management strategies that can cope with the range of impacts to the system caused by internal climate variability. Some examples of these management strategies could be incorporating seasonal or interannual forecasts into management, or developing mechanisms to mitigate the expected risks. However, the reason for this differential response, which is only observed for the maximum storage metric, is still not fully understood. For climate models, Hawkins & Sutton (2009) show that internal climate variability is increasingly important the smaller the spatial scale. The case may be similar here, although further work is needed to fully generalize these results to water resources systems. While the reservoir location in the basin has already been discussed, the ratio of storage to inflow variability and demand for each reservoir are likely key factors that are not currently accounted for in this analysis.

A second key insight is that a basin-wide ANOVA can be biased by the component which has the largest magnitude of uncertainty. The basin-wide ANOVA indicates that the whole basin is impacted by uncertainty in the projections of M&I demand; however, in the individual reservoir analyses, this result is only observed for Lake Lanier, which has the largest magnitudes of uncertainty among all the reservoirs. This insight underscores the importance of not only performing a basin-wide analysis, but also investigating the individual components of the basin. Furthermore, because the effect of M&I demand is only observed at Lake Lanier, this implies that uncertainty in M&I demand projections does not affect the uncertainty of reservoir response at West Point Lake, Walter F. George and Lake Seminole, a conclusion which is relevant to the consideration of reservoir operations in the ACF in light of the “Tri-State Water Wars.”
9 Discussion and Conclusion
This study extends the decision-scaling methodology to a complex, transboundary river basin stressed by climate variability and change as well as changing water demand. The central experiment of this study is a system stress test to examine the system response to a wide range of transient stressor changes. While changes in water demand are a function of historic demand, changes in climate are generated by a stochastic weather generator that preserves historic climate statistics and allows simulation of internal climate variability. Climate is then translated to streamflow through hydrologic models. From the stress test results, this study examines the marginal and combined sensitivities of the system to changes in stressors, creates ex-post scenarios based on the system response, determines model-projections-based estimates of the likelihood of future system performance, and investigates significant spatiotemporal variability in system performance. This study also extends the use of ANOVA to characterize how uncertainty in ranges of stressor projections causes uncertainty in system response.

All performance metrics of the system, including water supply reliability and vulnerability, are sensitive to stressor changes; this finding contrasts with Lettenmaier et al. (1999) which stated that M&I water supply was insensitive to the range of climate changes evaluated. In that study, the range of climate changes were based on three downscaled GCM outputs and a doubled CO₂ scenario, where at the end of the planning period temperature change ranged from 2.4 to 3.9 °C, and precipitation changed ranged from -4% to 14%. The climate changes explored in that study do not span the ranges projected by the bias-corrected and spatially downscaled CMIP3 and CMIP5 projections for the region, and limited the analysis to ranges that were shown to have less severe impacts by this study. In general, the results of this study show that precipitation is the dominant stressor throughout the basin, especially when considering the frequency of shortfalls, independent of the projections of the stressors. When considering the magnitude of shortfalls, however, M&I demand has as much of an impact as precipitation in the upper portion of the basin. This implies that the on-going dispute over the magnitude of Atlanta water withdrawals could be
overshadowed by the more pressing issue of a finite amount of water, but does not negate the role of increasing water withdrawals (especially at Atlanta) to affect the magnitude of future shortfalls. While all demand sites have nearly the same reliability (likely due to the model configuration and the method of implementing curtailments), Atlanta is the most vulnerable and most affected by changes in stressors; therefore, it is in Atlanta’s own best interests to implement effective planning strategies. Thus, effective planning needs to address the basin-wide issue of planning for precipitation changes, a stressor that is not controlled by humans, in conjunction with local water use management.

The *ex-post* scenarios, defined by the system response and the current system performance, when coupled with the stressor projections, indicate there is a high likelihood of risk to the system, especially at certain locations in the basin. However, it is noted that these likelihoods are based on the model response to projections of the stressors, and are thus limited by the assumptions of the projections and the model. In addition to the high likelihood of risk, the magnitude of risk is also high at certain locations in the basin, where system performance will be lower than historically. This should generate further incentive for the region to develop planning strategies to cope with this risk and to make the system more robust. Different projection sets, such as the two different iterations of GCM outputs (CMIP3 and CMIP5), result in different estimates of future system performance; new iterations of projections will likely change those estimates yet again. Across and within projections, the range of the magnitude of risk also varies; for example, using “worst” or “best” case scenarios from CMIP3 or CMIP5 could lead to vastly different results. Overall, the *ex-post* scenarios, which are defined based on the system, not on projections, are a useful way of analyzing system risk. Projections can then be used to estimate the likelihood of future system performance, with the caveat that the limitations and uncertainties inherent in projections are recognized.
The ANOVA results for whole basin and for individual reservoirs reveal the spatiotemporal variation in uncertainty in system response as a result of uncertainty in the projections of different stressors. The results clearly indicate that the uncertainty in system response from internal climate variability increases the further downstream and decreases the further ahead in time. The importance of internal climate variability is a call for improved representations of extreme events to understand current system performance and for improved methods for managing the resulting uncertainty in system response. At the same time, uncertainty in projections of trends in precipitation becomes very important by the end of the time period, meaning that reservoir planning must consider a wide range of system performance given the current precipitation projections. Therefore, improved precipitation forecasts, a weakness of GCM forecasts, could lead to large improvements in estimates of future performance by reducing or better defining the range of projection uncertainty. The ANOVA results also show that uncertainty in M&I demand projections is only important for Lake Lanier, which suggests that when only considering M&I demand as a stressor, all reservoir management planning except for Lake Lanier can be the same regardless of what M&I demand projection proves true (across the range of projections explored here). The uncertainty in the future response of individual reservoirs is due to multiple factors including the location in the basin and the operations of that reservoir; large uncertainty is associated with upstream reservoirs with flood control storage. Finally, the ANOVA results can be biased by the component, one reservoir, that has the greatest uncertainty, emphasizing the importance of both basin-wide and component specific analysis.

In summary, this study illustrates the important insights that can be gained by using the decision-scaling methodology, by examining marginal and joint sensitivities of the system to changes in multiple stressors with a variety of performance metrics, by using the system response to define \textit{ex post} scenarios of risk and no risk and using projections to indicate likelihoods, by investigating spatiotemporal variations, and by quantifying sources of uncertainty. While this study focuses on water quantity for water resources
planning, water quality is a significant concern in the basin as well; future work could examine the relative impacts of stressors on water quality in the basin. This study also shows the importance of metric choice – vulnerability, reliability, and reservoir storages (both maximum and minimum) reveal different but important impacts throughout the basin – future studies could further investigate metrics specifically tailored to individual stakeholder interests. In this study, water supply performance is a key performance indicator; however, changes in thermoelectric cooling for power generation were not considered and the model does not differentiate between supply to different demand sectors at a given demand site. Future work could develop a more nuanced understanding of impacts to water supply by prioritizing water delivery at each site. Finally, this study evaluated impacts on the current system; future work could evaluate the effectiveness of adaptation measures, or evaluate the impact of changes in water management that are being considered for the basin.
10 References


Environmental Protection Division. (1997a). Chattahoochee River basin management plan. Atlanta, GA: Georgia Department of Natural Resources.

Environmental Protection Division. (1997b). Flint River basin management plan. Atlanta, GA: Georgia Department of Natural Resources.


MDL-1824 Tri-State water rights litigation, 09-14657 (United States Court of Appeals for the Eleventh District 2011).


Appendix A. Likelihood of Future System Performance

Section 7.3 describes the evaluation of model-projection-based risk for the three key locations in the ACF basin. For water supply reliability, the section focuses on Atlanta, and describes the deviations from the Atlanta results for Columbus and Blountstown. Here, the full results for water supply reliability for Columbus and Blountstown are shown in Figure 17 and Figure 18.

Figure 17: A visual representation of the likelihood of future system performance for Columbus water supply reliability. The control case reliability (the black line) defines the threshold between the risk and no risk scenarios. The other lines are the empirical density functions of the projections. For the independent groups of projections (CMIP3, CMIP5, M&I demand, and agriculture demand) the densities are calculated at baseline levels of the other groups as necessary. For example, CMIP3 was evaluated at the baseline levels of M&I and agriculture demands. Note that the empirical density function for agriculture demand is infinity because there is only one projection.
Figure 18: A visual representation of the likelihood of future system performance for Blountstown water supply reliability. The control case reliability (the black line) defines the threshold between the risk and no risk scenarios. The other lines are the empirical density functions of the projections. For the independent groups of projections (CMIP3, CMIP5, M&I demand, and agriculture demand) the densities are calculated at baseline levels of the other groups as necessary. For example, CMIP3 was evaluated at the baseline levels of M&I and agriculture demands. Note that the empirical density function for agriculture demand is infinity because there is only one projection.
Appendix B. Testing ANOVA Assumptions

The ANOVA assumptions were tested for each 5 year time period for the basin-wide ANOVA results for the maximum storage metric using the Shapiro-Wilk normality test (Shapiro & Wilk, 1965) and by plotting the residuals against the fitted results. Another way to analyze normality is to examine the normal quantile-quantile plots for each time period, shown in Figure 19.

![Normal quantile-quantile plots for the basin-wide ANOVA results for maximum reservoir storage at each 5 year time period.](image)

Each time period represented in Figure 19 fails the Shapiro-Wilk normality test (Shapiro & Wilk, 1965); however, some time periods, such as periods 8 and 9, visually appear relatively normal compared to other time periods, such as periods 10 and 11. The heteroskedasticity of the residuals, especially in later time periods, is shown in Figure 20.
Figure 20: The residuals of the basin-wide ANOVA results for maximum reservoir storage versus the fitted storage values. The units are % of maximum reservoir storage.
Appendix C. Spatial Variation in ANOVA Results

The ANOVA results for the maximum storage metric for the two reservoirs in the system not shown in Section 8.3 are shown in Figure 21.

Figure 21: The ANOVA results based on the maximum storage metric for West Point Lake in a) magnitude (units of squared %) and b) fraction (unitless), and for Walter F. George Lake in c) magnitude (units of squared %) and d) fraction (unitless).