

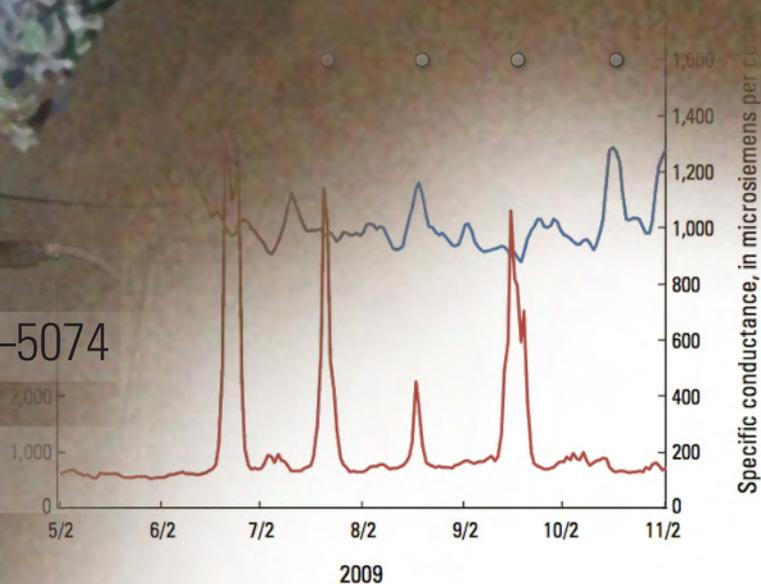
Prepared in cooperation with the U.S. Army Corps of Engineers—Savannah District

# Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009



Scientific Investigations Report 2011–5074

U.S. Department of the Interior  
U.S. Geological Survey





## COVER DESIGN

### *Photographs:*

Alan Cressler and John Join  
USGS Georgia Water Science Center

Base satellite image from U.S. Geological Survey,  
USDA Farm Service Agency, GeoEye, DigitalGlobe,  
and Google Maps,™ 2011

### *USGS Monitoring Gage:*

Abercorn Creek Near Savannah, GA

# **Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009**

By Paul A. Conrads, Edwin A. Roehl, Jr., and Steven R. Davie

Prepared in cooperation with the U.S. Army Corps of Engineers–Savannah District

Scientific Investigations Report 2011–5074

**U.S. Department of the Interior**  
**U.S. Geological Survey**

**U.S. Department of the Interior**  
KEN SALAZAR, Secretary

**U.S. Geological Survey**  
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2011

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit <http://www.usgs.gov> or call 1-888-ASK-USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Conrads, P.A., Roehl, E.A., Jr., Davie, S.R., 2011, Simulation of specific conductance and chloride concentration in Abercorn Creek, Georgia, 2000–2009: U.S. Geological Survey Scientific Investigations Report 2011–5074, 46 p.

## **Acknowledgments**

The authors thank Joe Hoke of the U.S. Corps of Engineers–Savannah District for his coordination of the project and John Sawyer, Heath Lloyd, and Tony Tucker of the City of Savannah for providing data for the intake and sharing insight on the operations of the water treatment facility.



## Contents

Acknowledgments .....	iii
Abstract .....	1
Introduction.....	2
Purpose and Scope .....	2
Description of the Study Area .....	5
Previous Studies .....	7
Approach.....	7
Data-Collection Network.....	7
Limitation of the Datasets.....	9
Data Preparation.....	10
Characterization of Specific Conductance and Chloride Concentration .....	10
Simulation of Specific Conductance and Chloride Concentrations .....	17
Artificial Neural Network Models.....	17
Development of Specific Conductance and Chloride Models .....	18
Statistical Measures of Prediction Accuracy.....	20
Specific Conductance Models .....	20
Development of Extrapolated Specific Conductance Dataset.....	24
Chloride Models .....	26
Development of a Decision Support System.....	31
Architecture.....	31
Model Simulation Control and Streaming Graphics .....	32
Application of the Savannah Chloride Model Decision Support System.....	32
User-Defined Hydrograph .....	33
Inputs from Three-Dimensional Model Output .....	35
Summary and Conclusions.....	38
References Cited.....	39
Appendix 1. Savannah Chloride Model Decision Support System User’s Manual.....	41

## Figures

1. Map showing the Abercorn Creek study area, Effingham County, Georgia .....	3
2. Map showing U.S. Geological Survey continuous streamflow monitoring network for the lower Savannah River .....	4
3. Conceptual model of the location of the freshwater-saltwater interface and salinity stratification-de-stratification cycle in estuarine rivers .....	5
4. Graph showing gage heights and streamflow for Abercorn Creek near Savannah, Georgia (station 02198810), for the period June 14 to August 12, 2010 .....	6
5. Map showing Mitigation Plan 6A.....	9
6. Map showing locations of U.S. Geological Survey water-quality sampling stations used for surface-water assessment in Tetra Tech, 2006 .....	11

7.	Graph showing daily streamflow at Savannah River near Clyo (station 02198500) and daily maximum specific conductance at the Savannah River near Port Wentworth, Georgia (station 02198840, I-95 bridge), for the period May to October 2009, including dates of the new moon .....	12
8.	Graph showing flows from Savannah River at Clyo, Georgia, and specific conductance at Savannah River near Port Wentworth, Georgia (I-95 bridge), and water-supply intake on Abercorn Creek for the period February 2000 to October 2009 .....	13
9.	Three-dimensional scatter plots of specific conductance at the intake ( $SC_{Intake}$ ) on Abercorn Creek, Savannah River at the I-95 bridge ( $SC_{I-95}$ ), and flows at Savannah River at Clyo, Georgia ( $Q_{Clyo}$ ) .....	13
10.	Scatter plots of the specific conductance at the City of Savannah's intake and at the Savannah River near Port Wentworth, Georgia (I-95), gage .....	14
11.	Graph showing specific conductance at two stations on the Savannah River (I-95 and Sav-3) and two stations on Abercorn Creek (AB-5 and Intake) for September 1 to 15, 2010 .....	15
12.	Aerial photograph of the confluence of Abercorn Creek and the Savannah River showing the difference in water color of the two streams .....	15
13.	Graph showing specific conductance and chloride values at the intake on Abercorn Creek for the period January 2003 to October 2009 .....	16
14.	Graph showing scatter plot of chloride and specific conductance values at the intake on Abercorn Creek .....	16
15.	Schematic showing multilayer perceptron artificial network architecture .....	17
16.	Schematic showing the two-stage model architecture to predict chloride concentrations ( $Cl_{Intake}$ ) at the intake using flow ( $Q_{Clyo}$ ) and specific conductance ( $SC_{I-95}$ ) inputs .....	19
17.	Three dimensional scatter plots of specific conductance at the City of Savannah's intake on Abercorn Creek ( $SC_{Intake}$ ), Savannah River near Port Wentworth, Georgia (I-95 bridge, $SC_{I-95}$ ), flows at Savannah River at Clyo, Georgia ( $Q_{Clyo}$ ) .....	19
18.	Schematic showing modifications to the two-stage model architecture to accommodate large extrapolation (ANN M2) beyond the range of historical conditions to predict chloride concentrations ( $Cl_{Intake}$ ) at the Intake using flow ( $Q_{Clyo}$ ) and specific conductance ( $SC_{I-95}$ ) inputs .....	20
19.	Specific conductance (SC) decorrelation model .....	21
20.	Measured, predicted, and residual error from the Single input Single Output Artificial Neural Network model used to decorrelate specific conductance at I-95 from the streamflow at Clyo .....	21
21.	Inputs and outputs for the models to predict specific conductance at the intake.....	22
22.	Measured and simulated specific conductance at the City of Savannah's intake from the ANN M1 model and ANN M2 model for the period February 1, 2000, to October 31, 2009 .....	23
23.	Three-dimensional response surfaces for the ANN M1 and ANN M2 models.....	24
24.	Salinity predictions at the I-95 bridge for actual conditions and a simulated 6-foot harbor deepening scenario using the EFDC model.....	24
25.	Three-dimensional scatter plots of 4-day average flow, 2-day average specific conductance (decorrelated from flow) and predicted specific conductance at the intake, and linear extrapolations of flow and specific conductance beyond the measured historical values for the ANN M1e and ANN M2e models.....	25

26.	Measured and predicted chloride concentrations at the intake and the residual error .....	26
27.	Schematic showing the model architecture for the chloride model and error correction model .....	27
28.	Computed and predicted residual error from the chloride model.....	29
29.	Measured and predicted chloride concentrations at the City of Savannah's intake for the period January 14, 2003, to October 31, 2009 .....	29
30.	Specific conductance and chloride model architecture .....	30
31.	Architectural elements of the Savannah Chloride Model Decision Support System .....	32
32.	Screen capture showing the model simulator controls used to run a simulation in the Savannah Chloride Model Decision Support System.....	33
33.	Actual and user-defined flow conditions at Clyo and actual chloride concentrations at the intake for the period October 1, 2006, to October 1, 2009 .....	34
34.	Simulated actual and user-defined specific conductance at the intake for the period October 1, 2006, to October 1, 2009.....	34
35.	Simulated actual and user-defined chloride concentration at the intake for the period October 1, 2006, to October 1, 2009.....	35
36.	Measured specific conductance, EFDC model predicted change in specific conductance at I-95, and the SCM DSS specific conductance input for a 4-foot harbor deepening and mitigation scenario for the period January 14, 2003, to October 31, 2009 .....	36
37.	Simulated actual and user-defined chloride concentrations at the intake using the ANN M1e and ANN m2e models for the period January 14, 2003, to October 31, 2009 .....	36
38.	Frequency distribution of simulated actual and user-defined chloride concentrations at the intake using the ANN M1e and ANN M2e models for the period January 14, 2003, to October 31, 2009 .....	37

## Tables

1.	U.S. Geological Survey continuous river gaging network (2011) for the lower Savannah River .....	8
2.	Water-quality constituents analyzed daily by the City of Savannah at the Abercorn Creek intake .....	8
3.	U.S. Geological Survey stations used for chloride surface-water assessment in Tetra Tech, 2005 .....	10
4.	Minimum, maximum, and mean chloride concentration at selected U.S. Geological Survey stations.....	12
5.	Model performance statistics for the specific conductance decorrelation model and the ANN M1 and ANN M2 models .....	23
6.	Model inputs, outputs, and performance statistics for the chloride model and error correction models.....	28
7.	Number of days that chloride concentrations are at or greater than 15 milligrams per liter for measured conditions and simulated conditions by two ANN models at the City of Savannah's water plant intake on Abercorn Creek.....	37

## Conversion Factors

Inch/Pound to SI

<b>Multiply</b>	<b>By</b>	<b>To obtain</b>
Length		
inch (in.)	2.54	centimeter (cm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
acre	4,047	square meter (m <sup>2</sup> )
acre	0.4047	hectare (ha)
square mile (mi <sup>2</sup> )	259.0	hectare (ha)
square mile (mi <sup>2</sup> )	2.590	square kilometer (km <sup>2</sup> )
Flow rate		
cubic foot per second (ft <sup>3</sup> /s)	0.02832	cubic meter per second (m <sup>3</sup> /s)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$^{\circ}\text{F} = (1.8 \times ^{\circ}\text{C}) + 32$$

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$^{\circ}\text{C} = (^{\circ}\text{F} - 32) / 1.8$$

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Elevation, as used in this report, refers to distance above the vertical datum.

Acronyms and abbreviations used in the report

AI	artificial intelligence
ANN	artificial neural network
CRADA	Cooperative Research and Development Agreement
DSS	Decision Support System
EFDC	Environmental Fluid Dynamics Code
I&D	industrial and domestic
I-95	Interstate 95
μS/cm	microsiemens per centimeter
mg/L	milligram per liter
Mgal/d	million gallons per day
psu	practical salinity units
RMSE	root mean square error
SCM DSS	Savannah Chloride Model Decision Support System
SISO	Single Input Single Output
USACE	U.S. Army Corps of Engineers
USGS	U.S. Geological Survey
3D	three dimensional

# Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009

By Paul A. Conrads,<sup>1</sup> Edwin A. Roehl, Jr.,<sup>2</sup> and Steven R. Davie<sup>3</sup>

## Abstract

The City of Savannah operates an industrial and domestic water-supply intake on Abercorn Creek approximately 2 miles from the confluence with the Savannah River upstream from the Interstate 95 bridge. Chloride concentrations are a major concern for the city because industrial customers require water with low chloride concentrations, and elevated chloride concentrations require additional water treatment in order to meet those needs. The proposed deepening of Savannah Harbor could increase chloride concentrations (the major ion in seawater) in the upper reaches of the lower Savannah River estuary, including Abercorn Creek.

To address this concern, mechanistic and empirical modeling approaches were used to simulate chloride concentrations at the city's intake to evaluate potential effects from deepening the Savannah Harbor. The first approach modified the mechanistic Environmental Fluid Dynamics Code (EFDC) model developed by Tetra Tech and used for evaluating proposed harbor deepening effects for the Environmental Impact Statement. Chloride concentrations were modeled directly with the EFDC model as a conservative tracer. This effort was done by Tetra Tech under a separate funding agreement with the U.S. Army Corps of Engineers and documented in a separate report. The second approach, described in this report, was to simulate chloride concentrations by developing empirical models from the available data using artificial neural network (ANN) and linear regression models. The empirical models used daily streamflow, specific conductance (field measurement for salinity), water temperature, and water color time series for inputs.

Because there are only a few data points that describe the relation between high specific conductance values at the Savannah River at Interstate 95 and the water plant intake, there was a concern that these few data points would determine the extrapolation of the empirical model and potentially underestimate the effect of deepening the harbor on chloride

concentrations at the intake. To accommodate these concerns, two ANN chloride models were developed for the intake. The first model (ANN M1e) used all the data. The second model (ANN M2e) only used data when specific conductance at Interstate 95 was less than 175 microsiemens per centimeter at 25 degrees Celsius. Deleting the conductivity data greater than 175 microsiemens per centimeter removed the "plateau" effect observed in the data. The chloride simulations with the ANN M1 model have a low sensitivity to specific conductance (salinity) at Interstate 95, whereas the chloride simulations with the ANN M2 model have a high sensitivity to salinity at Interstate 95.

The two modeling approaches (Tetra Tech's EFDC model and the one described in this report) were integrated into a decision support system (DSS) that combines the historical database, output from EFDC, ANN models, ANN model simulation controls, streaming graphics, and model output. The DSS was developed as a Microsoft Excel™/Visual Basic for Applications program, which allowed the DSS to be prototyped, easily modified, and distributed in a familiar spreadsheet format. The EFDC and ANN models were used to simulate various harbor deepening scenarios. To accommodate the geometry changes in the harbor, the ANN models used the EFDC model-simulated salinity changes for a historical condition as input. The DSS uses a graphical user interface and allows the user to interrogate the ANN models and EFDC output.

Two scenarios were simulated using the Savannah Chloride Model DSS to demonstrate different input options. One scenario decreased winter streamflows to a constant streamflow for 45 days. Streamflows during the period January 1 to February 15 were set to a constant 3,600 cubic feet per second for the simulation period of October 1, 2006, to October 1, 2009. The decreased winter streamflow resulted in predictions of increased specific conductance by as much as 50 microsiemens per centimeter and chloride concentrations by as much as 4.8 milligrams per liter during the periods of decreased streamflows. The second scenario used EFDC output for a 4-foot deepening of the harbor and streamflow configurations to mitigate for salinity increases in the vicinity of an extensive freshwater tidal marsh. A 4-foot harbor deepening scenario was simulated for the 7-year period from

<sup>1</sup> U.S. Geological Survey South Carolina Water Science Center, Columbia, South Carolina.

<sup>2</sup> Advance Data Mining, Greenville, South Carolina.

<sup>3</sup> Tetra Tech, Atlanta, Georgia.

January 2003 to October 2009. The ANN M2e model is more sensitive than the ANN M1e model to changes in specific conductance resulting from a 4-foot deepening and simulates chloride concentrations as high as 40 milligrams per liter. The ANN M1e model, which used all the data, simulated chloride concentration as high as 20.3 milligrams per liter.

## Introduction

The City of Savannah owns and operates an industrial and domestic (I&D) water treatment plant. The raw water intake for this facility is on Abercorn Creek, a tributary of the Savannah River (fig. 1). The conventional water treatment plant was constructed in 1947 to treat 35 million gallons of water per day (Mgal/d). The plant originally served many industries in the area that required water with a chloride concentration of less than 12 milligrams per liter (mg/L). Historically, the water plant withdrew groundwater and surface water to meet its quantity and quality water demands. As the population of Savannah increased, the treatment plant and its associated processes have been upgraded to its current 75-Mgal/d maximum capacity. Because of groundwater salinity intrusion in the Savannah area, a capacity-use restriction has been imposed on groundwater withdrawals, and the plant has expanded its use of surface-water supplies (Georgia Environmental Protection Division, 2001).

Abercorn Creek is tidally affected by the Savannah River estuary (fig. 2); thus, the daily tidal fluctuations cause the continuous change in the physical and chemical characteristics of the source water and add to the complexity of the treatment process. Chloride is the major ion in seawater, and although Abercorn Creek is a freshwater system with salinity less than 0.5 practical salinity unit (psu), there is concern that the proposed deepening of Savannah Harbor will increase salinity in the lower Savannah River estuary and ultimately chloride concentrations at the intake.

The U.S. Geological Survey (USGS) and the U.S. Army Corps of Engineers (USACE) determined that this concern presented an opportunity to develop an empirical model using data-mining techniques, including artificial neural network (ANN) models, to simulate specific conductance and chloride concentrations in Abercorn Creek using the real-time gaging network and water-quality data for Abercorn Creek and the lower Savannah River estuary. The USGS, in cooperation with the USACE–Savannah District, initiated a study to (1) develop empirical models to simulate chloride concentrations at the City of Savannah intake and (2) develop a spreadsheet application that integrates historical data, empirical chloride models, and output from the three-dimensional (3D) mechanistic model of Savannah Harbor that is easy to use and can be readily disseminated. The USGS collaborated with Tetra Tech and Advanced Data Mining (contractors to USACE–Savannah District) on this study.

The USGS entered into a Cooperative Research and Development Agreement (CRADA) with Advanced Data Mining in 2002 to collaborate on applying data-mining techniques and ANN models to water-resources investigations. The emerging field of data mining addresses the issue of extracting information from large databases (Weiss and Indurkha, 1998). Data-mining methods come from different technical fields, such as signal processing, statistics, artificial intelligence, and advanced visualization. Data mining uses methods for maximizing the information content of data, determining which variables have the strongest correlations to the problems of interest, and developing models that predict future outcomes. This knowledge encompasses both understanding of cause-effect relations and predicting the consequences of alternative actions.

## Purpose and Scope

This report presents the results of an investigation in which the relation between specific conductance and chloride concentrations in the Abercorn Creek as a result of changing streamflow and tidal conditions was analyzed. This report documents the development of the Savannah Chloride Model Decision Support System (SCM DSS) and provides examples of applying the SCM DSS to simulate chloride response caused by modifications to Savannah Harbor.

An important part of the USGS mission is to provide scientific information for the effective water-resources management of the Nation. To assess the quantity and quality of the Nation's surface water, the USGS collects hydrologic and water-quality data from rivers, lakes, and estuaries by using standardized methods (Rantz and other, 1982; Wagner and others, 2000) and maintains the data from these stations in national databases. Often these databases are underutilized for addressing contemporary hydrologic issues. The techniques presented in this report demonstrate how information can be extracted from disparate databases and used to assist local, State, and Federal agencies. The application of data-mining techniques, including the application of ANN models, to simulate chloride concentrations in Abercorn Creek demonstrates how empirical models of complex hydrologic systems can be developed, disparate databases and models can be integrated, and study results can be easily disseminated to meet the needs of a broad range of end users. The results of this investigation also demonstrate how the extrapolation of models to conditions much greater than historical conditions can have substantial effects on predicted results and how the need for extrapolation can be accommodated in the model development.

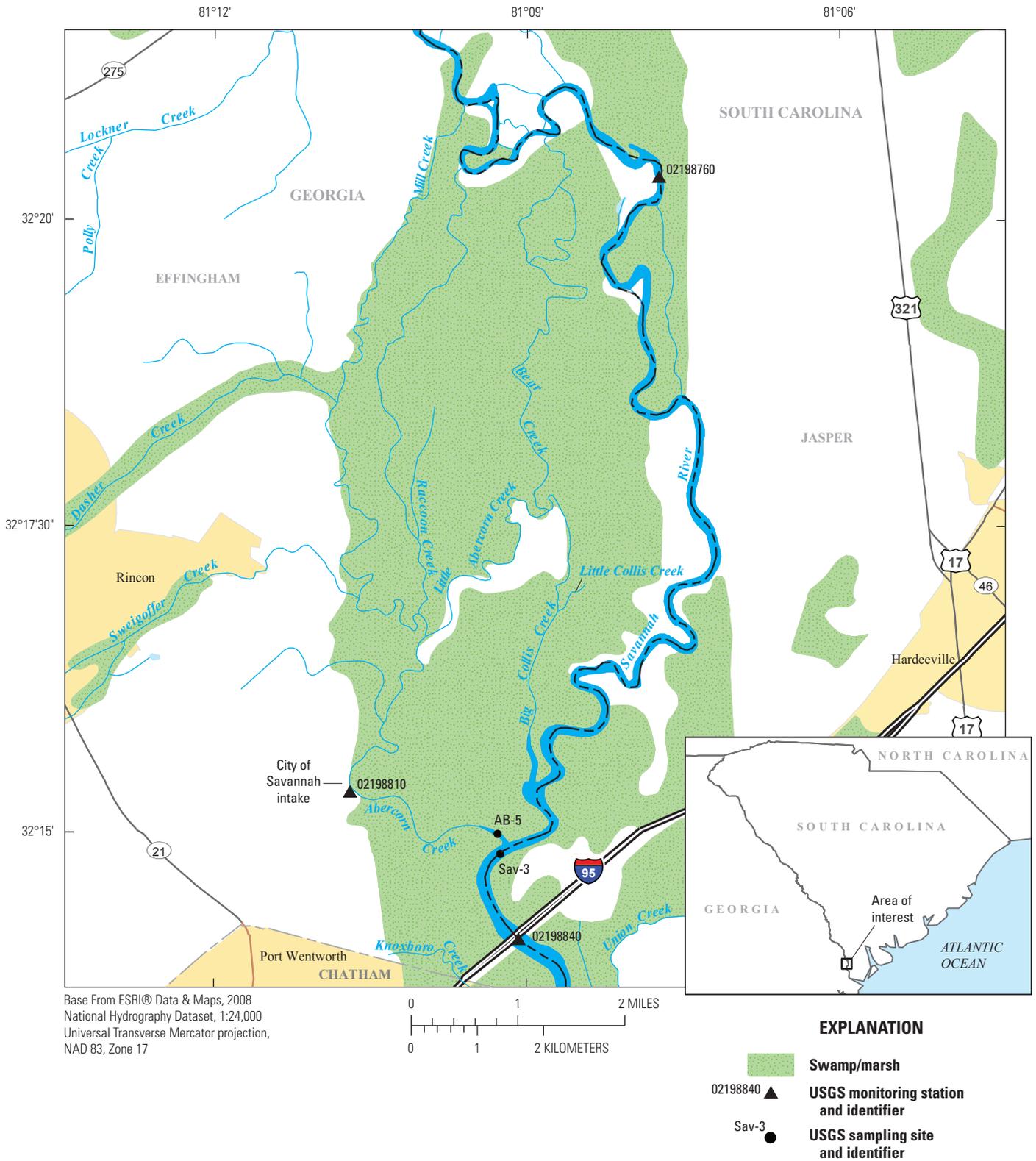
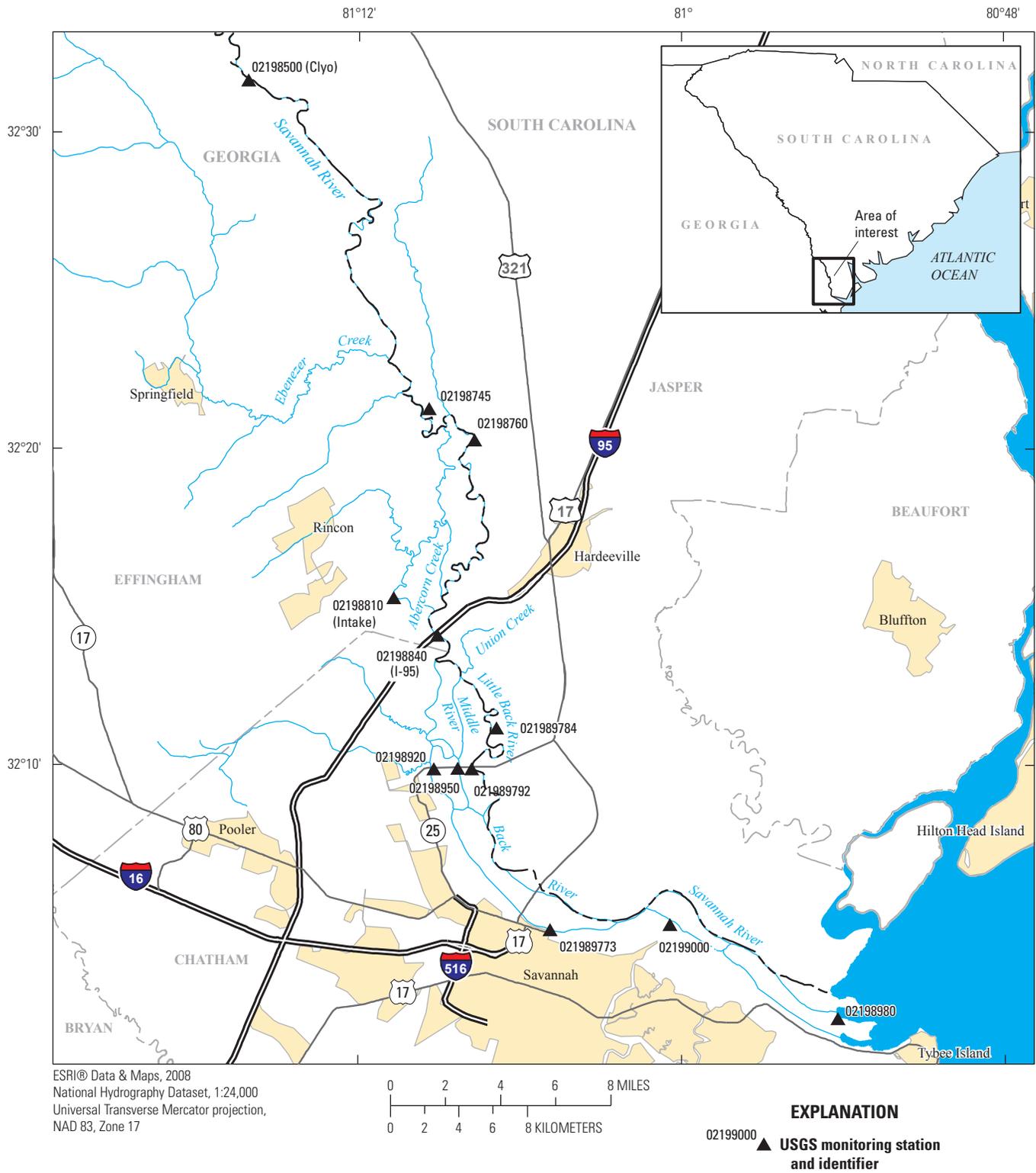


Figure 1. The Abercorn Creek study area, Effingham County, Georgia.

4 Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009



**Figure 2.** U.S. Geological Survey continuous streamflow monitoring network for the lower Savannah River. The lower Savannah River estuary extends from the I-95 bridge to the Atlantic Ocean.

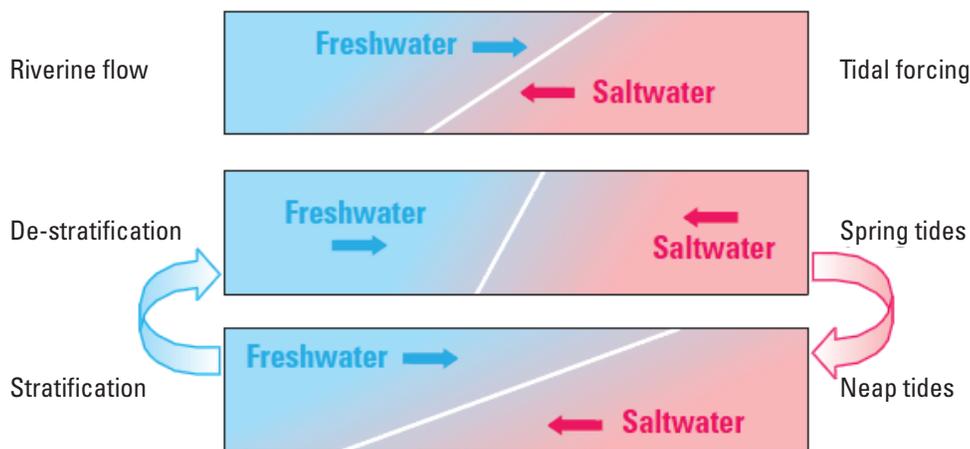
## Description of the Study Area

Complex estuarine and freshwater tidal systems are constantly responding to changing hydrologic, tidal, and meteorological conditions. Dyer (1997) stated that the challenge of studying estuaries is "... that river flow, tidal range, and sediment distribution are continually changing and this is exacerbated by the continually changing weather influences. Consequently, some estuaries may never really be steady-state systems; they may be trying to reach a balance they never achieve." The estuarine portions of the Savannah River (Interstate 95 [I-95] to the Atlantic Ocean) and the tidal freshwater Abercorn Creek are constantly integrating the changing streamflow of the Savannah River Basin, changing tidal conditions of the Atlantic Ocean, and changing meteorological conditions, including wind direction and speed, rainfall, low- and high-pressure systems, and hurricanes. The location of the saltwater-freshwater interface is a balance between upstream river flows and downstream tidal forcing (fig. 3). During periods of high streamflow, it is difficult for salinity to intrude upstream, and the saltwater-freshwater interface is moved downstream toward the ocean. During periods of low streamflow, salinity is able to intrude upstream, and the saltwater-freshwater interface is moved upstream by tidal forcing—either by an increase in mean water levels or a change in tidal range, or a combination of the two.

The Savannah River estuary is considered a partially stratified system with large differences in surface and bottom salinities occurring during neap and spring tides over the 14- and 28-day cycles (fig. 3). During spring tides (tides with the largest tidal range), there is increased energy in the system and mixing of less dense freshwater of the river and denser saltwater of the harbor. The mixing results in smaller variation in vertical salinity concentrations. During neap tides (tides with the smallest tidal range), there is decreased energy in the

system and less mixing between the freshwater and saltwater. The decreased mixing allows the freshwater to flow downstream over the saltwater intruding upstream. The decrease in mixing results in an increased salinity gradient from the surface to the bottom of the water column and increased salinity intrusion upstream. The partial stratification of the Savannah River estuary occurs downstream from the I-95 bridge gage. In the upper reaches of the estuary and upstream from I-95, the system is well mixed.

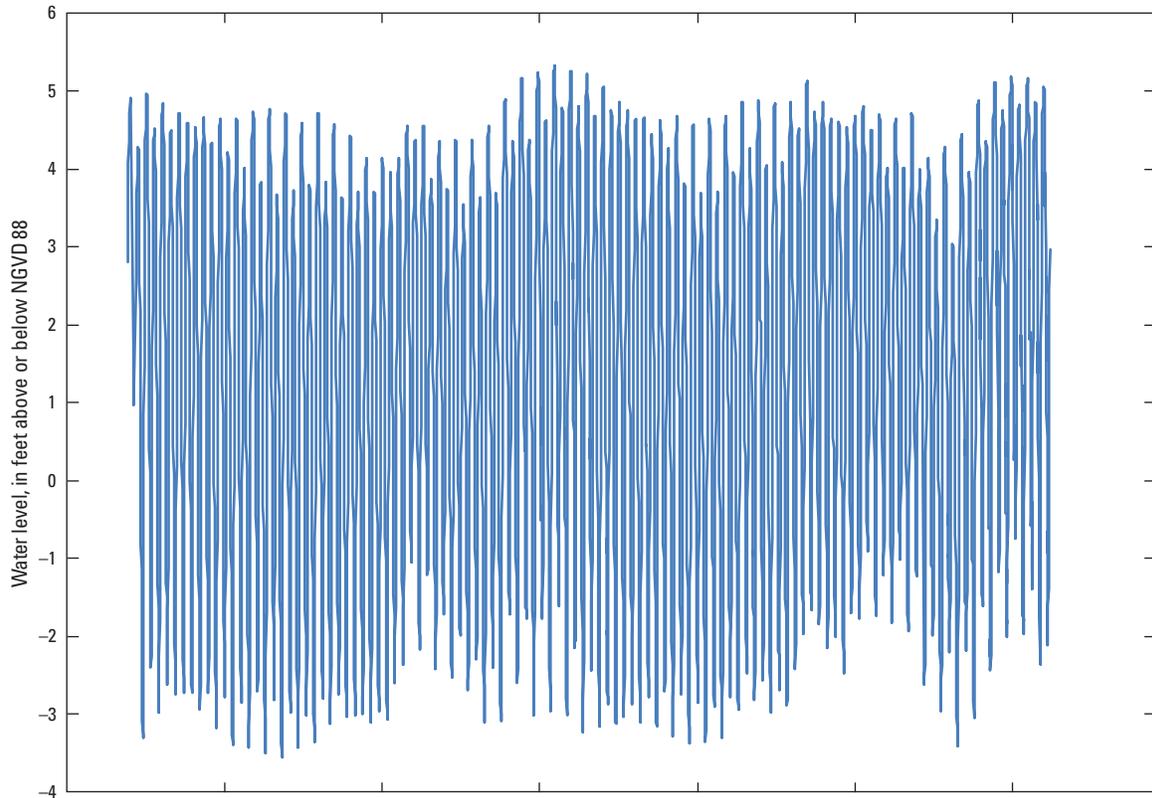
Abercorn Creek is a tributary to the Savannah River located in the coastal plain of Georgia (figs. 1, 2). The confluence of Abercorn Creek with the Savannah River is approximately 1 mile upstream from the I-95 bridge. Water enters Abercorn Creek from Bear Creek, which receives streamflow from the Savannah River, approximately 13 river miles upstream from the confluence of Abercorn Creek and the Savannah River (fig. 1). In the reach of the Savannah River between the headwaters of Bear Creek and the confluence of Abercorn Creek, the tidal fluctuations are dampened out. The intake for the City of Savannah Water Plant is located approximately 2 miles from the confluence with the Savannah River. Abercorn Creek is a tidal freshwater system with semidiurnal tides and salinity concentration less than 0.5 psu. The USGS maintains a real-time gaging station near the intake (station 02198810; figs. 1, 2) that records gage height, stream velocity, specific conductance, temperature, and precipitation at 15-minute intervals ([http://waterdata.usgs.gov/ga/nwis/uv/?site\\_no=02198810](http://waterdata.usgs.gov/ga/nwis/uv/?site_no=02198810)). Gage heights at the intake vary between approximately -3.5 and 5.5 feet (ft North American Vertical Datum 1988), and reversing tidal streamflows are between -3,800 and 3,500 cubic feet per second (ft<sup>3</sup>/s; fig. 4). Specific conductance, a field measurement that can be used to compute salinity, typically is less than 150 microsiemens per centimeter ( $\mu\text{S}/\text{cm}$ ) at 25 degrees Celsius.



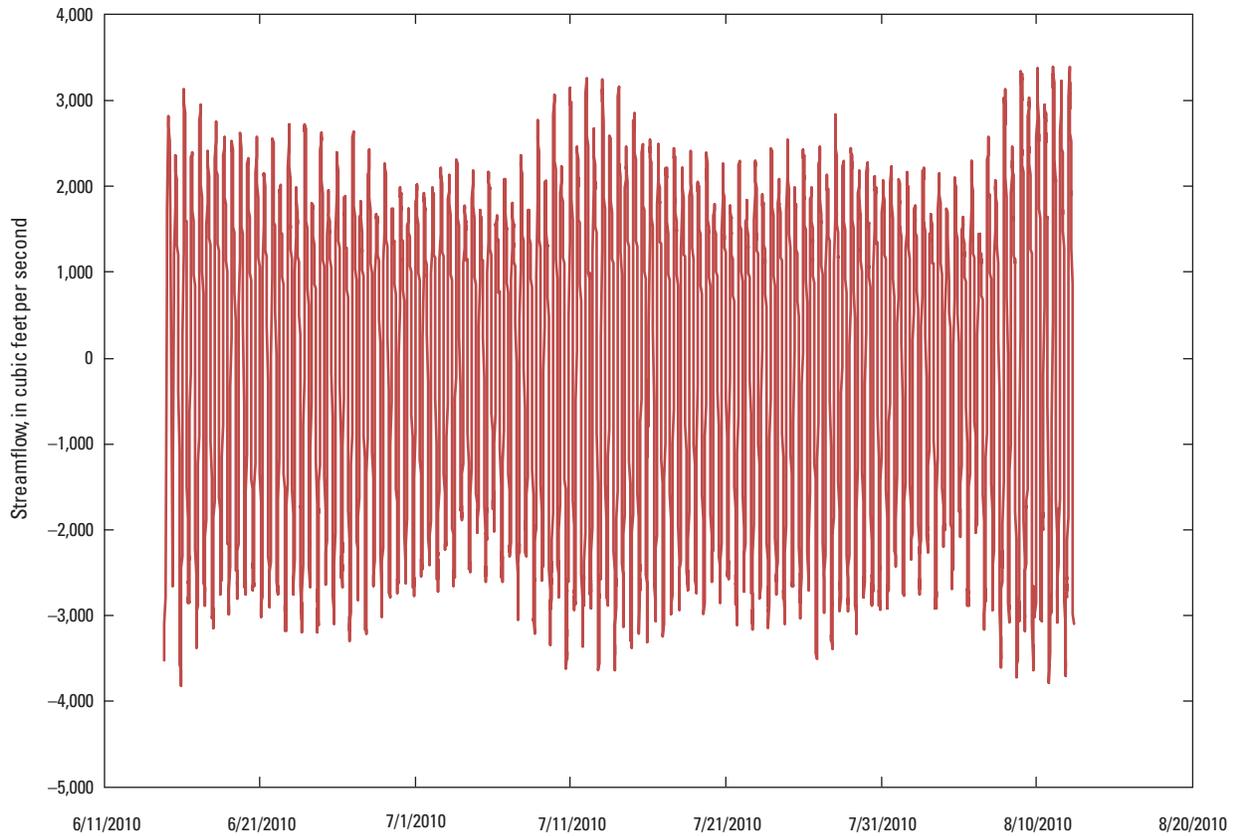
**Figure 3.** Conceptual model of the location of the freshwater-saltwater interface and salinity stratification-de-stratification cycle in estuarine rivers (from Conrads and others, 2006).

6 Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009

A.



B.



**Figure 4.** (A) Gage heights and (B) streamflow for Abercorn Creek near Savannah, Georgia (station 02198810, Intake), for the period June 14 to August 12, 2010.

## Previous Studies

Numerous ecological and hydrologic studies have been conducted to evaluate the potential effects of the proposed deepening of Savannah Harbor (Collins and others, 2001; Will and Jennings, 2001; Conrads and others, 2006; Tetra Tech, 2006a; Welch and Kitchens, 2006). Tetra Tech (2006b) analyzed the available chloride data to predict chloride concentrations at the intake on Abercorn Creek in response to downstream harbor modifications. Results from the Tetra Tech (2006b) study were used to develop a relation between upstream flow and chloride measurements at the water plant intake, investigate the effect of past harbor deepening on chloride levels at the intake, identify other potential sources of chloride, and develop a predictive model for chloride concentrations. The scarcity of data was noted during the Tetra Tech study, and associated uncertainties in the modeling results were recognized by technical reviewers.

Tetra Tech (2010) addressed technical reviewer's concerns with additional data and changes to the Environmental Fluid Dynamic Code (EFDC) model previously developed for evaluating harbor deepening scenarios (Hamrick, 1992; Tetra Tech, 2006a) to simulate chloride concentrations in Abercorn Creek. Tetra Tech (under a separate funding agreement with the USACE0–Savannah District) modified the EFDC model of the lower Savannah River to include Bear, Little Collis, Big Collis, Little Abercorn, and Abercorn Creeks (fig. 1). Chloride concentrations at the water plant intake were modeled as a conservative tracer directly with EFDC. The EFDC model uses boundary input data of streamflow, riverine and harbor chloride concentrations, and coastal water levels. Output from the EFDC model can be used as input for the ANN models in the DSS.

The emerging field of data mining involves extracting information from large databases. Data mining encompasses several technologies, including signal processing, advanced statistics, multidimensional visualization, chaos theory, and machine learning. Machine learning is a field of artificial intelligence in which computer programs are developed that automatically learn cause-effect relations from example cases and data. For numerical data, commonly used methods of machine learning include ANNs, genetic algorithms, multi-variate adaptive regression splines, and partial and ordinary least squares.

Previous studies by the authors and others have used data-mining techniques to predict hydrodynamic and water-quality behaviors in the Beaufort, Cooper, Savannah, and Waccamaw River estuaries of South Carolina and Georgia (Roehl and others, 2000; Conrads and others, 2003; Conrads and others, 2006; Conrads and Roehl, 2007) and stream temperatures in western Oregon (Risley and others, 2003). ANN models have also been used successfully over finite temporal resolutions (hourly or less) to simulate wind conditions and lake hydrodynamics (Buccola and Wood, 2010). These studies have demonstrated that ANN models, combined with data-mining techniques, can provide an effective approach for simulating complex hydrologic systems.

## Approach

The variability of specific conductance data and chloride concentrations in Abercorn Creek is a result of many factors, including streamflow and tidal conditions. Empirical and mechanistic modeling approaches were funded by the USACE to evaluate chloride dynamics on Abercorn Creek and to compare the results of the two modeling approaches. The empirical modeling approach was to develop ANN models to simulate specific conductance and chloride concentrations at the water plant intake. The empirical modeling approach used correlation functions that were synthesized directly from data to predict how specific conductance and chloride concentration at the City of Savannah intake respond to changing streamflow and tidal conditions. Continuous hydrologic datasets as well as sampling data collected at the water plant intake were available for Abercorn Creek and the Savannah River at the I–95 bridge. Empirical specific-conductance and chloride models were developed directly from these data by using data-mining techniques, linear regression, and ANN models.

The application of data-mining techniques to develop empirical models to simulate the specific conductance data and chloride concentration was undertaken in three phases: (1) obtaining and evaluating the suitability of the hydrologic and water-quality data for developing empirical models; (2) developing models to simulate the specific conductance and chloride concentration at the intake; and (3) developing a DSS that integrates historical databases, linear regression and ANN models, model controls, and model output into a spreadsheet application with a graphical user interface that allows the user to simulate scenarios of interest. Output from the EFDC model (Tetra Tech, 2010) can be used as input for the ANN models in the DSS.

## Data-Collection Networks

Many resource entities have collected data in the lower Savannah River, including the USGS, National Oceanic Atmospheric Administration, U.S. Environmental Protection Agency, Georgia Environmental Protection Division, South Carolina Department of Health and Environmental Control, the City of Savannah, the Georgia Ports Authority, and local colleges and universities. For this study, continuous data from the USGS network and discrete water-quality samples collected by the City of Savannah were used for analysis and to develop an empirical model and to calibrate mechanistic models.

The USGS has maintained a network of continuous streamflow, water-level, and water-quality monitors in the lower Savannah River since the mid-1980s. The current (2010) monitoring network is shown in figure 2, and stations are listed in table 1. In October 2009, a continuous monitoring station was installed on Abercorn Creek near the water plant intake (station 02198810) to record water level, velocity

## 8 Simulation of Specific Conductance and Chloride Concentration in Abercorn Creek, Georgia, 2000–2009

**Table 1.** U.S. Geological Survey continuous river gaging network (2011) for the lower Savannah River.

[NAD 83, North American Datum of 1983; Q, flow; DOR, date of report; WL, water level; SC, specific conductance; T, temperature; V, velocity; I, Interstate Highway; GA, Georgia; Precip, precipitation; USACE, U.S. Army Corps of Engineers; bold text are the station names used in this study]

Station number	Station name	Recorded physical properties	Period of record	Longitude (decimal degrees, NAD 83)	Latitude (decimal degrees, NAD 83)
02198500	Savannah River near Clyo, Georgia ( <b>Clyo</b> )	Q	October 1929 – DOR	–81.269	32.528
02198745	Savannah River near Rincon, Georgia	WL, SC, T	April 2009 – June 2010	–81.162	32.353
02198760	Savannah River above Hardeeville, South Carolina	WL	October 1987 – DOR	–81.129	32.339
02198810	Abercorn Creek near Savannah, Georgia ( <b>Intake</b> )	V, SC, WL, Q	October 2009 – DOR	–81.178	32.256
02198840	Savannah River near Port Wentworth, Georgia ( <b>I–95</b> )	WL, SC, T, Precip	June 1986 – DOR	–81.151	32.236
02198920	Savannah River at GA 25 at Port Wentworth, Georgia	WL, SC	October 1987 – DOR	–81.154	32.166
02198950	Middle River at GA 25 at Port Wentworth, Georgia	V, SC, WL, Q	November 2008 – DOR	–81.138	32.166
021989773	Savannah River at USACE Dock, at Savannah, Georgia	V, SC, WL, Q	May 2007 – DOR	–81.081	32.081
021989784	Little Back River above Lucknow Canal, near Limehouse, South Carolina	SC, WL	May 1990 – DOR	–81.118	32.186
021989792	Little Back River at GA 25 at Port Wentworth, Georgia	V, SC, WL, Q	November 2008 – DOR	–81.1300	32.166
02198980	Savannah River at Fort Pulaski, Georgia	WL	October 1987 – DOR	–80.903	32.034
02199000	South Channel Savannah River near Savannah, Georgia	WL	October 2009 – DOR	–80.003	32.083

(used to compute tidal streamflow), temperature, and specific conductance data. In addition to the continuous monitoring data, water samples for analysis of chloride concentrations and other water-quality ions have been collected. The City of Savannah has sampled the source water at the intake on a daily basis since the 1980s. Beginning in 2003, these samples were collected at hourly intervals and then composited as a daily sample for analysis of chloride concentrations and analytes (table 2) at the water treatment facility.

Discrete water-quality samples have been collected in the upper portion of the lower Savannah River by the USGS in response to the interest in the salinity and chloride dynamics near the water plant intake. Automatic samplers were installed in the spring of 2009 to collect water samples for chloride and specific conductance analysis at stations 02198745, 02198810, 02198840, and 02198920 (fig. 2). These data were used to evaluate the EFDC model chloride simulations (Tetra Tech, 2010).

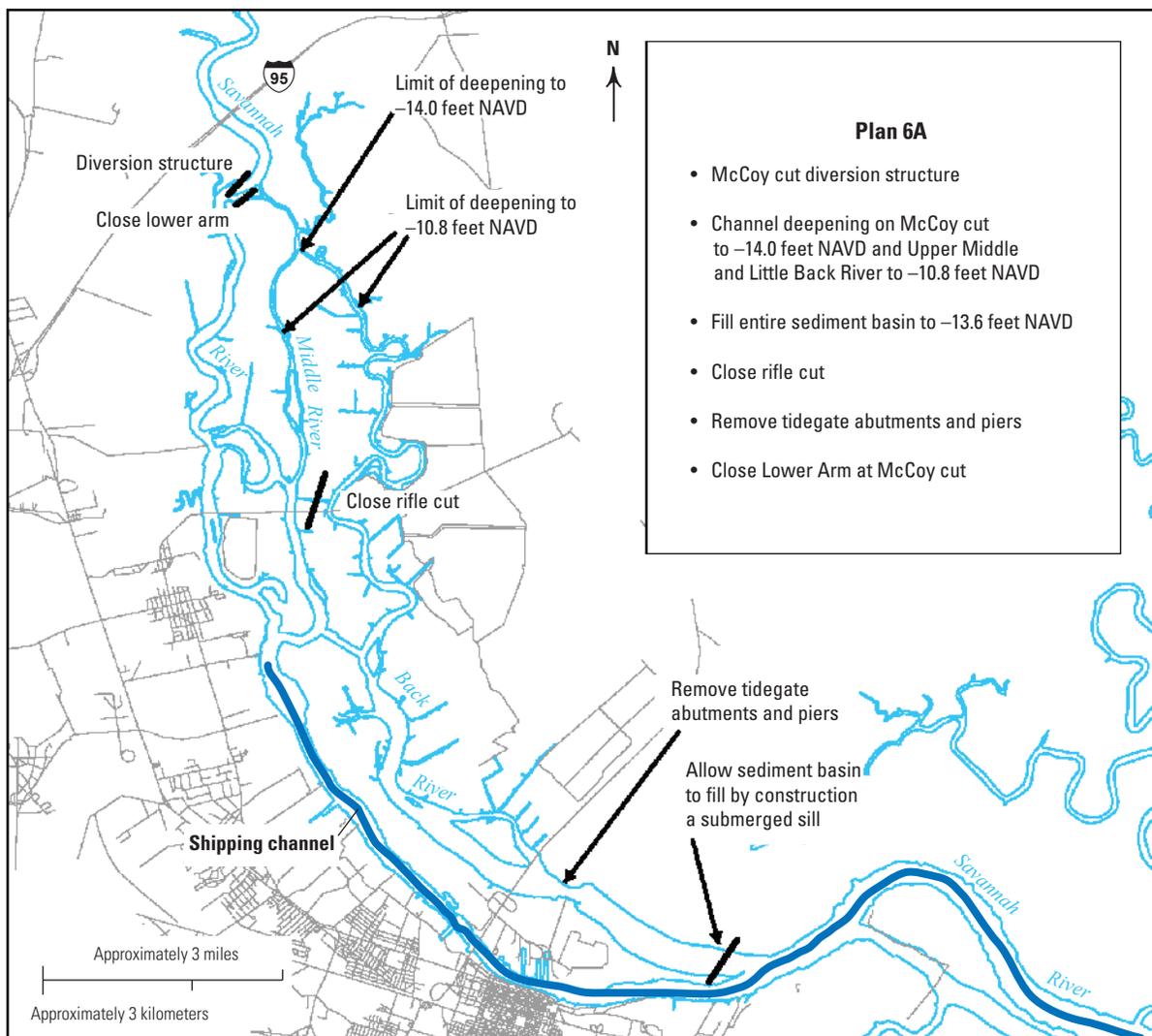
**Table 2.** Water-quality constituents analyzed daily by the City of Savannah at the Abercorn Creek intake.

Alkalinity	Nitrate
Aluminum	Nitrite
Bicarbonate alkalinity	pH
Calcium hardness	Phenolphthalein alkalinity
Carbon dioxide	Phosphate
Carbonate alkalinity	Silica
Chloride	Sodium
Color	Specific conductance
Dissolved oxygen	Sulfates
Fluoride	Total dissolved solids
Hydroxide alkalinity	Temperature
Iron	Total organic carbon
Langelier Saturation Index	Total hardness
Magnesium hardness	Total suspended sediment
Manganese	Turbidity

## Limitation of the Datasets

As with any modeling effort, empirical or mechanistic, the reliability of the model is dependent on the quality of the data and range of measured conditions used for training or calibrating the model. The available period of record for the river data-collection networks and the discrete water-quality data can limit the range of streamflow, water-level, tidal range, salinity, and chloride conditions that the models can accurately simulate. A long period of continuous record and large range of historical conditions are critical for developing accurate empirical models. For these reasons, the long-term, water-quality data collected at the intake by the City of Savannah were used for the empirical model development rather than the short-term data collected in 2009. Although the intake data provide the longest continuous record of chloride concentrations at the intake, these data were not collected for the purpose of analyzing chloride dynamics in the tidally affected Abercorn Creek.

Proposed depths for deepening the shipping channel of the harbor range from 2 to 6 feet. The proposed mitigation plan to minimize the salinity effects associated with the proposed deepening of Savannah Harbor involves major changes to the channel geometries and channel connections of the lower Savannah River estuary below the I-95 bridge (Plan 6A, fig. 5). Although the proposed channel deepening is 2–6 ft, the mitigation plan includes the deepening of channel connections to –14 feet NAVD 1988. These changes would likely affect the correlation between water-quality conditions below the I-95 bridge and the water plant intake. To avoid potential changes in the correlation between stations on the Savannah River and the water quality at the intake, only two long-term stations, which are at or upstream from the I-95 bridge, were used in the development of the empirical models. Streamflow data from station 02198500, Savannah River near Clio, GA, (referred to as Clio in this report) and water-level, specific conductance, temperature, and precipitation data from station 02198840, Savannah River near Port Wentworth, GA,



**Figure 5.** Mitigation Plan 6A (modified from U.S. Army Corps of Engineers, 2010).

(referred to as I–95 in this report) were used for developing empirical models to predict specific conductance data and chloride concentration at the intake.

## Data Preparation

The USGS and City of Savannah databases needed to be merged into one database for analysis and model development. The principal chloride dataset is the City of Savannah daily monitoring data for the water plant intake. The hourly and 15-minute streamflow data from the USGS Clyo gage and water level, specific conductance, and temperature from the USGS I–95 gage needed to be reduced to daily values and merged with the City of Savannah daily data. Tidal systems, such as Abercorn Creek and the lower Savannah River estuary, are highly dynamic and exhibit complex behaviors that evolve over multiple time scales. The complex behaviors of the variables in a natural system result from interactions between multiple physical forces. The semidiurnal tide is dominated by the lunar cycle, which is more influential than the 24-hour solar cycle; thus, a 24-hour average is inappropriate to use to reduce tidal data to daily values. For analysis and model development, the USGS data were digitally filtered to remove semidiurnal and diurnal tidal variability by using nested moving-window averages of 25 and 13 hours. Removing the semidiurnal tidal frequency allows a signal component that lies within a window of frequencies (for example, the 12.4-hour tidal cycle lies between periods of 12.0 to 13.0 hours) to be excised, analyzed, and modeled independently of other components. Digital filtering also can diminish the effect of noise in a signal to improve the amount of useful information that it contains. Working with filtered signals makes the modeling process more efficient, precise, and accurate. The Savannah River at Clyo is not tidally affected, and the daily mean streamflow values were used.

Tidal range (XWL) was computed from the field measurements of the physical properties. Tidal dynamics are a dominant force for estuarine systems, and the tidal range is a significant variable for determining the lunar phase of the tide and flushing dynamics of coastal rivers. Tidal range is calculated from water level (or gage height) and is defined as the water level at high tide minus the water level at low tide for each semidiurnal tidal cycle.

## Characterization of Specific Conductance and Chloride Concentration

Chloride is a naturally occurring ion deposited in the earth's soils or dissolved in the oceans. Concentrations found in freshwater can be from weathering of rocks containing chloride ion or from byproducts in anthropogenic sources, such as road salts, fertilizers, and industrial generation. Chloride makes up 1.9 percent of seawater by mass. Tetra Tech (2006b) obtained chloride data from 10 USGS sampling stations (table 3; fig. 6) in the Savannah River Basin to evaluate potential sources of chloride in Abercorn Creek. Chloride concentrations from the six freshwater sites varied from 0.2 to 6.5 mg/L (table 4). Chloride concentrations at two stations in the lower Savannah River estuary were as high as 11,000 mg/L.

Salinity intrudes into the lower Savannah River estuary from the ocean. The location of the saltwater–freshwater interface is determined by a balance between upstream river flows and downstream tidal forcing. During periods of high streamflow, it is difficult for salinity to intrude upstream, and thus, the saltwater–freshwater interface is moved downstream toward the ocean. During periods of low streamflow, salinity

**Table 3.** U.S. Geological Survey stations used for chloride surface-water assessment in Tetra Tech, 2005.

[Stations in bold font were used in this study. Station locations are shown in figure 6]

Station	Station description
02187500	Savannah River near Iva, South Carolina
02189000	Savannah River near Calhoun Falls, South Carolina
02192500	Little River near Mt. Carmel, South Carolina
02196000	Stevens Creek near Modoc, South Carolina
02196838	Butler Creek Reservoir at Fort Gordon, Georgia
02197300	Upper Three Runs near New Ellenton, South Carolina
<b>02198500</b>	<b>Savannah River near Clyo, Georgia (Clyo)</b>
<b>02198840</b>	<b>Savannah River near Port Wentworth, Georgia (I–95)</b>
02198920	Savannah River at GA 25 at Port Wentworth, Georgia
02198980	Savannah River at Fort Pulaski, Georgia



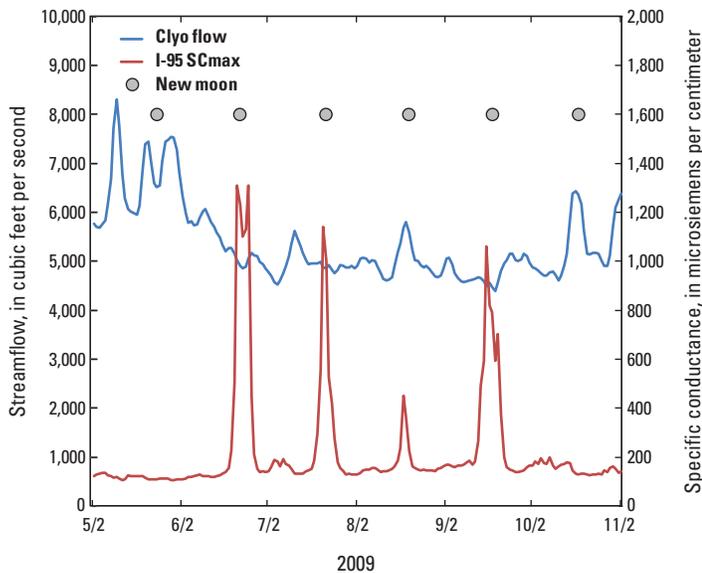
**Table 4.** Minimum, maximum, and mean chloride concentration at selected U.S. Geological Survey stations.

[Modified from Tetra Tech, 2005]

Station number (fig. 6)	Number of observations	First year of collection	Last year of collection	Minimum	Maximum	Mean
				milligrams per liter		
02187500	34	1957	1972	1.4	4.7	2.44
02189000	21	1956	1972	1.2	4.4	2.54
02192500	2	1959	1961	2	2	2
02196000	1	1961	1961	6.5	6.5	6.5
02196838	1	1999	1999	3.68	3.68	3.68
02197300	150	1967	1993	0.2	3.7	2.22
02198920*	33	1958	2003	3.2	6,900	2,239
02198980*	209	1960	1960	3.8	11,000	4,848

\* Stations are located in the estuary.

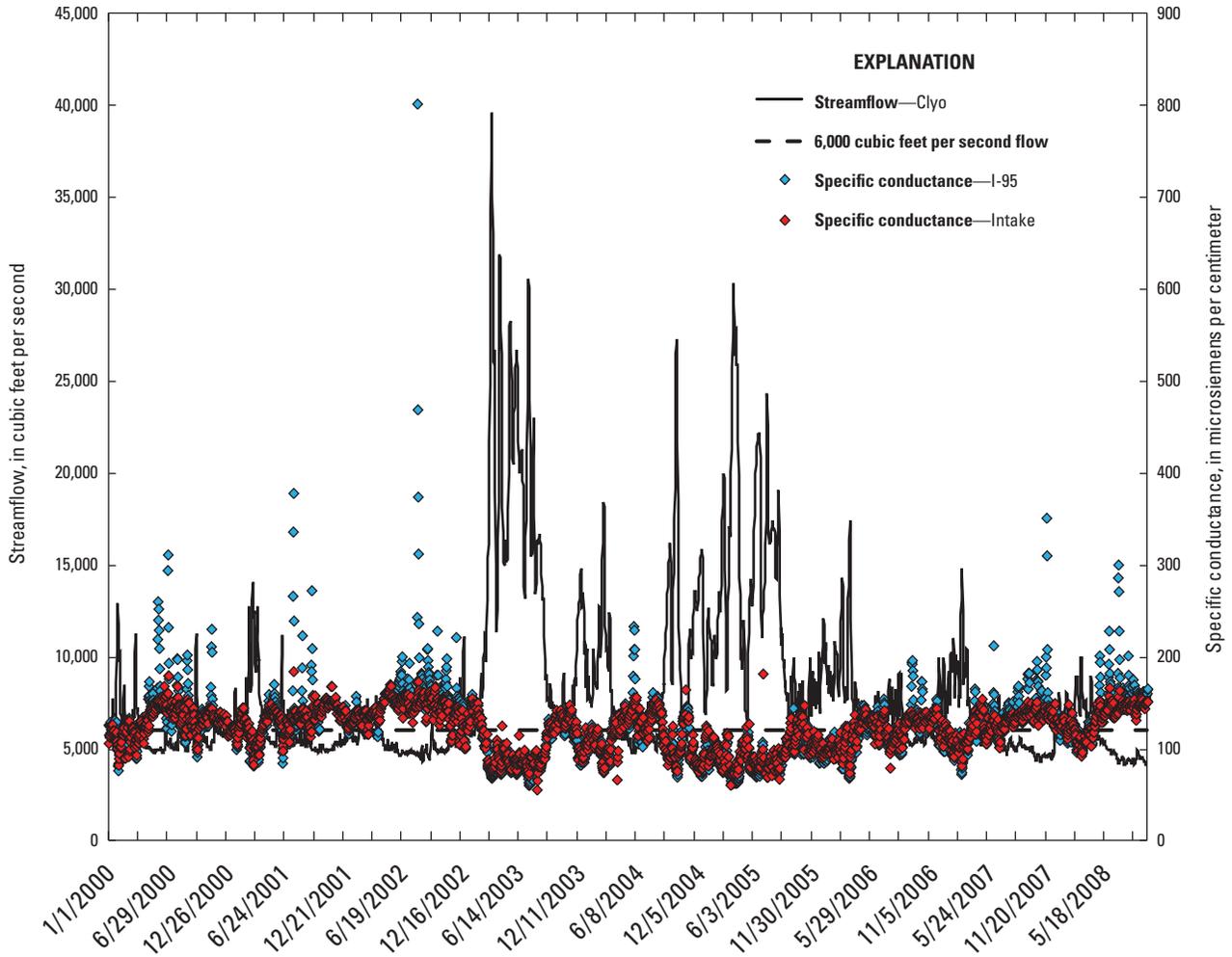
is able to intrude upstream, and subsequently, the saltwater-freshwater interface is moved upstream. Historically, streamflows at Clio range from 5,000 to 50,000 ft<sup>3</sup>/s (U.S. Geological Survey, 2011). Salinity in the Savannah River estuary varies in response to changing streamflow and tidal conditions. The daily maximum specific conductance at the I-95 gage and daily mean streamflow for the Clio gage and the dates of the new moon for the summer of 2009 are shown in figure 7. The data show that there is a convergence of conditions needed for the elevated specific conductance (the field measurement for computing salinity): Savannah River streamflow must be less than 6,000 ft<sup>3</sup>/s at the occurrence of spring tide of the



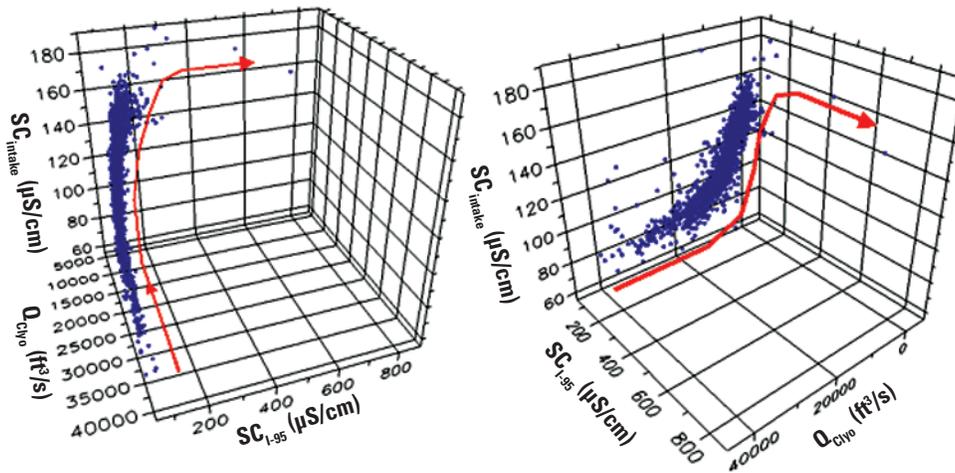
**Figure 7.** Daily streamflow at Savannah River near Clio (station 02198500) and daily maximum specific conductance at the Savannah River near Port Wentworth, Georgia (station 02198840, I-95 bridge), for the period May to October 2009, including dates of the new moon.

new moon. During the new moon, tidal ranges are greatest when the gravitation attraction of the sun and moon are aligned. The increased tidal energy under these conditions typically leads to a salinity intrusion event. Meteorological conditions, such as tropical storms and high winds, may exacerbate salinity intrusion.

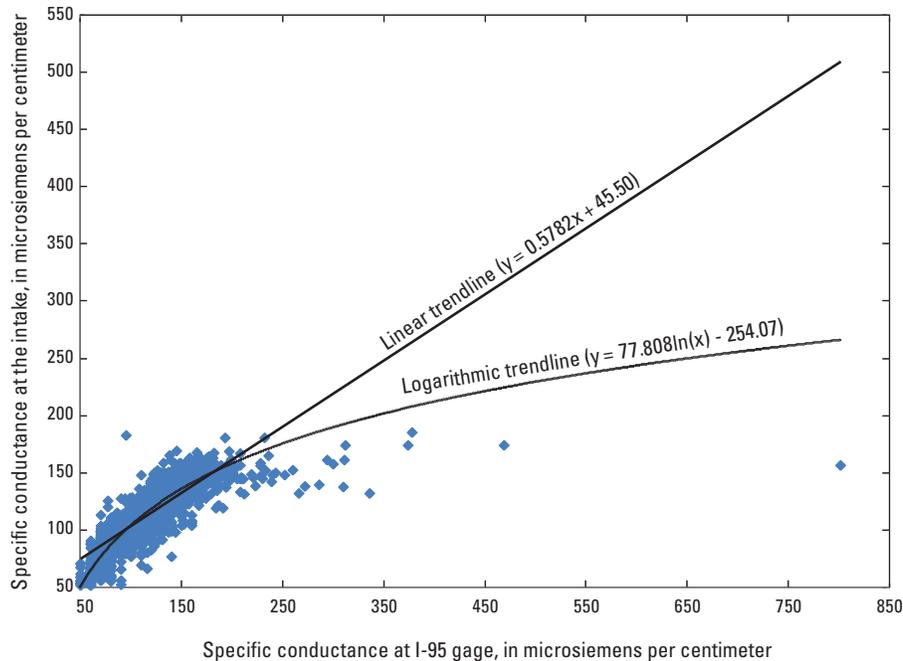
The salinity dynamics at the City of Savannah’s intake, approximately 3 miles from the I-95 gage, differ substantially from the dynamics at the I-95 bridge. The spikes in salinity intrusions at I-95 are dampened at the intake. Figure 8 shows 9 years of streamflow at the Clio gage and specific conductance data at the I-95 gage and the intake. The specific conductance at the I-95 gage shows rapid increases when streamflow decreases below 6,000 ft<sup>3</sup>/s. The specific conductance at the intake does not show rapid increases but rather gradual increases. The dampening of the salinity intrusion also can be seen in a 3D scatter plot of streamflow at the Clio gage and specific conductance at the I-95 gage and the intake (fig. 9). Specific conductance at the intake is shown on the vertical z-axis, and Clio streamflow and specific conductance at I-95 are shown on the horizontal x- and y-axes. The scatter plot on the right is a rotation of the scatter plot on the left. The plots show that as streamflow at Clio decreases, specific conductance at I-95 increases to nearly 800 μS/cm, whereas at the intake the specific conductance values are less than 200 μS/cm. Visually, the specific conductance response to low flows at the intake also shows a substantial nonlinear component with a “plateau” of specific conductance values with decreased streamflow. The red arrows in figure 9 are a visual 3D approximation of the nonlinear component. A scatter plot of the specific conductance at I-95 and the intake provide a two-dimensional view of the nonlinear relation between specific conductance at the two sites (fig. 10). Linear and nonlinear (logarithmic) trendlines were fit to the data and show the



**Figure 8.** Flows from Savannah River at Clyo, Georgia, and specific conductance at Savannah River near Port Wentworth, Georgia (I-95 bridge), and water-supply intake on Abercorn Creek for the period February 2000 to October 2009.



**Figure 9.** Three-dimensional scatter plots of specific conductance at the intake ( $SC_{Intake}$ ) on Abercorn Creek, Savannah River at the I-95 bridge ( $SC_{I-95}$ ), and flows at Savannah River at Clyo, Georgia ( $Q_{Clyo}$ ). Scatter plot on the right is a 90-degree rotation of the scatterplot on the left. The red arrow is an approximation of the trend of the specific conductance trend in relation to specific conductance at the I-95 bridge and streamflow. [ $\mu\text{S}/\text{cm}$ , microsiemens per centimeter;  $\text{ft}^3/\text{s}$ , cubic feet per second]



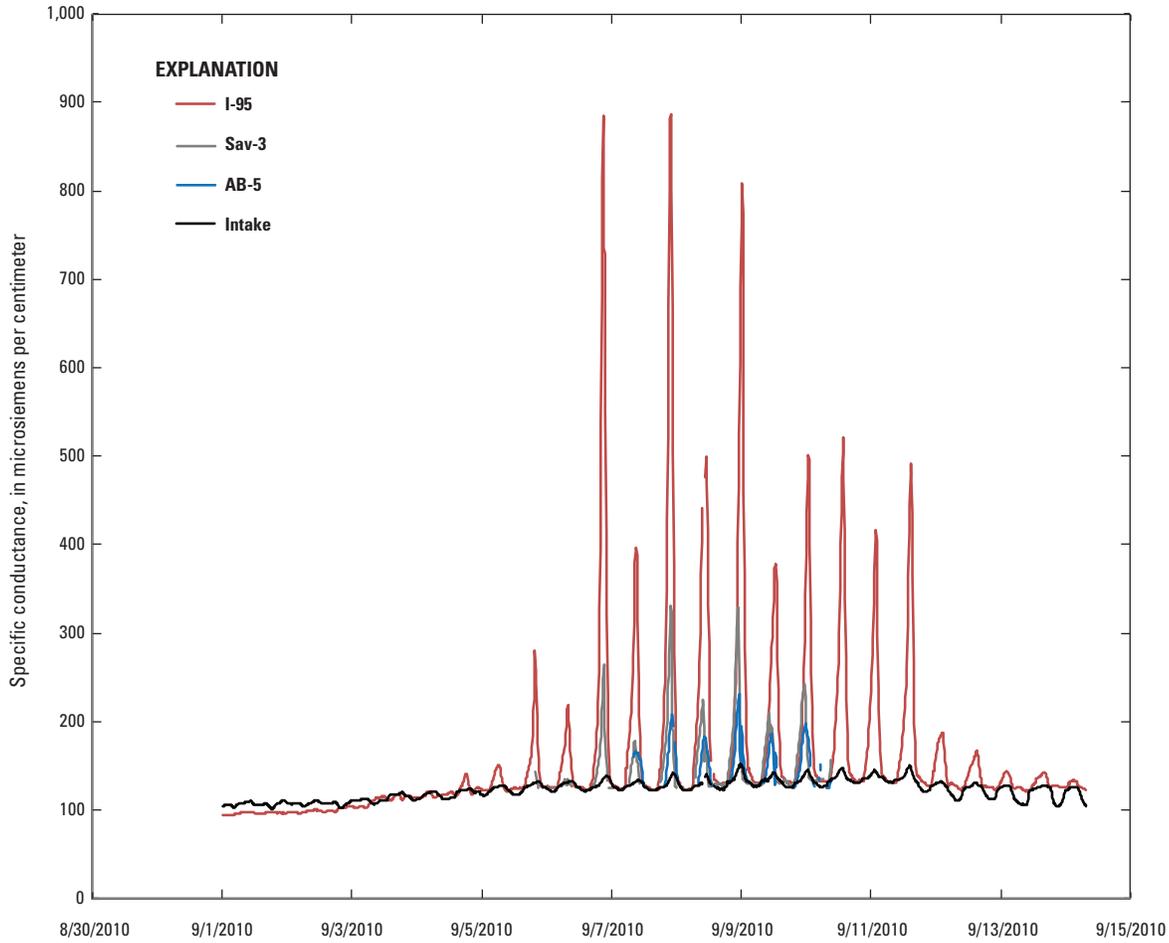
**Figure 10.** Scatter plots of the specific conductance at the City of Savannah’s intake and at the Savannah River near Port Wentworth, Georgia (I–95) gage.

influence of the few data points where specific conductance is greater than 300  $\mu\text{S}/\text{cm}$  at the I–95 gage.

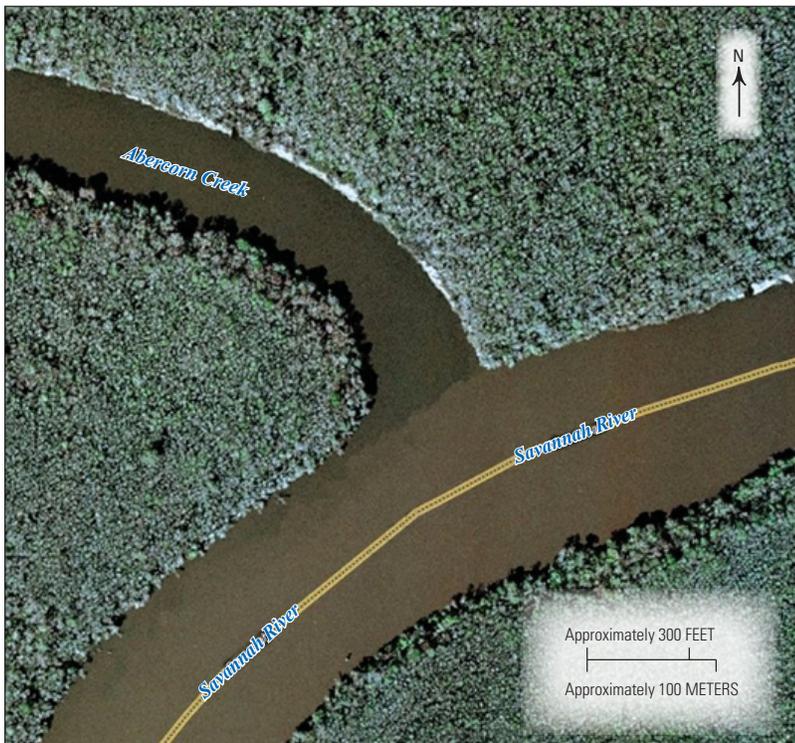
Specific conductance probes were deployed at two sites near the confluence of Abercorn Creek and the Savannah River to collect data during the salinity intrusion event in September 2010 (fig. 1; John Joiner, U.S. Geological Survey, unpub. data, 2010). The data show that there is substantial dampening of the specific conductance intrusion in the Savannah River between I–95 and Station Sav-3 and additional dampening between the Savannah River and the mouth of Abercorn Creek (fig. 11). An aerial photograph (date of photograph unknown) of the confluence of Abercorn Creek and the Savannah River shows a color difference between the two reaches and that there is not complete mixing of Abercorn Creek and Savannah River waters (fig. 12). The natural dampening of the salinity intrusion in Abercorn Creek may be due to differences in channel geometries and channel depths at the confluence of Abercorn Creek and the Savannah River, watershed dynamics

in Abercorn and Bear Creeks, differences in slopes between Abercorn Creek and the Savannah River, or differences in water temperature between Abercorn Creek and the Savannah River.

Chloride concentration at the water treatment plant appears to follow a similar dynamic to specific conductance with increasing concentration during periods of low streamflow and decreasing concentration during periods of high streamflow (fig. 13). However, a scatter plot of specific conductance and chloride concentration at the intake shows substantial variability between the two parameters (fig. 14). The correlation between the two parameters is low with a coefficient of determination ( $R^2$ ) of 0.57, indicating that only 57 percent of the variability in chloride concentrations is explained by specific conductance and a large portion of the variability in chloride concentration is not explained by salinity intrusion in the upper reaches of the lower Savannah River estuary.



**Figure 11.** Specific conductance at two stations on the Savannah River (I-95 and Sav-3) and two stations on Abercorn Creek (AB-5 and Intake) for September 1 to 15, 2010. See figure 1 for station locations.



**Figure 12.** Aerial photograph of the confluence of Abercorn Creek and the Savannah River showing the difference in water color of the two streams.

Base satellite image from U.S. Geological Survey, USDA Farm Service Agency, GeoEye, DigitalGlobe, and Google Maps, 2011

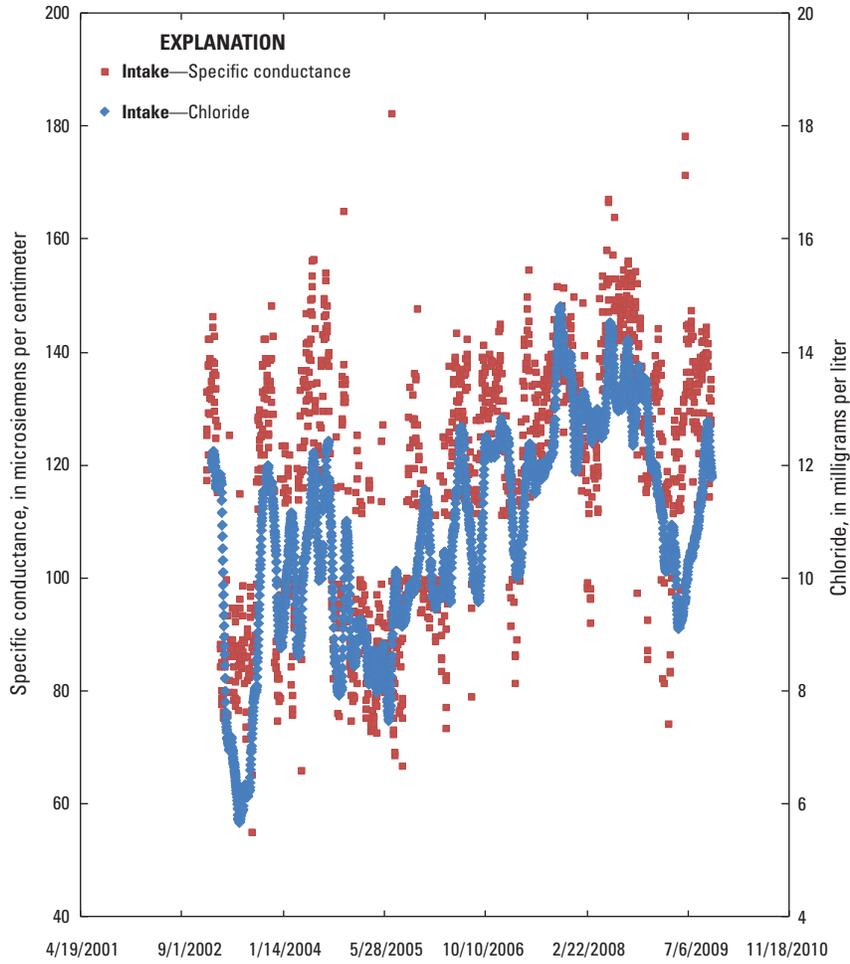


Figure 13. Specific conductance and chloride values at the intake on Abercorn Creek for the period January 2003 to October 2009.

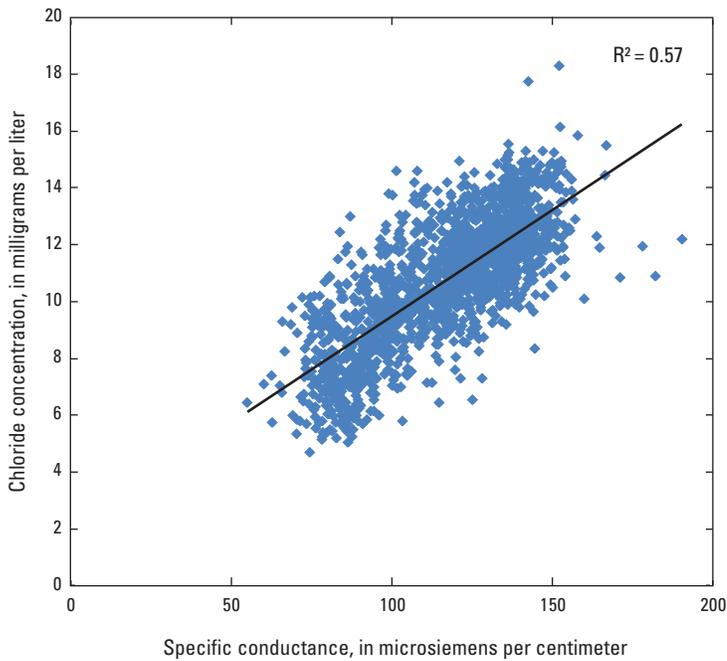


Figure 14. Scatter plot of chloride and specific conductance values at the intake on Abercorn Creek.

## Simulation of Specific Conductance and Chloride Concentrations

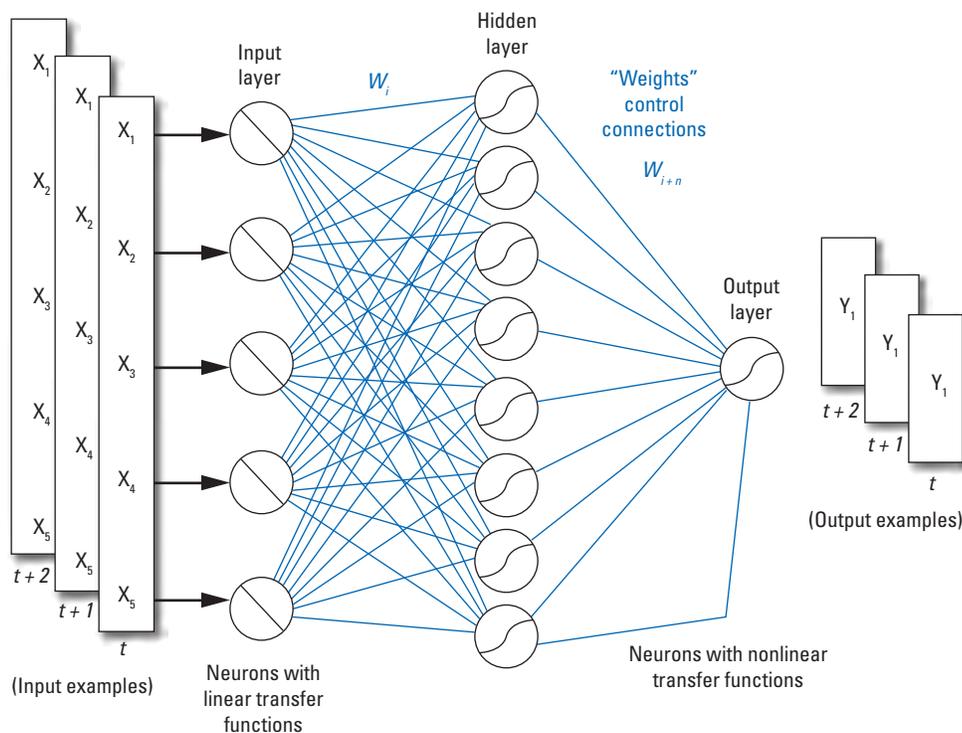
Simulating riverine and estuarine systems often is done by using dynamic mechanistic (or deterministic) models that incorporate the mathematical descriptions of the physics of the hydrodynamics and water chemistry of surface-water systems. These 1-, 2-, or 3D models generally require extensive data collection and are time consuming to apply to estuarine systems. Although mechanistic models have been the state of the practice for regulatory evaluations of anthropogenic effects on hydrologic systems, developments in the field of advanced statistics, machine learning, and data mining offer opportunities to develop empirical ANN models that often are more accurate than multidimensional mechanistic models. Conrads and Roehl (1999) compared the application of a mechanistic model and an ANN model to simulate dissolved-oxygen concentrations in the tidally affected Cooper River in South Carolina. They found that the ANN models offer some advantages, including faster development time, utilization of larger amounts of data, the incorporation of optimization routines, and model dissemination in spreadsheet applications.

### Artificial Neural Network Models

Models generally fall into one of two categories: mechanistic (or deterministic) or empirical. Mechanistic models are created from first-principles equations, whereas empirical modeling adapts generalized mathematical functions to fit a

line or surface through data from one or more variables. The most common empirical approach is ordinary least squares, which relates variables using straight lines, planes, or hyperplanes, whether the actual relations are linear or not. Calibration of either type of model attempts to optimally synthesize a line or surface through the measured data. Calibrating models is difficult when data have substantial measurement error or are incomplete, or when the variables for which data are available provide only a partial explanation of the sources of variability. The principal advantages of empirical models, such as ANN models, over mechanistic models are that they can be developed faster and are more accurate provided that the modeled systems are well characterized by data. Empirical models, however, are prone to problems when poorly applied. Overfitting and multicollinearity caused by correlated input variables can lead to invalid mappings between input and output variables (Roehl and others, 2003).

An ANN model is a flexible mathematical structure capable of describing complex nonlinear relations between input and output datasets. The structure of ANN models is loosely based on the biological nervous system with interconnections of neurons and synapses (Hinton, 1992). Although numerous types of ANN models exist, the most commonly used type of ANN is the multilayer perceptron (Rosenblatt, 1958). As shown in figure 15, multilayer perceptron ANNs are constructed from layers of interconnected processing elements called neurons with each executing a simple “transfer function.” All input layer neurons are connected to all hidden layer neurons, and all hidden layer neurons are connected to all output neurons. Multiple hidden layers are possible, but a single layer is sufficient for most problems.



**Figure 15.** Multilayer perceptron artificial network architecture (from Conrads and Roehl, 2007).

Typically, linear transfer functions are used to scale input values from the input layer to the hidden layer and generally fall within the range that corresponds to the most linear part of the s-shaped sigmoid transfer functions used from the hidden layer to the output layer (fig. 15). Each connection has a “weight”  $w_i$  associated with it, which scales the output received by a neuron from a neuron in an antecedent layer. The output of a neuron is a simple combination of the values it receives through its input connections and the associated weights, as well as the neuron’s transfer function.

An ANN is “trained” by iteratively adjusting its weights to minimize the error by which it maps inputs to outputs for a dataset composed of input/output vector pairs. Prediction accuracy during and after training can be measured by a number of metrics, including coefficient of determination ( $R^2$ ) and root mean square error (RMSE). An algorithm that is commonly used to train multilayer perceptron ANNs is the back error propagation training algorithm (Rumelhart and others, 1986). Jensen (1994) describes the details of the multilayer perceptron ANN, the type of ANN used in this study. Multilayer perceptron ANNs can synthesize functions to fit high-dimension, nonlinear multivariate data. Devine and others (2003) and Conrads and Roehl (2005) describe the use of multilayer perceptron ANN in multiple applications to model and control combined manmade and natural systems, including disinfection byproduct formation, industrial air emissions monitoring, and surface-water systems affected by point- and nonpoint-source pollution.

Experimentation with a number of ANN model architectural and training parameters is a routine part of the modeling process. For correlation analysis or predictive modeling applications, a number of potential ANN models are trained and evaluated for their statistical accuracy and their representation of process physics. Interactions between combinations of variables also are considered in addition to the selection of the training dataset from the overall dataset. For models with a large dataset with good representation over the range of historical behaviors, a small percentage of the dataset (10–25 percent) may be selected for the training dataset. For models with limited data, a larger percentage (75–100 percent) may be used in the training dataset. In general, a high-quality predictive model can be obtained when:

- The data ranges are well distributed throughout the state space of variables describing the physical system of interest,
- The input variables selected by the modeler share “mutual information” about the output variables,
- The functional form “prescribed” or “synthesized” by the model to “map” (correlate) input variables to output variables is a good one. Machine-learning techniques, like ANN models, synthesize a best fit to the data. Techniques such as ordinary least squares and finite-difference models prescribe the functional form of the model’s fit of the calibration data.

Subdividing a complex modeling problem into sub-problems and then addressing each is an effective means to achieving the best possible results. A collection of submodels whose calculations are coordinated by a computer program constitutes a “super-model.” For the Abercorn Creek study, daily ANN models (submodels) were developed for specific conductance and (or) chloride at the I–95 gage and at the water plant intake. These submodels were then incorporated into a super-model application that integrates the model controls, model database, and model outputs. The super-model for the project is Savannah River Chloride Decision Support System (SCM DSS). The ANN models described in this report were developed using the iQuest™ data-mining software<sup>4</sup> (Version 2.03C DM Rev31). The ANN models were deployed in the DSS using the Visual Basic run-time library of the iQuest R/T™ software.

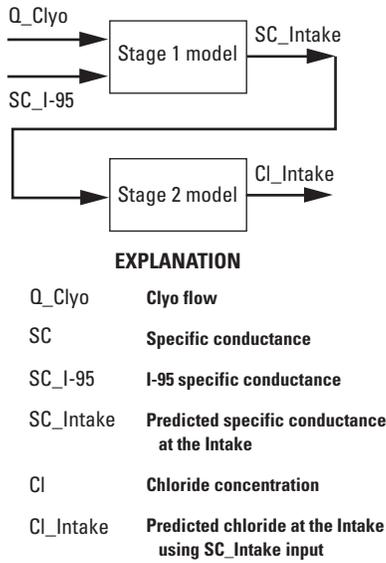
## Development of Specific Conductance and Chloride Models

The USGS and City of Savannah data were used to develop empirical process models of Abercorn Creek and are included in SCM DSS to allow the user to run long-term simulations to evaluate permutations of the actual historical record. The SCM DSS models predict how Savannah River flow and specific conductance affect specific conductance and chloride concentrations at the water plant intake. A two-stage model was used (fig. 16). The first stage predicts specific conductance at the intake, and the second stage predicts chloride concentrations at the intake using the predicted specific conductance at the intake.

Both empirical and mechanistic models are more accurate when interpolating within the historical range of the data used to develop the model than when extrapolating to conditions beyond the range of the data used to develop the model. Salinity simulations with the EFDC model of proposed deepening scenario simulations indicated that estimated maximum salinity concentration at the I–95 gage could be three times greater than the historical maximum. Thus, the development of empirical models to predict chloride concentrations at the intake needed to accommodate large extrapolation from historical conditions.

The 3D scatter plots of streamflow at the Cloy gage and specific conductance at the I–95 gage and the water plant intake provide an indication of how an empirical model may extrapolate (figs. 9, 10). If an ANN model were to fit the data well, it would capture the “plateau” effect wherein there is a large dampening of specific conductance values between I–95 and the intake. In this case, large increases in specific

<sup>4</sup> The iQuest™ software is exclusively distributed by Advanced Data Mining, LLC, 3620 Pelham Road, PMB 351, Greenville, SC 29615–5044, Phone: (864) 201–8679, e-mail: [info@advdatamining.com](mailto:info@advdatamining.com), <http://www.advdatamining.com>.



**Figure 16.** The two-stage model architecture to predict chloride concentrations (Cl\_Intake) at the intake using flow (Q\_Clyo) and specific conductance (SC\_I-95) inputs.

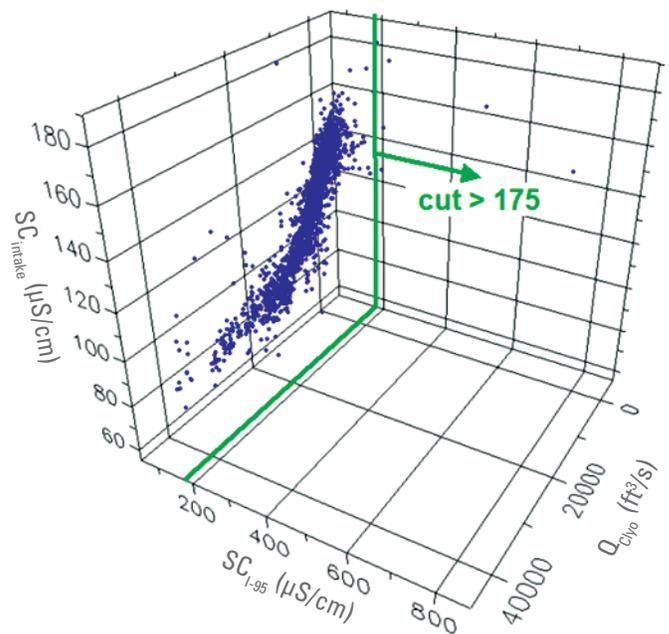
conductance (salinity) at the I-95 gage resulting from a proposed deepening of the harbor would not equate to large increases in salinity and chloride concentrations at the intake.

Only a few data points describe the relation between high specific conductance values at the I-95 gage and the intake (figs. 9, 10). For example, there are less than 10 data points for which specific conductance values at the I-95 gage exceed 300  $\mu\text{S}/\text{cm}$ , and these points exhibit a large degree of scatter. The USACE was concerned that these few data points would determine the extrapolation of the model and potentially underestimate the effect that deepening the harbor could have on chloride concentrations at the intake. To accommodate these concerns, it was decided to develop two specific conductance models for the intake. The first model (ANN M1) would use all the data. The second model (ANN M2) would only use data when specific conductance at I-95 was less than 175  $\mu\text{S}/\text{cm}$  (figs. 17 and 18). Deleting these data from the training dataset for the second model removes the “plateau” effect, and extrapolation by a well-fitted ANN model would show a large increase in specific conductance at the intake with large increases of specific conductance at the I-95 gage.

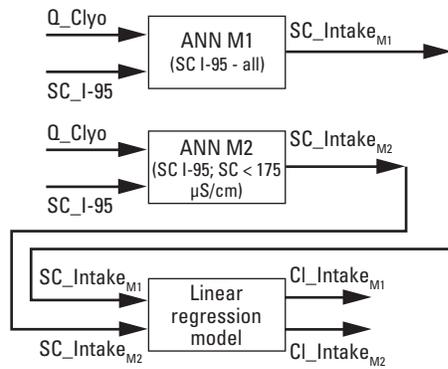
The trendlines shown in figure 10 are an analogous example of how the two models will differ in the results they provide. The ANN M1 model is analogous to the logarithmic trendline and captures the nonlinear relation between the specific conductance at the two sites. The ANN M2 model is analogous to the linear trendline and does not simulate the plateau effect. The ANN M1 model will have a low sensitivity to specific conductance at I-95, and the ANN M2 model will have a high sensitivity to specific conductance at I-95. The

two models will provide two different results for simulating the effect of deepening Savannah Harbor. The models will have to extrapolate to conditions greater than the historical range of conditions. The two models provide a range of specific conductance outcome depending on the validity of the limited number of data points defining the plateau effect and the dampening of specific conductance between I-95 and the intake.

The ANN models used to predict specific conductance values at the intake are effective due to the ability to fit the nonlinear character of the specific conductance and streamflow data. Because of the sigmoid transformation used in ANN models (fig. 15), the ANN models are ineffective in extrapolating linear behaviors. The sigmoid transformation can produce nonlinearities at the predicted extremes of the range of data. The chloride model (stage 2 model, fig. 16) also will be used to extrapolate conditions exceeding the historical range of the data. Because the relation between chloride concentrations and specific conductance at the intake is linear (fig. 14), a linear regression model (fig. 18) was used to predict chloride concentrations at the intake. The linear model allows chloride concentrations to be predicted at highly extrapolated specific conductance values. Details on the inputs to the specific conductance and chloride models are described in subsequent sections of the report.



**Figure 17.** Three-dimensional scatter plots of specific conductance at the City of Savannah’s intake on Abercorn Creek ( $SC_{Intake}$ ), Savannah River near Port Wentworth, Georgia (I-95 bridge,  $SC_{I-95}$ ), flows at Savannah River at Clyo, Georgia ( $Q_{Clyo}$ ). Green line shows the removed specific conductance data (cut >175) for the development of the second ANN model (ANN M2). [ $\mu\text{S}/\text{cm}$ , microsiemens per centimeter;  $\text{ft}^3/\text{s}$ , cubic feet per second]

**EXPLANATION**

Q_Clyo	Clyo flow
SC	Specific conductance
SC_I-95	I-95 specific conductance
ANN M1	Artificial neural network Model 1
ANN M2	Artificial neural network Model 2
SC_Intake <sub>M1</sub>	Predicted specific conductance at the Intake by ANN M1
SC_Intake <sub>M2</sub>	Predicted specific conductance at the Intake by ANN M2
Cl	Chloride concentration
Cl_Intake <sub>M1</sub>	Predicted chloride at the Intake using SC_Intake <sub>M1</sub> input
Cl_Intake <sub>M2</sub>	Predicted chloride at the Intake using SC_Intake <sub>M2</sub> input

**Figure 18.** Modifications to the two-stage model architecture (fig. 16) to accommodate large extrapolation (ANN M2) beyond the range of historical conditions to predict chloride concentrations (Cl\_Intake) at the intake using flow (Q\_Clyo) and specific conductance (SC\_I-95) inputs. [ $\mu\text{S}/\text{cm}$ , microsiemens per centimeter]

## Statistical Measures of Prediction Accuracy

Statistical measures of prediction accuracy were computed for the specific conductance models (ANN M1 and ANN M2) and the chloride model. The statistics for the stage 1 and stage 2 models provide model performance measures of the individual models in the cascading modeling approach to simulate chloride concentrations. Because several models are used, the statistics for the individual models may not be an indication of the quality of the final estimates. Thus, the specific conductance and chloride simulations should be evaluated by the statistics for the final simulation.

Model accuracy typically is reported in terms of  $R^2$  and commonly is interpreted as the “goodness of the fit” of a model. A different interpretation poses the question, “How much information does one variable or a group of variables provide about the behavior of another variable?” For example, in the first context, an  $R^2 = 0.6$  might be disappointing,

whereas in the latter, it is merely an accounting of how much information is shared by the variables being used. The mean error and RMSE statistics provide a measure of the prediction accuracy of the ANN models. The mean error is a measure of the bias of model predictions—whether the model over- or underpredicts the measured data. The mean error is presented as the adjustment to the simulated values to equal the measured values; therefore, a negative mean error indicates an oversimulation by the model, and a positive mean error indicates an underprediction by the ANN model. Mean errors near zero may be misleading because negative and positive discrepancies in the simulations can cancel each other. The root mean square error addresses the limitations of mean error by computing the magnitude, rather than the direction (sign) of the discrepancies. The units of the mean error and RMSE statistics are the same as the simulated variable of the model.

The accuracy of the models, as given by RMSE, should be evaluated with respect to the measured range of the output variable. The percent model error is the ratio of the RMSE to the range of the output measured data. A model may have a low RMSE, but if the range of the output variable is small, the model may only be accurate for a small range of conditions and the model error may be a relatively large percentage of the model response. For example, if the RMSE for a model is 0.5 ft and the measured range is 0 to 2 ft, the percentage model error would be 25 percent. Likewise, a model may have a large RMSE, but if the range of the output variable is large, the model error may be a relatively small percentage of the total model response. For example, if the RMSE for a model is 2 ft and the measured range is 0 to 20 ft, the percentage model error would be 10 percent.

## Specific Conductance Models

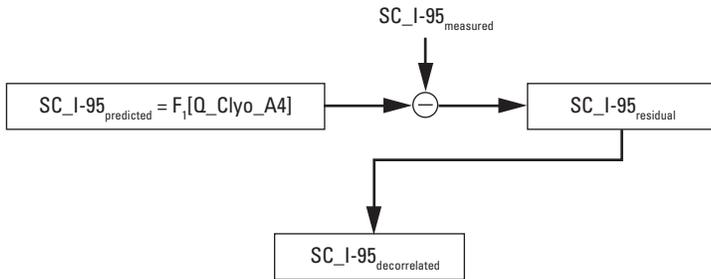
Explanatory variables typically have a strong relation to the behavior of a response variable. For example, specific conductance at the I-95 gage and the water plant intake, and chloride concentrations at the intake have a relation to the Savannah River streamflows at the Clyo gage. When using multiple explanatory variables, it is difficult, if not impossible, to understand the individual effects of variables (sometime referred to as confounded or correlated variables) on a response variable. Empirical models have no notion of process physics, nor the nature of interrelations between input variables. To be able to clearly analyze the effects of confounded variables, the unique informational content of each variable must be determined by “decorrelating” the confounded variables.

The specific conductance data at the I-95 gage and the streamflow data at Clyo are highly correlated. The correlation coefficient of various moving window sizes of Clyo streamflows and I-95 specific conductance were computed. The window size yielding the highest correlation between the two signals, as measured by the Pearson coefficient (Helsel and Hirsch, 1995), was a 4-day moving window average (Q-Clyo\_A4). Decorrelation was accomplished by using a Single Input

Single Output (SISO) ANN model with the 4-day moving average Clyo streamflow as the input and specific conductance data from the I-95 gage as the output (fig. 19). The residual error (the difference between predicted and measured values) is the “unshared” information between the two signals and the decorrelated signal for specific conductance at I-95. Figure 20 shows the measured and predicted specific conductance

at I-95 as well as the residual error from the decorrelation model. The SISO model simulated the low-frequency portion of the measured time series well but did not capture the higher frequencies as seen in the spikes in the measured time series. The spikes are due to the salinity intrusion events resulting from the combination of tidal forcing and low flow and are maintained in the decorrelated time series.

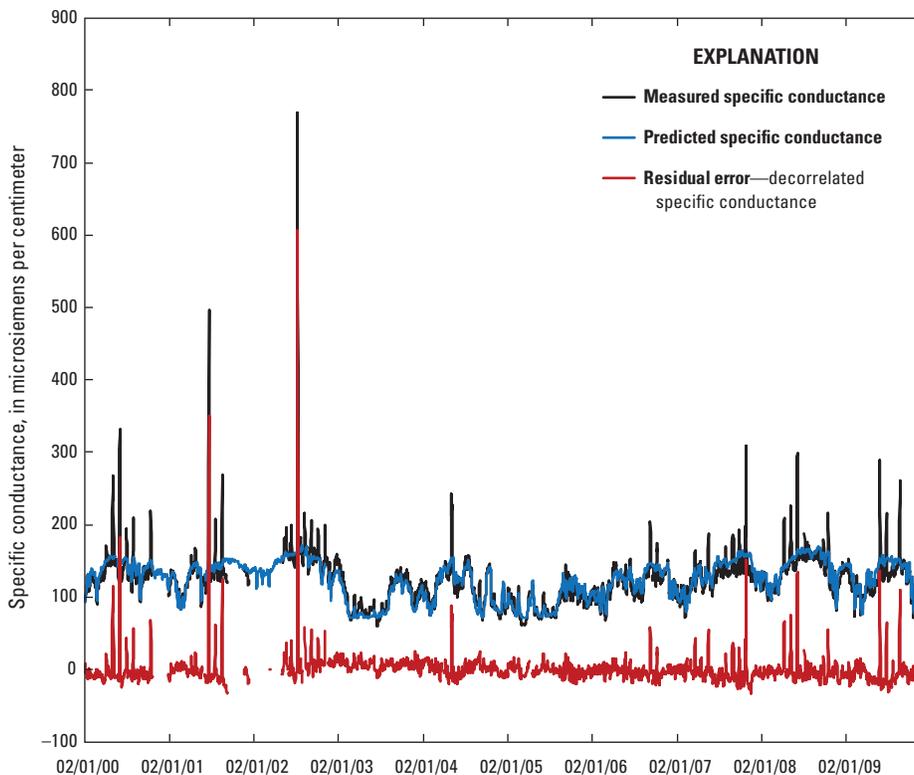
**Decorrelation model**



**EXPLANATION**

- SC                      **Specific conductance**
- SC\_I-95\_predicted    **Predicted SC at I-95**
- Q\_Clyo\_A4            **Four-day moving window average Clyo flow**
- SC\_I-95\_measured    **Measured SC at I-95**
- SC\_I-95\_residual     **Difference between measured and predicted SC at I-95**
- SC\_I-95\_decorrelated **Decorrelated SC at I-95**

**Figure 19.** Specific conductance (SC) decorrelation model. SC\_I-95 is the specific conductance at the I-95 bridge, and Q\_Clyo\_A4 is the 4-day moving window of streamflow at Clyo, Georgia. F1 is a Single Input Single Output Artificial Neural Network Model (SISO ANN). The residual error from the SISO ANN is computed by subtracting the SC\_I-95 predicted values from the measured values.



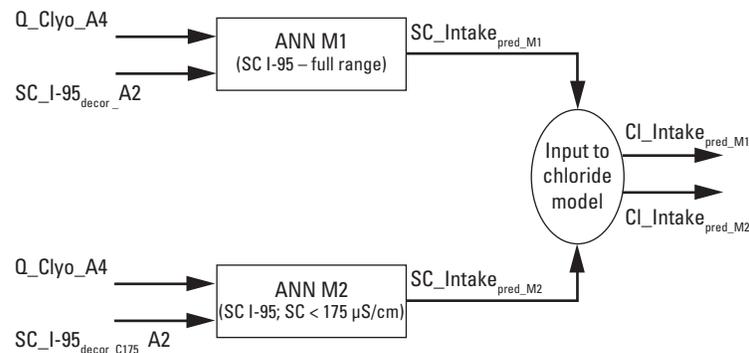
**Figure 20.** Measured, predicted, and residual error from the Single input Single Output Artificial Neural Network model used to decorrelate specific conductance at I-95 from the streamflow at Clyo. The decorrelated specific conductance at I-95 is the residual error time series.

As discussed previously, two ANN models were developed to predict specific conductance at the intake: the ANN M1 model used the full range of specific conductance values at I-95, and the ANN M2 model only used specific conductance values at I-95 less than 175  $\mu\text{S}/\text{cm}$  (fig. 18). The optimal moving window average of the decorrelated specific conductance time series was determined (2-day) by evaluating the correlation between the decorrelated time series and the specific conductance time series at the intake. The 2-day moving window average was used as input to the specific conductance ANN models for the intake (fig. 21). The complete decorrelated specific conductance dataset was used for developing the ANN M1 model using the full range of specific conductance data and the computed 2-day moving window average computed ( $\text{SC\_I-95}_{\text{decor\_A2}}$ ). For the ANN M2 model, the dataset was truncated to include only specific conductance values less than 175  $\mu\text{S}/\text{cm}$  ( $\text{SC\_I-95}_{\text{decor\_C175\_A2}}$ ).

The performance of the ANN M1 and M2 models are similar as seen in the plots of the measured and simulated specific conductance values and performance statistics (fig. 22; table 5). The ANN M1 model captured more of the rapid increase in specific conductance during the last 3 years of the simulation (2006–2009) than the ANN M2 model, but also simulated rapid increases in specific conductance that were not in the measured record at the intake. The differences in the model simulations can be seen in the 3D response surfaces for each model (fig. 23). Three-dimensional response surfaces can

be generated to display two explanatory variables (streamflow at Clyo and specific conductance at I-95) with a response variable (specific conductance at the intake). The data for the response surface were computed by the ANN model across the full range of the displayed input variables.

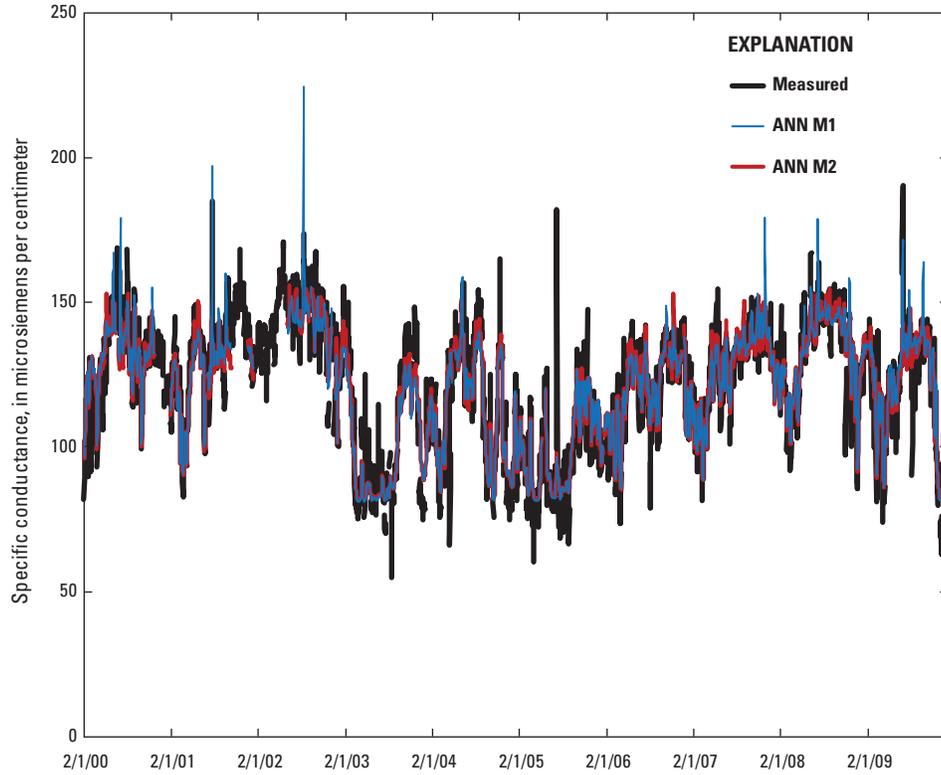
Response surfaces are a valuable tool for understanding the dynamics of riverine and estuarine systems as simulated by the ANN models and for projecting how the models will extrapolate hydrologic conditions beyond the range of data used to train the models. The upper right edge of the response surface for the ANN M1 model (fig. 23A) shows smaller increases in the specific conductance at the intake with increases in flow and specific conductance than the response surface for the ANN M2 model (fig. 23B). The bending of the response surface at these high chloride conditions (fig. 23A) is similar to the plateau effect seen in the 3D scatter plot of the data in figure 9. The response surface for the ANN M2 (fig. 23B) does not show the bending surface at these high conditions. Extrapolations with the ANN M1 model will, therefore, result in smaller increases in specific conductance at the intake with decreases in Clyo streamflow and increases in I-95 specific conductance conditions; whereas, the ANN M2 model has greater increases in specific conductance at the intake. These differences in the response surfaces for these conditions are a result of truncating the specific conductance data at I-95 to include only the values below 175  $\mu\text{S}/\text{cm}$  for the ANN M2 model.



**EXPLANATION**

ANN M1	Artificial neural network Model 1	Q_Clyo_A4	Four-day moving window average Clyo flow
ANN M2	Artificial neural network Model 2	SC	Specific conductance
SC_Intake <sub>pred_M1</sub>	Predicted specific conductance at the Intake by ANN M1	SC_I-95	I-95 specific conductance
SC_Intake <sub>pred_M2</sub>	Predicted specific conductance at the Intake by ANN M2	SC_I-95 <sub>decor_A2</sub>	Two-day moving window average of decorrelated specific conductance at I-95
Cl	Chloride concentration	SC_I-95 <sub>decor_C175_A2</sub>	Two-day moving window average of decorrelated specific conductance at I-95 using only the SC data less than 175 microsiemens per centimeter
Cl_Intake <sub>pred_M1</sub>	Predicted chloride at the Intake using SC_Intake <sub>pred_M1</sub> input		
Cl_Intake <sub>pred_M2</sub>	Predicted chloride at the Intake using SC_Intake <sub>pred_M2</sub> input		

**Figure 21.** Inputs and outputs for the models to predict specific conductance at the intake.



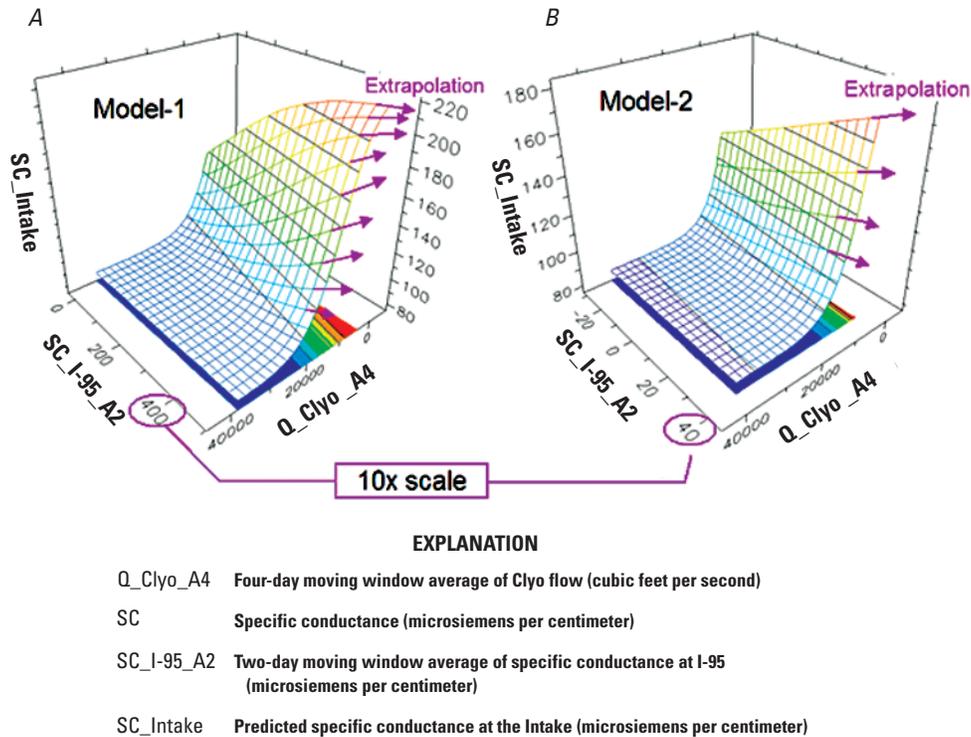
**Figure 22.** Measured and simulated specific conductance at the City of Savannah’s intake from the ANN M1 model and ANN M2 model for the period February 1, 2000, to October 31, 2009.

**Table 5.** Model performance statistics for the specific conductance decorrelation model and the ANN M1 and ANN M2 models.

[R<sup>2</sup>, coefficient of determination; RMSE, root mean square error; µS/cm, microsiemens per centimeter]

Model	Hidden layer neurons	Training dataset (number of observations)	Testing dataset (number of observations)	R <sup>2</sup> (training/testing)	RMSE (µS/cm; training/testing)	Range <sup>a</sup> of measured data (µS/cm)	Percent model error
Specific conductance decorrelation	1	2,631	617	0.62/0.59	22/24	719	3.2
ANN M1	2	1,312	900	0.82/0.84	9.2/9.2	136	6.8
ANN M2	5	1,247	851	0.83/0.85	8.7/8.5	127	6.8

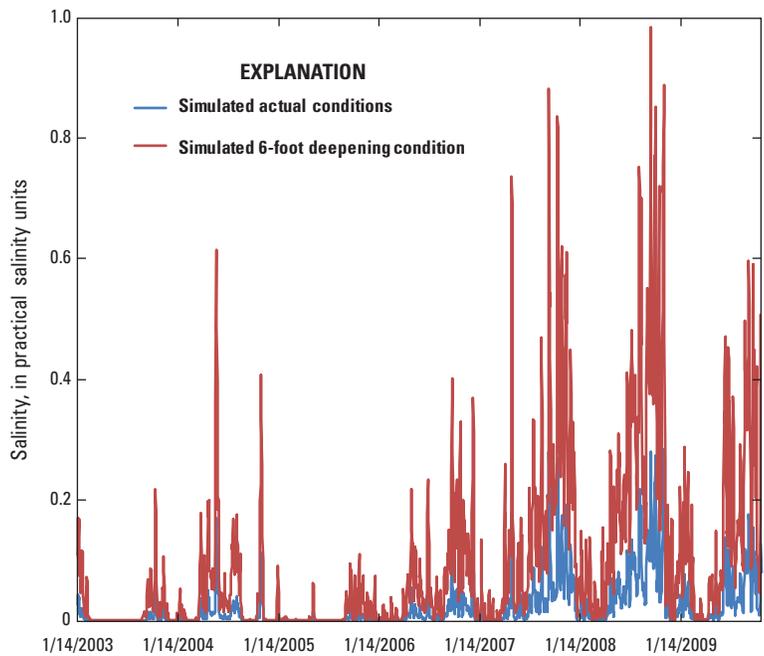
<sup>a</sup> Difference in maximum and minimum measured values.



**Figure 23.** Three-dimensional response surfaces for the (A) ANN M1 and (B) ANN M2 models. ANN M1 and ANN M2 architecture shown in figure 18.

### Development of Extrapolated Specific Conductance Dataset

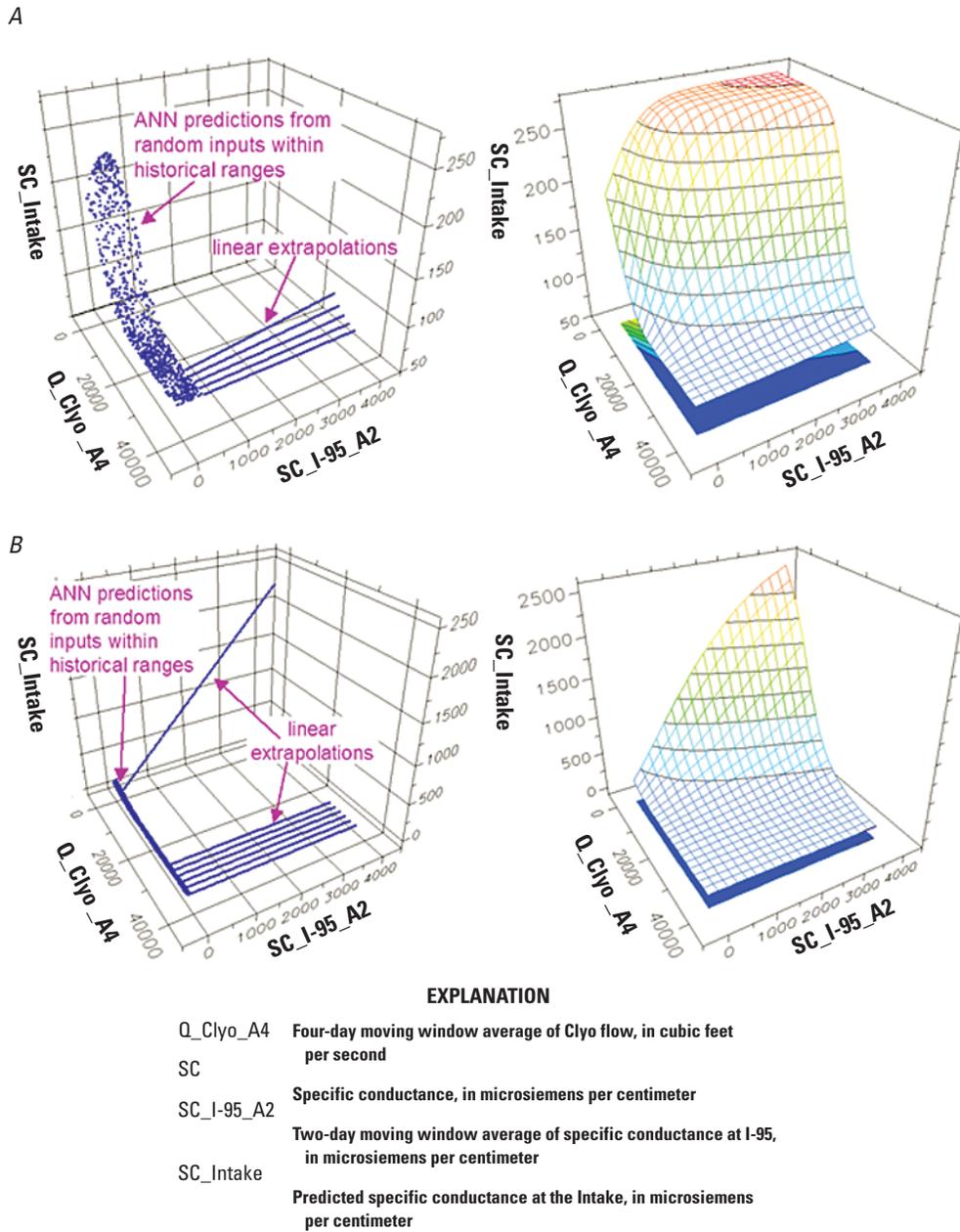
Preliminary simulations of a proposed deepening of the Savannah Harbor using the EFDC model showed increases in salinity concentrations at the I-95 gage more than three times as great as the historical maximum (fig. 24). The iQuest™ software used to develop the ANN model allows for some extrapolation beyond the historical range of the data but not to the extent necessary for this application. To accommodate this constraint, a synthetic dataset was generated to define how specific conductance at the intake would respond to specific conductance at I-95 above the historical maximums. The ANN M1 and ANN M2 models were used to generate a set of input-output vectors within the historical range of the data. From these data, linear extrapolations were generated by extending the line between model predictions at the endpoint of the extremes of the historical range of streamflow at Clyo and specific conductance at I-95. These extrapolated arrays of values were then used to retrain the ANN M1 and ANN M2 models. Selecting the number of arrays to use was an iterative process to ensure that a smooth model surface was generated. The new “extrapolated” versions of the models are named ANN M1e and ANN M2e.



**Figure 24.** Salinity predictions at the I-95 bridge for actual conditions and a simulated 6-foot harbor deepening scenario using the EFDC model.

The simulated specific conductance values for the water plant intake and the inputs of the 4-day average streamflow at Clyo and 2-day average specific conductance (decorrelated) at the I-95 gage are shown in 3D scatter plots in figure 25 for the ANN M1e and ANN M2e models. The points on the scatter plots are the simulated specific conductance values at the intake. Also shown are the points of the linear extrapolations of the specific conductance at the intake to changes in specific conductance at I-95 beyond the historical condition.

For the ANN M1 model, the upper edge of the response surface (fig. 23) indicates an extrapolated model will capture the plateau effect shown in figures 9 and 10. For ANN M1e, only linear extrapolations were computed for the higher streamflow ranges (fig. 25A). For the ANN M2e model, linear extrapolation was generated for the higher streamflow ranges and also for a low-flow condition (fig. 25B). Figure 25 also shows the ANN model response surfaces generated from ANN models trained using the extrapolated dataset, ANN M1e and ANN M2e, respectively. The response surfaces in figure 25 are the extrapolations of the response surfaces shown in figure 23.



**Figure 25.** Three-dimensional scatter plots of 4-day average flow, 2-day average specific conductance (decorrelated from flow), and predicted specific conductance at the intake, and linear extrapolations of flow and specific conductance beyond the measured historical values for the (A) ANN M1e and (B) ANN M2e models. Three-dimensional response surfaces to the right show response surfaces using the extrapolation models (A) ANN M1e and (B) ANN M2e.

## Chloride Models

The chloride model, like the specific conductance models, requires extrapolation to conditions greater than the historical range of the data. The linear regression model allows chloride concentrations to be predicted at highly extrapolated specific conductance values. The relation between specific conductance and chloride concentrations at the intake is linear (fig. 14); therefore, a linear regression model was used (eq. 1).

$$\text{Ch\_Intake} = 0.0078 * \text{SC\_Intake}_{\text{pred}} + 1.639 \quad (1)$$

where Ch\_Intake is the chloride estimate at the intake, and SC\_Intake<sub>pred</sub> is the estimated specific conductance value at the intake. Correlation analysis showed that the 3-day moving window average of predicted specific conductance from the ANN M1e and ANN M2e models had the highest Pearson coefficient to the daily chloride concentration at the intake. The measured and predicted chloride concentrations are shown in figure 26 as well as the residual error from the linear regression model.

To improve the chloride concentration predictions, an error correction model was developed to predict the residual error of the linear regression model (fig. 27). The final chloride concentration prediction is the summation of the prediction from the linear regression model and the error correction model. The chloride concentrations at the intake

are correlated to streamflow, water color, and maximum daily water temperature, and these inputs are used in the error correction model. Water color is a measure of organic material in the water and is an indicator of source of the water at the intake, either freshwater upstream from the intake, tidal water from downstream from the intake, or a combination. Water color is correlated to streamflow and was decorrelated from streamflow using the same approach for decorrelation of specific conductance at I-95 from streamflow (fig. 19). Two inputs from the decorrelated water color variable were used as inputs to the error correction model—the 1-day lag and the 1-day change in water color.

The residual chloride error is at least partially and nonlinearly dependent on the chloride concentrations at the intake. To determine the chloride concentration at the intake that is independent of streamflow (or decorrelated from streamflow), a linear regression to describe the chloride concentration attributable to streamflow was computed (eq. 2).

$$\text{Ch\_Q} = -12.572 * \text{Q\_Clyo\_A6\_Mod} + 20.481 \quad (2)$$

where Ch\_Q is the chloride concentration independent of streamflow, and Q\_Clyo\_A6\_Mod is transformation of the 6-day moving window average Clyo streamflow (eq. 3). The streamflow input into the equation should be extrapolated to streamflow conditions beyond the historical range of conditions. An inverse exponential function was used to transform

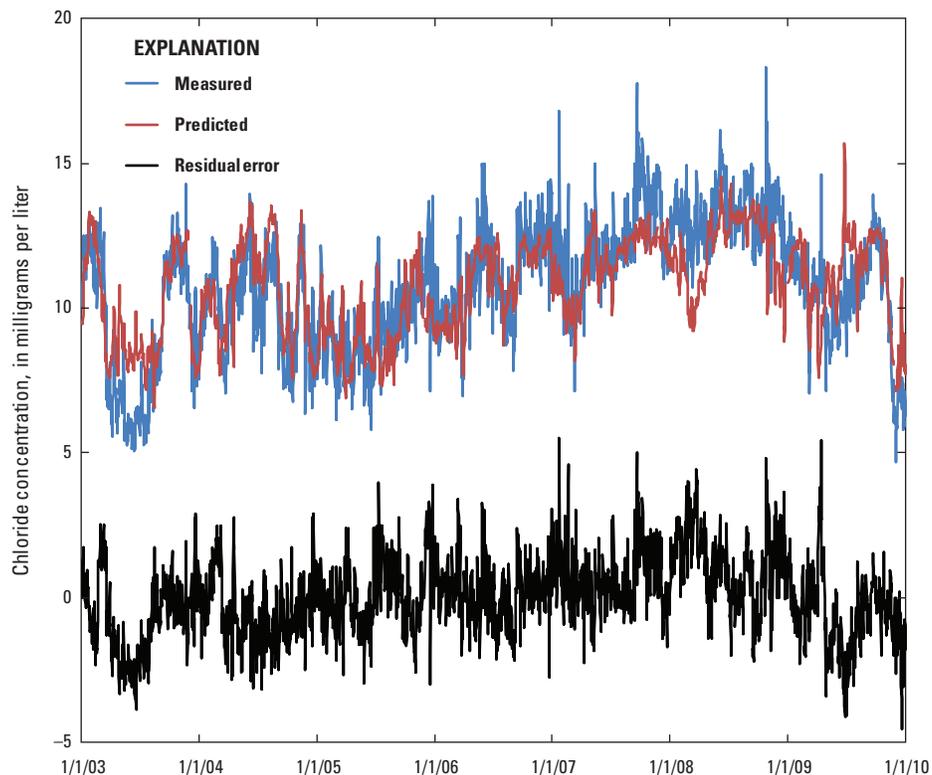
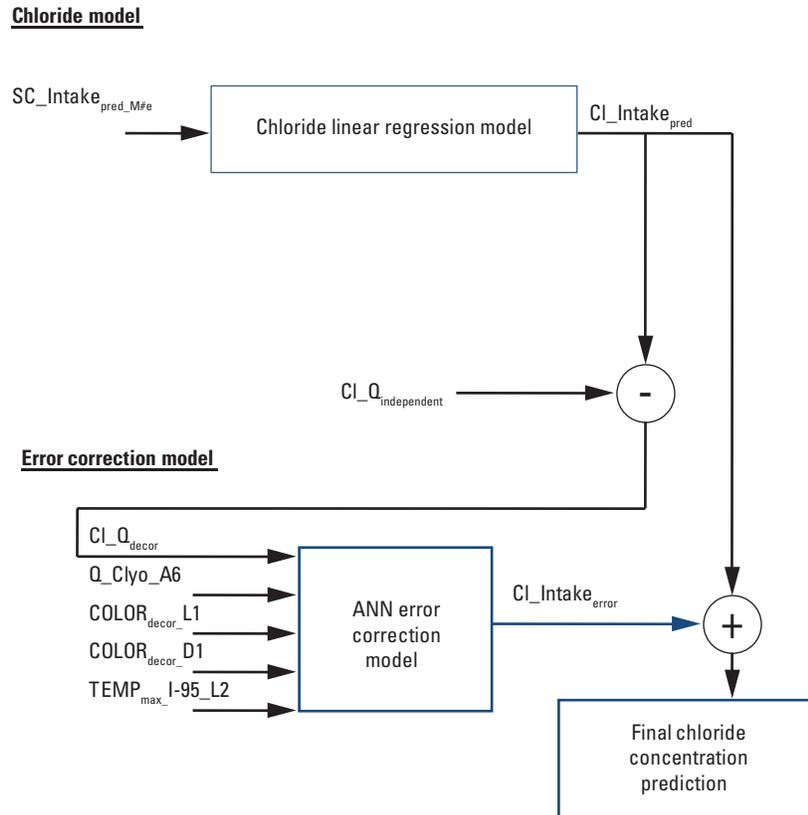


Figure 26. Measured and predicted chloride concentrations at the intake and the residual error.



**Figure 27.** The model architecture for the chloride model and error correction model. The dash and cross in the circles are subtraction and addition signs, respectively.

streamflow values for input into the regression equation. The inverse exponential transformation can be extrapolated and captures the nonlinear relation between streamflow and chloride concentrations.

$$Q\_Clyo\_A6\_Mod = 1 - e^{-(0.002 * Q\_Clyo\_A6)} \quad (3)$$

where  $Q\_Clyo\_A6$  is the 6-day moving window average of Clyo streamflow.

Similar to the decorrelation of specific conductance at I-95 from streamflow (fig. 19), the chloride concentrations predicted by streamflow are subtracted from the chloride predictions from the linear regression model of chloride concentrations as a function of specific conductance values at the intake. Figure 27 shows a schematic of the chloride and error correction models, and table 6 lists model performance statistics for the models used to make the final chloride concentration predictions.

The error correction model estimates the central tendency of the residual error from the chloride regression model (fig. 28) but does not predict the large variability in the error signal. The final chloride concentration predictions are slightly better than the linear regression model predictions (fig. 29). The error correction model increases  $R^2$  of the final chloride concentration predictions by 19 percent and decreases the percent model error by 31 percent (table 6). Although predictions of the final chloride concentrations are improved, the final model predictions do not simulate the full range of the measured chloride concentrations.

Although the basic architecture of the chloride concentration model is a two-stage approach (fig. 16), there are many submodels to decorrelate input variables, simulate a subset of the specific conductance datasets, and correct chloride model error. A schematic of the overall model architecture with the inputs and outputs interaction is shown in figure 30.

**Table 6.** Model inputs, outputs, and performance statistics for the chloride model and error correction models.

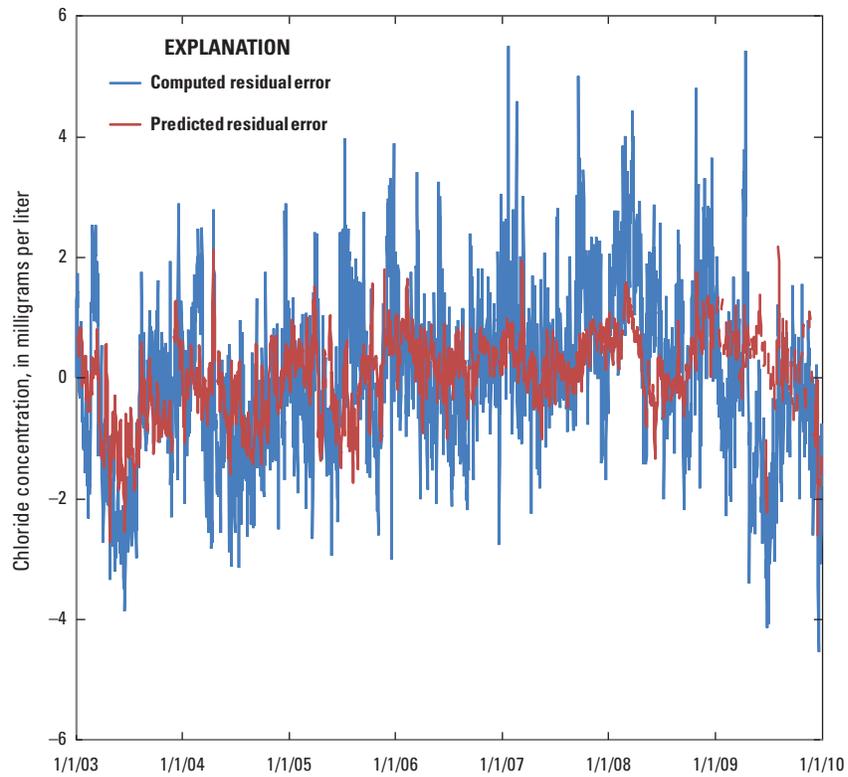
[n, number in dataset; R<sup>2</sup>, coefficient of determination; RMSE, root mean square error; ANN, artificial neural network; SC, specific conductance; predM#, change in SC predicted by either the M1e or M2e model; NA, not applicable; resid, residual error; max, maximum; decor, decorrelation]

Model	Type of model	Inputs	Output	Hidden layer neurons	n	Training dataset count	Testing dataset count	R <sup>2</sup> (training/testing)	RMSE <sup>a</sup> (Training/testing)	Range <sup>b</sup> of measured data	Percent model error
Color decorrelation	ANN	6-day average flow	Color <sub>decor</sub>	1	2,436	1,460	976	0.22/0.17	26/26	267	9.7
Chloride	linear regression	SC_Intake pred_M1#e	Chloride <sub>inake</sub>	NA	2,510	NA	NA	0.59	1.7	13	13.4
Chloride correction model	ANN	6-day average flow Temp <sub>max</sub> Color <sub>decor</sub> - 1-day lag Color <sub>decor</sub> - 1-day change Chloride <sub>inake-decor</sub>	Chloride <sub>resid</sub>	1	2,294	1,380	914	0.25/0.23	1.22/1.28	9.28	13.1
Final chloride predictions <sup>c</sup>	NA			NA	2,294	NA	NA	0.70	1.2	13	9.3

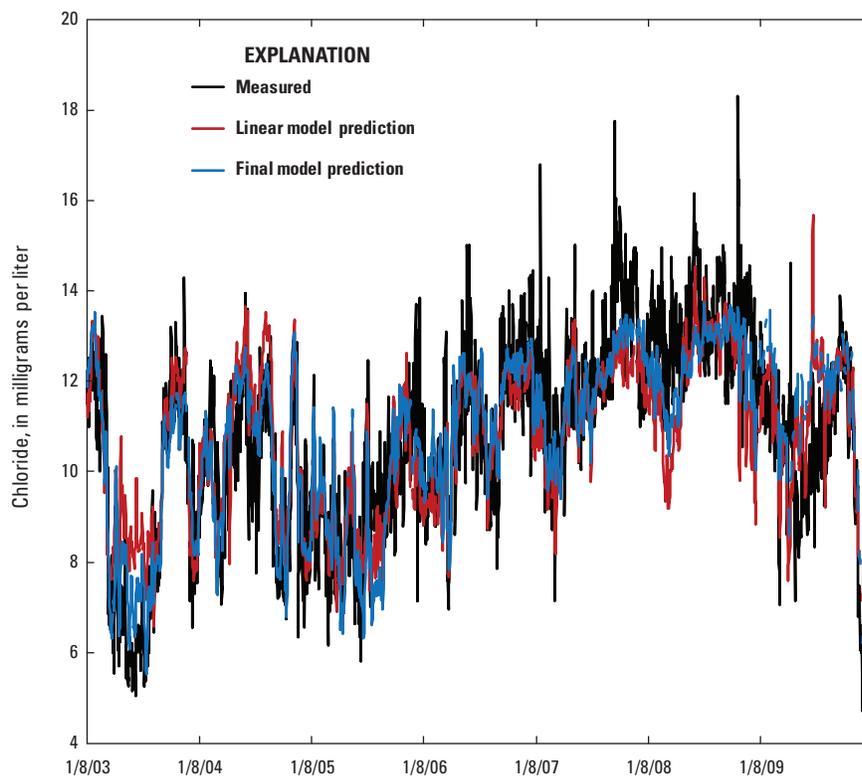
<sup>a</sup> Units for RMSE are color units for color and milligrams per liter for chloride.

<sup>b</sup> Difference in maximum and minimum measured values.

<sup>c</sup> Final chloride predictions is a summation (additive) of the predictions from the linear regression chloride model and the ANN chloride correction model.

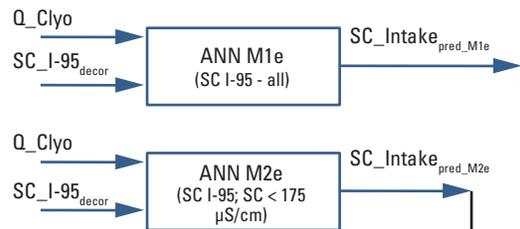


**Figure 28.** Computed and predicted residual error from the chloride model.

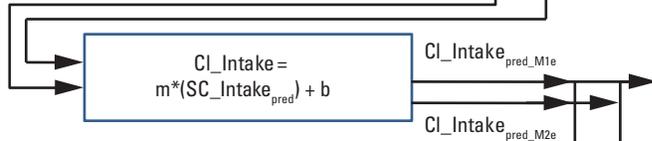


**Figure 29.** Measured and predicted chloride concentrations at the City of Savannah's intake for the period January 14, 2003, to October 31, 2009.

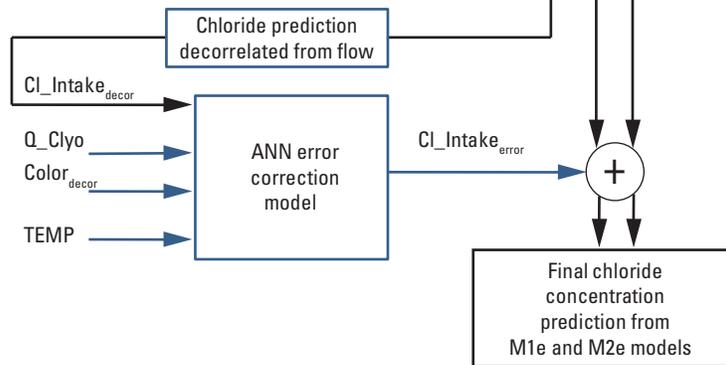
**Intake specific conductance models**



**Intake chloride concentration model**



**Chloride error correction model**



**EXPLANATION**

Q_Clyo	Clyo flow	m	Slope
SC	Specific conductance	b	y-intercept
SC_I-95 <sub>decor</sub>	Decorrelated specific conductance at I-95	Cl_Intake <sub>pred_M1e</sub>	Predicted chloride at the Intake using SC_Intake <sub>M1</sub> input
ANN M1e	Artificial neural network Model 1 trained extrapolated dataset	Cl_Intake <sub>pred_M2e</sub>	Predicted chloride at the Intake using SC_Intake <sub>M2</sub> input
ANN M2e	Artificial neural network Model 2 trained extrapolated dataset	Cl_Intake <sub>decor</sub>	Predicted chloride at the intake decorrelated from flow input
SC_Intake <sub>pred_M1e</sub>	Predicted specific conductance at the Intake by ANN M1e	COLOR <sub>decor</sub>	Decorrelated water color value inputs at the Intake
SC_Intake <sub>pred_M2e</sub>	Predicted specific conductance at the Intake by ANN M2e	TEMP	Daily maximum water temperature input at I-95
Cl	Chloride concentration	Cl_Intake <sub>error</sub>	Predicted error in chloride regression model
Cl_Intake	Chloride at the Intake		
SC_Intake <sub>pred</sub>	Predicted specific conductance at the Intake slope		

**Figure 30.** Specific conductance and chloride model architecture. [ $\mu\text{S}/\text{cm}$ , microsiemens per centimeter]

## Development of a Decision Support System

Resource managers and stakeholders face difficult challenges when managing interactions between natural and manmade systems. At considerable cost, complex mechanistic models that are based on first principles physical equations, are developed and operated by scientists to evaluate options for using a resource while minimizing harm. However, varying technical abilities and financial constraints among stakeholders effectively restrict access to relevant scientific knowledge and analytical tools. Decision support system (DSS) technology can help meet the need to provide equal access to the knowledge and tools required for informed decisionmaking. Even though the collective interests and computer skills within the community of managers, scientists, and other stakeholders are quite varied, equal access to the scientific knowledge is needed to make the best possible decisions. Dutta and others (1997) define DSSs as, “systems helping decision-makers to solve various semi-structured and unstructured problems involving multiple attributes, objectives, and goals... Historically, the majority of DSSs have been either computer implementations of mathematical models or extensions of database systems and traditional management information systems.” While there appears to be no strict criteria that distinguish a DSS from other types of programs, Dutta and others (1997) suggest that artificial intelligence (AI) is a characteristic of more advanced DSSs: “With the help of AI techniques DSSs have incorporated the heuristic models of decision makers and provided increasingly richer support for decision making. Artificial intelligence systems also have benefited from DSS research as they have scaled down their goal from replacing to supporting decision makers.”

The authors of this report have previously developed four DSSs in South Carolina, Georgia, and Florida to evaluate (1) wastewater discharges and dissolved-oxygen concentration in the Beaufort River estuary (Roehl and others, 2006; Conrads and others, 2003); (2) salinity effects on freshwater tidal wetlands and a proposed deepening of the Savannah Harbor (Conrads and others, 2006); (3) the effect of controlled streamflow releases from reservoirs on the Pee Dee River in North Carolina and on salinity dynamics along the South Carolina coast (Conrads and Roehl, 2007); and, (4) the effect of controlled flow releases on the water levels, specific conductance, and total phosphorus of the freshwater marsh of the Loxahatchee National Wildlife Refuge (Conrads and Roehl,

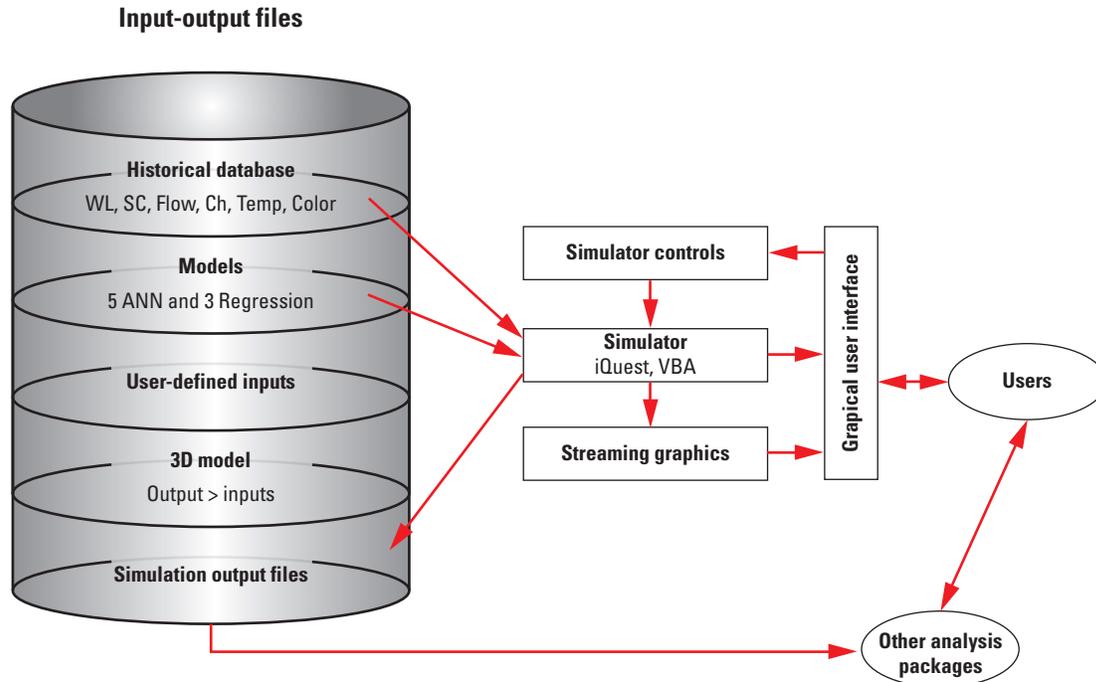
2010). These DSSs are spreadsheet applications that provide predictive models with databases for ANN model simulation, graphical user interfaces, and displays of results. Additional features, including optimizers, integrations with other models and software tools, and color contouring of simulation output data, make the DSSs easily distributable and immediately usable by all resource managers and stakeholders.

The development of a Savannah Chloride Model Decision Support System (SCM DSS) for Abercorn Creek required a number of steps, including: (1) merging all the data into a single comprehensive database; (2) developing specific-conductance and chloride models; (3) incorporating output from the EFDC model; and (3) developing a Microsoft Office Excel® application that integrates the new database, submodels, model inputs and outputs, and visualization routines into a single package that is easy to use and disseminate. The user’s manual for the installation and operation of the SCM DSS is provided in appendix 1.

### Architecture

The basic architectural elements of the SCM DSS are shown in figure 31. The DSS reads and writes files for the various run-time options that can be selected by the user through the system’s graphical user interface. A historical database, containing 7 years of hydrologic and water-quality data, is read into the simulator along with the linear regression and ANN submodels at the start of a simulation. By using graphical user interface controls, the user can evaluate scenarios for alternative flow, specific conductance, and harbor geometry changes (as simulated by EFDC). The outputs generated by the simulator are written to files for post processing in Microsoft Office Excel™ or other analysis software packages. The DSS also provides streaming graphics for each gage during simulations of specific conductance and chloride concentrations response for the models.

Linking the 3D hydrodynamic EFDC model to the simulator is accomplished by loading in a file of simulated differences in specific conductance values for the gage at I-95. The simulated EFDC salinity values are converted to specific conductance values. The changes in specific conductance at I-95 reflect hypothetical channel geometry, and mitigation scenarios run in the hydrodynamic model are then used as inputs to the ANN M1e and ANN M2e models to simulate the specific conductance at the intake.



**Figure 31.** Architectural elements of the Savannah Chloride Model Decision Support System.

## Model Simulation Control and Streaming Graphics

The simulator in the SCM DSS integrates the historical database with the ANN and linear regression models. The date/time controls on the user control panel (fig. 32) are used to adjust start and end dates and graphical and tabular output for a simulation. The simulator allows the user to run “what-if” simulations by varying the streamflow and specific conductance values from their historical values. The user has three simulation options for changing streamflow inputs and four options for changing specific conductance inputs:

- as a percentage of historical streamflow or specific conductance values,
- as constant streamflow or specific conductance values,
- as user-defined time series of streamflow or specific conductance, and
- as specific conductance input from an EFDC model simulation output.

Explanations of how to use each of the options in the SCM DSS are provided in the user’s manual in appendix 1.

The top of the SCM DSS control panel (fig. 32) shows the simulation period, output options, the selected input options, and the measured value, simulated measured value, and simulated user-defined options values for the current time step. The SCM DSS also shows streaming graphics on

the control panel while a simulation is running are shown. The graphs display the historical measured data, simulated historical conditions (to show model accuracy), and the simulated output using the input option set by the person using the graphical user interface controls or an input file.

## Application of the Savannah Chloride Model Decision Support System

The development of the ANN and linear regression and the SCM DSS application provides resource managers with a tool for evaluating specific conductance and chloride concentration dynamics at the water plant intake in Abercorn Creek. The SCM DSS allows users to simulate various streamflow and specific conductance conditions and analyze the specific conductance and chloride concentration response at the intake. In the SCM DSS, the user is able to set Clio streamflows as a constant flow, a percentage of historical flow, or as a user-defined hydrograph. Specific conductance values at I–95 can be set by the user as constant values, a percentage of historical values, a user-defined time series, or as output from an EFDC model simulation. The following section describes applications of the SCM DSS to two management scenarios. The results from these scenarios are intended to demonstrate the utility of the SCM DSS and are not intended to be interpreted as proposed hydraulic operations or a regulatory application of the DSS.

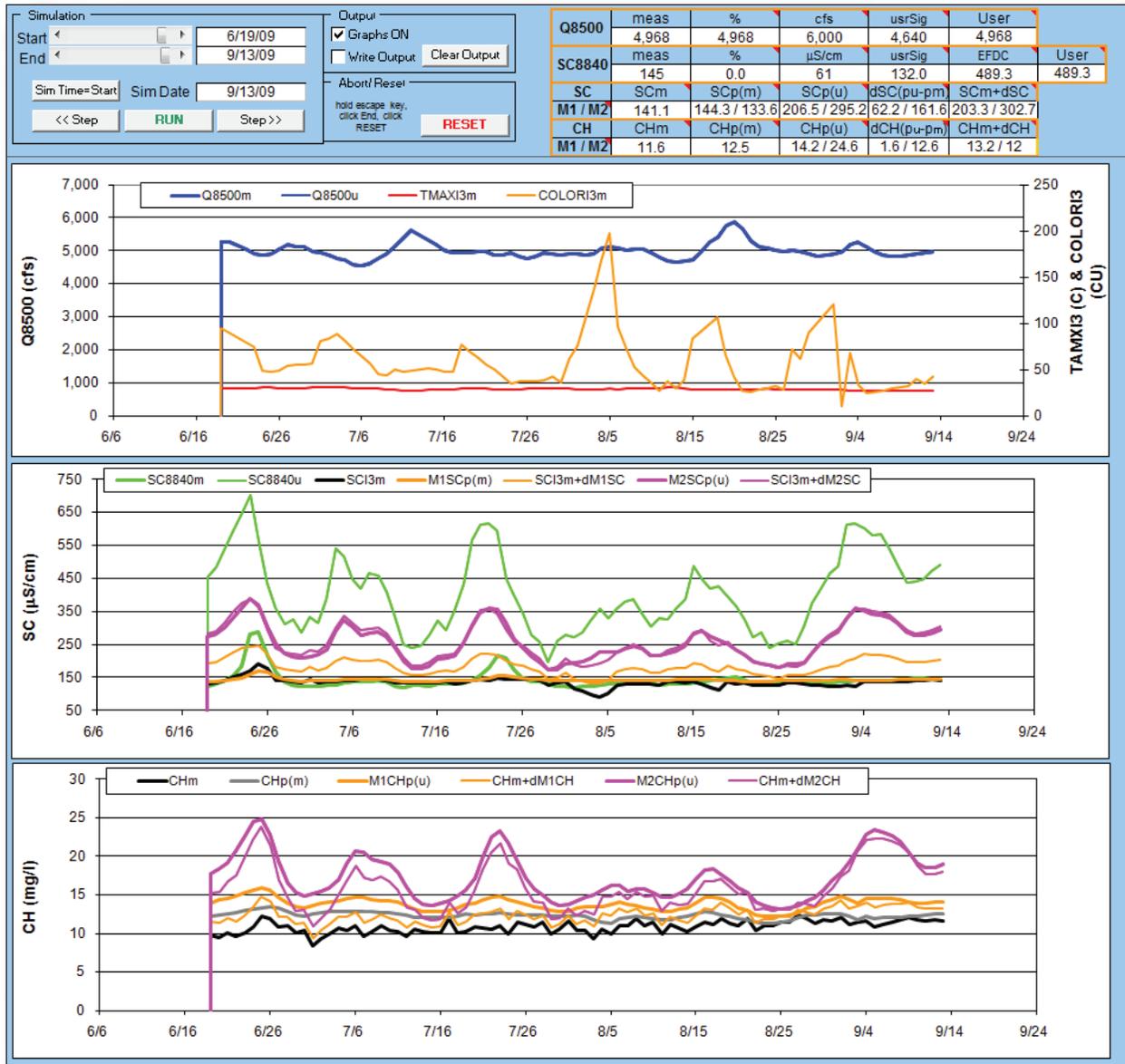


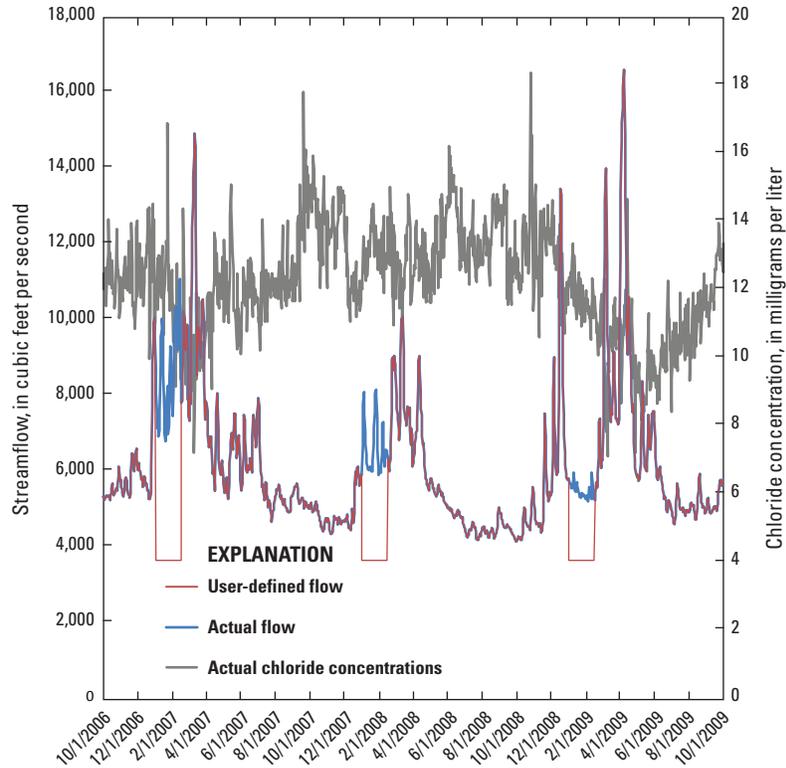
Figure 32. The model simulator controls used to run a simulation in the Savannah Chloride Model Decision Support System.

### User-Defined Hydrograph

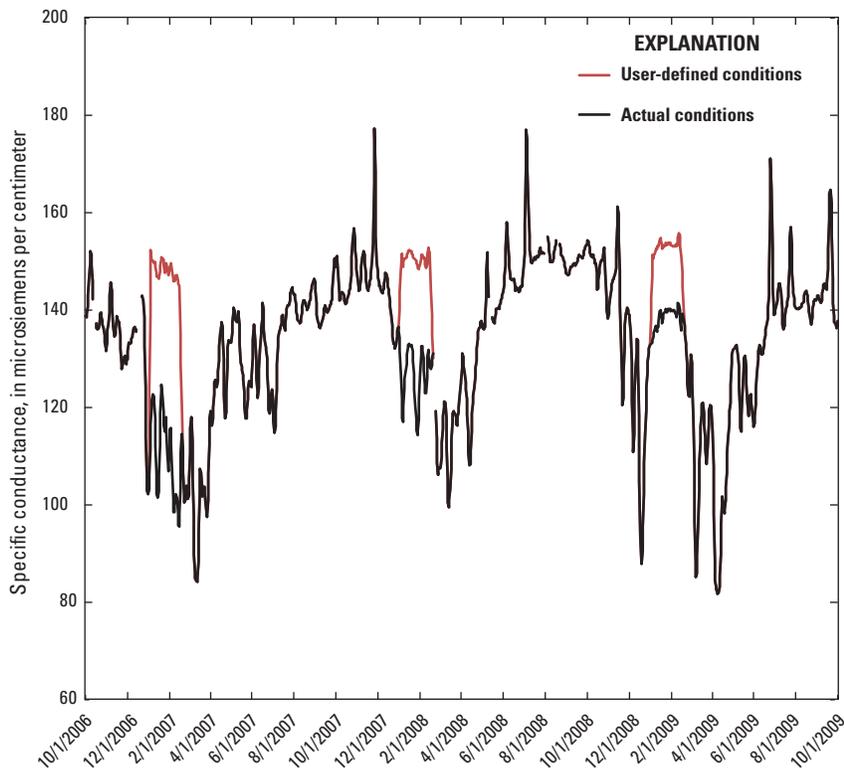
One user-specified option is to input the Savannah River flow at Clyo with a user-defined hydrograph. With this option, a user-defined daily hydrograph is created outside of the SCM DSS. The simulation time period is selected, and the user-defined hydrograph is used as input. A scenario using this option was simulated to evaluate the effect of reducing controlled flow releases to the Savannah River in the winter during drought periods to store more water in upstream reservoirs. For the period October 1, 2006, to October 1, 2009, the January 1 to February 15 daily streamflows were reduced to 3,600 ft<sup>3</sup>/s (fig. 33). The largest streamflow reduction for the

simulation period occurred in 2007 when average streamflow for the winter period was 8,500 ft<sup>3</sup>/s. The average streamflow periods for 2008 and 2009 were 6,590 and 5,420 ft<sup>3</sup>/s, respectively. The measured streamflow conditions for the period and the user-defined streamflow conditions and the actual chloride concentrations are shown in figure 33.

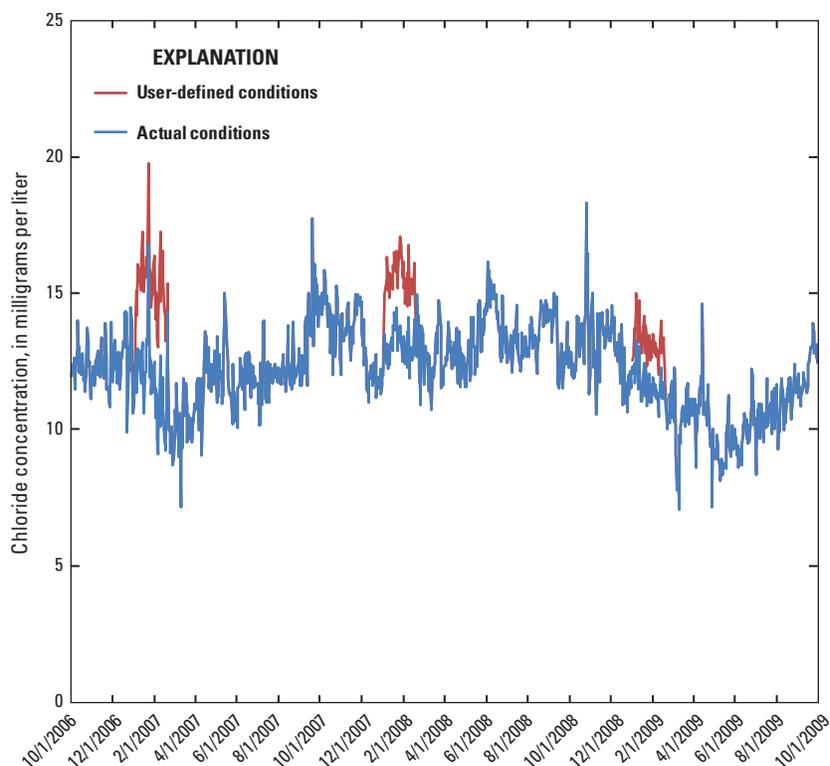
Results of the ANN M1e model are shown in figures 34 and 35. Simulations indicated that decreasing the streamflows during winter periods had the effect of increasing the specific conductance and chloride concentration at the intake for those periods. The predicted maximum increase in specific conductance of 50 µS/cm occurred on February 13, 2007 (fig. 34). The decreased streamflows and predicted increase in



**Figure 33.** Actual and user-defined flow conditions at Clys and actual chloride concentrations at the intake for the period October 1, 2006, to October 1, 2009.



**Figure 34.** Simulated actual and user-defined specific conductance at the intake for the period October 1, 2006, to October 1, 2009. The user-defined flow conditions are shown in figure 33.



**Figure 35.** Simulated actual and user-defined chloride concentration at the intake for the period October 1, 2006, to October 1, 2009. The user-defined flow conditions are shown in figure 33.

specific conductance resulted in an average predicted increase in the chloride concentration at the intake of 2.4 mg/L for the January 1 to February 15 periods with a maximum increase of 4.8 mg/L occurring on January 14, 2007 (fig. 35).

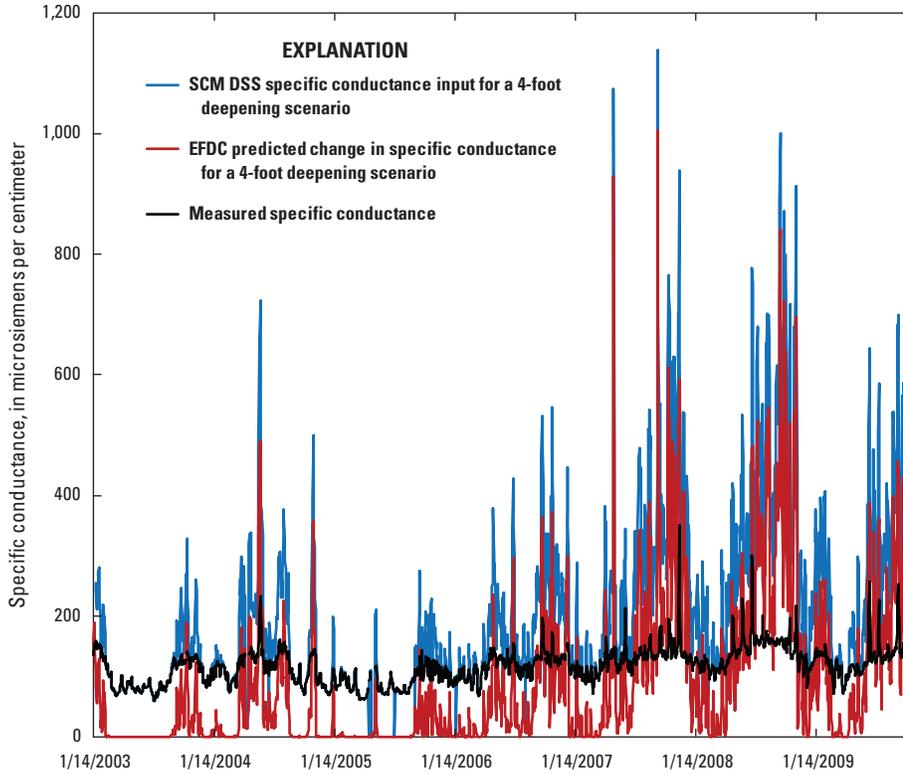
### Inputs from Three-Dimensional Model Output

Another option for user-defined inputs to the SCM DSS is to use output from the 3D hydrodynamic EFDC model of the Savannah River estuary. Using this option, the differences between a historical baseline and alternative harbor geometry specific conductance simulations at the I-95 gage are used for input to the SCM DSS. A harbor deepening scenario was simulated for the period January 14, 2003, to October 31, 2009, using the historical streamflow condition and the EFDC specific conductance inputs. Two simulations were generated with the EFDC model. The first was the actual historical conditions for the simulation period. The second simulation used the same boundary input conditions but a different channel geometry file. The change in channel geometry and configuration represented a 4-ft deepening of the harbor and implementation of mitigation actions to limit salinity increases in the Little Back and Middle Rivers (fig. 5). The differences between the two EFDC salinity simulations are post-processed and converted to specific conductance for the I-95 gage. The final specific conductance input for the ANN M1e and

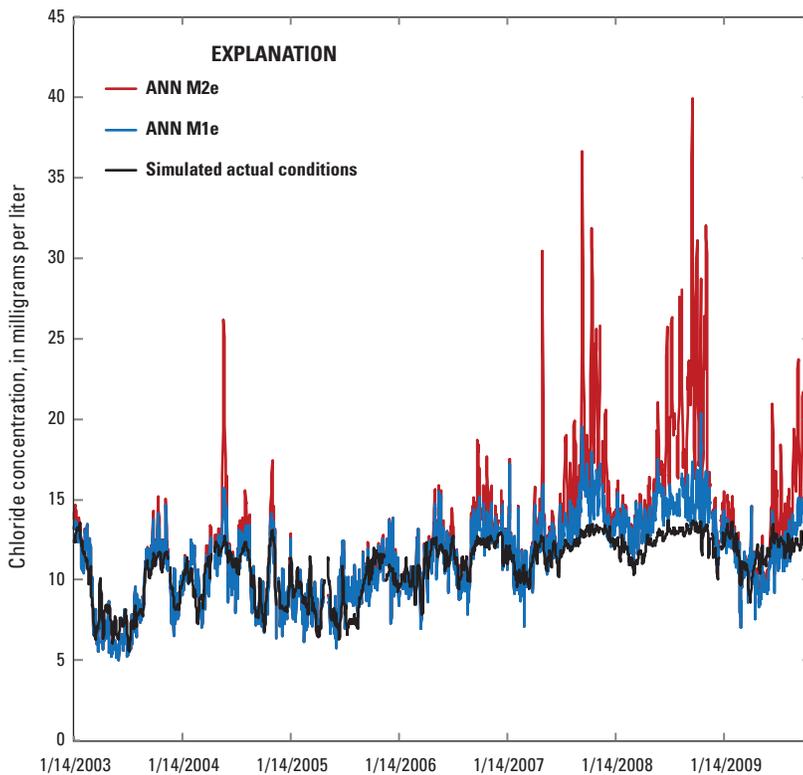
ANN M2e models is the summation of the measured specific conductance at I-95 and the EFDC predicted change in specific conductance at I-95. The measured specific conductance, EFDC model predicted change in specific conductance, and the specific conductance inputs to the ANN M1e and M2e models for a 4-ft harbor deepening and mitigation scenario for the period January 14, 2003, to October 31, 2009, are shown in figure 36.

The two ANN models (M1e and M2e) were developed to extrapolate differently for the hypothetical conditions that are greater than the historical conditions. The specific conductance inputs for this scenario are much greater than the historical conditions. The average and maximum measured specific conductance values are 117 and 351  $\mu\text{S}/\text{cm}$ , respectively, whereas the average and maximum EFDC simulated inputs are 203 and 1,140  $\mu\text{S}/\text{cm}$ , respectively (fig. 36). The simulated chloride concentration values at the intake show the difference in the model extrapolations (fig. 37). The ANN M1e model, trained on the complete dataset, predicts a smaller change in the chloride concentration values than the ANN M2e model, which was trained on a subset of the complete dataset. The M1e model predicted a relative increase (difference between simulated measured conditions and user-defined conditions) in daily average chloride concentrations of 6.6 mg/L, whereas the M2e model predicted an increase of 26.3 mg/L.

The magnitude of chloride concentrations in source water is of concern to operators of municipal water treatment plants



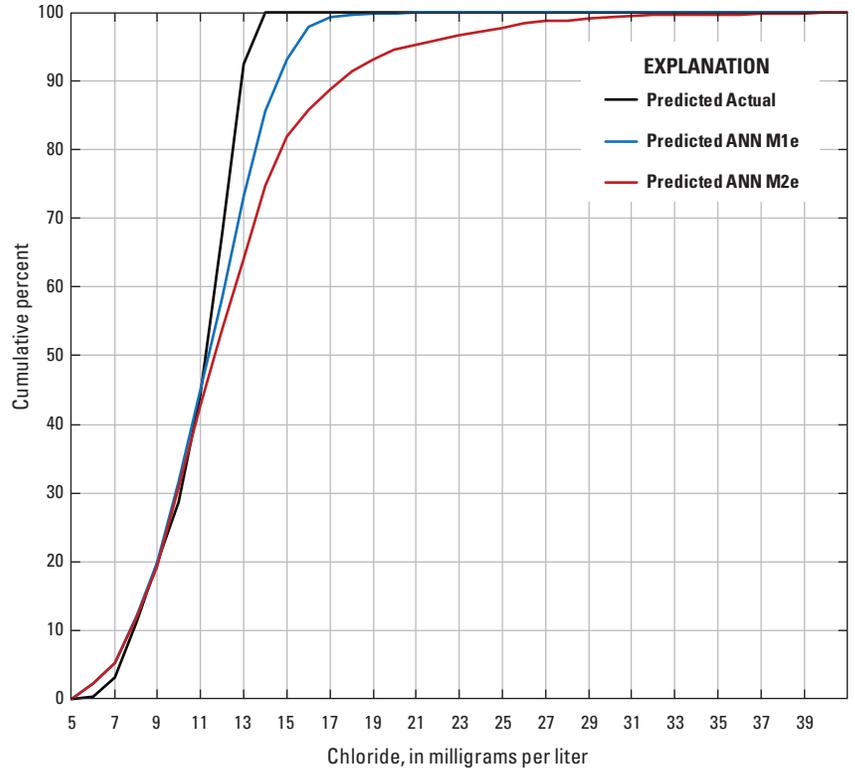
**Figure 36.** Measured specific conductance, EFDC model predicted change in specific conductance at I-95, and the SCM DSS specific conductance input for a 4-foot harbor deepening and mitigation scenario for the period January 14, 2003, to October 31, 2009.



**Figure 37.** Simulated actual and user-defined chloride concentrations at the intake using the ANN M1e and ANN M2e models for the period January 14, 2003, to October 31, 2009.

that supply water for industrial users, but the frequency and duration of high chloride concentrations in source water can be of greater concern. The cumulative frequency distribution of the chloride concentration response predicted by the two ANN models for the deepening scenario for the simulation period is shown in figure 38. Fifty percent of the time the relative difference between chloride concentrations for simulated historical conditions and the simulated conditions of harbor deepening is not large. The effect of the harbor deepening and the mitigation effort is evident in the cumulative percent differences above 50 percent where there is a divergence of the ANN M1e and M2e models from the simulated historical condition. As seen in the time series plot (fig. 37), the M2e model is predicting larger chloride concentration values than the M1e model for the 4-ft deepening scenario. The cumulative frequency distribution shows that the M2e model predicts chloride concentrations of 21.0 mg/L or less 95 percent of the time, whereas the M1e model predicts chloride concentrations of less than 15.5 mg/L 95 percent of the time.

The number of days that source water with high chloride concentration occurs is important for plant operations and planning. Table 7 lists the number of days that source water with a chloride concentration of 15 mg/L or greater occurred during the simulation period. For historical (actual) conditions, there was only one occurrence when the concentrations were greater than 15 mg/L for 3 consecutive days. The ANN models predicted increases in the frequency and duration of days with chloride concentration greater than 15 mg/L. For example, for chloride concentration of 15 mg/L or greater occurring for 7 consecutive days, the ANN M1e and ANN M2e models predicted that there would be 44 and 224 days, respectively, when these conditions occur.



**Figure 38.** Frequency distribution of simulated actual and user-defined chloride concentrations at the intake using the ANN M1e and ANN M2e models for the period January 14, 2003, to October 31, 2009.

**Table 7.** Number of days that chloride concentrations are at or greater than 15 milligrams per liter for measured conditions and simulated conditions by two ANN models at the City of Savannah’s water plant intake on Abercorn Creek.

[>, greater than]

Number of days	Measured	ANN M1e	ANN M2e
>3	0	76	301
>7	0	44	224
>14	0	17	151
>30	0	0	71
>60	0	0	11
>120	0	0	0
Maximum	3	24	71

## Summary and Conclusions

There is a concern that the proposed deepening of Savannah Harbor could increase salinity and chloride concentrations in the upper reaches of the lower Savannah River estuary, including Abercorn Creek, where the City of Savannah operates an intake for their water treatment plant. Elevated chloride concentrations would require additional water treatment to meet the needs of their industrial customers. The U.S. Geological Survey and the U.S. Army Corps of Engineers–Savannah District determined that an opportunity existed to develop an empirical model using data-mining techniques artificial neural network (ANN) models in order to simulate chloride concentrations at the City of Savannah’s water plant intake. Hydrologic and water-quality data have been collected in the lower Savannah River estuary for many years. Data characterizing the hydrology of the system—streamflows, water levels, specific conductance (field measurement used for computing salinity), water temperature, dissolved oxygen, precipitation—have been collected at numerous gaging stations in the Savannah River Basin since the mid-1990s. The City of Savannah has collected daily water-quality measurements at the water plant intake since the late 1980s.

To evaluate the potential effects of the proposed deepening of Savannah Harbor, the U.S. Army Corps of Engineers funded the development of mechanistic and empirical modeling approaches to simulate chloride concentrations at the City of Savannah’s intake. Both modeling approaches used data from the U.S. Geological Survey real-time network and data collected by the City of Savannah at the intake for the period 2000–2009. One modeling approach (funded separately from this study) modified the mechanistic Environmental Fluid Dynamics Code (EFDC) model used for evaluating proposed harbor deepening effects for the Environmental Impact Statement. Chloride concentrations were modeled directly with EFDC as a conservative tracer. The EFDC model uses boundary input data of streamflow, riverine and harbor chloride concentrations, and coastal water levels. The empirical modeling approach (the study described in this report) developed models directly from available data using ANN and linear regression models that used streamflow at Cloy, specific conductance and temperature at I–95, and water color time series at the water intake for inputs.

Few data points describe the relation between high specific conductance values at I–95 and the intake. Salinity simulations based on the EFDC model of proposed deepening scenario simulations predicted that the salinity concentration at I–95 could be three times the historical concentrations. The development of empirical models for predicting chloride concentrations at the intake required accommodation of large extrapolations from historical conditions. To accommodate these concerns, two ANN chloride models were developed for the intake. The first model (ANN M1) used all the data. The second model (ANN M2) only used data when specific conductance values at I–95 were less than 175 microsiemens per centimeter ( $\mu\text{S}/\text{cm}$ ). The chloride simulations with ANN

M1 have a low sensitivity to specific conductance (salinity) at I–95, whereas the chloride simulations with ANN M2 have a high sensitivity to salinity at I–95. To accommodate the geometry changes in the harbor, the ANN models use the EFDC model-simulated salinity changes from a historical condition as input.

The specific conductance and chloride models, historical database, model simulation controls, streaming graphics, and model output were integrated into a decision support system named the Savannah Chloride Model Decision Support System (SCM DSS). The SCM DSS allows the user to manipulate the streamflow and specific conductance inputs to the system. Three options are available to the user in setting the streamflows or specific conductance: percentage of historical flow, constant flow, and a user-defined hydrograph. A fourth option is available for setting specific conductance by using output from an EFDC model simulation. Output from the SCM DSS includes tabular time series of predictions of the measured data and predictions of the user-specified conditions. The SCM DSS is a spreadsheet application that facilitates the dissemination and utility of the DSS.

Two scenarios were simulated with the SCM DSS to demonstrate different input options. One scenario increased winter streamflows to a constant streamflow for 45 days. Streamflows during the period January 1 to February 15 were set to a constant 3,600 cubic feet per second for the simulation period of October 1, 2006, to October 1, 2009. These decreased winter streamflows resulted in simulated increases in specific conductance of as much as 50  $\mu\text{S}/\text{cm}$  and chloride concentrations of as much as 4.8 milligrams per liter (mg/L) during the periods of decreased streamflows.

The second scenario used EFDC generated output for a 4-foot (ft) deepening of the harbor. The 4-ft deepening scenario included changes in the channel and flow configuration to mitigate for salinity increases in the vicinity of an extensive freshwater tidal marsh. A 4-ft harbor deepening scenario was simulated with the models for the 7-year period from January 2003 to October 2009. The ANN M2e model is more sensitive than the ANN M1e model to changes in specific conductance resulting from a 4-ft deepening and simulated chloride concentrations as high as 40 mg/L. The ANN M1e model, trained on all the data, is less sensitive than the ANN M2e model to the changes in specific conductance and simulated chloride concentrations greater than 20.3 mg/L.

The ANN models predicted increases in the frequency and duration of days with chloride concentration greater than 15 mg/L. The ANN M1e and ANN M2e models predicted 44 and 224 days, respectively, with concentrations greater than 15 mg/L for the previous 7 days. For historical (actual) conditions, there was only one occurrence when the concentrations were greater than 15 mg/L for 3 consecutive days.

The simulation of harbor deepening and mitigation scenarios is a large extrapolation of the salinity and chloride dynamics of the lower Savannah River estuary to conditions that have never been measured in the system. The extrapolations are dependent on the sensitivity of the models to salinity/

chloride dynamics at the intake to the flow and salinity dynamics in the Savannah River. The historical data show a low sensitivity between salinity at I-95 and salinity at the intake. One of the models, ANN M1e, has a similar low sensitivity and a dampening of salinity intrusion in Abercorn Creek. If the dynamics of the system are maintained between a pre- and post-deepening, including background chloride and specific conductance conditions, the chloride response at the intake should be closer to that predicted by the ANN M1e model. If the deepening changes the intrusion dynamics between I-95 and the intake and there is not a dampening of the intrusion in Abercorn Creek, the chloride response at the intake will be closer to predictions by the ANN M2e model. For either case, only detailed post-construction monitoring of salinity and chloride concentrations in Abercorn Creek will show how the system is responding to a deepening of Savannah Harbor.

## References Cited

- Buccola, N.L., and Wood, T.M., 2010, Empirical models of wind conditions on Upper Klamath Lake, Oregon: U.S. Geological Survey Scientific Investigations Report 2010-5201, 26 p.
- Collins, M.R., Post, W.C., and Russ, D.C., 2001, Distribution of shortnose sturgeon in the lower Savannah River: South Carolina Department of Natural Resources, final report to Georgia Ports Authority, at <http://sav-harbor.com/>.
- Conrads, P.A., and Roehl, E.A., Jr., 1999, Comparing physics-based and neural network models for simulating salinity, temperature, and dissolved oxygen in a complex, tidally affected river basin, *in* Proceedings of the 1999 South Carolina Environmental Conference, March 15-16, 1999, Myrtle Beach, SC.
- Conrads, P.A., and Roehl, E.A. Jr., 2005, Integration of data mining techniques with mechanistic models to determine the impacts of non-point source loading on dissolved oxygen in tidal waters, South Carolina Environmental Conference, March 2005, Myrtle Beach, SC.
- Conrads, P.A., and Roehl, E.A., Jr., 2007, Analysis of salinity intrusion in the Waccamaw River and the Atlantic Intracoastal Waterway near Myrtle Beach, South Carolina, 1995-2002: U.S. Geological Survey Scientific Investigations Report 2007-5110, 41 p., 2 apps.
- Conrads, P.A., and Roehl, E.A., Jr., 2010, Analysis and simulation of water-level, specific conductance, and total phosphorus dynamics of the Loxahatchee National Wildlife Refuge, Florida, 1995-2006: U.S. Geological Survey Scientific Investigations Report 2010-5244, 64 p.
- Conrads, P.A., Roehl, E.A., Jr., Daamen, R.C., and Kitchens, W.M., 2006, Simulation of water levels and salinity in the rivers and tidal marshes in the vicinity of the Savannah National Wildlife Refuge, Coastal South Carolina and Georgia: U.S. Geological Survey Scientific Investigations Report 2006-5187, 134 p.
- Conrads, P.A., Roehl, E.A., Jr., and Martello, W.P., 2003, Development of an empirical model of a complex, tidally affected river using artificial neural networks: Water Environment Federation, National TMDL Science and Policy 2003 Specialty Conference, November 2003, Chicago, IL.
- Devine, T.W., Roehl, E.A., Jr., and Busby, J.B., 2003, Virtual sensors—Cost effective monitoring, *in* Proceedings of the Air and Waste Management Association Annual Conference, June 2003, San Diego, CA.
- Dutta, S., Wierenga, B., and Dalebout, A., 1997, Case-based reasoning systems—From automation to decision-aiding and stimulation: IEEE Transactions on Knowledge and Data Engineering, v. 9, no. 6, p. 911-922.
- Dyer, K.R., 1997, Estuaries—A physical introduction (2d ed.): New York, John Wiley and Sons, Ltd, 195 p.
- Georgia Environmental Protection Division, 2001, Savannah River Basin Watershed Protection Plan: Georgia Department of Natural Resources, Environmental Protection Division, accessed May 9, 2011, at <http://www.gaepd.org/Documents/savannah.html>.
- Hamrick, J.M., 1992, A three-dimensional environmental fluid dynamics computer code—Theoretical and computational aspects: The College of William and Mary, Virginia Institute of Marine Science, Special Report 317, 63 p.
- Helsel, D.R., and Hirsch, R.M., 1995, Studies in environmental science 49—Statistical methods in water resources: Amsterdam, Elsevier Sciences, 529 p.
- Hinton, G.E., 1992, How neural networks learn from experience: Scientific American, v. 267, no. 3, p. 145-151.
- Jensen, B.A., 1994, Expert systems—Neural networks, *in* Lipták, B.G., ed., Process control, Instrument engineers' handbook (3d ed.): Boca Raton, FL, CRC Press, p. 48-54.
- Miller, R.L., Bradford, W.L., and Peters, N.E., 1988, Specific conductance—Theoretical considerations and application to analytical quality control: U.S. Geological Survey Water-Supply Paper 2311, 16 p.
- Rantz, S.E., and others, 1982, Measurement and computation of streamflow, volumes 1 and 2: U.S. Geological Survey Water-Supply Paper 2175, 631 p.

- Risley, J.C., Roehl, E.A., Jr., and Conrads, P.A., 2003, Estimating water temperatures in small streams in western Oregon using neural network models: U.S. Geological Survey Water-Resources Investigations Report 02–4218, 59 p.
- Roehl, E.A., Jr., Conrads, P.A., and Cook, J.B., 2003, Discussion of “Using complex permittivity and artificial neural networks for contaminant prediction” by John B. Lindsay, Julie Q. Shang, and R. Kelly Rowe: *Journal of Environmental Engineering*, v. 129, no. 11, p. 1069–1071.
- Roehl, E.A., Jr., Conrads, P.A., and Daamen, R.C., 2006, Features of advanced decision support systems for environmental studies, management, and regulation, in Voinov, A., Jakeman, A.J., Rizzoli, A.E., eds., *Proceedings of the iEMS Third Biennial Meeting—Summit on Environmental Modelling and Software*, July 9–13, 2006, Burlington, VT: International Environmental Modelling and Software Society, CD-ROM, accessed May 9, 2011, at <http://www.iemss.org/iemss2006/sessions/all.html>.
- Roehl, E.A., Jr., Conrads, P.A., and Roehl, T.A.S., 2000, Real-time control of the salt front in a complex, tidally affected river basin, *Smart Engineering System Design: Volume 10, Proceedings of the Artificial Neural Networks in Engineering Conference*, New York, ASME Press, p. 947–954.
- Rosenblatt, Frank, 1958, The perceptron—A probabilistic model for information storage and organization in the brain: *Psychological Review*, v. 65, no. 6, p. 386–408.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986, Learning internal representations by error Propagation, in Rumelhart D.E., McClelland, J.L., and the PDP Research Group, *Parallel distributed processing—Explorations in the microstructure of cognition, Volume 1—Foundations*: Cambridge, MA, MIT Press, p. 318–362.
- Tetra Tech, 2006a, Development of the Hydrodynamic and Water Quality Models for the Savannah Harbor Expansion Project, Prepared for U.S. Army Corps of Engineers—Savannah District, Contract No.: DACA65-99-D-0065, at <http://www.sas.usace.army.mil/shexpan/SHEPreportSupplStudies.html>.
- Tetra Tech, 2006b, Savannah Harbor Expansion Project—Chloride data analysis and model development, Prepared for U.S. Army Corps of Engineers—Savannah District, Contract No.: W912-HN-05-D-0014 Work Order 0018, at <http://www.sas.usace.army.mil/shexpan/SHEPreportSupplStudies.html>.
- Tetra Tech, 2010, Chloride Modeling Savannah Harbor Expansion Project Savannah, Georgia, Prepared for U.S. Army Corps of Engineers – Savannah District, Contract No.: W912-HN-05-D-0014, at <http://www.sas.usace.army.mil/shexpan/SHEPreportSupplStudies.html>.
- U.S. Army Corps of Engineers, 2010, Draft Tier II Environmental Impact Statement for Savannah Harbor Expansion Chatham County, Georgia and Jasper County, South Carolina, accessed May 9, 2011, at <http://www.sas.usace.army.mil/shexpan/SHEPTierII.html>.
- U.S. Geological Survey, 2011 Water-resources data for the United States, water year 2010: U.S. Geological Survey Water-Data Report WDR-US-2010, accessed May 11, 2011, at <http://wdr.water.usgs.gov/>.
- Wagner, R.J., Matraw, H.C., Ritz, G.F., and Smith, B.A., 2000, Guidelines and standard procedures for continuous water-quality monitors—Site selection, field operation, calibration, record computation, and reporting: U.S. Geological Survey Water-Resources Investigations Report 00–4252, 53 p.
- Weiss, S.M., and Indurkha, Nitin, 1998, *Predictive data mining—A practical guide*: San Francisco, CA, Morgan Kaufmann Publishers, Inc., 228 p.
- Welch, Z.C., and Kitchens, W.M., 2006, Predicting freshwater and oligohaline tidal marsh vegetation communities, Florida Cooperative Fish and Wildlife Research Unit, University of Florida, Report Submitted to U.S. Army Corps of Engineers, Solicitation: W91278-06-T-0012, at <http://sav-harbor.com/>.
- Will, T.A., and Jennings, C.A., 2001, Assessment of spawning sites and reproductive status of striped bass, *Morone saxatilis*, in the Savannah River estuary, Georgia Cooperative Fish and Wildlife Research Unit, University of Georgia, Final Report for Project 10-21-RR251-144 prepared for The Georgia Ports Authority, accessed May 9, 2011, at <http://sav-harbor.com/>.

# Appendix 1: User's manual for the Savannah Chloride Model Decision Support System (SCM DSS)

## 1. Introduction

This document describes how to install and operate the Savannah Chloride Model Decision Support System (SCM DSS) of Abercorn Creek. The SCM DSS is a decision support system (DSS) built around a suite of empirical hydrologic, specific conductance, and chloride models.

## 2. SCM DSS Installation, Removal, and Technical Assistance

**NOTE:** *SCM DSS will not run on 64-bit Windows XP® and Vista® operating systems because of incompatibility of the NNCALC32.xll Add-in with these operating systems. NNCALC32.xll is used to execute the Artificial Neural Network (ANN) models.*

### 2.1 Installation

1. Create a folder called SCM at the top level of your C: drive.
2. Extract all files from the distributed SCM-yyyymmdd.zip<sup>1</sup> file. The zip file contains the following application files:
  - SCM-yyyymmdd.xls – a Microsoft (MS) Excel® spreadsheet application.
  - 6 files with an “enn” extension – these are the ANN files.
  - NNCALC32.xll – a custom MS Excel add-in used to execute the \*.enn files.
  - SCMUserGuide-yyyymmdd.doc – the MS Word file that you are reading right now.
3. Open your copy of MS Excel for MS Office 2000® (or newer). Ensure that the standard Excel Add-ins listed below are installed and checked “available.”

Analysis Toolpak  
Analysis Toolpak – VBA

Add-ins are accessed from Excel's Tools menu. If any are missing, it may be necessary to install them from your MS Office CD-ROM.

4. Set the macro security level of Excel to either medium or low using Tools > Macro > Security. SCM DSS uses VBA macros for a variety of purposes and must be able to execute them to operate correctly.
5. Install the NNCALC32 Add-in that resides in the NNCALC folder described in Step 1. This may be accomplished by clicking on Tools > Add-ins > Browse, browse to the SCM folder you created, click on the NNCALC32 icon, then click OK.
6. Open the SCM-yyyymmdd.xls Excel spreadsheet application. When Excel asks if you want to run macros click “Enable Macros,” otherwise SCM DSS will not operate correctly.

Select the “Controls” worksheet (fig. A1). “Controls” is the worksheet that lets the user set up and run simulations. At the top-left corner of “Controls” is a text box labeled “Where Model Files Are Located.” The model files are the \*.enn files of the ANNs. Type in the fully qualified path name of the folder setup in (1) above and save the Excel application using File > Save for the setup changes to be permanent.

To check that the models are connected and operating correctly, select the “Controls” worksheet (fig. A2). At upper right are fields with the row headers “SC M1 / M2” and “CH M1 / M2,” SC for specific conductivity and CH for chlorides. If these fields show numerical values and not an MS Excel or NNCALC32 error code, the application is properly configured and ready to use. If all of these fields show “?” or an error code then try exiting MS Excel and then reloading MS Excel and the SCM application.

An error code indicates that an ANN cannot execute because either the NNCALC32 Add-In is not installed per (4) or NNCALC32 cannot find \*.enn files because the folder path name in the “Where model files are located” text box is incorrect. If you cannot get SCM to operate, re-check the configuration items in (3)–(6) above.

<sup>1</sup> yyyymmdd is the version date of the SCM image to be installed.

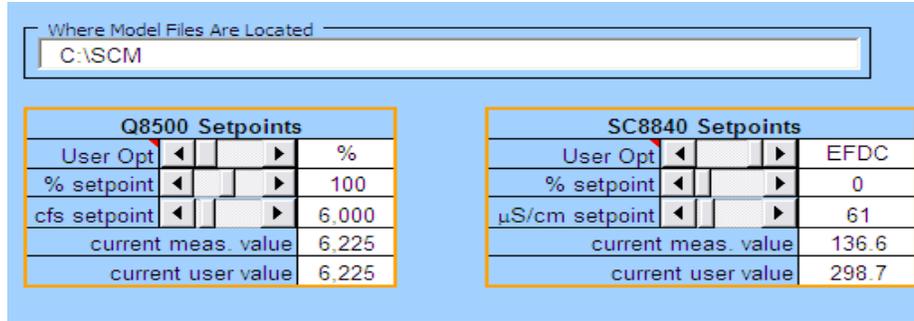


Figure A1. "Controls" worksheet screenshot.

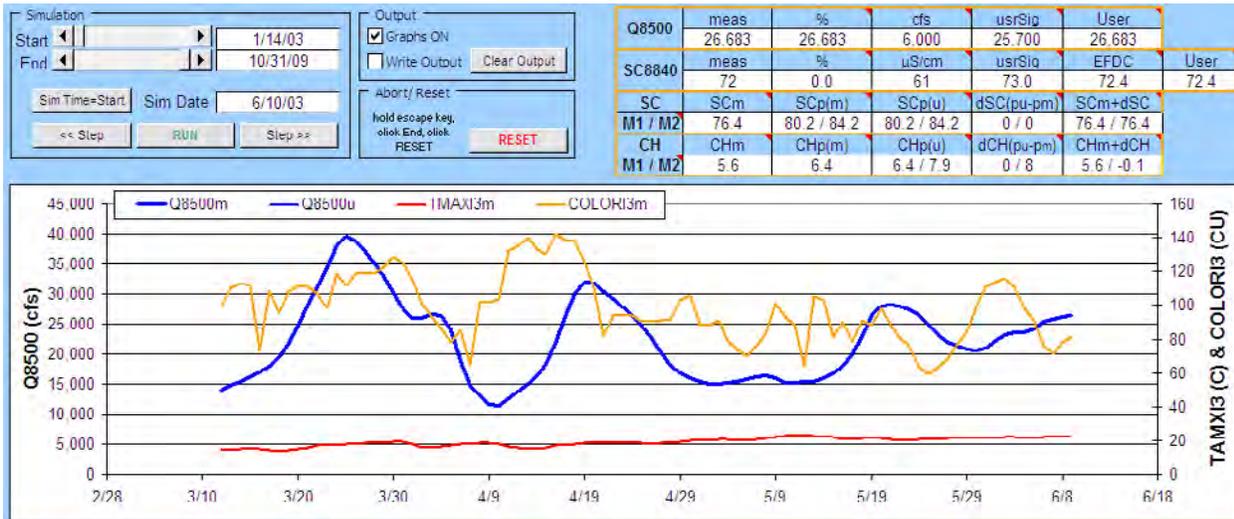


Figure A2. Upper portion of "Controls" worksheet showing "SC M1 / M2" and "CH M1 / M2" fields at upper right.

## 2.2 Removal

Simply delete the folder created to hold the SCM DSS files and its contents. Consider removing the Add-ins and reverting to the default MS Excel security settings.

## 2.3 Technical Assistance

Please contact Paul Conrads of the USGS at (803) 750–6140, pconrads@usgs.gov, if you have questions or problems with SCM DSS.

## 3. SCM DSS Features and Operation

SCM DSS is opened like any standard Excel workbook. Simply open the SCM-yyyyymmdd.xls file and begin. The SCM DSS and its graphical user interface (GUI) are made up of a number of worksheets that are detailed below.

## 3.1 "Info" Worksheet

The "Info" worksheet is automatically displayed when SCM is first loaded (fig. A3). It shows a map of the study area, and gives the application's version date and the contact information of its developers.

## 3.2 Variable Descriptions and "ReleaseNotes" Worksheet

SCM refers to many input and output variables in the form of row and column headers (fig. A4). Moving the mouse over a header marked with a red caret immediately above and to the right of the header will provide a description of the header variable. Descriptions of variables also are provided in the "ReleaseNotes" worksheet (fig. A5). This worksheet also describes SCM's development history and any new features or changes.

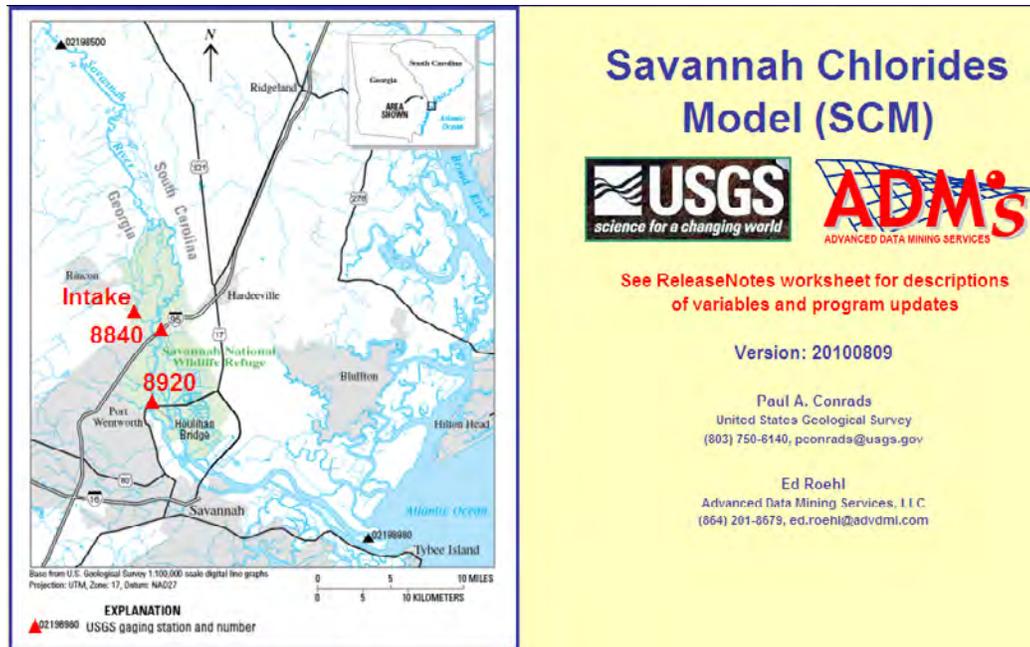


Figure A3. "Info" worksheet.

Q8500	meas	%	cfs	usrSig	User	
	6,225	6,225	6,000	6,040	6,225	
SC8840	meas	%	µS/cm	usrSig	EFDC	User
	137	0.0	61	422.0	200.7	298.7
SC	SCm	SCp(m)	SCp(u)	d	Model's predicted Daily and Hourly SC using user setpoints	
M1 / M2	132.7	131.1 / 124.1	162.2 / 186.3	3		
CH	CHm	CHp(m)	CHp(u)	d		
M1 / M2	12.5	12.3	13.2 / 15.1	0.8 / 11.3	13.3 / 3.7	

Figure A4. Online description of variable SCp(u) on "Controls" worksheet.

VARIABLE	DESCRIPTIONS OF OUTPUT VARIABLES
DATE	Date stamp
ROW	Database worksheet for identifier
Q8500 Input Opt	Q8500 user-input option selected on the "Setpoints" worksheet
Q8500m	Measured daily average Q8500 (cfs)
Q8500u	User input daily average Q8500 (cfs)
WL8840m	Measured daily average WL8840 (ft) - in "Database" worksheet for reference only
XWL8840m	Measured daily average tidal range XWL8840 (ft) - in "Database" worksheet for reference only
TMAXI3m	Measured 8840 daily average air temperature, interpolated up to 3 days to fill missing values
RELEASE NOTES	
DATE	CHANGE
20100325	First Release of SCM
"	Functionality to read EFDC output is not yet implemented pending further specification of how it is to operate; however, EFDC output can be appropriately configured outside of SCM, pasted into SCM's "EFDC" worksheet, and run.
20100414	Modified output to provide more decimal places.
"	Added user controls and userdefined signals for SC8840MIN to Controls, UserInputSignals (formerly usrHyds), Output, and the new Setpoints worksheets.

Figure A5. "ReleaseNotes" worksheet.

### 3.3 “Controls” Worksheet

The “Controls” worksheet (fig. A2) is the GUI component that lets the user setup and run simulations.

At the top is a text box labeled “Where Model Files Are Located” and is used to configure SCM when it is first installed on a user’s computer and is described further in Section 2.0. As shown in figure A6, “Start” and “End” dates for simulations can be set using the controls at upper left. The end date must be more recent than the start date. The “Sim Date” text box indicates the time stamp that is providing the current input values to SCM’s models. The “<<Step” and “Step>>” move the current time stamp backwards or forwards one time step each time they are clicked. “Sim Time=Start” sets the current time stamp to the Simulation “Start” date. “RUN” will start and run a simulation between the dates indicated by the Simulation “Start” and “End” dates.

The “Controls” worksheet provides numerical and streaming graphical information that can be observed during simulations or when incrementally stepping through time. This allows the user to examine specific periods and behaviors of interest in detail. The SCM also will write inputs and output data to the “Output” worksheet. Because of the added computational load, simulations are slowed when streaming graphics and simulation output are generated. The “Graphs ON” and “Write Output” check boxes of the “Output” controls at the upper right in figure A6 allow the user to toggle the streaming graphics “on” or “off”. The “Clear Output” button erases all data in the “Output” worksheet to allow data from a new simulation to be recorded.

A simulation may be stopped at any time during an execution by holding down the “Esc” key, after which a pop-up window will appear like that shown in figure A7. Click on the “End” button to stop the simulation, then click the “Reset”

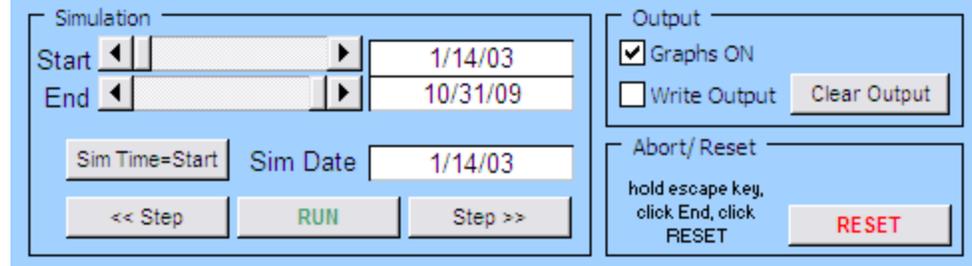


Figure A6. Simulation controls on “Controls” worksheet.

button shown at lower right in figure A6. The “Reset” button activates MS Excel’s automatic calculation feature (autocalc). Because the model programmatically manipulates autocalc for performance reasons, aborting a simulation can sometimes leave the model in a state where autocalc is not activated. This is remedied by clicking the “Reset” button.

### 3.4 “Setpoints,” “UserInputSignals,” and “EFDC” Worksheets

Figure A8 shows the “Setpoints” worksheet. The input parameters that can be manipulated by the user are Q8500 and SC8840, and there are several options for doing so. The following Q8500 inputs options are selected using its “User Opt” control.

- “%” – percent of historical flow. The “% setpoint” control is used to set the percentage.
- “cfs” – fixed flow rate. The “cfs setpoint” control is used to set the flow rate.
- usrSig – user-defined signal, which the user can paste into the “UserInputSignals” worksheet (fig. A8). The “Clear usrSigs” button clears the “user Q8500” and “User SC8840” fields.

The following SC8840 inputs options are selected using its “User Opt” control.

- “%” – percent of SC8840. The “% setpoint” control is used to set the percentage.
- “uS/cm” – fixed specific conductivity. The “uS/cm setpoint” control is used to set the specific conductivity.
- usrSig – user-defined signal, which the user can paste into the “UserInputSignals” worksheet (fig. A8).
- “EFDC” – uses EFDC output data to bias the SC8840 historical data, which is then input to the models. The EFDC data is loaded into SCM as a comma separated value (CSV) file with a specified format. The EFDC data are loaded into the “EFDC” worksheet (fig. A9). The user must type the path-

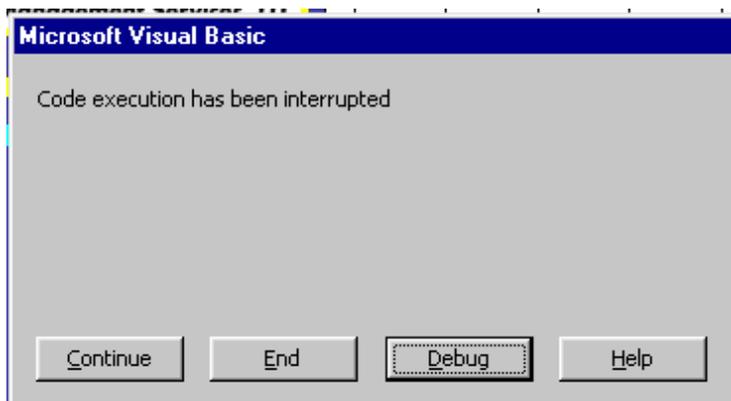


Figure A7. Pop-up window that appears when a simulation is interrupted by using the “Esc” key.

DATE	User Q8500	User SC8840
1/1/03	7560	104
1/2/03	7190	113
1/3/03	7000	125
1/4/03	6910	117
1/5/03	6850	118
1/6/03	7040	131
1/7/03	6920	129
1/8/03	6580	128
1/9/03	6370	124
1/10/03	6380	124
1/11/03	6430	132
1/12/03	6230	134
1/13/03	6060	134
1/14/03	6040	133

Clear usrSigs

**NOTES**  
 Paste the data for user-defined Q8500 and SC8840 signals into the columns at left.

It is left to the user to check the correctness of the user-defined input data and its synchronization with the given time stamps.

Figure A8. "UserInputSignals" worksheet.

Load File Data

Clear File Data

EFDC CSV File Path  
 C:\SCM-2\Channel5ftdeeperMitigation6AScenario-SCM.csv

**NOTE** - Hourly parameters predicted by EFDC are loaded by either pasting the data into the cyan-colored columns or loading it from a file having the prescribed format. The file loader will crash at the cell where any inappropriate data is encountered. It is left to the user to check the correctness of the user-defined input data and its synchronization with the given time stamps.

DATE	TIME	SALT8840-EFDCp(m)	SALT8840-EFDCp(u)	SC8840-EFDCp(m)	SC8840-EFDCp(u)	CH8840-EFDCp(m)	CH8840-EFDCp(u)
1/1/03	0:00	5.3E-05	0.000221	132.4199982	132.7400055	9.412433624	9.469901085
1/1/03	1:00	3.1E-05	0.000111	132.3699951	132.5299988	9.439320564	9.487821579
1/1/03	2:00	1.7E-05	6.3E-05	132.3500061	132.4400024	9.502119064	9.551641464
1/1/03	3:00	1E-05	4.2E-05	132.3300018	132.3999939	9.576375008	9.651046753
1/1/03	4:00	4E-05	4.2E-05	132.3899994	132.3999939	9.557168961	9.539047241
1/1/03	5:00	0.000334	0.002873	132.9499969	137.7700043	9.483275414	9.584448814
1/1/03	6:00	0.002017	0.044436	136.1399994	216.5500031	9.535814285	9.853849411
1/1/03	7:00	0.013364	0.117285997	157.6600037	354.3500061	9.600149155	10.11502266
1/1/03	8:00	0.020153999	0.082001999	170.5399933	287.6600037	9.612464905	10.02068138
1/1/03	9:00	0.007557	0.041283999	146.6499939	210.5800018	9.597166061	9.898773193
1/1/03	10:00	0.002049	0.01757	136.1999969	165.6399994	9.559149742	9.768639565
1/1/03	11:00	0.000741	0.007566	133.7200012	146.6699982	9.510479927	9.619592667
1/1/03	12:00	0.00034	0.003394	132.9600067	138.75	9.503297806	9.564650536
1/1/03	13:00	0.00017	0.001589	132.6399994	135.3300018	9.513816833	9.568307877
1/1/03	14:00	9.1E-05	0.00079	132.4900055	133.8099976	9.531633377	9.621137619
1/1/03	15:00	4.7E-05	0.000441	132.3999939	133.1499939	9.564516068	9.707937241

Figure A9. "EFDC" worksheet. The "Clear usrSigs" button clears the "user Q8500" and "User SC8840" fields.

name of the EFDC file to be loaded into the "EFDC CSV File Path" text box. The "Load File Data" button loads the file, and the loaded data can be inspected in the cyan-colored fields of the worksheet. The "Clear File Data" button clears the data in the cyan fields. Note: only the data in the "SC8840-EFDCp(m)" and "SC8840-EFDCp(u)" columns is used by SCM.

### 3.5 "Database" and "Output" Worksheets

The "Database" worksheet contains the time series data used by SCM to run simulations (fig. A10). These data are described in the "ReleaseNotes" worksheet, and are derived from the raw field measurements. They are augmented by calculated variables whose values are calculated on-the-fly by SCM's computer code. The user should not alter data in the "Database" worksheet.

The “Output” worksheet contains a record of key variables for a particular simulation (fig. A11). The “Write Output” check box on the “Controls” worksheet must be checked for output to be written. The variables written to the “Output”

worksheet are explained in “ReleaseNotes” worksheet. The user can copy output values into another MS Excel workbook for further analysis.

DATE	ROW	Q8500m	WL8840m	XWL8840m	SC8840m	TMAXI3m	COLORI3m	SCI3m	CHm
1/1/03	2	unk	unk	unk	unk	unk	unk	100.5	unk
1/2/03	3	7418	unk	unk	unk	10.3	91	100.35	10.05
1/3/03	4	7554	1.55	7.89	116.8	10.7	88	100.2	10.7
1/4/03	5	7180	1.37	7.84	126.6	10.9	62	102.4	10.8
1/5/03	6	6993	1.09	8.31	123.6	10.7	69	109.63	11.2
1/6/03	7	6910	1.81	7.81	124.3	10.3	58	116.87	11.35
1/7/03	8	6854	1.46	7.38	132.1	10.1	70	124.1	12.5
1/8/03	9	7043	1.00	6.98	131.4	9.9	70	127	11.85
1/9/03	10	6915	0.63	6.58	129.5	9.6	49	121.7	11.5
1/10/03	11	6575	0.48	6.42	129.5	10	74	138.8	12.25
1/11/03	12	6363	0.57	5.97	126.8	10.5	76	120	11.45
1/12/03	13	6381	0.65	5.86	132.9	10.3	79	124.23	11.4
1/13/03	14	6430	1.01	5.67	134.9	10.3	69	128.47	11
1/14/03	15	6225	1.18	5.87	136.6	10	64	132.7	12.5
1/15/03	16	6060	0.93	6.25	135.1	9.7	65	130.2	12.25
1/16/03	17	6043	0.80	6.52	139.4	9.5	68	126.4	12.15
1/17/03	18	6168	1.18	7.14	151.2	9.5	59	124.7	12.5
1/18/03	19	6049	0.77	7.36	149.4	9.5	60	130.1	12.2

Figure A10. Example measured data from the “Database” worksheet.

DATE	ROW	Q8500 Input Opt	Q8500m	Q8500u	WL8840m	XWL8840m	TMAXI3m	COLORI3m	SC8840-EFDC(pu-pm)	SC8840 Input Opt	SC8840m	SC8840u
1/14/03 0:00	15	usrSig	6225	6040	1.18	5.87	10	64	162.11	uS/cm	136.62	61
1/15/03 0:00	16	usrSig	6060	6170	0.93	6.25	9.7	65	190.46	uS/cm	135.06	61
1/16/03 0:00	17	usrSig	6043	6050	0.8	6.52	9.5	68	239.59	uS/cm	139.38	61
1/17/03 0:00	18	usrSig	6168	5870	1.18	7.14	9.5	59	167.53	uS/cm	151.18	61
1/18/03 0:00	19	usrSig	6049	5710	0.77	7.36	9.5	60	227.41	uS/cm	149.42	61
1/19/03 0:00	20	usrSig	5863	5710	1.04	7.96	9.1	59	189.09	uS/cm	149.36	61
1/20/03 0:00	21	usrSig	5709	6040	0.71	7.91	8.5	62	141.39	uS/cm	149.29	61
1/21/03 0:00	22	usrSig	5715	6040	0.24	8.2	8.1	57	128.24	uS/cm	150.14	61
1/22/03 0:00	23	usrSig	6045	5870	0.37	8.45	9	52	105.79	uS/cm	154.66	61
1/23/03 0:00	24	usrSig	6038	5860	1.1	8.21	9.3	43	73.1	uS/cm	158.98	61
1/24/03 0:00	25	usrSig	5869	5900	1.33	7.01	9.3	49	112.01	uS/cm	156.62	61
1/25/03 0:00	26	usrSig	5856	6010	0.64	7.01	8.7	45	115.49	uS/cm	145.38	61
1/26/03 0:00	27	usrSig	5898	5860	0.81	7.61	7.7	55	95.78	uS/cm	151.89	61
1/27/03 0:00	28	usrSig	6010	5790	0.99	7.43	7.6	54	99.05	uS/cm	151.5	61
1/28/03 0:00	29	usrSig	5852	6020	0.62	7.46	7.6	48	144.8	uS/cm	151.63	61
1/29/03 0:00	30	usrSig	5796	5860	1.08	7.54	7.9	44	156.15	uS/cm	152.78	61
1/30/03 0:00	31	usrSig	6023	5730	1.13	7.9	8.5	44	157	uS/cm	148.8	61
1/31/03 0:00	32	usrSig	5851	5870	1.07	8.03	9.4	44	160.93	uS/cm	151.38	61
2/1/03 0:00	33	usrSig	5734	6270	1.09	8.08	9.4	74	138.2	uS/cm	151.7	61
2/2/03 0:00	34	usrSig	5873	6740	1.08	7.91	10.1	78	96.37	uS/cm	148.37	61
2/3/03 0:00	35	usrSig	6280	7060	1	7.91	10.5	82	61	uS/cm	151.92	61
2/4/03 0:00	36	usrSig	6748	6640	0.94	7.97	10.7	45	27.63	uS/cm	151.66	61

Figure A11. Example output from the “Output” worksheet.

**Prepared by:**

USGS Science Publishing Network  
Raleigh Publishing Service Center  
3916 Sunset Ridge Road  
Raleigh, NC 27607

Manuscript approved for publication on May 23, 2011.  
Editorial review and layout by Kay P. Naugle.  
Illustrations prepared by Jimmy Banton.  
Cover design by Ray Douglas.

**For additional information regarding this publication, contact:**

Director  
USGS South Carolina Water Science Center  
720 Gracern Road, Suite 129  
Columbia, SC 29210  
phone: 803-750-6100  
e-mail: dc\_sc@usgs.gov

**Or visit the South Carolina Water Science Center Web site at:**

*<http://sc.water.usgs.gov>*

