Simulating Riverine and Marsh Water Levels and Salinity

Simulating salinity for estuarine systems typically is done using dynamic deterministic models that incorporate the mathematical descriptions of the physics of coastal hydrodynamics. These one-, two-, or three-dimensional models generally are expensive and time consuming to apply to complex coastal systems with satisfactory results. Conrads and Roehl (2005) assert that in estuaries, mechanistic model calibration is "… particularly difficult due to low watershed gradients, poorly defined drainage areas, tidal complexities, and a lack of understanding of watershed and marsh processes." Although mechanistic models have been the state of the practice for regulatory evaluations of anthropogenic effects on estuarine systems, developments in the field of advanced statistics, machine learning, and data mining offer opportunities to develop empirical ANN models that are often more accurate. Conrads and Roehl (1999) compared the application of a deterministic model and an ANN model to simulate dissolved oxygen on the tidally affected Cooper River in South Carolina. Results of their study indicated that the ANN models offer important advantages, including faster development time, use of larger amounts of data, the incorporation of optimization routines, and model dissemination in spreadsheet applications. With the real-time gaging network on the Savannah River and the availability of large databases of hydrologic and water-quality data, the GPA realized an opportunity to develop an empirical model using data-mining techniques, including ANN models, to simulate the water level and pore-water salinity of the tidal marshes in the vicinity of SNWR.

The emerging field of data mining addresses the issue of extracting information from large databases. Data mining comprises several technologies that include signal processing, advanced statistics, multidimensional visualization, chaos theory, and machine learning. Machine learning is a field of artificial intelligence (AI) in which computer programs are developed that automatically learn cause-effect relations from example cases and data. For numerical data, commonly used methods include ANN models, genetic algorithms, multivariate adaptive regression splines, and partial and ordinary least squares (OLS).

Data mining can solve complex problems that are unsolvable by any other means. Weiss and Indurkhya (1998) define data mining as "… the search for valuable information in large volumes of data. It is a cooperative effort of humans and computers." A number of previous studies by the authors and others have used data mining to simulate hydrodynamic and water-quality behaviors in the Beaufort, Cooper, and Savannah River estuaries (Roehl and Conrads, 1999; Conrads and Roehl, 1999; Roehl and others, 2000; Conrads and others, 2002a; 2002b) and stream temperatures in western Oregon (Risley and others, 2002). These studies have demonstrated the performance of data mining to simulate water level, water temperature, dissolved oxygen, and specific conductance, and for assessing the effects of reservoir releases and point and nonpoint sources on receiving streams.

The ultimate goal of this study is to produce an effective model to simulate water level and specific conductance in the tidal marsh for a given set of streamflow, water-level, and tidal range conditions. The approach taken uses all available streamflow, water-level, and specific-conductance measurements since the last major changes in channel configuration in 1994. The modeling approach uses correlation functions that were synthesized directly from data to simulate how the change in water level and specific conductance at each station location is affected by streamflow and tidal conditions over time.

Limitations of the Data Sets

As with any modeling effort, empirical or deterministic, 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 and marsh data-collection networks can limit the range of streamflow, water-level, tidal range, and salinity conditions that the ANN model can accurately simulate. As noted previously, substantial changes in the salinity response of the system can occur due to a small change in streamflow (fig. 12). Using the specific-conductance record for the Little Back River at USFW Dock (station 021989791) as representative of the salinity dynamics of the system, scatter plots of daily streamflow and specific-conductance data were generated for each network for the period that it was active (fig. 18). The period of record for the USGS river network (fig. 18*A*) covers the full range of historical conditions from extreme low to high flows. The GPA river network (fig. 18*A*), active during summer 1997 and 1999, covers a much smaller range of hydrologic conditions and did not measure streamflow and the corresponding salinity response for streamflows less than 5,440 ft^3/s or greater than 11,600 ft^3/s. The USGS marsh network (fig. 18*B*) was established in 2000 during the drought and measured a large range of marsh salinity conditions for streamflow ranging from 4,310 to 39,600 ft^3/s. The GPA marsh network (fig. 18*B*) also was established during the drought, but was discontinued in 2002 and only measured low-flow conditions less than 14,100 ft^3/s. Data networks and the range of flows measured for each gaging network are summarized in table 2.

Signal Decomposition, Correlation Analysis, and State-Space Reconstruction

The behavior, or dynamics, of a natural system results from interactions between multiple physical forces. For example, the specific conductance at a fixed location is subject to daily, seasonal, and annual streamflow conditions and semidiurnal, fortnightly, seasonal, and annual tidal water-level conditions. For the application of the ANN models to the

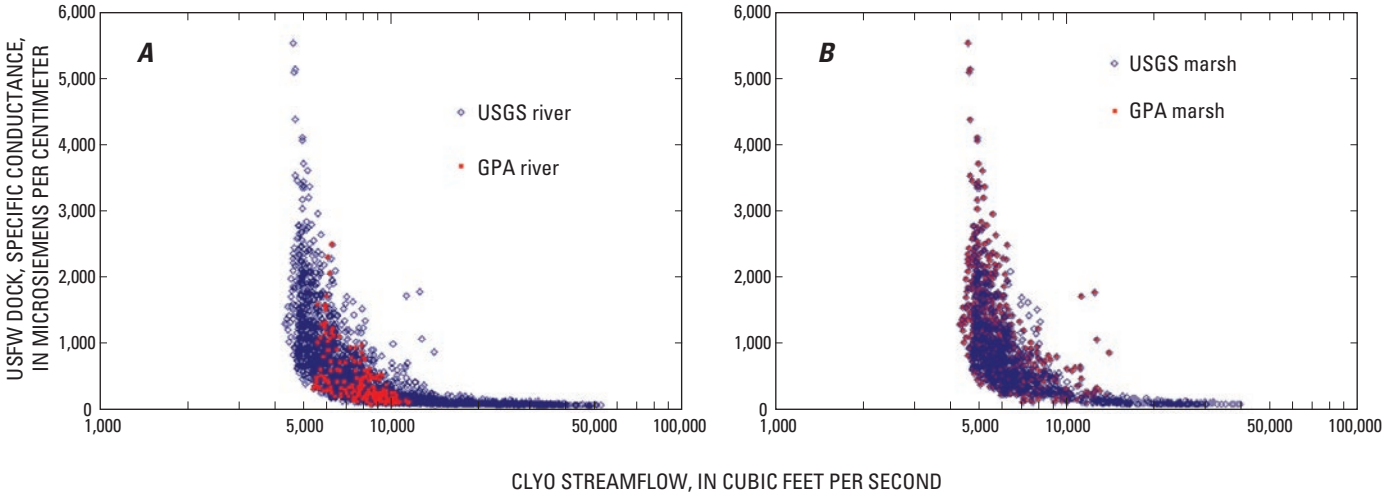

Figure 18. The daily mean streamflow at Savannah River at Clyo, Ga., and daily mean specific conductance at U.S. Fish and Wildlife Service Dock, Little Back River, for the conditions during which the gaging network was active. The two river networks are shown in the plot on the left (*A*), and two marsh networks are shown on the plot on the right (*B*).

Savannah River, data-mining methods are applied to maximize the information content in raw data while diminishing the influence of poor or missing measurements. Methods include digital filtering using fast Fourier Transforms, time derivatives, time delays, running averages, and differences between stations. Signals, or time series, manifest three types of behavior: periodic, chaotic, and noise. Periodic behavior is perfectly predictable. Examples of periodic behavior are the diurnal sunlight and temperature patterns caused by the rising and setting sun or tidal water levels attributed to orbital mechanics. Noise refers to random components, usually attributed to measurement error, and is unpredictable. Chaotic behavior neither is totally periodic nor noise, and always has a physical cause. Weather is an example of chaotic behavior. Chaotic behavior is somewhat predictable, especially for small time frames and prediction horizons.

Signal Decomposition

Signal decomposition involves splitting a signal into subsignals, called "components," which are independently attributable to different physical forces. To analyze and model these time series, the periodic and chaotic components of the signals need to be separated. Digital filtering can separate out the chaotic component in the water-level time series. Computation of the tidal range time series from the water-level time series separates out the periodic components of the water-level time series. Digital filtering can also diminish the effect of noise in a signal to improve the amount of useful information that it contains.

Time derivatives are a common analytical method used in the sciences to analyze the dynamics of a system. Time derivatives are also computed from the measured, computed, and filtered variables on the Savannah River to further understand the dynamics of the system. The 1-day derivative of the low-pass filtered water-level time series for a 90-day period was plotted with the original time series and the low-pass filtered data (fig. 19). The 1-day derivatives show the rate of change of the chaotic component of the water-level time series. For the 90-day period, the daily change in filtered water level is as high as 1.5 ft.

Often there are time delays between when an event is measured and the time that the response is observed in a system. Modeling a system is more complicated when two events of interest, a cause and an effect, do not occur simultaneously. The time between cause and effect is called the "time delay" or "delay." Each input variable of a model has its own delay. Determining the correct time delays for pulses and system response is critical to accurately simulating a dynamic system. For the Savannah River, there is a time delay between the measured streamflow at Clyo, Ga., and the response in specific conductance at the stations near the SNWF. Time delays between when the flow enters the system and the river response of the specific conductance were determined for each station by analyzing the correlation between lagged flow values for various time delays and the salinity response values at a station.

Table 2. Period of record and range of daily flow conditions measured for the river and marsh gaging networks used in the study.

[ft^3/s, cubic foot per second; USGS, U.S. Geological Survey; GPA, Georgia Ports Authority]

Gaging network	Period of record used in this study	Range of flow conditions (ft^3/s)
USGS River	Jan 1994–May 2005	4,320 – 52,600
GPA River	July–Sept 1997 July–Oct 1999	5,440 – 11,600
USGS Marsh	June 1999–May 2005	4,320 – 39,600
GPA Marsh	June 1999–Oct 2002	4,320 – 14,100

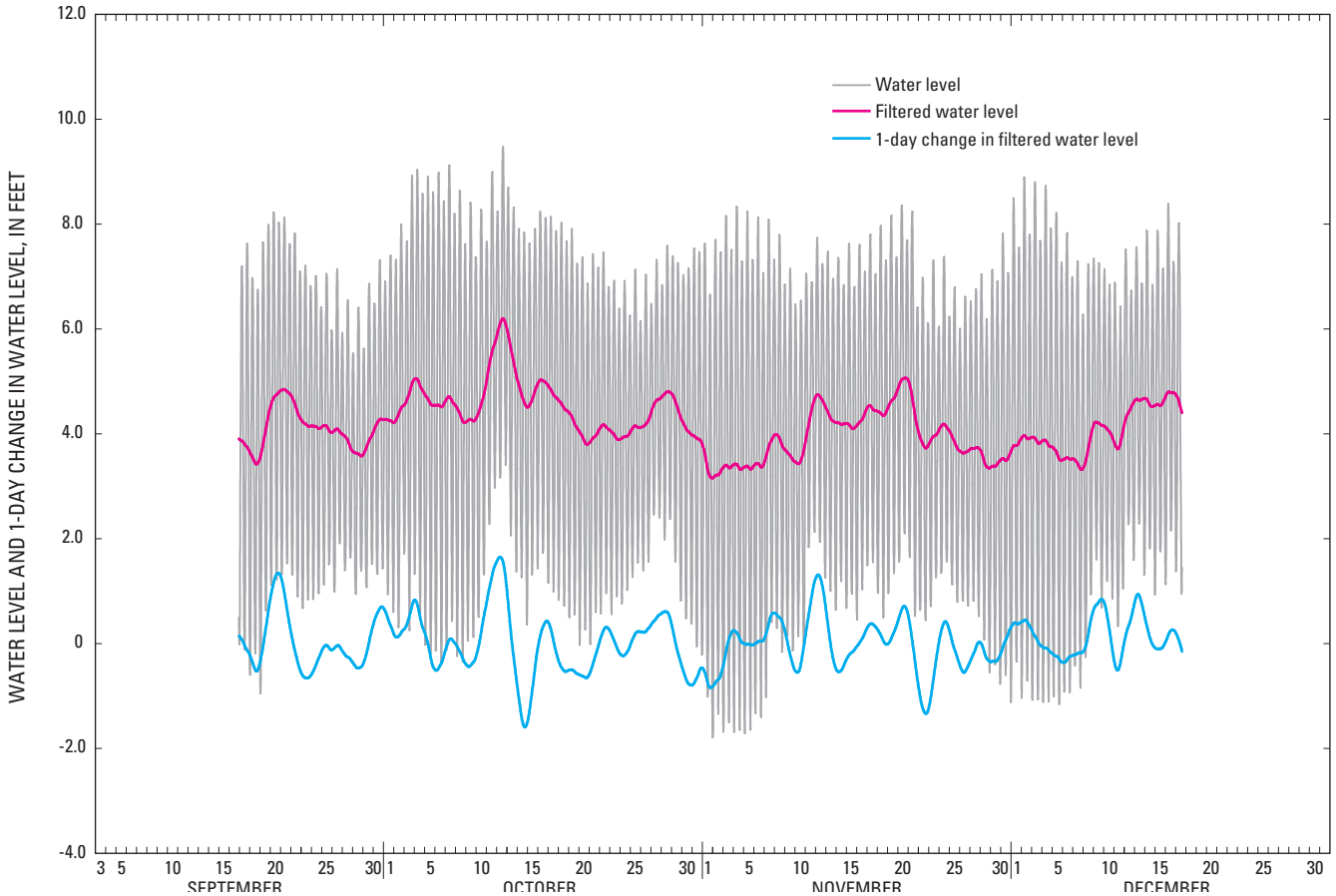

Figure 19. Hourly water levels at Fort Pulaski (station 02198980), low-pass filtered water levels, and 1-day change in filtered water levels for the 90-day period September 13 to December 13, 1994.

Averages and running averages are commonly used to remove the variability of measurements and to represent prevailing behaviors of the system. To capture the effect of the extended low-flow condition as a result of the extended drought, time derivatives of running average flow conditions were computed. The 14-day difference (time derivative) of the 14-day running average flow conditions can capture the change in average flow conditions over a 28-day period. For example, if the 14-day running average for June 28, 2002 (June 15–28) is 4,743 ft^3/s and the 14-day running average for June 15, 2002 (June 2–15) is 4,900 ft^3/s, then the 14-day difference (time derivative) for June 28, 2002 is –157 ft^3/s. The daily streamflow at Clyo, Ga., the 14-day average streamflow, and the 14-day difference in 14-day average flows are shown in figure 20. The time derivative signal shows the 14-day trend of the change in 14-day average streamflow conditions. There is a significant difference in the time derivative signal for the extended drought after the high flows of the El Niño in 1998. Prior to the drought, 14-day differences in 14-day average flows fluctuate by more than 10,000 ft^3/s. During the drought, fluctuations were less than 5,000 ft^3/s.

Correlation Analysis

The relations between the many variables and their various components are ascertained through correlation analyses to provide deeper understanding of system dynamics. For example, salinity intrusion is dependent on streamflow and tides, and correlation analysis provides a measure of relative contribution of each variable. Sensitivity analysis quantifies the relations between a dependent variable of interest and causal variables. Computing sensitivities requires defining the relation between variables through modeling.

Using statistical and(or) ANN software, the computer systematically correlates factors that most influence parameters of interest (for example, specific conductance) to candidate combinations of controlled and uncontrolled variables (for example, streamflow and tidal conditions). Correlation methods based on statistics and ANNs are applied in combination. Promising results identified by the analysis are validated by comparing them to known patterns of behavior.

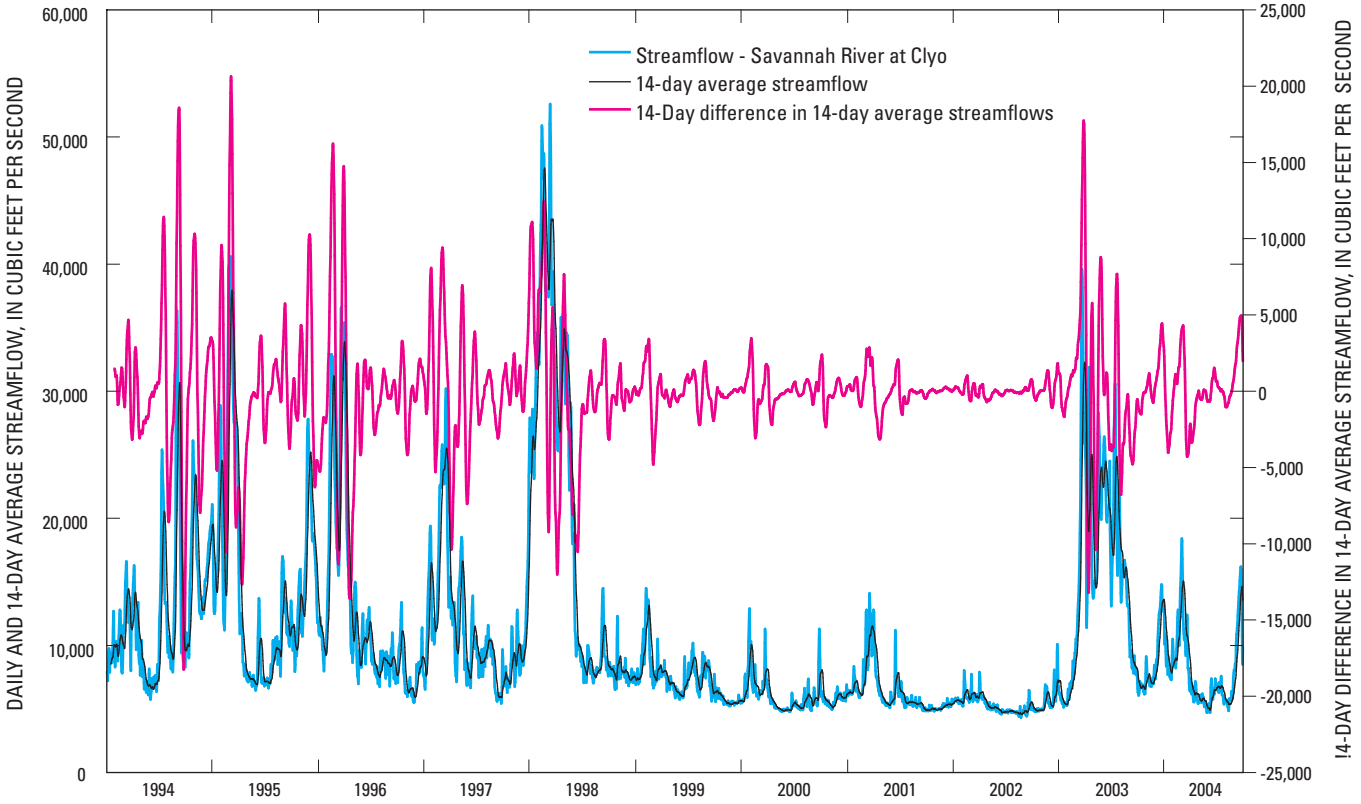

Figure 20. Daily flows, 14-day average flows, and 14-day differences in 14-day average flows for the Savannah River at Clyo, Ga., for the period 1994–2004. Note the scale for the 14-day differences in 14-day average flows is on the right.

State-Space Reconstruction

Chaos theory provides a conceptual framework called "state-space reconstruction" (SSR) for representing dynamic relations. Data collected at a point in time can be organized as a vector of measurements; for example, element one of the vector might be the water level, element two the streamflow, and so on. Engineers say that a process evolves from one state to another, in time, and that a vector of measurements, also referred to as a "state vector," represents the process state at the moment the measurements were taken. A sequence of state vectors represents a "state history." Mathematicians say that the state vector is a point in a "state space" having a number of dimensions equal to the number of elements in the vector. For example, eight vector elements equates to eight dimensions. Empirical modeling is the fitting of a multidimensional surface to the points arrayed in state space.

Chaos theory proposes that a process can be optimally represented (reconstructed) by a collection of state vectors, Y(t), using an optimal number of measurements equal to "local dimension" d_L that are spaced in time by integer multiples of an optimal time delay, τ_d (Abarbanel, 1996)[1]. For a multivariate process of k independent variables:

$$Y(t) = \{[x_1(t), x_1(t-\tau_{d1}),\ldots, x_1(t-(d_{L1}-1)\tau_{d1})], \ldots, [x_k(t), x_k(t-\tau_{dk}),\ldots, x_k(t-(d_{Lk}-1)\tau_{dk})]\} \quad (1)$$

where each $x(t,\tau_{di})$ represents a different dimension in state space, and therefore, a different element in a state vector. Values of d_{Li} and τ_{di} are estimated analytically or experimentally from the data. The mathematical formulations for models are derived from those for state vectors. To predict a dependent variable of interest y(t) from prior measurements (forecasting) of k independent variables (Roehl and others, 2000):

$$y(t) = F\{[x_1(t-\tau_{p1}), x_1(t-\tau_{p1}-\tau_{d1}), \ldots, x_1(t-\tau_{p1}-(d_{M1}-1)\tau_{d1})], \ldots,[x_k(t-\tau_{pk}), x_k(t-\tau_{pk}-\tau_{dk}), \ldots, x_k(t-\tau_{pk}-(d_{Mk}-1)\tau_{dk})]\} \quad (2)$$

where F is an empirical function such as an ANN, each $x(t,\tau_{pi},\tau_{di})$ is a different input to F, and τ_{pi} is yet another time delay. For each variable, τ_{pi} is specified according to one of the following constraints: time delay at which an input variable becomes uncorrelated to all other inputs, but can still provide useful information about y(t); time delay of the most recent available measurement of x_i; or time delay at which an input variable is most highly correlated to y(t). Here, the state space local dimension d_L of Equation 1 is replaced with a model

[1] In Chaos theory, d_L and t_d are called "dynamical invariants," and are analogous to the amplitude, frequency, and phase angle of periodic time series.

input variable dimension d_M, which is determined experimentally. It is noted that $d_M \leq d_L$, and tends to decrease with increasing k.

Input-Output Mapping and Problem Representation

The development of ANN models to predict the water level and pore-water salinity of the tidal marsh was undertaken in two phases. The first phase was to train the ANN models to simulate the water level and specific conductance at the USGS and the GPA riverine sites. Inputs to the ANN models of the USGS river network include time series, or signals, of streamflow, tidal water level, and tidal range. Because of the limited data set, the GPA river network sites were modeled using differences in water level and specific conductance with the USGS river network stations. Outputs from these models are water level and salinity at the river network stations. Simulated specific-conductance values are post-processed to salinity values using the equation documented by Miller and others (1988). The second phase was to train the ANN models to simulate water level and pore-water specific conductance at the USGS and the GPA marsh sites. Inputs for these models include the water-level and specific-conductance signals from the USGS and GPA river networks near the marsh gaging sites. Outputs from these models are water level and salinity at the marsh network stations.

One of the ultimate applications of the ANN models will be to simulate the change in water level and specific conductance in the marsh likely to result from a potential harbor deepening. As discussed previously, a 3D model, EFDC, has been applied and will be used to simulate water-level and salinity changes in the river resulting from geometric changes expected from proposed physical changes to the harbor. Using the USGS river network time series as input for the marsh, the ANN models can accommodate the integration of output from other models of the river. Rather than use the ANN river predictions as input to the ANN marsh models, EFDC predictions were used. The final application of the ANN models can be run under two different modes. Mode 1 uses the streamflow and tidal condition inputs and evaluates the effects of changing hydrologic conditions on water level and specific conductance of the river and marsh sites. Mode 2 uses EFDC predictions of changes from a base-case scenario at the USGS river network as inputs to evaluate the effects of changes in the channel geometry on water level and specific conductance of the marsh sites.

Decorrelation of Variables

Often, explanatory variables share information about the behavior of a response variable. It is difficult, if not impossible, to understand the individual effects of these variables (sometimes known 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 "de-correlating" the confounded variables. For the Savannah River application, the boundary input data, streamflow from Clyo, Ga., and water levels from Fort Pulaski, are both distant from the river and marsh gaging network near the SNWF. The Pearson correlation coefficient, R, between Clyo streamflows and the Savannah Harbor water level is -0.03, indicating very little correlation between the two time series.

The only variables in the application that needed to be decorrelated were data from the GPA specific-conductance stations (GPA10S, GPA11, GPA11R, and GPA12) used as inputs for the USGS marsh salinity models. Decorrelation is accomplished by generating an empirical correlation function and computing its residual error by subtracting the function's predicted values from the actual measurement. The residual error is the "unshared" information between the two signals. Single Input Single Ouput (SISO) ANN models were built to decorrelate data from GPA11, GPA11R, and GPA12 from GPA10S. The residual error signal was then used in the USGS marsh salinity models as the decorrelated inputs for those stations.

Artificial Neural Network Models

Models generally fall into one of two categories, deterministic (or mechanistic) or empirical. Deterministic models are created from first-principles equations, whereas empirical modeling adapts generalized mathematical functions to fit a line or surface through data from two or more variables. The most common empirical approach is OLS, which relates variables using straight lines, planes, or hyper-planes, whether the actual relations are linear or not. Calibrating either type of model attempts to optimally synthesize a line or surface through the observed data. Calibrating models is made difficult when data have substantial measurement error or are incomplete, and the variables for which data are available may only be able to provide a partial explanation of the causes of variability. The principal advantages that empirical models have over deterministic models are that they can be developed much faster and are often more accurate when 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 data sets. The architecture of ANN models is loosely based on the biological nervous system (Hinton, 1992). Although there are numerous types of ANNs, the most commonly used type is the multilayer perceptron (MLP) (Rosenblatt, 1958). As shown in figure 21, MLP ANN's are constructed from layers of interconnected processing elements called neurons, each executing a simple "transfer function."

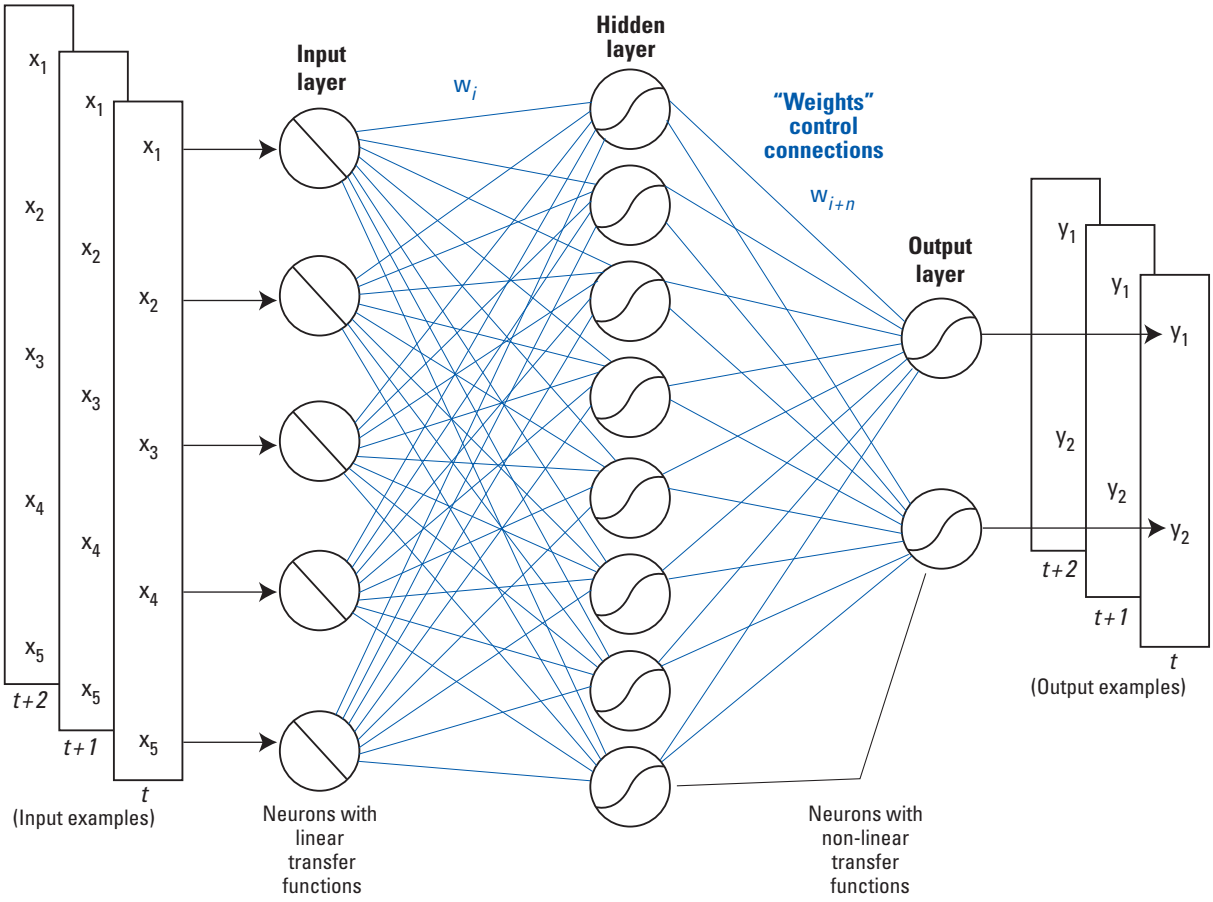

Figure 21. Multilayer perceptron artificial neural network architecture.

All input layer neurons are connected to every hidden layer neuron, and every hidden layer neuron is connected to every output neuron. There can be multiple hidden layers, but a single layer is sufficient for most problems.

Typically, linear transfer functions are used to simply scale input values to fall within the range that corresponds to the most linear part of the s-shaped sigmoid transfer functions used in the hidden and output layers. Each connection has a "weight," w_i, associated with it that 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 their weights, and 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 data set comprising "input/output vector pairs." Simulation 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 MLP ANNs is the back error propagation (BEP) training algorithm (Rumelhart and others, 1986). Jensen (1994) describes the details of the MLP ANN, the type of ANN used in this study. MLP ANNs can synthesize functions to fit high-dimension, nonlinear multivariate data. Devine and others (2003) and Conrads and Roehl (2005) describe their use of MLP ANN in multiple applications to model and control combined man-made and natural systems including disinfection byproduct formation, industrial air emissions monitoring, and surface-water systems affected by point and nonpoint source contaminants.

Experimentation with a number of ANN architectural and training parameters is a normal part of the modeling process. For correlation analysis or predictive modeling applications, a number of candidate ANNs are trained and evaluated for both their statistical accuracy and their representation of process physics. Interactions between combinations of variables also are considered. Finally, a satisfactory model can be exported for end-user deployment. In general, a high-quality predictive model can be obtained when:

- the data are well distributed throughout the state space (historical range of conditions) of interest,
- the input variables selected by the modeler share "mutual information" about the output variables, and
- the form "prescribed" or "synthesized" for the model used to "map" (correlate) input variables to output variables is a good one. Techniques such as OLS and physics-based finite-difference models prescribe the functional form of the model's fit of the calibration data. Machine learning techniques like ANN's synthesize a best fit to the data.

Subdividing a complex modeling problem into subproblems and then addressing each is a 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 Savannah study, individual ANN models (submodels) were developed for the river and marsh water level and salinity at each continuous station. 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 the Model-to-Marsh (M2M) DSS described later in this report. The ANN models and plots described herein were developed using the iQuest™ data-mining software[2] (Version 2.03C DM Rev31). The ANN models were deployed in the DSS using the Visual Basic runtime library of the iQuest R/T™ software.

Statistical Measures of Prediction Accuracy

The R^2, the mean error (ME), root mean square error (RMSE), and percent model error (PME) have been computed for the training and testing data sets for each model and are listed in Appendix I. Model accuracy usually is reported in terms of R^2 and commonly is interpreted as the "goodness of the fit" of a model. A second interpretation is one of answering the question, "How much information does one variable or a group of variables have about the behavior of another variable?" 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 developers believe that the river and marsh water-level and salinity models are unusually accurate relative to one-dimensional, two-dimensional, and three-dimensional finite-difference models developed for comparably complex estuaries and tidal marsh systems.

The ME and RMSE statistics provide a measure of the prediction accuracy of the ANN models. The ME is a measure of the bias of model predictions—whether the model over or under predicts the measured data. The ME is presented as the adjustment to the simulated values to equal the measured values. Therefore, a negative ME indicates an over simulation by the model and a positive ME indicates an under prediction by the ANN model. Mean errors near zero may be misleading because negative and positive discrepancies in the simulations can cancel each other. RMSE addresses the limitations of ME by computing the magnitude, rather than the direction (sign) of the discrepancies. The units of the ME and RMSE statistic are the same as the simulated variable of the model.

The minimum and maximum values of the measured output are listed in Appendix I. The accuracy of the models, as given by RMSE, should be evaluated with respect to the range of the output variable. 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. 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. The PME was computed by dividing the RMSE by the range of the measured data. The models of the USGS river and marsh networks have the greatest range of the output variables.

Generally, the USGS river water-level and specific-conductance models have R^2 values in the 0.95 to 0.99 and 0.57 to 0.90 ranges, respectively, and PME values of 1.5 to 4.1 percent and 1.0 to 6.3 percent, respectively. (The statistics for the river models are based on the test data sets of the hourly models.) The USGS marsh water-level and specific-conductance models' range of R^2 values are 0.72 to 0.87 for water levels, 0.74 to 0.93 for specific conductance, and 3.4 to 5.4 percent and 4.0 to 10.6 percent, respectively, for PME. (Statistics for the marsh water level are based on the test data set of the hourly models, and statistics for the marsh specific-conductance predictions are based on all measured data.)

Development of Artificial Neural Network Models

The following sections describe how the water-level and salinity models were developed for the river and marsh sites. All the stations from the USGS and the GPA river and marsh networks were modeled. The river stations used to model the marshes were generally limited to the USGS river stations because their long period of record covers the largest range of hydrologic conditions. The model developments for each type of station—river water level, river salinity, marsh water level, and marsh pore-water salinity—follow a similar approach. One example is given for each type of model and is followed by a general description of the performance of all the stations of that particular type. Figures are shown for each type of model. A graph of measured and simulated data for a short period (1 to 2 months) is shown for the example station to show model performance. The graph of the example station is followed by graphs showing the performance of the daily river water-level and specific-conductance models for the period of record to show how well the models simulate the long-term trends of the system. The graphs of measured and simulated daily data are followed by graphs showing the performance of the hourly models for a short period (2 to 3 months). For the marsh models, the graph of the measured and simulated values for the example station is followed by graphs showing the performance of the hourly models. Model summaries for the example models (the input and output variables, size of training and testing data sets, and their respective R^2s) are given in table 3, and model summaries of all the models can be found in Appendix II. Summary statistics for all the models can be found in Appendix I.

[2] The iQuest™ software is exclusively distributed by Advanced Data Mining, LLC, 3620 Pelham Road, PMB 351, Greenville, SC 29615-5044 Phone: 864–201–8679, email: info@advdatamining.com: *http://www.advdatamining.com*.

The river models were developed in two stages. The first stage simulated the low-pass filtered daily water-level or specific-conductance signals to capture the long-term dynamics of the system. The second stage simulated the higher-frequency hourly water level or specific conductance, using the simulated water level and specific conductance as a carrier signal. Each river water-level model uses three general types of input signals, or time series: streamflow signal(s), water-level signal(s), and tidal range signal(s). The signals may be of the measured series values, filtered values, and/or a time derivative of the signals. The specific-conductance (salinity) models use an additional streamflow input: a time derivative of a moving window average, such as the 14-day difference in the 14-day average streamflows (fig. 19). The available data set for developing the models was randomly bifurcated into training and testing data sets. For the large data sets, such as the USGS riverine network stations dating back to 1994, a zone averaging, or box, filter of the data, was used to separate the data into training and testing data sets. Using the zone average filter, all the data are used in the test data set and a small selected sample of the data is used for the training data set. The filter separates the data set into user-specified number of zones or boxes and determines the input vectors with the highest information content and reserves these vectors for the training data set. The percentage of training and testing data depended on the length of the data set and the range of hydrologic conditions in the data set. Typically, the zone averaging filter uses a small percentage of the data (less than 10 percent) for the training data set. Model development summaries and variable descriptions can be found in Appendixes II and III.

Riverine Water-Level Model at Little Back River near Limehouse

The daily water-level model (wl8979a-2005-1) for Little Back River near Limehouse (station 02198979) uses streamflow, water-level, and tidal range inputs (table 3). The streamflow inputs are daily average streamflow (Q8500A) and the 1-day derivative of streamflow (DQ8500A). The water-level data inputs are daily water level from Fort Pulaski (FWL8980A) and the 1-day time derivative of daily water level lagged 1 day (DWLAD1). Gaps in the time series for Savannah River at Fort Pulaski (station 02198980) were filled by correlating to upstream water levels such as Front River at Broad Street (station 02198920). The tidal range inputs are daily tidal range (XWL8980A) and a 1-day time derivative of daily tidal range lagged 1 day (DXWLAD1). For testing and training the daily model, there were 78,980 data values available. Ten percent of the data was used for training and 90 percent was used for testing. The coefficient of determination, R^2, for the training and testing were 0.95 and 0.96, respectively (table 3, Appendices I and II). The daily model used two hidden-layer neurons.

The hourly water-level model (wl8979h-2005-1) uses the simulated daily water level from the daily model and Fort Pulaski water-level inputs. The simulated daily water-level input (PWL8979A) captures the long-term movement of the water level that is characterized by the streamflow and tidal range data in the daily model. The six water-level inputs; LG1NWL, LG1D3NWL, LG4D3NWL, LG7D3NWL, LG10D3NWL, LG13D3NWL, are the 1-hour lagged water level (LG1NWL), and 3-hour time derivatives of the water level lagged 1, 4, 7, 10, and 13 hours, respectively. The lagged water-level inputs capture the periodic semidiurnal tidal signal. For testing and training the hourly model, there were 79,216 data values available (approximately 11 years of hourly data from 1994 to 2005). The data set was bifurcated into training and testing data sets. Ten percent of the data was used for training and 90 percent was used for testing. The R^2 for the training and testing data were 0.98 and 0.98, respectively (table 3, Appendices I and II). The hourly model used two hidden-layer neurons.

The measured and simulated hourly water levels are shown in figure 22 for the first quarter of 2002. The model simulates the measured data well, but does not capture the full range of the tide and slightly under simulates the maximum and minimum water levels. The PME for the period of record is less than 3 percent (Appendix I)).

The daily water levels represent the chaotic portion of the water-level signal and the long-term trend of the system. The models are able to simulate this portion of the signal quite well (R^2 from 0.88 to 0.96) but miss the extreme high and low water levels (figs. 23 and 24). The simulations for the Front River at Houlihan Bridge (station 02198920, fig. 23*B*), under simulates the daily water levels more in the first half of the record than the second half. This may be the result of either a systematic change in the data collection or a change in the water-level behavior at the station. The under simulation for the first half of the time series explained the lower R^2 for the daily water-level model as compared to the other sites. The hourly water levels represent the periodic portion of the water-level signal and, generally, the models are able to simulate this portion of the signal quite well (R^2 from 0.98 to 0.995, figure 24).

Riverine Specific-Conductance Model at Little Back River at USFW Dock

The daily salinity model (sc89791a-2005-2) for Little Back River at USFW Dock uses six streamflow, three water-level, and three tidal range inputs. The streamflow inputs are daily average streamflow (Q8500A), the 1-day derivative of streamflow (D8500A), daily streamflow lagged 2 days (LAQ2), the 2-day change in streamflow (DAQ2), the 16-day change in streamflow (DAQ16), and the 30-day change in streamflow (DAQ30). The water-level data inputs are filtered daily water level from Fort Pulaski (FWL8980A), the 1-day time derivative of filtered daily water level (DWLA), and the 1-day time derivative of filtered daily water level lagged 2 days (LG2DWL). The tidal range daily inputs are daily tidal range (XWL8980A), the 1-day time derivative of daily tidal range (DXWLA), and the 1-day time derivative of daily tidal range lagged 2 days (LG2DXWLA). For testing and training

Table 3. Model name and model summary for four types of models used in the study.

[WL, water level; SC, specific conductance; XWL, tidal range; MWA, moving window average]

Model name	Model type	Input variables	Variable description	Training/testing data points	Hidden layer neurons
wl8979a-2005-1	Daily	Q8500A	daily average flow	7,987/70,993	2
		DQ8500A	1-day change in daily average flow		
		FWL8980A	filled, filtered daily WL		
		XWL8980A	daily tidal range		
		DXWLAD1	1-day change in tidal range, lagged 1 day		
		DWLAD1	1-day change in WL, lagged 1 day		
wl8979h-2005-1	Hourly	LG1NWL	1-hour lag in the in WL	8,029/71,187	2
		LG1D3NWL	1-hour lag in the 3-hour change in WL		
		LG4D3NWL	4-hour lag in the 3-hour change in WL		
		LG7D3NWL	7-hour lag in the 3-hour change in WL		
		LG10D3NWL	10-hour lag in the 3-hour change in WL		
		LG13D3NWL	13-hour lag in the 3-hour change in WL		
		PWL8979A	Predicted daily WL at 02198979		
sc89791a-2005-1	Daily	Q8500A	daily average flow	9,660/75,782	3
		DQ8500A	1-day change in daily average flow		
		LAQ2	2-day lag of the daily flow		
		DAQ2	2-day change in daily flow		
		DAQ16	16-day change in daily flow		
		DAQ30	30-day change in daily flow		
		FWL8980A	filled, filtered daily WL		
		XWL8980A	daily tidal range		
		DWLA	1-day change in WL		
		DXWLA	1-day change in tidal range		
		LG2DWLA	1-day change in WL, lagged 2 days		
		LG2DXWLA	1-day change in XWL, lagged 2 days		
		LG2DXWLA	1-day change in tidal range, lagged 2 days		
sc89791h-2005-1	Hourly	FWL8980A	filled, filtered daily WL	9,736/76,366	3
		XWL8980A	daily tidal range		
		NXWL	hourly WL		
		LG1NWL	1-hour lag in the in hourly WL		
		LG1D3NWL	1-hour lag in the 3-hour change in WL		
		LG4D3NWL	4-hour lag in the 3-hour change in WL		
		LG7D3NWL	7-hour lag in the 3-hour change in WL		
		LG10D3NWL	10-hour lag in the 3-hour change in WL		
		LG13D3NWL	13-hour lag in the 3-hour change in WL		
		PSC89791A	Predicted daily SC at 021989791		

Table 3. Model name and model summary for four types of models used in the study.—Continued

[WL, water level; SC, specific conductance; XWL, tidal range; MWA, moving window average]

Model name	Model type	Input variables	Variable description	Training/testing data points	Hidden layer neurons
pb2mwl-2005	Hourly	FWL8840	WL at station 02198840	3,284/18,228	2
		DFWL8840	difference with WL at 02198840		
		LG3DFWL8840	difference with WL at 02198840 lagged 3 days		
		LG6DFWL8840	difference with WL at 02198840 lagged 6 days		
		FWLDIF8977	WL at station 02198977		
		FWLDIF8979	WL at station 02198979		
		FWLDIF8920	WL at station 02198920		
pb2msc-2005-2	Hourly	SCDIF8840A	difference with SC between 02198840 and 021989791	2,142/18,770	1
		SCDIF8920A	difference with SC with difference with SC between 021988920 and 021989791		
		FSC89791A4WK	4-week MWAs of SC at 021989791		
		LG672FSC89791A4WKD4WK	difference between 4-week and lagged 4-week MWAs of SC at 021989791		
		FSC89791A2WKD4WK	difference between 2- and 4-week and lagged MWAs of SC at 021989791		
		FSC89791A1WKD2WK	difference between 1- and 2-week and lagged MWAs of SC at 021989791		
		FSC89791A48D1WK	difference between 2-day and 1-week and lagged MWAs of SC at 021989791		
		FSC89791DA48	difference between hourly and 2-day MWA of SC at 021989791		
		DFSC89791DA48	3-hour time derivative of SC at 021989791		
		LG3DFSC89791DA48	time derivative lagged 3 hours of SC at 021989791		
prb2msc	Hourly	RSC10S_12RS_A1WK	residual error of predicted weekly average SCGPA12RS (for decorrelation)	4,176/16,736	1
		RSC10S_11RB_A1WK	residual error of predicted weekly average SCGPA11RB (for decorrelation)		
		RSC10S_11B_A1WK	residual error of predicted weekly average SCGPA11B (for decorrelation)		
		PSCGPA10S_FLR_A1WK	1-week MWA of predicted hourly SC at GPA10S		

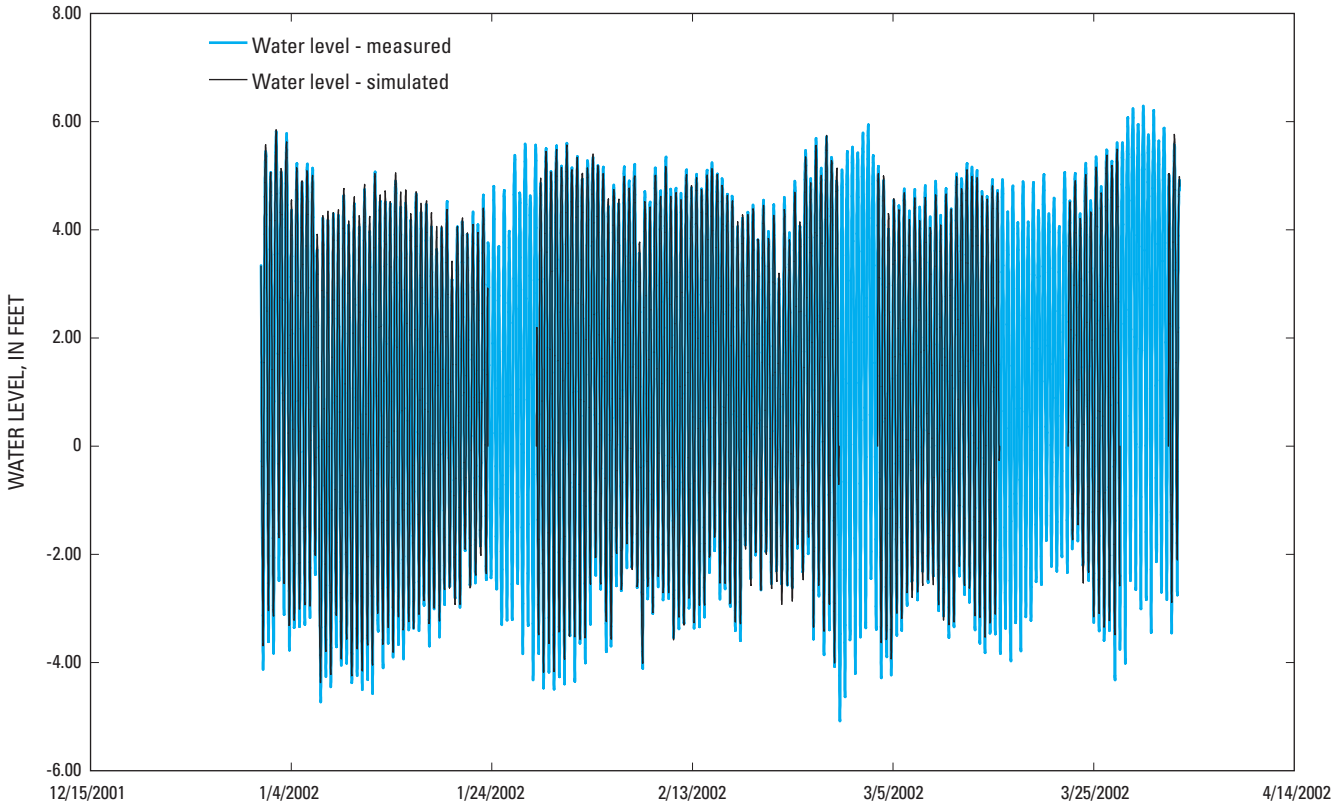

Figure 22. Measured and simulated hourly water levels at Little Back River near Limestone (station 02198979) for the period January 1 to March 31, 2002. Breaks in the prediction time series are caused by incomplete time series for one or more of the model inputs.

the daily model, there were 85,442 data values available. Eleven percent of the data was used for training, and 89 percent was used for testing. The R^2 for the training and testing were 0.89 and 0.87, respectively. The daily model used three hidden-layer neurons.

The hourly salinity model (sc89791h) uses simulated daily specific-conductance values from the daily model, two tidal range inputs, and six water-level inputs. The simulated daily specific-conductance input (PSC8979A) captures the long-term movement of the specific conductance that is characterized by the streamflow and tidal range data in the daily model. The two tidal range inputs are the daily tidal range (XWL8980A) and hourly tidal range (NXWL). The six water-level inputs—LG1NWL, LG1D3NWL, LG4D3NWL, LG7D3NWL, LG10D3NWL, LG13D3NWL—are 1-hour lagged water level, and 3-hour time derivatives of the water level lagged 1, 4, 7, 10, and 13 hours, respectively. The lagged water-level inputs capture the periodic semidiurnal tidal cycle signal. For testing and training the hourly model, there were 86,102 data values. Eleven percent of the data was used for training, and 89 percent was used for testing. The R^2 for the training and testing were 0.89 and 0.83, respectively. The hourly model used three hidden-layer neurons. A summary of the input variables, size of the training and testing data sets, number of hidden-layer neurons, and coefficient of determination for the hourly model can be found in table 3 and Appendix I.

The measured and simulated hourly salinities are shown in figure 25 for the summer of 2002. This period was the end of the 5-year drought that began in 1998 and had the highest salinity intrusions of the drought. The model simulates the measured data and captures the salinity intrusions occurring on a 14- and 28-day cycle. The model also is able to simulate the full range of the large salinity intrusion on about August 10, 2002.

The measured and simulated daily and hourly specific conductances for 1994 to 2005 for the four USGS river water-level gages are shown in figures 26 and 27. As with the water-level models, the daily specific conductance represents the chaotic portion of the signal, and the models are able to simulate this portion of the signal well (R^2 from 0.85 to 0.87, fig. 26) but do not capture the variability as well as the daily water-level models. The quality of hourly models, in terms of the coefficient of determination, vary from a R^2 of 0.57 at Savannah River at I-95 (station 02198840, fig. 27) to 0.87 at Front River at Houlihan Bridge (station 02198920, fig. 27). Although the specific-conductance model for I-95 explains only 57 percent of the variability, the results are quite satisfactory, in terms of timing and magnitude of response, for a station on the leading edge of the saltwater-freshwater interface where the salinity response is less than 0.5 psu.

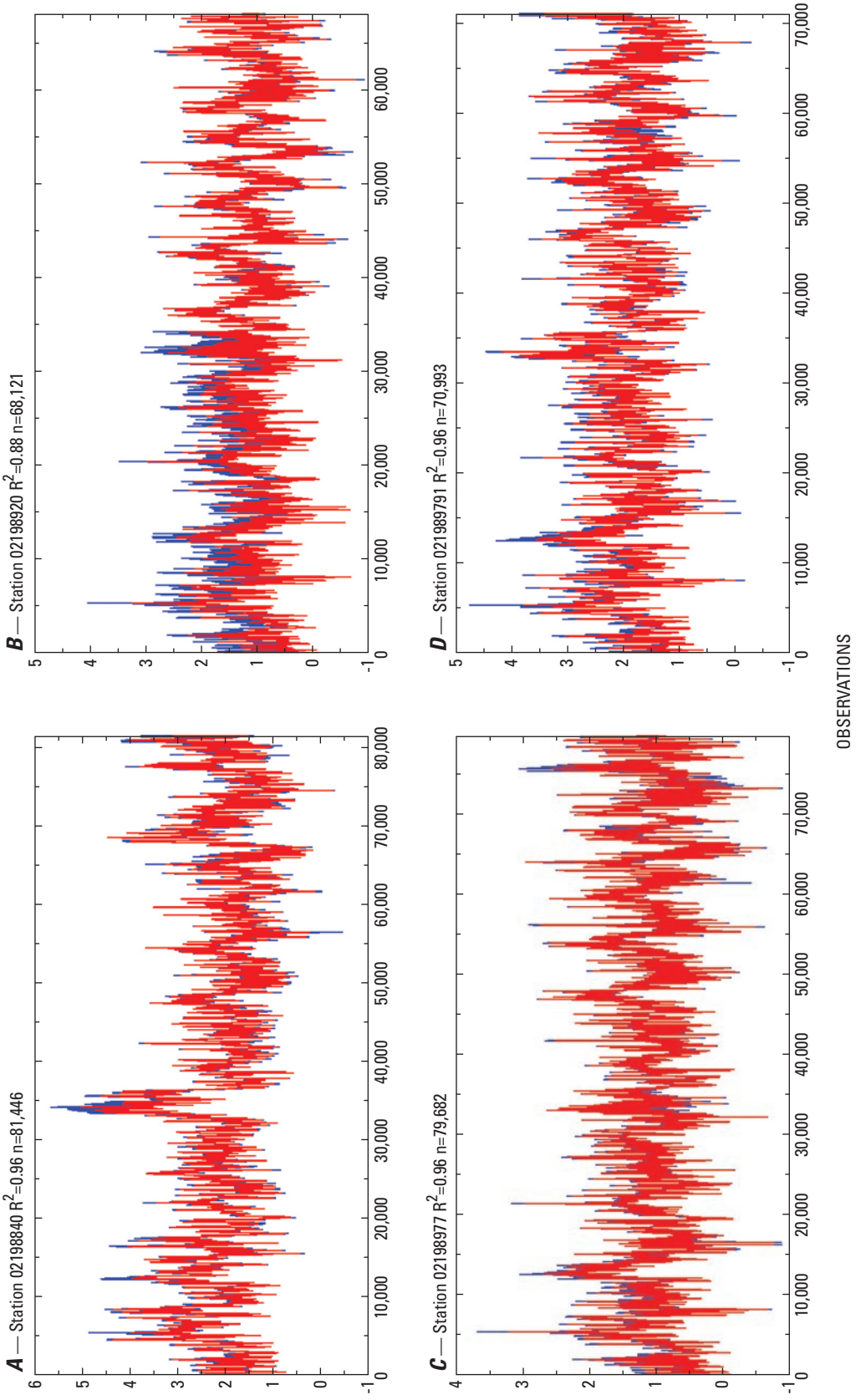

Figure 23. Measured (blue trace) and simulated (red trace) daily water levels for four USGS river gages for the period 1994 to 2004. Data to train the ANN models were bifurcated into training and testing sets. Graphs show data from the testing data set along with the coefficient of determination (R^2) and number (n) of data points.

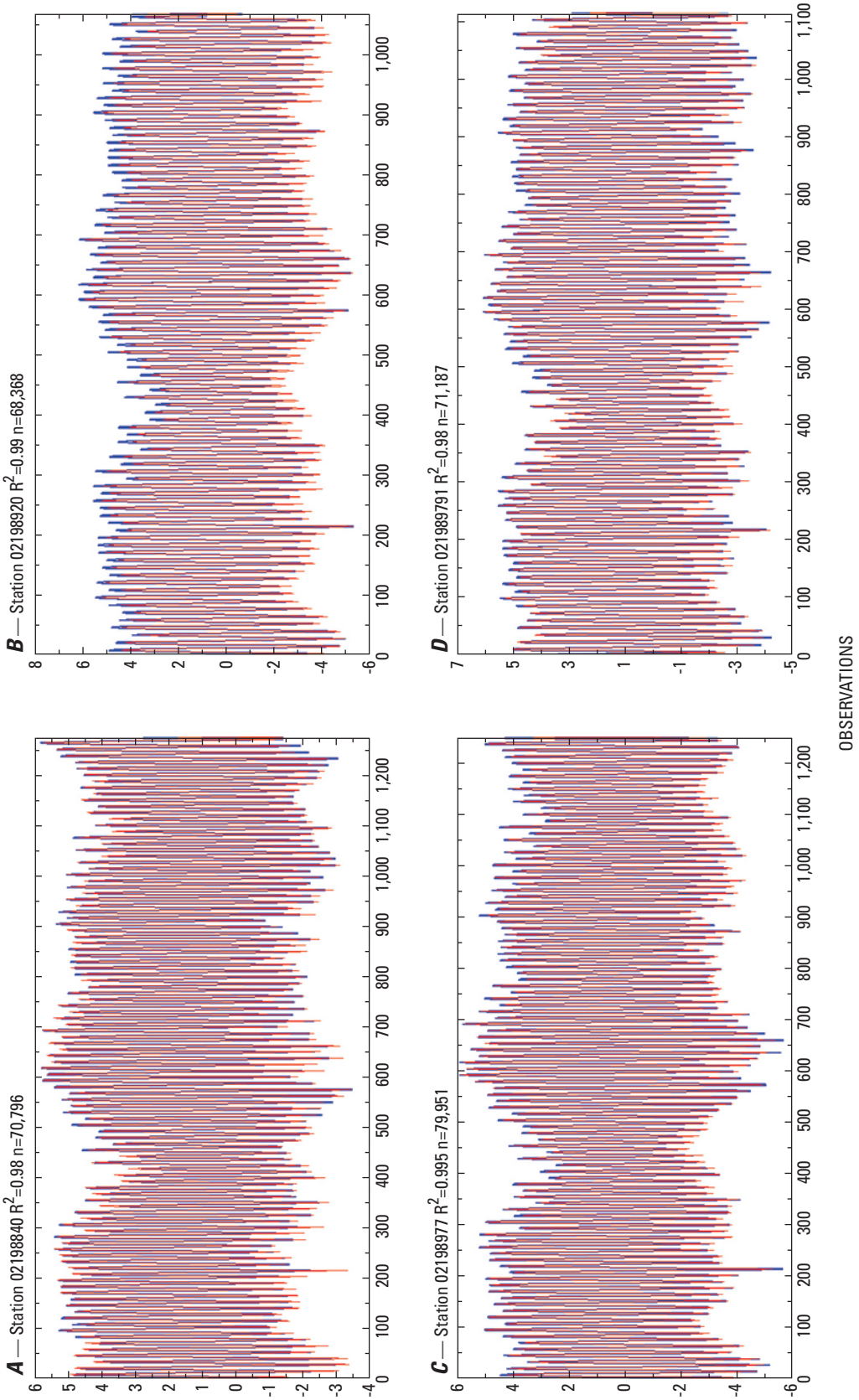

Figure 24. Measured (blue trace) and simulated (red trace) hourly water levels for four USGS river gages. Models were trained and tested on data from the period 1994 to 2004. Data to train the ANN models were bifurcated into training and testing sets. For clarity, only 1,100 observations (approximately 50 days) are shown from the testing data set along with the coefficient of determination (R^2) and number (n) of data points for the testing data set.

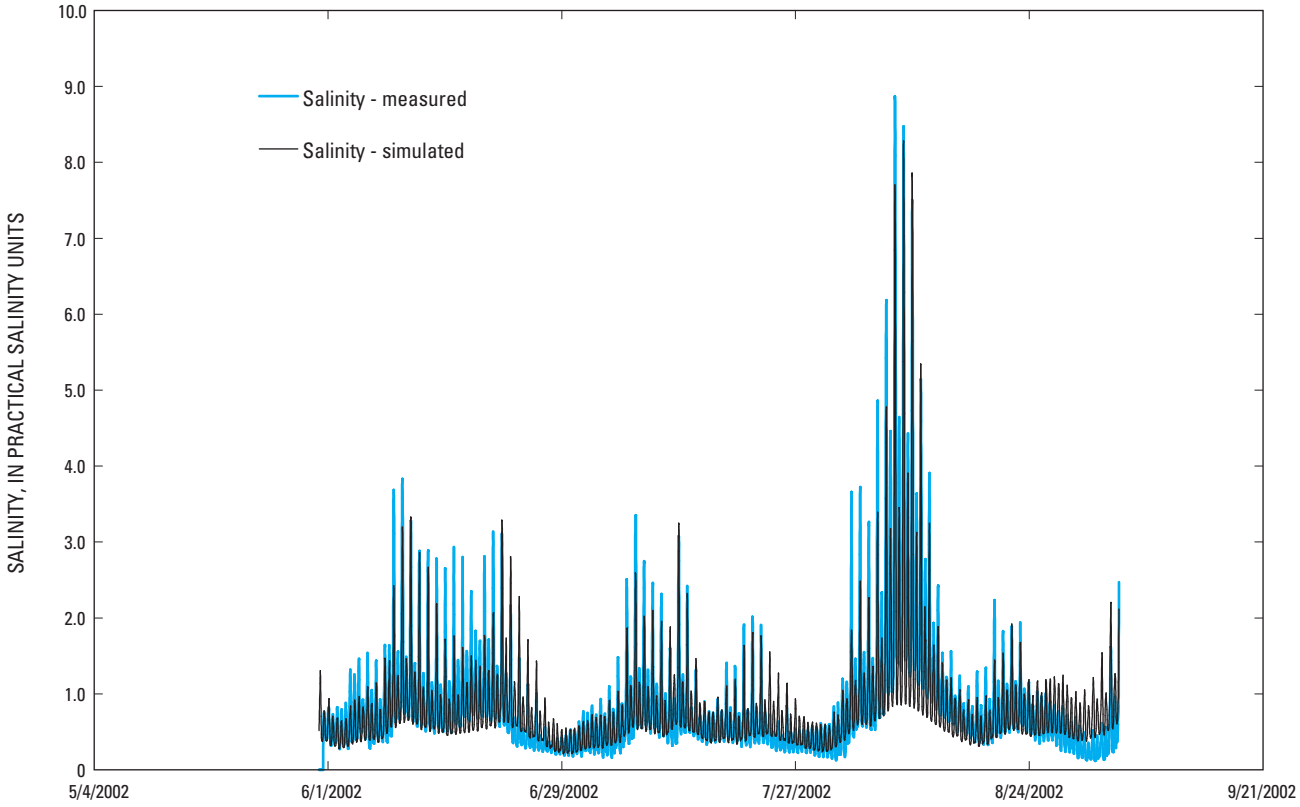

Figure 25. Measured and simulated hourly salinity at Little Back River at U.S. Fish and Wildlife Service Dock (station 02198979) for the period June 1 to August 31, 2002.

Marsh Water-Level Model at Little Back River at Site B2

The hourly water-level model (pb2mwl-2005) for Little Back River at Site B2 uses seven water-level inputs and differences in water levels from the four USGS riverine gages in the vicinity of SNWF. The water-level inputs include the water level at I-95 (station 02198840), 1-day change in water level, 1-day change in water level lagged 3 days, and 1-day change in water level lagged 6 days (variables FWL8840, DFWL8840, LG3DFWL8840, and LG6DFWL8840, respectively). Differences in water level at Site B2 and the USGS riverine gages are also used for inputs. These inputs include the difference in water level with stations 02198977, 02198979, and 02198920 (variables FWLDIF8977, FWLDIF8979, and FWLDIF8920, respectively). For testing and training the daily model, there were 21,512 data values available (approximately 2.5 years of hourly data from 1999 to 2005). Fifteen percent of the data was used for training and 85 percent was used for testing. The R^2 for the training and testing were 0.80 and 0.77, respectively. The daily model used two hidden-layer neurons.

The measured and simulated hourly water levels are shown in figure 28 for the first quarter of 2002. The model simulates the measured data well but does not capture the full range of the tide and slightly under simulates the maximum and minimum water levels.

The measured and simulated hourly water levels for the seven USGS marsh stations are shown in figures 29 and 30. The quality of marsh water-level models, in terms of the coefficient of determination, varies from a R^2 of 0.72 to 0.87. Generally, the models for the stations closer to the harbor (Sites B3, B4, and F1) explain more of the water-level variability (R^2 vary from 0.84 to 0.87) than the models for the stations farther from the harbor (R^2 vary from 0.72 to 0.79).

Marsh Specific-Conductance Model at Site B2

A potential mitigation scenario to ameliorate an increase in salinity in the vicinity of the SNWR would be to divert additional streamflow to the Middle and Little Back Rivers. One concern with using only the long-term USGS riverine gages to simulate the pore-water salinity in the marsh is that there would not be any data on the Middle Back River that may be important for evaluating mitigation alternatives. Data in this area are available from the GPA river networks. Although the data from these stations are limited to conditions from either the summer of 1997 or 1999, data from GPA10, GPA11, GPA11R, and GPA12 were used as inputs to

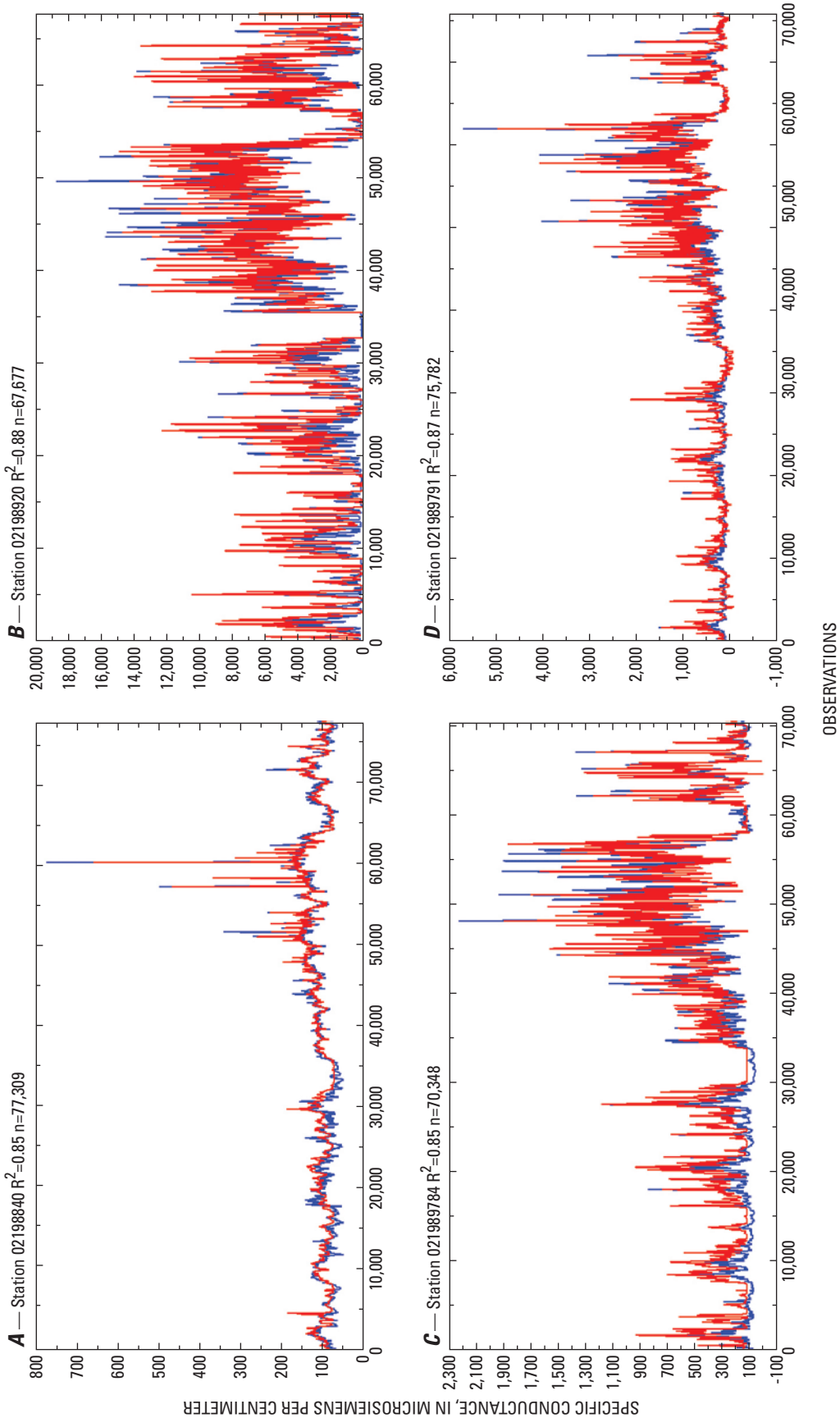

Figure 26. Measured and simulated daily specific conductance for four USGS river gages for the period 1994 to 2004. Data to train the ANN models were bifurcated into training and testing sets. Graphs show data from the testing data set along with the coefficient of determination (R^2) and number (n) of data points.

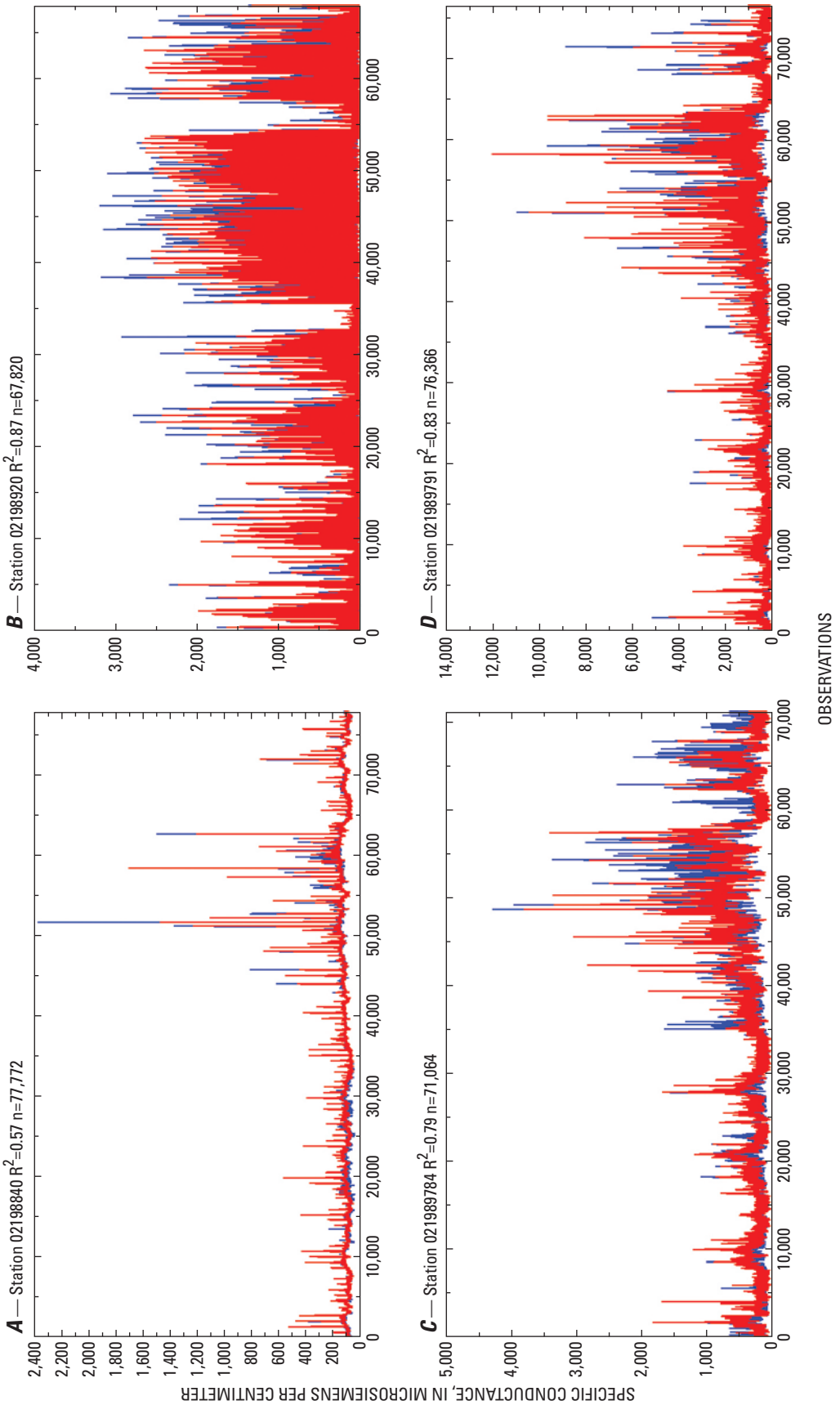

Figure 27. Measured (blue trace) and simulated (red trace) hourly specific conductance for four USGS river gages for the period 1994 to 2004. Data to train the ANN models were bifurcated into training and testing sets. Graphs show data from the testing data set along with the coefficient of determination (R^2) and number (n) of data points.

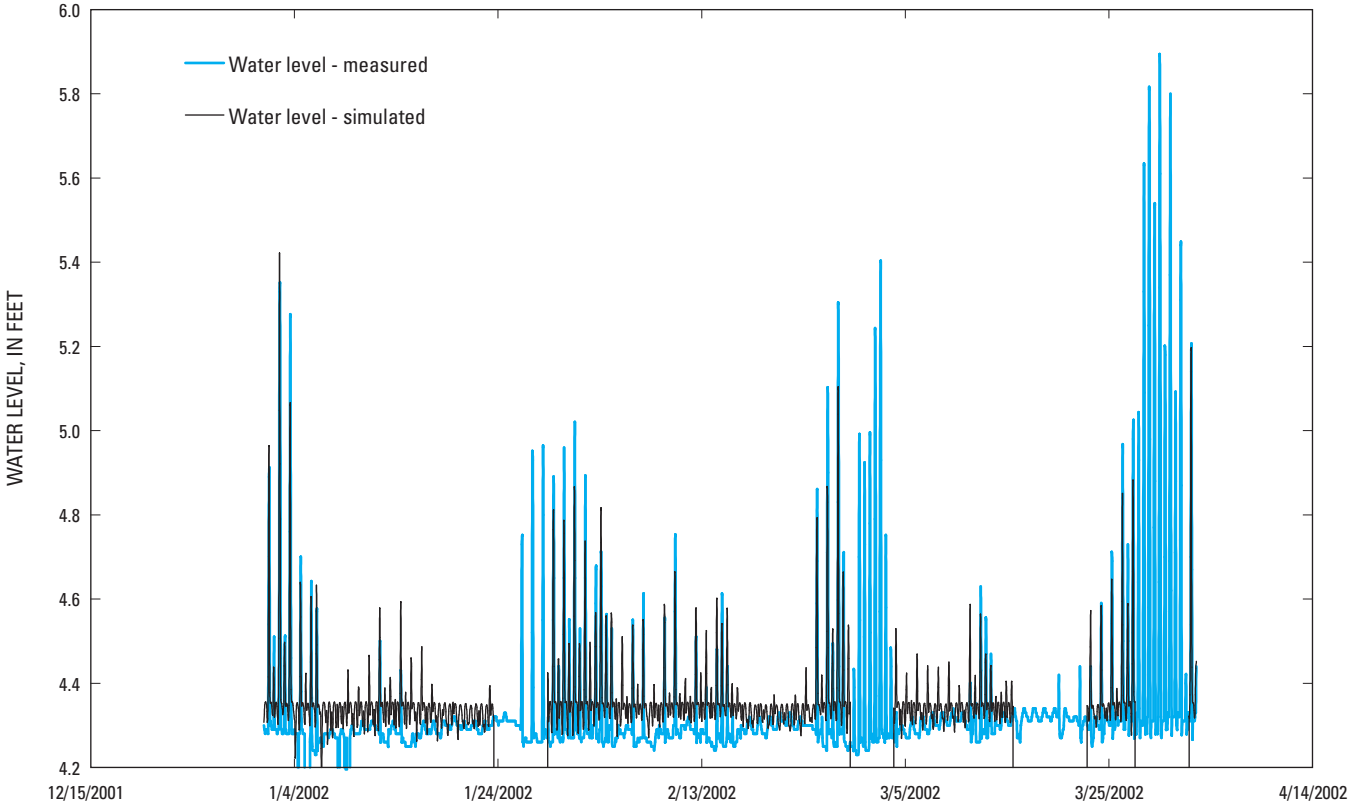

Figure 28. Measured and simulated hourly water levels at Site B2 on the Little Back River for the period January 1 to March 31, 2002. Breaks in the prediction time series are caused by incomplete time series for one or more of the model inputs.

the marsh salinity models to address spatial distribution of the river salinity inputs for the marsh salinity models.

There are two technical issues with using the short-term GPA data that need to be addressed. The first technical issue is that the GPA data are not concurrent with the USGS marsh data (summers of 1997 and 1999 compared to 2000–2005). The models for the four GPA stations were used to simulate the time series of specific conductance for the period 2000–2005. To generate the data for this period, the GPA models were used to make salinity predictions for a range of streamflow conditions much larger than measured during the summers of 1997 and 1999 (fig. 18*A*). The second issue is that the GPA data are highly correlated. To ensure that the data from the GPA site inputs unique information for the marsh predictions, the data for stations GPA11, GPA11R, and GPA12 were systematically decorrelated from GPA10.

The marsh pore-water salinities were modeled in two stages. The first stage used the USGS river data to simulate the marsh salinity. The second stage used the GPA river data to simulate the residual error (difference between the simulated and measured marsh salinity) from the first stage model. The final marsh salinity predictions are the sum of the predictions from the two models.

The first stage pore-water salinity model (pb2msc-2005-2) for Little Back River at Site B2 uses 10 specific-conductance inputs and differences in specific conductance from two USGS riverine gages in the vicinity of SNWF. Two inputs are the difference in specific conductance at the Little Back River at USFW Dock (021989791) and the Savannah River at I-95 (station 02198840) and Front River at Houlihan Bridge (station 02198920), variables SCDIF8840A and SCDIF8920A, respectively. There are four moving window average inputs of 48 hours, 1-, 2-, and 4-weeks (variables FSC89791A48, FSC89791D1WK, FSC89791D2WK, FSC89791D4WK) and four time derivative inputs (variables LG672FSC89791A4WKD4WK, FSC89791DA48, DFSC89791DA48, and LG3DFSC89791DA48). For testing and training the daily model, there were 20,912 data values available (approximately 5 years of hourly data from 1999 to 2005). Ten percent of the data was used for training, and 90 percent was used for testing. The R^2 for the training and testing were 0.83. The model used one hidden layer neuron.

The second stage model (prb2msc) simulates the residual error from the first stage marsh pore-water salinity model. Four inputs used to simulate the residual error included weekly averages of the decorrelated specific-conductance variable from each GPA site and a 1-week moving window average of specific conductance at GPA10 (variables RSC10S_11B_A1WK, RSC10S_11BR_A1WK, RSC10S_12RS_A1WK, and PSCGPA10S_FLR_A1WK, respectively). A summary of the

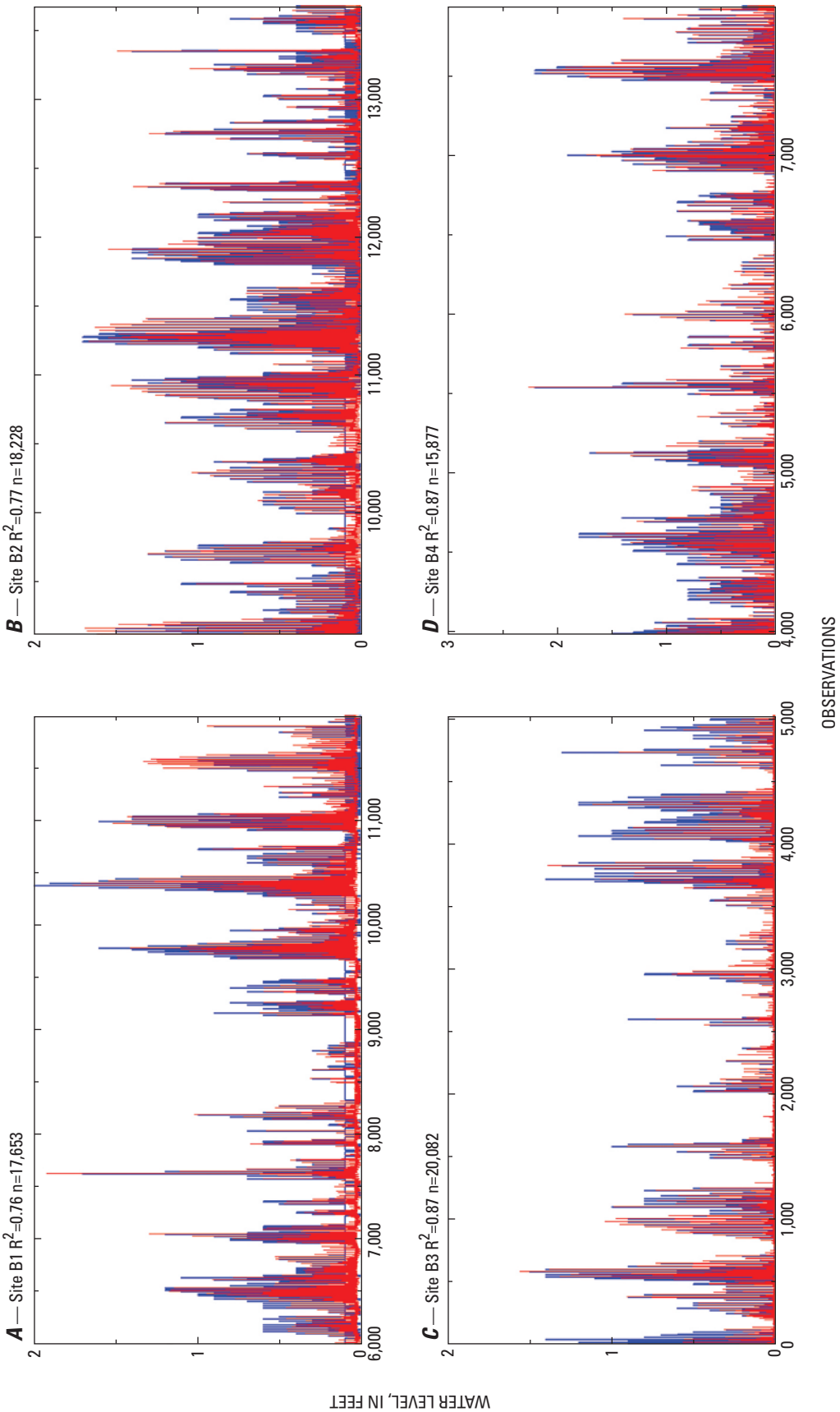

Figure 29. Measured (blue trace) and simulated (red trace) hourly water levels for four USGS marsh gages along the Little Back and Back Rivers for the period June 2000 to May 2005. Data to train the ANN models were bifurcated into training and testing sets. For clarity, only 4,500 to 5,000 observations (approximately 180 days) are shown from the testing data set along with the coefficient of determination (R^2) and number (n) of data points for the testing data set.

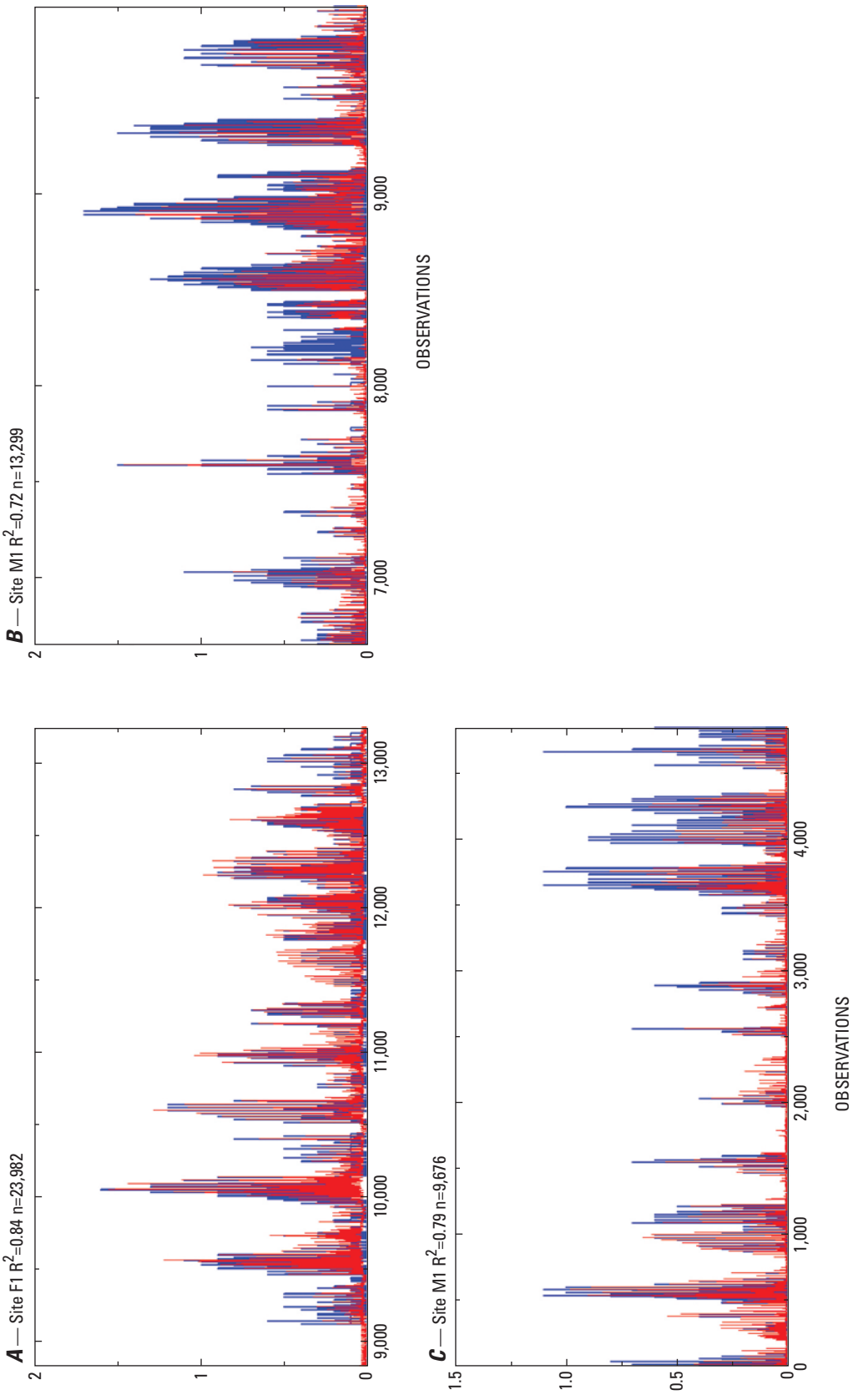

Figure 30. Measured (blue trace) and simulated (red trace) hourly water levels for the three USGS gages along the Middle Back and Front Rivers for the period June 2000 to May 2005. Data to train the ANN models were bifurcated into training and testing sets. For clarity, only 4,500 to 5,000 observations (approximately 180 days) are shown from the testing data set along with the coefficient of determination (R^2) and number (n) of data points for the testing data set.

input variables, size of the training and testing data sets, number of hidden-layer neurons, and coefficient of determination for the daily model can be found in table 3 and Appendix I.

The measured and simulated hourly pore-water salinities are shown in figure 31 for summer 2002. This period was the end of the 5-year drought that began in 1998 and had the highest salinity intrusions of the drought. The model simulates the salinity intrusions occurring on a 14- and 28-day cycle and generally under simulates the salinity concentrations.

The measured and simulated hourly salinity for the seven USGS marsh stations are shown in figures 32 and 33. The quality of marsh water-level models, in term of the coefficient of determination, varies from a R^2 of 0.74 to 0.93. Generally, the models follow the trend of the measured data, as reflected in the high R^2, but miss the variability over short time periods.

Analysis of Estuary Dynamics Using Three-Dimensional Response Surfaces

The river and marsh ANN models can be used to examine the effect of water level and tidal range in the harbor, Savannah River streamflow at Clyo on water levels, and specific conductance in the primary and secondary river channels and the marshes. Three-dimensional response surfaces can be generated to display two explanatory variables (water level, tidal range, or streamflow) with a response variable (specific-conductance or water level). The data for the surface are computed by the ANN model across the full range of the displayed input variables, while the "unshown" inputs (all the models have more than two inputs) are set to a constant value, such as a historical mid-range or mean value. Response surfaces are a valuable tool for understanding the dynamics of an estuary and for comparing how variable interactions differ throughout the system.

The conceptual model of salinity intrusion occurring during de-stratified conditions during neap tides (fig. 13) can be seen in the plot of the time series at GPA04 for summer 1999 (fig. 14). Salinity intrusion dynamics can also be seen in the response surfaces showing the interaction of daily streamflow, water level, and specific conductance for neap and spring tidal conditions (fig. 34). For example, the salinity intrusion on the Front River at Houlihan Bridge site (station 02198920) is shown in figure 34. The response surfaces show increasing streamflow along the x-axis (horizontal, out of the page), increasing water level along the y-axis (horizontal, into

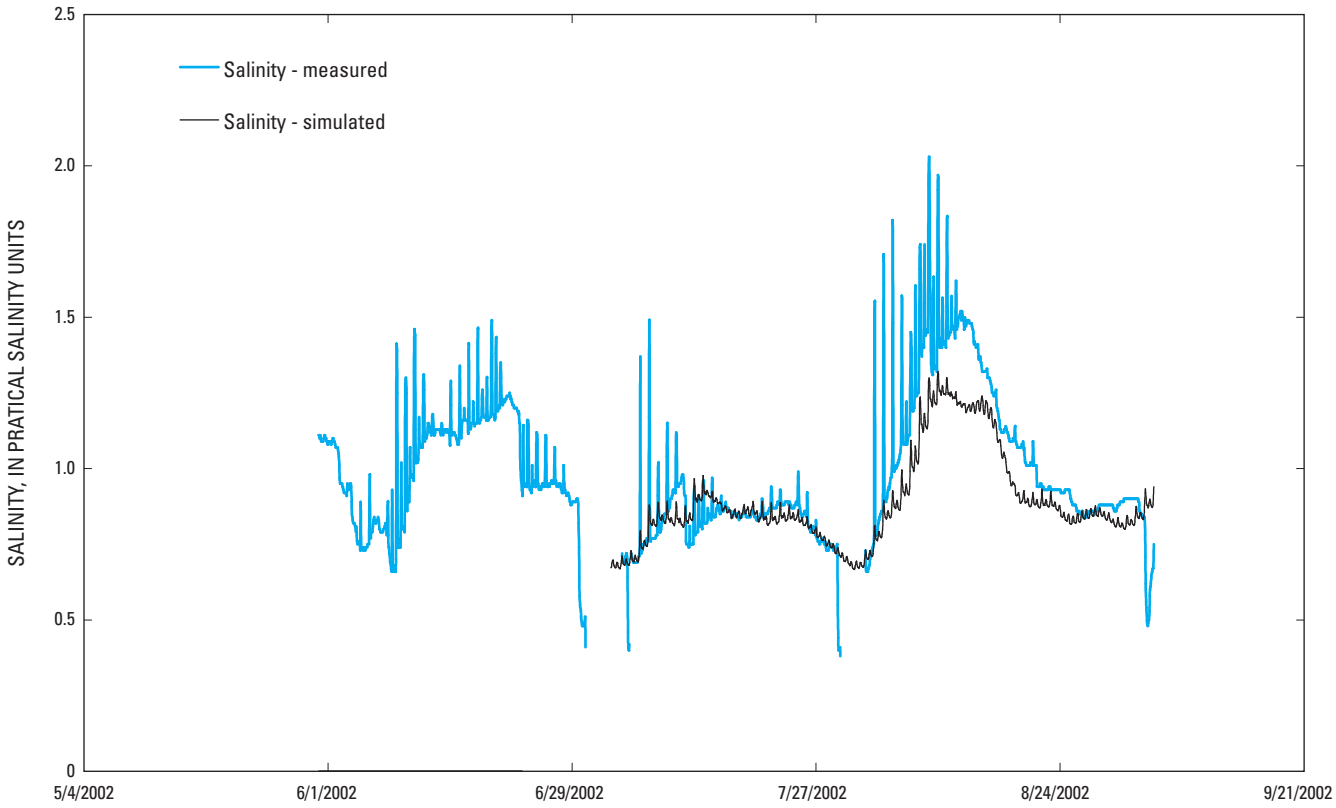

Figure 31. Measured and simulated hourly salinities at Little Back River at Site B2 for the period June 1 to August 31, 2002. Breaks in the prediction time series are caused by incomplete time series for one or more of the model inputs.

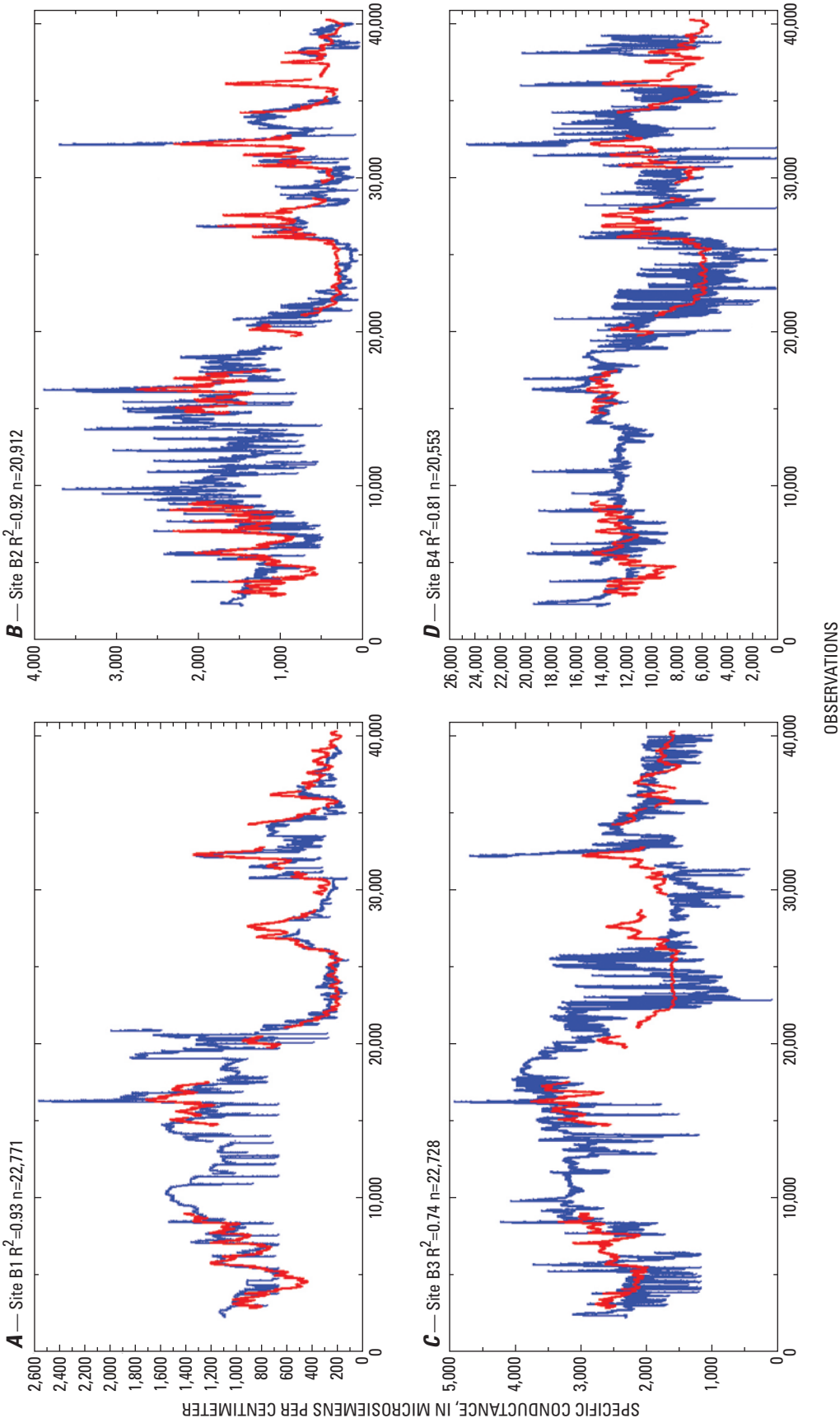

Figure 32. Measured (blue trace) and simulated (red trace) hourly specific conductance for four USGS gages along the Little Back and Back Rivers for the period June 2000 to May 2005. Data to train the ANN models were bifurcated into training and testing sets. Graphs show the complete data set along with the coefficient of determination (R^2) and number (n) of data points. Breaks in the simulated values are caused by one or more missing input values.

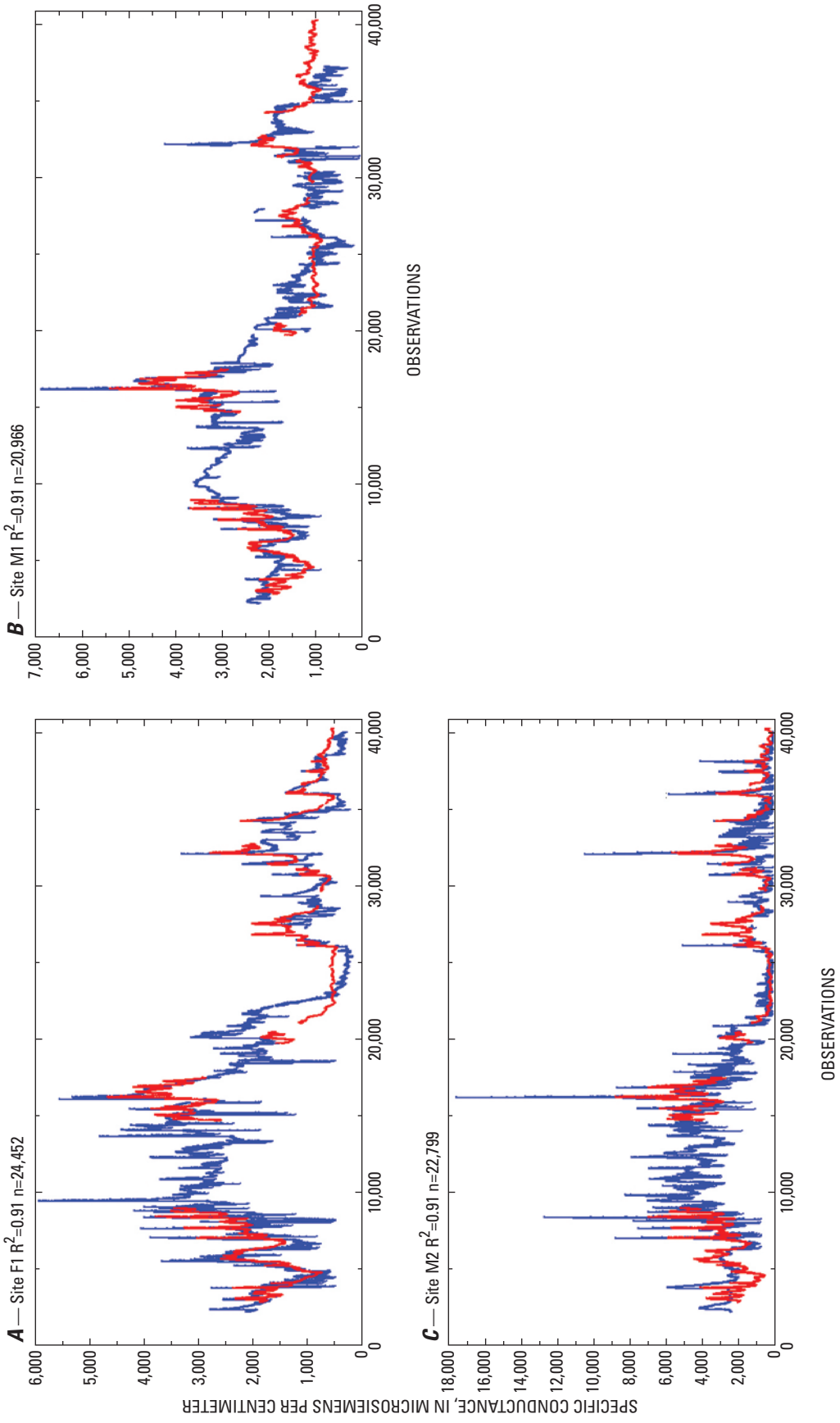

Figure 33. Measured (blue trace) and simulated (red trace) hourly specific conductance for four USGS gages along the Front and Middle Back Rivers for the period June 2000 to May 2005. Data to train the ANN models were bifurcated into training and testing sets. Graphs show data from the testing data set along with the coefficient of determination (R^2) and number (n) of data points. Breaks in the simulated values are caused by one or more missing input values.

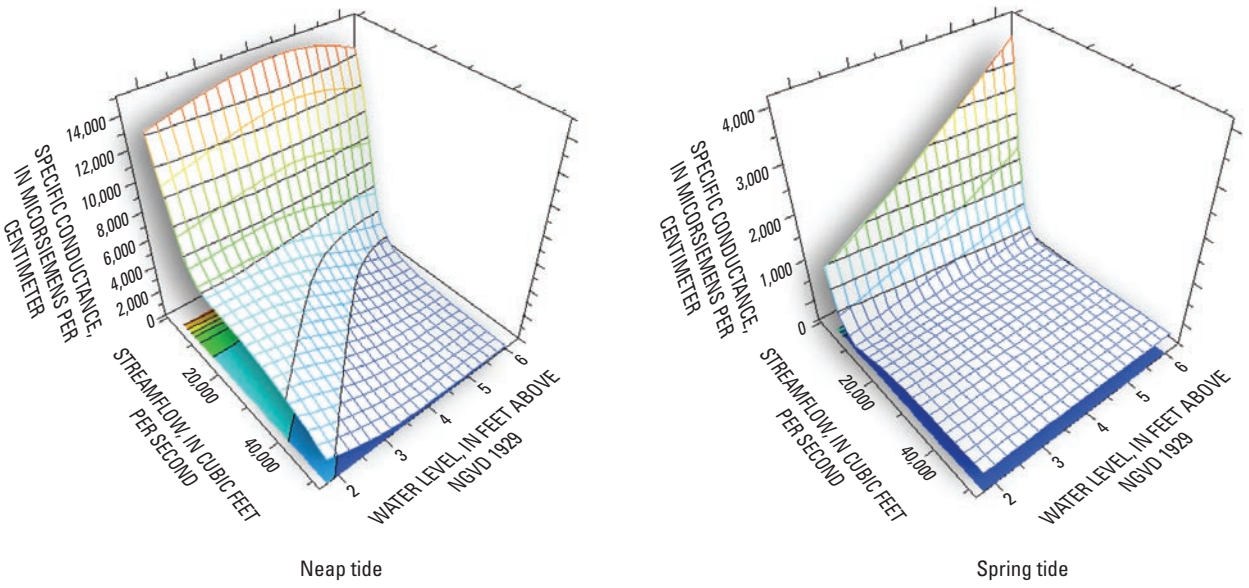

Figure 34. The interaction of water level (FWL8980A) and streamflow (Q8500A) on specific conductance for Front River at Houlihan Bridge (station 02198920) for neap and spring tide conditions. The response surface on the left shows specific conductance response at the gage under neap tide conditions (all other input values set to the mean historical condition for both response surfaces). The response surface on the right shows the gage response under spring tide conditions. Note differences in the z-axis (vertical) scale. During neap tides, salinity intrusion at Front Street is greatest during low-flow conditions through the full range of Harbor water-level conditions. During spring tides, salinity intrusion is greatest during low-flow conditions and increasing water-level conditions in the Harbor.

the page), and increasing specific-conductance (salinity) along the z-axis (vertical) for neap and spring tides (left and right response surfaces, respectively). All other variables were set to their historical means. In figure 34, the neap tide surface on the left shows that salinity intrusion occurs during low-flow conditions for the entire range of water levels in the harbor. Note the change across the response surface when flows are below 10,000 ft^3/s. The spring tide surface on the right in figure 34 shows that salinity intrusion during spring tides occurs only during low flow and increasing harbor water-level conditions. Note the slope of the edge of the response surface at low flows as water levels increase and specific conductance increases from 1,000 to 4,000 µS/cm.

The response surfaces at the I-95 Bridge (station 02198840) show different salinity intrusion dynamics (fig. 35), with the greatest salinity intrusions occurring on spring tides. The response surface at left shows that during neap tides, increases in salinity occur during low-flow conditions and all water-level conditions in the harbor. The response surface at right shows that during spring tides, similar specific-conductance levels occur during low-flow conditions and for most water levels. During very high harbor water-level conditions, salinity intrusion can be more than twice as high as during neap tides. The increased salinity intrusion is possibly due to decreased stratification and fully mixed conditions during neap and spring tides in the upper reaches of the river, which would limit intrusion to periods with increased volumes of saltwater (high water-level conditions), high mixing energy (spring tides), and limited upstream flow resistance to intruding saltwater (low-flow conditions). It should be remembered that the response surfaces show daily values. Maximum hourly specific-conductance intrusions during the drought ranged from 100 to 8,000 µS/cm.

The increased daily specific-conductance, or salinity, values during spring tides at I-95 (station 02198840) can also be seen by plotting streamflows, tidal ranges, and specific conductance with the other inputs set to their historical means. Figure 36 shows that during low flow and historical mean harbor water levels, daily specific conductance can increase as much as 100 µS/cm from neap to spring tides.

Figure 37 shows specific-conductance response surfaces for two sites on the Little Back River. The response surface at left for Little Back River at the USFW Dock (Station 021989791) shows that at mean harbor water levels, salinity intrusion is greatest during low flows and spring tides. Downstream at Little Back River near the Lucknow Canal (Station 021989784), higher specific conductance under similar water-level conditions occurs during low flows and neap tides. Although neither response surfaces show dramatic changes in specific-conductance response under different tidal

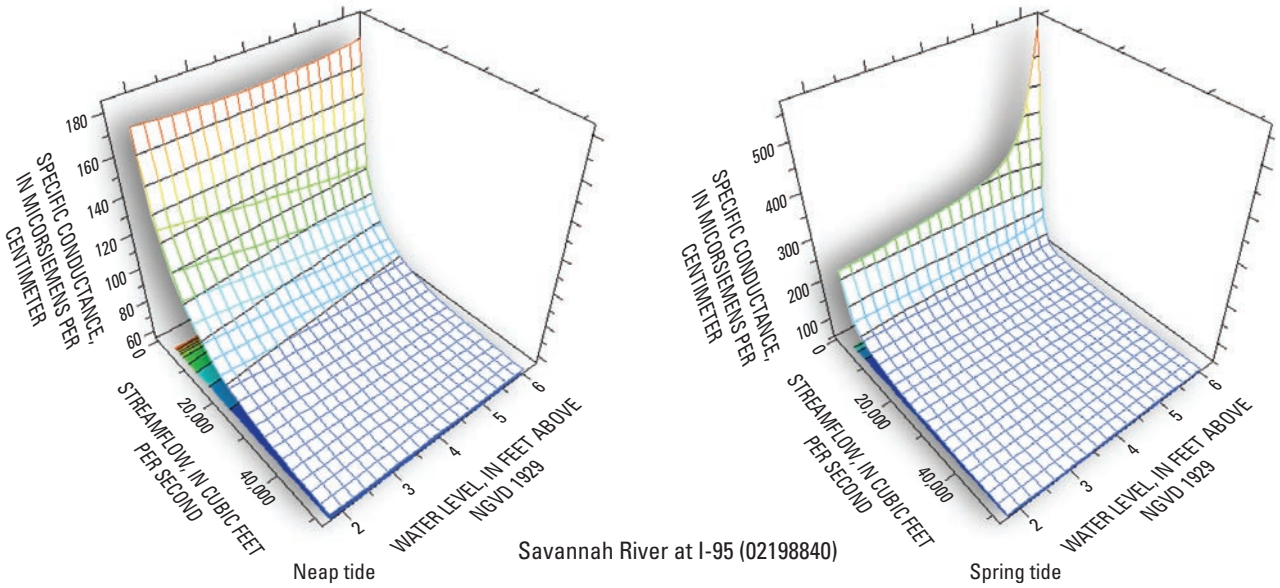

Figure 35. The interaction of water level (FWL8980A) and streamflow (Q8500A) on specific conductance at Savannah River at I-95 (station 02198840) for neap and spring tide conditions. The response surface on the left shows specific conductance response at the gage under neap tide conditions (all other input values set to the mean historical condition for both response surfaces). The response surface on the right shows the specific conductance response under spring tide conditions. Note differences in the z-axis (vertical) scale. During neap tides, salinity intrusion at I-95 is greatest during low-flow conditions through the full range of water-level conditions. During spring tides, salinity intrusion can be over twice as high as neap tide during low-flow and high water-level conditions.

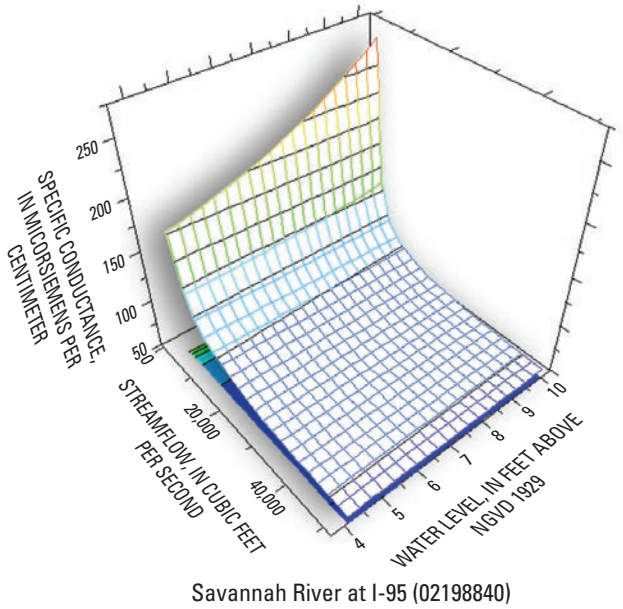

Figure 36. The interaction of tidal range (XWL8980A) and streamflow (Q8500A) on specific conductance at Savannah River at I-95 (station 02198840). The response surface shows the specific conductance response at the gage under mean water-level conditions (all other input values set to the mean historical condition for both surfaces). During mean water-level conditions, specific conductance intrusion at I-95 is greatest during low-flow and spring-tide conditions.

range conditions, they do indicate a gradual change in salinity intrusion behavior from the I-95 Bridge to the station near Lucknow Canal.

Response surfaces can also be used to analyze the influence of Savannah River streamflow and harbor water levels on the local water level in the vicinity of the SNWR. As noted above, streamflow at Clyo, Ga., (station 02198500) and harbor water levels (station 02198980) are virtually uncorrelated (R=-0.03). The effect of streamflow and water level on local water levels depends on their proximity to the harbor. Water levels in the upper reaches exhibit more riverine influence than the downstream stations. The riverine influence diminishes farther downstream, and the influence of the tide is more prominent. Figures 38 and 39 show response surfaces for four water-level gages. The upstream station, Savannah River at I-95 (fig. 38, left response surface), shows a significant contribution of streamflow to the local water level. The downstream station, the Front River at Broad Street (fig. 39, right response surface), shows that streamflow has a negligible influence on its water level. The other two stations, Front River at the Houlihan Bridge (station 02198920, fig. 38, right response surface) and Little Back River at Limehouse (station 02198979, fig. 39, left response surface), show large contributions of harbor water level and small contributions of streamflow to the local water levels.

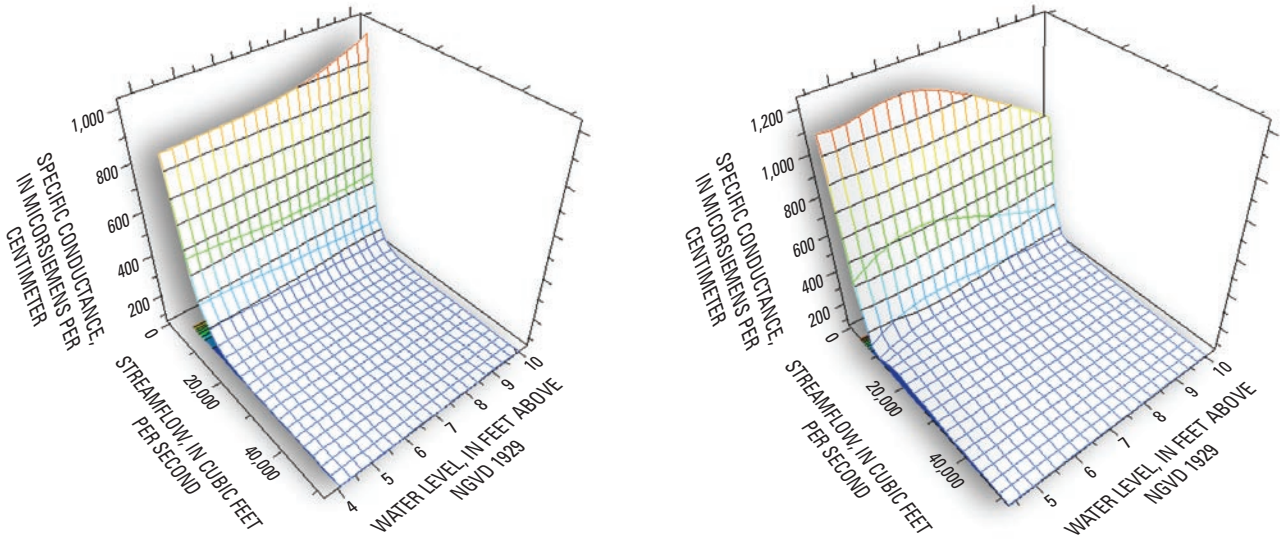

Figure 37. The interaction of tidal range (XWL8980A) and streamflow (Q8500A) on specific conductance at Little Back River at U.S. Fish and Wildlife Service Dock (station 021989791, left response surface) and for Little Back River at Lucknow Canal (station 021989784, right response surface). The response surface on the left shows higher salinity intrusion at the upstream station occurs during spring tides. The response surface on the right shows that high salinity intrusions occur during neap tides.

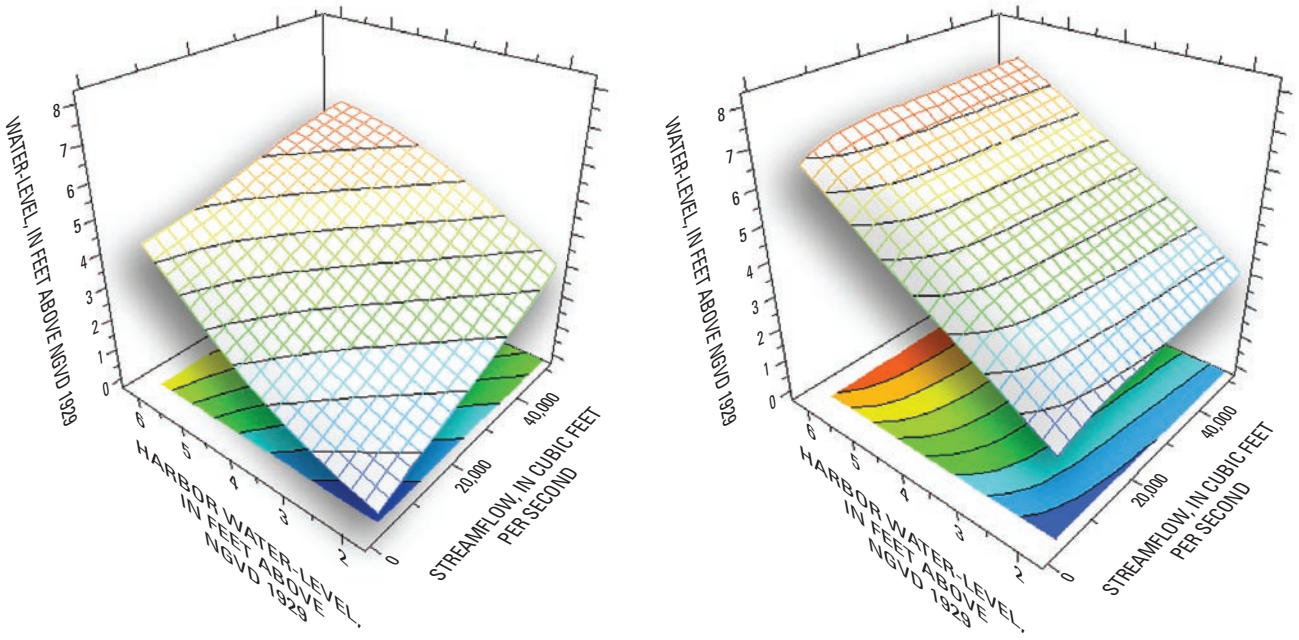

Figure 38. The interaction of harbor water level (FWL8980A) and streamflow (Q8500A) on local water levels at Savannah River at I-95 (station 02198840, left response surface) and for Front River at Houlihan Bridge (station 02198920, right response surface). The response surface of water levels on the left shows higher contribution of streamflow than the downstream station on the right.