SIMULATION OF SALINITY IN THE TIDAL MARSHES IN THE VICINITY OF THE SAVANNAH NATIONAL WILDLIFE REFUGE USING ARTIFICIAL NEURAL NETWORKS

Paul A. Conrads[1], Edwin A. Roehl[2], Ruby C. Daamen[2], Wiley Kitchens[3]

AUTHORS: Hydrologist[1], U.S. Geological Survey, 720 Gracern Road, Suite 129, Stephenson Center, Columbia, SC 29210
Principals[2], Advanced Data Mining Services, 3620 Pelham Road, PMB 351, Greenville, SC 29515
Assistant Unit Leader[3], Florida Cooperative Fish and Wildlife Research Unit, Building 810, University of Florida, Gainesville, Fl 32611
REFERENCE: *Proceedings of the 2007 Georgia Water Resources Conference*, held March 27–29, 2007, at the University of Georgia.

Abstract. The Savannah Harbor is one of the busiest ports on the East Coast of the United States and is located downstream from the Savannah National Wildlife Refuge, which is one of the Nation's largest freshwater tidal marshes. The Georgia Ports Authority and the U.S. Army Corps of Engineers funded hydrodynamic and ecological studies to evaluate the potential effects of a proposed deepening of Savannah Harbor as part of the Environmental Impact Statement. These studies included a three-dimensional (3D) model of the Savannah River estuary system, which was developed to simulate changes in water levels and interstitial (or pore-water) salinity in the system in response to geometry changes as a result of the deepening of Savannah Harbor, and a marsh-succession model that predicts plant distribution in the tidal marshes in response to changes in the water-level and interstitial salinity conditions in the marsh. Beginning in May 2001, the U.S. Geological Survey entered into cooperative agreements with the Georgia Ports Authority to develop empirical models to simulate the water level and salinity of the rivers and tidal marshes in the vicinity of the Savannah National Wildlife Refuge and to link the 3D hydrodynamic river-estuary model and the marsh-succession model.

Understanding freshwater inflows, tidal water levels, and specific conductance in the rivers and marshes is critical to enhancing the predictive capabilities of a successful marsh succession model. Data-mining techniques, including artificial neural network (ANN) models, were applied to address various needs of the ecology study and to integrate the riverine predictions from the 3D model to the marsh-succession model. ANN models were developed to simulate riverine water levels and specific conductance in the vicinity of the tidal marshes for the full range of historical conditions using data from the river gaging networks. ANN models also were developed to simulate the marsh water levels and interstitial salinities using data from the marsh gaging networks. Using the marsh ANN models, the continuous marsh network was hindcasted to be concurrent with the long-term riverine network. The hindcasted data allow ecologists to compute hydrologic parameters—such as hydroperiods and exposure frequency—to help analyze historical vegetation data.

INTRODUCTION

Under sponsorship from the U.S. Army Corps of Engineers and the Georgia Ports Authority (GPA), the Lower Savannah River Estuary and the surrounding freshwater tidal marshes of the Savannah National Wildlife Refuge (SNWR) have been studied for years by a variety of governmental agencies, water users, universities, and consultants. Their interests are in maintaining water quality and predicting the potential impacts of a proposed harbor deepening on the estuary and tidal wetlands. Two major initiatives were the development of a three-dimensional hydrodynamic model (3DM) by a team of hydrologists, and the development of a marsh succession model (MSM) by a team of plant ecologists. The 3DM predicted changes in riverine water levels and salinity in the system in response to potential harbor changes. The MSM predicted plant distribution in the tidal marshes in response to changes in the water-level and salinity conditions in the marsh. A mechanism for linking riverine and marsh behaviors was needed.

To support 3DM and MSM development, many disparate databases had been created that described the natural system's complexity and behaviors, but these databases had not been compiled into a usable form. Variables having particular relevance include those describing bathymetry, meteorology, discharge (Q), water level (WL), specific conductance (SC), water temperature, and dissolved oxygen concentration. Most of the databases were composed of time series that varied by variable type, periods of record, measurement frequency, location, and reliability. Scientists recognized that data mining techniques, which include artificial neural networks (ANN), could be used to link riverine and marsh behaviors. This paper presents a brief description of a few of the technical aspects of the study. Readers interested in a more detailed description of the study are referred to Conrads and others (2006).

MODELING

The authors had previously developed ANN-based models of estuaries in Charleston and Beaufort, South Carolina (Conrads and Roehl, 1999; Conrads and others, 2003). The type of ANN used in these cases was the multi-layered perceptron (MLP) described by Jensen (1994) which is a multivariate, non-linear regression method based on machine learning. For the Savannah River Estuary Study, linking the riverine predictions of the 3DM to the MSM required the development of a *super-model*, called the "model-to-marsh" or M2M model. The M2M needed to simulate riverine and marsh water levels and salinity in the vicinity of the SNWR for the full range of historical conditions using data from the riverine and marsh gaging networks. The M2M integrates cascading *sub-models* that are used to decorrelating variables, simulating river responses, and simulating tidal marsh responses.

Linking the 3DM to the MSM is accomplished by using predicted differences in WL and SC values for the river generated by the 3DM as input to the M2M. Using the predicted difference for the river, the M2M predicts the change in WL and SC in the tidal marshes. These predictions are then used by the MSM to predict changes in the plant communities in the tidal marshes.

Historical Databases

The available data from four disparate databases required extensive clean up for problems such as erroneous and missing values and timing errors. The resulting database includes 11½ years of half-hourly data (200,000+ time stamps) for 110 variables. The original sources of data were:

- Q_{Clyo} and WL_{Harbor} – 11½ years of half-hourly WL signals in Savannah Harbor and river flows measured 50 miles inland at Clyo, Georgia, by the U.S. Geological Survey (USGS).
- USGS riverine WL and SC – 11½ years of half-hourly signals collected from four stations in the Lower Savannah River by the USGS.
- GPA riverine WL and SC - half-hourly signals collected on behalf of the GPA from 14 stations over 3 months each in 1997 and 1999. Some stations recorded surface and bottom SC measurements (SC_{top}, SC_{bottom}).
- USGS marsh WL and SC – 4½ years of half-hourly signals collected from seven stations (2000-2005).
- GPA marsh WL and SC – 19 months of half hourly SC and WL data collected from 10 stations.

Much of the field data was collected during a record-setting 4½-year drought, raising concerns that the data was not representative of "normal" hydrodynamic conditions. It was expected that the ANNs could reasonably interpolate from the field data by "learning" the full range of behaviors exhibited over 11½ years, which also included two El Niño events when flows were substantially above average, and presumably periods of normal conditions.

Signal Decomposition

The hydrodynamic and water-quality behaviors observed in estuaries are superpositions of behaviors forced by periodic planetary motions and chaotic meteorological disturbances. Theoretically, periodic behaviors are perfectly predictable, and chaotic behaviors are only somewhat so; therefore, the real problem with modeling estuaries is to empirically synthesize chaotic output signals from multiple chaotic input signals. Signals are easily decomposed into periodic and chaotic components using spectral filtering. The primary chaotic inputs to the Lower Savannah River are the flows measured at Clyo, Georgia, and the chaotic oceanic disturbances represented in the chaotic component of WL_{Harbor}.

The empirical representations of the dynamic behaviors that underlie periodic and chaotic signals are different. Multiple periodic signals are superpositions of individual periodic signals that are represented by three constants — phase, amplitude, and frequency. Abarbanel (1996) describes how chaotic univariate systems can be optimally represented by *dynamical invariants* — characteristic *time*

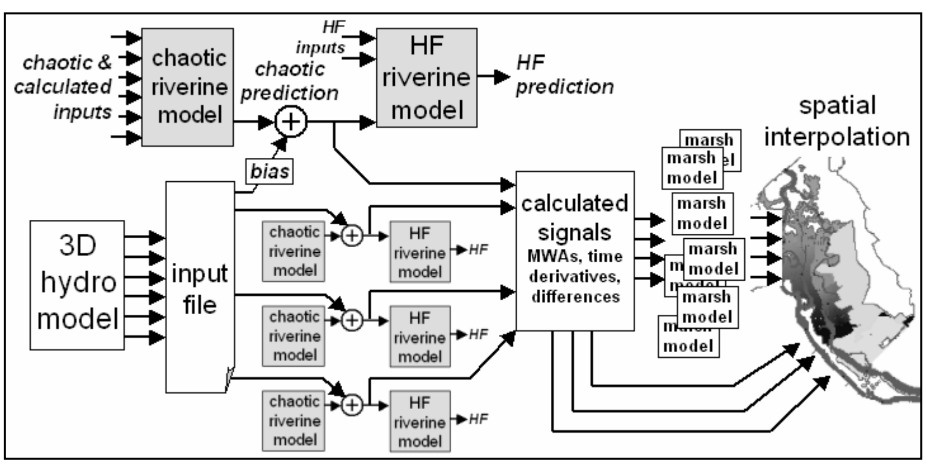

Figure 1. Data flow through the super model. Separate sub-models were used for each WL and SC prediction.

delays and *dimensions*. Roehl and others (2000) describe an ANN model that predicted the salt-front location in the Cooper River, which incorporated signal decomposition and extended the univariate representation of chaotic behaviors to a multivariate system.

Chaotic components were extracted from raw signals by applying a low-pass spectral filter to remove high-frequency (HF) diurnal and semi-diurnal variability. The periodic tidal range (XWL) was computed from WL_{Harbor}. The chaotic component of Q_{Clyo} was further processed with moving window averages of up to 2 weeks, so that when input to an ANN with multiple time delays, flow histories of up to 44 days were represented.

M2M

The M2M super-model includes 127 sub-models. Figure 1 shows that cascading sub-models predicted chaotic WL and SC signal components at riverine and marsh gaging sites. Using low-pass filtered Q_{Clyo}, WL_{Harbor}, and XWL signal components for inputs, "chaotic sub-models" predicted chaotic WL and SC behaviors at four USGS gaging sites in the main channel. These outputs were input to "HF sub-models" that also used HF WL_{Harbor} and XWL component inputs to obtain HF WL and SC predictions at the four gaging sites.

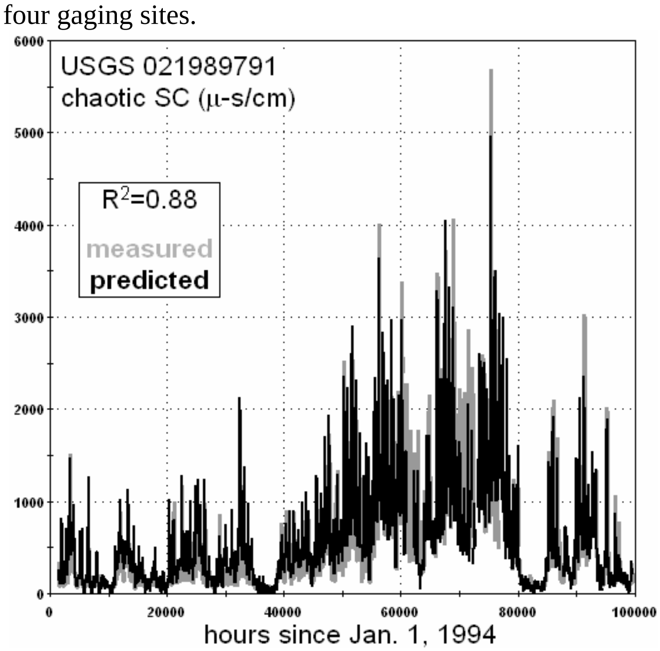

Figure 2. Measured and predicted chaotic riverine SC. Increased SC at center right occurred during the drought.

The chaotic predictions at the main channel sites were then transformed into calculated signals to decorrelate them and to represent dynamic behaviors that evolve over weeks. The calculated signals were used as inputs to model the historically shorter signals at the many remaining riverine and marsh stations. This provided one set of ANNs that linked the river's main channel behaviors to tidal forcing and freshwater flows, and a second set that linked main channel behaviors to those in the marsh. Figures 2, 3, and 4 show SC predictions at a riverine site and at a nearby marsh site. The R^2 of the SC predictions at most of the gaging sites ranged between 0.8 and 0.9. The R^2 of the marsh interstitial SC predictions generally were greater than 0.7.

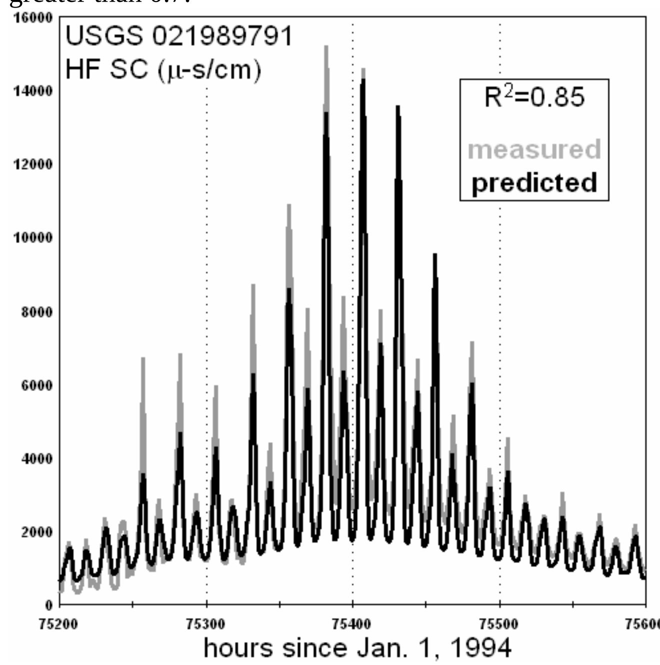

Figure 3. Measured and predicted HF riverine SC. 16.6 days are shown during the drought.

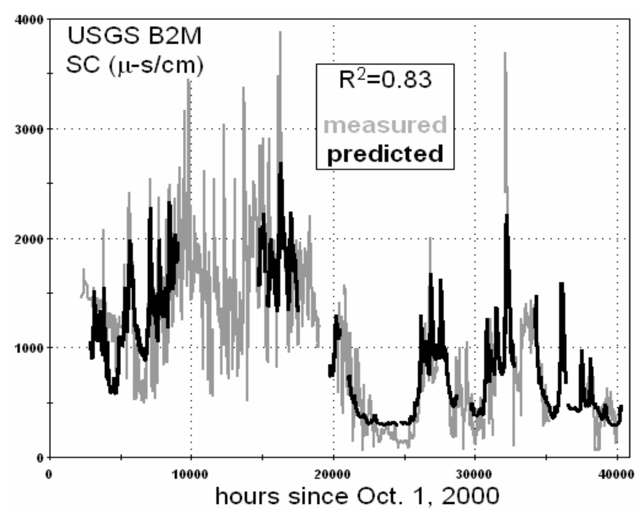

Figure 4. Measured and predicted marsh SC. Gaps mark missing input data.

Simulation and Decision Support

Conrads and others (2006) describe how the execution of the large number of Savannah area sub-models was orchestrated by a custom decision support system (DSS). The DSS integrates the super-model with an 11½-year database, comprising more that 200,000 records of half-hourly measurements, for running long-term simulations. The DSS also provides a graphical user interface, streaming graphics, several freshwater flow input options, and output file generation to allow stakeholders of varying technical backgrounds to evaluate alternative scenarios under the widely ranging conditions that manifest in such a long historical record.

3DM Integration

Figure 1 shows that the 3DM is linked to the M2M super-model through an output file. The file contains WL and SC biases for the main gaging sites. The biases are calculated by subtracting 3DM predictions representing proposed channel geometries from predictions generated using the actual historical conditions.

MSM Integration

Figure 1 shows that riverine and marsh predictions at gaged sites are interpolated to generate a 2D contour map of SC on a grid of the study area. The interpolation is performed using rules written for each grid cell. The rules accommodate the different transport mechanisms of channels and marshes. The interpolation, extrapolation, and visualization are performed in a custom post-processor that imports output from the DSS and writes interpolated values to an output file. The post-processor converts SCs to salinities, and provides different options for time-averaging the predictions. Output from the post-processor can be imported into the MSM so that plant ecologists can evaluate the impacts of predicted salinity changes.

CONCLUSIONS

The M2M leverages and integrates millions of dollars of field data collection and modeling performed over more than a decade by several scientific organizations. A divide-and-conquer super-model solution, enabled by signal decomposition and accurate ANN sub-models, allowed a large amount of disparate data and intermediate works to be optimally used in their entirety. The packaging of the super-model and data in a DSS makes the scientific products immediately accessible and useful to all stakeholders.

LITERATURE CITED

Abarbanel, H.D.I., 1996 *Analysis of Observed Chaotic Data*, Springer-Verlag New York, Inc., New York, 4-12, 1996.

Conrads, P.A. and E.A. Roehl, 1999,Comparing physics-based and neural network models for predicting salinity, water temperature, and dissolved oxygen concentration in a complex tidally affected river basin, South Carolina Environmental Conference, Myrtle Beach, March 1999.

Conrads, P.A., E.A. Roehl, and W.P. Martello, 2003 Development of an empirical model of a complex, tidally affected river using artificial neural networks," Water Environment Federation TMDL Specialty Conference, Chicago, Illinois, November 2003.

Conrads, P.A., Roehl, E.A., 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

Jensen, B.A., 1994, Expert systems - neural networks, Instrument Engineers' Handbook Third Edition, Chilton, Radnor PA

Roehl, E.A., P.A. Conrads, and T.A. Roehl, 2000, Real-time control of the salt front in a complex, tidally affected river basin, Proceedings of the Artificial Neural Networks in Engineering Conference, St. Louis, 947-954,