7th International Conference on Hydroinformatics
HIC 2006, Nice, FRANCE

FEATURES OF ADVANCED DECISION SUPPORT SYSTEMS FOR ENVIRONMENTAL STUDIES, MANAGEMENT, AND REGULATION

EDWIN ROEHL
Advanced Data Mining, 3620 Pelham Road, PMB 351
Greenville, SC 29615, United States

RUBY C. DAAMEN
Advanced Data Mining, 3620 Pelham Road, PMB 351
Greenville, SC 29615, United States

PAUL A. CONRADS
U.S. Geological Survey, 720 Gracern Road, Suite 129
Columbia, SC 29210, United States

Natural resource managers and stakeholders face difficult challenges when managing interactions between natural and man-made systems. Even though the collective interests and computer skills of the community of managers, scientists, and other stakeholders are quite varied, there is an overarching need for equal access by all to the scientific knowledge needed to make the best possible decisions. A decision support system (DSS) can meet this need. DSSs have been described as, *"computer-based systems (for) 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."* This paper describes DSS developed for three hydrologic systems in South Carolina and Georgia, USA. The goals of the three DSSs were – the regulatory permitting of wastewater treatment plants on the Beaufort River, evaluating the environmental impact of a proposed deepening of Savannah Harbor, and evaluating hydroelectric generation on the Pee Dee River to protect Myrtle Beach-area freshwater intakes from salinity intrusions. These DSSs provide predictive models with real-time databases for simulation, graphical user interfaces, and streaming displays of results. Additional features include optimizers, integrations with other models and software tools; and color contouring of simulation output data.

INTRODUCTION

Natural resource managers and stakeholders face difficult challenges when managing the interactions between natural and man-made systems. Complex mathematical (mechanistic) models based on first principles physical equations are often developed and operated by scientists to evaluate options for using a resource while minimizing damage. The interests and computer skills of the actual decision-makers and other stakeholders, however, can be quite varied. There is an overarching need for equal access by all to the body of scientific knowledge needed to make the best possible decisions. A decision support system (DSS) can meet this need.

Dutta and others [3] describe DSSs as, "*computer-based systems (for) helping decision-makers to solve various semistructured and unstructured problems involving multiple attributes, objectives, and goals… With the help of AI (Artificial Intelligence) techniques DSS have incorporated the heuristic models of decision-makers and provided increasingly richer support for decision-making.*"

This paper describes DSSs developed to address hydrologic issues in three estuaries in South Carolina, USA, and shown in Figure 1. The issues are regulatory permitting of three wastewater treatment plants (WWTP) on the Beaufort River, evaluating the environmental impact of a proposed deepening of Savannah Harbor, and regulating hydroelectric generation on the Pee Dee River to protect coastal freshwater intakes from salinity intrusions.

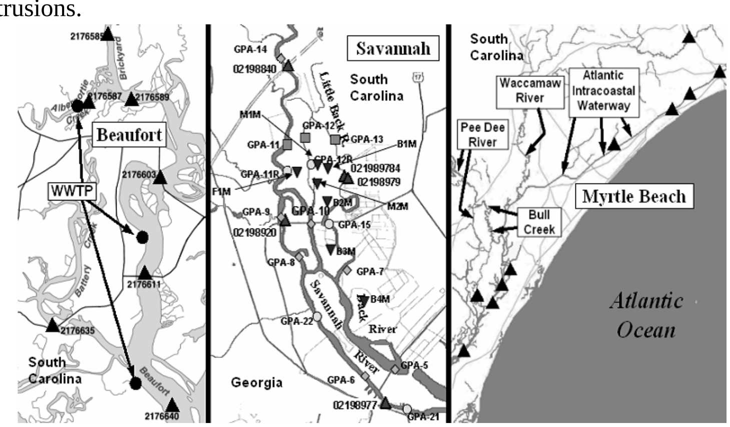

Figure 1. Beaufort, Savannah, and Myrtle Beach study areas. Markers denote gaging sites (triangles or squares) or wastewater treatment plant outfalls (circles).

MODELS

Commonly, a DSS is a software package built around a model, making the model the DSS's most important component because ostensibly, it can correctly predict, "*What will happen if we do A instead of B?*" Models are often developed at considerable expense; therefore, the packaging around a DSS is done only to maximize its usefulness to the broadest possible community of users. Regardless of the packaging, a model that lacks scientific credibility can delay the natural resource management process indefinitely.

Models used to simulate hydrologic systems are usually either mechanistic or empirical models. Calibrating these models is a process of fitting a line or surface (function) through data from two or more variables. Functions predicting system response are either prescribed or synthesized. The functions prescribed by mechanistic models are physical equations, which incorporate tunable coefficients that are adjusted by modelers to match calibration data. Linear regression is the most common empirical modeling technique. It prescribes straight lines, planes, or hyper-planes to fit calibration data. The insurmountable problem with prescriptive modeling techniques is that if their functional form is inherently unable to fit the variable relations that are manifested in the data, a representative model is unobtainable. Some mechanistic modeling projects have consumed millions of dollars and many years of effort, with some models never accepted by the regulatory agencies and stakeholders.

Artificial neural networks (ANN) are a machine learning technique from AI. Unlike mechanistic and linear models, ANNs synthesize rather than prescribe non-linear functions to fit multivariate data. Conrads and Roehl [1] found that ANN models had prediction errors that were significantly lower than those of a state-of-the-practice mechanistic model when predicting water temperature (WT), specific conductance (SC), and dissolved oxygen (DO) on Charleston's Cooper River estuary. Other benefits of ANN models included short development time and fast execution that allows for numerical optimization, spreadsheet integration, and integration of ANN models with mechanistic models.

DSS FEATURES

All three of the DSSs were developed as Microsoft Excel™/Visual Basic for Applications[1] (VBA) programs. This allowed the DSS to be prototyped, easily modified, and distributed in a familiar form. The DSSs are operated through a graphical user interface (GUI). Other common elements of the DSS are described below.

Data

Predictive models were developed to represent complex, non-linear, dynamic behaviors manifested in years of time series. Spectral filtering was applied to decompose the hydrodynamic, water-quality, and meteorological signals into components that differentiate *periodic* and *chaotic* behaviors. Moving window averages (MWA) of varying window sizes are applied to augment these components with calculated variables that represent behaviors evolving on different time scales, for example, it takes months of data to represent an extended drought condition.

Modeling and Simulation

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

ANN *sub-models* are used to systematically decorrelate input variables and predict individual signal components. The sub-models are then assembled into a *super-model* that represents an entire system. This produces predictive models that are customized to the unique circumstances and data of a particular system. Each DSS has at least two executions of the super-model. One generates predictions using actual historical input conditions, which are used to compute prediction errors and graphically depict accuracy. The second instantiation generates "*What if?"* predictions using user-defined *controllable* inputs. Two of the applications provide optimizers that modulate controllable inputs during simulations to obtain predictions that match user-defined setpoints.. Each DSS incorporates a database of measured, filtered, and calculated time-series variables for running long-term simulations. Under user control, a VBA program loops through database records, assembles input vectors, executes super-model instantiations, post-processes and writes model output, and drives graphics.

BEAUFORT RIVER WATER QUALITY

The Beaufort River, South Carolina, is a complex estuarine river system that supports a variety of uses including shellfish grounds, fisheries nursery habitats, shipping access to Port Royal, receiving waters for wastewater effluent, and an 18-mile reach of the Intracoastal Waterway. The river was on the Section 303(d) list of impaired waters of South Carolina for low DO concentrations [4]. The Clean Water Act stipulates that a Total Maximum Daily Load (TMDL) must be determined for impaired waters. A data-collection and modeling program was launched in response to the TMDL mandate, and to support the permitting of two existing facilities and a new municipal WTTP constructed (2006) by Beaufort-Jasper Water and Sewer Authority.

Data

A network of seven real-time gaging stations was operated by the U.S. Geological Survey (USGS) on the Beaufort River and its tributaries. Water level (WL), WT, SC, and DO were measured at 15-minute intervals for 34 months. The *DO-difference from saturation* (DOD) was calculated per USGS guidelines [5] to extract the component of DO variability that was unrelated to gas-in-liquid solubility. Three stations were equipped with acoustic velocity meters (AVM) that were used to compute tidal streamflow, and daily rainfall was measured at two locations. Biochemical oxygen demand (BOD) and ammonia (NH_3) loads generally were measured only once per week at each WWTP, and not concurrently plant-to-plant. The low sample frequency of the BOD and NH_3 loads dictated a 1-day time step for the model.

Modeling and Simulation

The Beaufort super-model was composed of 118 separate ANN sub-models. DOD at each gaging site was modeled using inputs representing WL, SC, WT, rainfall, BOD and NH_3. Conrads and others [2] detail how cascaded sub-models were used to decorrelate input variables and predict dynamic point and non-point source load responses. A cubic-spline was used to predict DOD at river locations between the gaging sites. Bathymetric data were used to construct a geometric model having 90- by 90-meter cells. A medial axis transform was fitted to a two-dimensional planar view of the waterways to provide the lengthwise spatial coordinate.

Special Features

A constrained optimizer was configured to represent South Carolina State law that governed the maximum allowable impact that nutrient loads from the three WWTPs could have on riverine DO. Water-resource regulators evaluate receiving waters for seasonal impact limits and segment the river for volume-averaging the impacts. Users can allocate the TMDL load among the BOD and NH_3 discharges from each WWTP. At each time step, the optimizer iterates load inputs as the assimilative capacity changes. The GUI provides controls for exploring different load and segmentation scenarios. The overall TMDL was found to be sensitive to these parameters. The DSS also allows for evaluation of rainfall impacts. Rainfall can be adjusted as a percentage of historical conditions.

Status

In terms of acceptance by stakeholders, the Beaufort DSS was particularly successful when compared to similar coastal initiatives in South Carolina that used state-of-the-practice mechanistic models. Permits were issued only 35 months after model development began, as compared to 10 or more years for similar mechanistic modeling projects in Myrtle Beach and Charleston. The shortened timeframe was due, in part, to demonstrably better prediction accuracy, a modeling process that continuously engaged stakeholders, and DSS packaging that directly addressed the permitting problem.

SAVANNAH HARBOR DEEPENING

The Savannah Harbor is one of the busiest ports on the East Coast of the USA. The harbor is located downstream from the Savannah National Wildlife Refuge (SNWR), an important freshwater marsh. Under sponsorship from the U.S. Army Corps of Engineers and the Georgia Ports Authority (GPA), the Lower Savannah River estuary has been studied for years to evaluate the potential impacts of a proposed harbor deepening. Many databases have been created that describe the natural system’s complexity and behaviors. A *three-dimensional finite-element hydrodynamic model* (3DM) is being developed to predict changes in riverine WL and salinity (S) in response to harbor geometry changes. A *marsh succession model* (MSM) also is being developed to predict how plant distributions in the marshes would respond to WL and S changes. This created a need for a third model, the *model to marsh* (M2M), which would link river and marsh WL and S behaviors. As a result, there was a need for a DSS to integrate all of the models and data for stakeholders.

Data

Figure 1 shows the extensive network of real-time gaging sites operated for the Savannah River study. The WL and SC data included: 11½ years from five USGS gaging sites in the harbor and river; 6 years from seven USGS marsh sites; 3 months from 14 riverine backwater sites operated on behalf of the GPA for 2 different years; and 19 months from 10 GPA marsh sites; and 11½ years of flow (Q) from an upstream river gaging site (not shown in fig. 1). The resulting database was composed of 11½ years of half-hourly data (200,000+ time stamps) for 110 measured variables. Further processing extracted chaotic signal components and calculated the tidal range and various MWA.

Modeling and Simulation

The M2M super-model comprises 127 sub-models. Chaotic sub-models predicted chaotic WL and SC at four USGS gaging sites in the main channels using inputs for Q and harbor WL. These outputs were input to "high frequency" (HF) sub-models that also used HF harbor WL inputs to obtain HF WL and SC predictions at the four gaging sites. The chaotic predictions in the main channel were input to sub-models for the remaining riverine and marsh sites. 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 backwaters and the marsh.

The Savannah DSS provides for simulations of up to 11½ years at hourly, or half-hourly time steps. The variable Q can be set by the user to be a constant or a percentage of the historical flow. User-defined hydrographs also can be run.

Special Features

The 3DM is a complicated program, limiting its accessibility; however, the impacts of different harbor change scenarios can be evaluated using a file generated by the 3DM and imported into the DSS. The file contains WL and SC biases that are calculated by subtracting 3DM predictions (representing proposed channel geometries) from predictions generated using historical conditions.

A custom post-processor imports simulation output and interpolates predictions at gaged sites to generate a 2D contour map of S on a grid of the study area. The interpolation is performed using expert rules written for each grid cell that accommodate the area's topological features and the different transport mechanisms of channels and marshes. The post-processor provides options for time-averaging the predictions, and writes interpolated values to an output file that can be imported into the MSM.

Status

The Savannah DSS was first prototyped in 2002 and a production version was delivered in 2004. Delays in the completion of the 3DM and MSM have postponed its widespread deployment. Most of the original marsh data was collected during a record-setting 4½-

year drought between 1998 and 2002; therefore, the M2M's ANN models were recently retrained with an additional 2½ years of non-drought data.

PEE DEE RIVER SALINITY INTRUSION

Six reservoirs in North Carolina discharge into the Pee Dee River, which flows 160 miles through South Carolina to the coastal communities near Myrtle Beach. During the drought between 1998 and 2002, salinity intrusions inundated a coastal municipal freshwater intake near Myrtle Beach, South Carolina. The North Carolina reservoirs are currently being re-licensed by the Federal Energy Regulatory Commission (FERC) for a 50-year operating permit. The water has important commercial value for generating electric power and for waterfront property development. A coalition composed of Alcoa Power, Progress Energy, the Pee Dee River Coalition, and the South Carolina Department of Natural Resources sought to model the system's hydrodynamics and determine the minimum flows needed to protect coastal intakes.

Data

Nine USGS gaging sites provided the WL and SC data used in the study. The data spanned 17½ years, but not all of the sites were operated concurrently. The quality of the data improved with time. Inflows were obtained from an additional seven USGS gaging sites, and rainfall was obtained from six regional meteorological stations. Coastal wind speed and direction were obtained from one additional meteorological station. The resulting database comprises 17½ years of hourly data (150,000+ time stamps) for 27 measured variables. Further processing extracted chaotic signal components and calculated the tidal range and various MWA.

Modeling and Simulation

The Pee Dee super-model is similar to that of the Savannah DSS; however, only SC is predicted. The super-model employs 18 sub-models, a chaotic and HF sub-model pair for each gaging site. Tidal forcing was input from the easternmost site along the Atlantic Intracoastal Waterway, which was found to be largely unaffected by river flows. The controllable input to the model is the flow from the most-downstream dam (Q_d) on the Pee Dee River, which is summed with the other measured flows with adjustments made for transport delays. The Q_d generally is much larger than the other combined flows in the Pee Dee basin. Rainfall was found to be well accounted for in the inflows, and that wind speed and direction are influential at the southernmost gaging sites.

The Pee Dee DSS provides for simulations corresponding to the most recent and higher quality 6½ years of data, at daily or hourly time steps. The Q_d can be set by the user to be a constant or a percent of the historical measurements. User-defined hydrographs also can be run.

Special Features

The Pee Dee DSS also provides a constrained optimizer that automatically modulates Q_d to match user-set maximum-SC setpoints. The setpoints can be applied on a daily or hourly basis. Higher Q_d is required to suppress hourly SC intrusions. The Pee Dee DSS also provides built-in documentation that describes the variables and user controls, and appears in pop-ups as the mouse is moved in the GUI.

Status

A number of technical review sessions were held where data and model issues were detailed, and successive prototypes were distributed to stakeholders. Feedback from the sessions dictated the DSS's final form, which was completed in 2005.

CONCLUSIONS

A DSS provides a means to effectively transform arcane databases and models into information that is equally accessible to all stakeholders for informed decision-making. Important features that the DSSs have in common include:

- *Predictive Models* - that reliably predict relevant behaviors.
- *Databases* – that contain data describing important historical behaviors.
- *Simulation* – programmatically time-step models.
- *GUIs* – that unite the DSS components with user controls and graphical output.

Features that are more specialized include:

- *Constrained Optimization* – greatly reduces the number of simulations needed.
- *Tool Integration* – the Savannah DSS integrates the 3DM and MSM
- *Expert Knowledge* – such as water quality standards and expert hydrology rules.

REFERENCES

[1] Conrads, P.A. and E.A. Roehl, 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.

[2] Conrads, P.A., E.A. Roehl, and W.P. Martello, 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.

[3] Dutta, S., Wierenga, B., Dalebout, A., Case-based reasoning systems: from automation to decision-aiding and stimulation," *IEEE Transactions on Knowledge and Data Engineering,* 9(6), 911-922, 1997.

[4] South Carolina Department of Health and Environmental Control, State of South Carolina 303d list for 1998, Bureau of Water, EPA Approved, June 1998

[5] U.S. Geological Survey, Technical Memorandum 81.11, Reston, VA, 1981.