ESTIMATION OF TIDAL MARSH LOADING EFFECTS IN A COMPLEX ESTUARY

John B. Cook[1], Paul A. Conrads[2], Edwin A. Roehl, Jr.[3]

AUTHORS: [1]Chief Executive Officer, Advanced Data Mining Intl, 3620 Pelham Rd., PMB 351, Greenville, SC 29615;843-513-2130; john.cook@advdmi.com; [2]Special Hydrologist, U.S. Geological Survey, Gracern Rd, Suite 129, Columbia, SC 29210; 803-750-6141; pconrads@usgs.gov; [3]Chief Technical Officer, Advanced Data Mining Intl, 3620Pelham Rd., PMB 351, Greenville, SC 29615; 864-201-8679; ed.roehl@advdmi.com.

REFERENCE: *Proceedings of the 2010 South Carolina Water Resources Conference,* held October 13-14, 2010, at the Columbia Metropolitan Convention Center.

Abstract. A critical element of a dissolved oxygen (DO) Total Maximum Daily Load (TMDL) is the determination of the relative impacts of point and nonpoint source impacts on the DO of an impaired stream. It is not uncommon that the ultimate oxygen demand for non-point sources during a rain event can be much greater than the effects of the fully permitted point-source loading to the stream. Traditionally, the loading of oxygen-consuming constituents is often estimated with a watershed model which is then coupled with a mechanistic model of the receiving stream. Calibrating coastal system applications of watershed models and mechanistic models to match the behavioral variability observed in actual field data is particularly difficult due to low watershed gradients, poorly defined drainage areas, chaotic forcing functions, and insufficient understanding of watershed and marsh process physics and chemistry.

Data mining offers an alternative approach for analyzing and modeling tidal DO signals to quantify their responses to point- and nonpoint-source loadings. Data mining can be used to extract DO signal components of nonpoint loading caused by rainfall events and tidally entrained organics from marshes and mudflats. At one gaging station on the Cooper River, rainfall was found to decrease DO concentrations at an average rate of approximately 0.25 milligrams per liter (mg/L) per inch of rainfall. Similarly, it was found that specific conductance, water level, and tidal range, which indicate tidal forcing, modulate DO in the range of 3.1 mg/L. This paper examines the approach of using data- mining techniques to improve models for coastal applications and more accurately quantifies the effects of nonpoint-source loading.

INTRODUCTION

The Cooper River is located in the lower Coastal Plain physiographic province in the lower part of the Santee-Cooper River Basin (fig. 1). The basin covers 21,700 square miles and is the second largest drainage basin on the East Coast. The Cooper River is formed by the confluence of the West and East Branches at an area referred to as the “Tee”. This area contains large amounts of poorly defined overbank storage and unmeasureable flows through broken levees between the main channel and rice fields. The West Branch extends from the tailrace of Pinopolis Dam to the Tee. The East Branch Cooper River is a tidal slough throughout its 8-mile reach. On the Cooper River, from the Tee to Flag Creek (Figure 1, just downstream of station USGS 021720675), industries are located along the west bank of the river and extensive *Spartina alterniflora* salt marshes dominate the east bank. Downstream of Flag Creek, the main channel has been dredged to a depth of 42 feet (ft) by the U.S. Army Corps of Engineers for navigational purposes.

The Cooper River is tidally affected throughout its entire reach, and has mean- and spring-tidal ranges of 5.27 and 6.11 ft, respectively, at the Customs House (Figure 1, station USGS 021720711). The objective of this project was to determine if data-mining techniques applied to the Cooper River time-series data could be used to determine the effects of rainfall and tidal flushing on Cooper River water quality. Data was collected over a three-year period of record.

METHODS

The variability of dissolved oxygen (DO) in the Cooper River is a result of many factors including the quality of the water from Lake Moultrie and Charleston Harbor, the loading of oxygen-consuming matter from the tidal marshes, abandoned rice fields, and other non-point sources, effluent from permitted point sources and physical characteristics of streamflow, tidal range, salinity, and especially temperature (Figure 2 shows the inverse relationship between water temperature [WT] and DO). To evaluate whether data mining could be used to determine the influence of tidal marsh loadings on DO, data from a gaging station that was near extensive marsh areas and relatively distal from point-source loadings was selected for evaluation. Of the nine stations in the database on the Cooper River and its tributaries, the gage on the East Branch (USGS station 02172037, fig. 2) was the most dominated by tidal marshes and abandoned

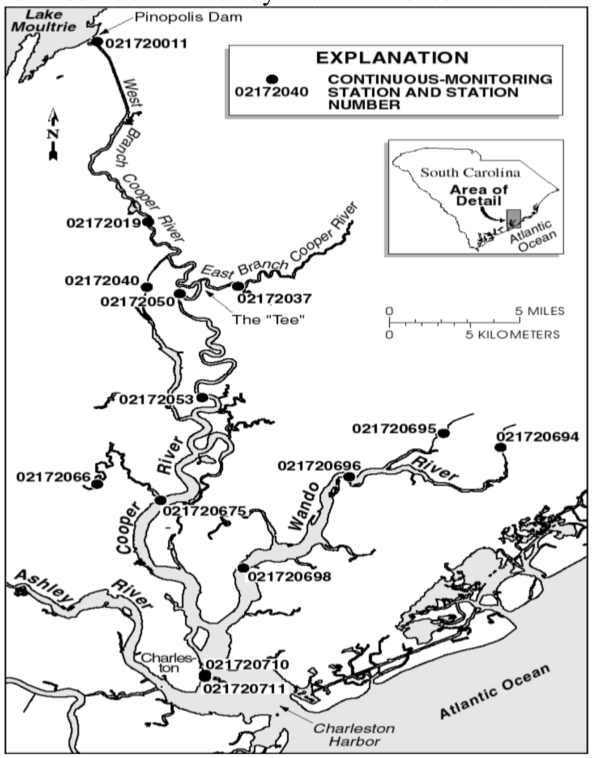

Figure 1: The Cooper and Wando River, SC.

rice fields, and farthest removed from the point-source discharges on the lower Cooper River.

The data used were comprised of hourly measurements for water level (WL), specific conductance (SC), WT, and DO. The effect on DO of the decay of organics can occur over a time scale of several days. This effect can be difficult to discern when coupled with high frequency forces such as diurnal and semi-diurnal ambient temperatures and tidal flow variability. Therefore, the hourly time series were digitally-filtered using frequency domain filtering (Press and others, 1993) to remove diurnal and semi-diurnal periodic signal components. (Filtered variables are denoted by an "f" subscript, for example, DO_f.) A further processing step was taken to decorrelate variables by systematically synthesizing cross-correlation functions and computing their residuals. This step was necessary to avoid the propensity of artificial neural network (ANN) models to over-fit when correlated variables are used as inputs. (Decorrelated variables are denoted by a "d" subscript, for example, $DO_{f,d}$.)

Rainfall data were collected from three National Weather Service stations located in the watershed. These measurements were averaged together, and the resulting signal was converted to a 2-day moving window average "RAINAA" (AA indicating average-of-the-average of three rainfall measurements).

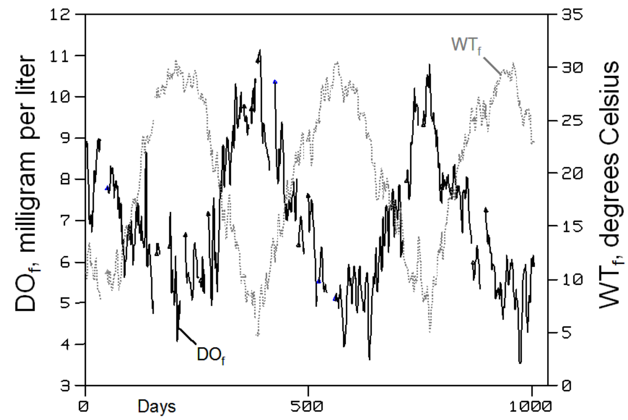

Figure 2: DO_f and WT_f for East Branch of the Cooper River (station 02172037).

The dataset was augmented with calculated variables. The dissolved-oxygen deficit "DOD" was computed as follows: $DOD = DO_{saturated} - DO_{measured}$. $DO_{saturated}$ was determined by adjusting the measured DO values for temperature and salinity (US Geological Survey, 1981). In addition, the daily tidal range (XWL) was computed, interpolated to produce hourly values, and then digitally-filtered as above.

RESULTS

RAINAA and $DOD_{f,d}$ vary seasonally, and upon close inspection of Figure 3, many incidents of high rainfall that coincide with spiking DOD can be seen. Figure 4 shows that an ANN model, having inputs for RAINAA at multiple time delays, τ, starting at 1 day, estimates a significant portion of the $DOD_{f,d}$ variability ($R^2 = 0.281$). $DOD_{f,d}$ was most sensitive to RAINAA @ τ = 3 days.

(The use of the "@" symbol is used by convention to match a model input variable with a time delay, τ, relative to a model output variable.)

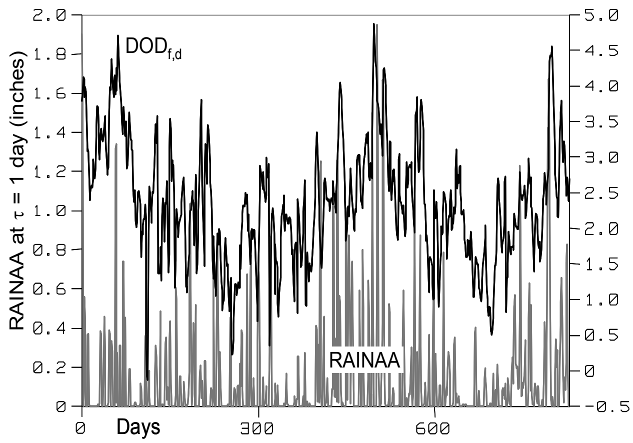

Figure 3: RAINAA and $DOD_{f,d}$ versus Days.

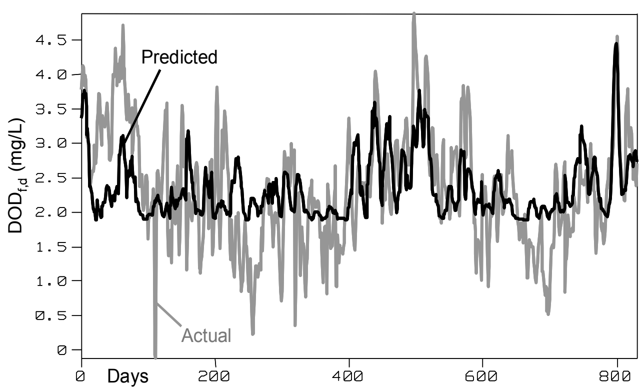

Figure 4: Actual and ANN prediction of $DOD_{f,d}$.

Figure 5 shows the ANN's predicted response for $DOD_{f,d}$ versus RAINAA @ τ = 1 and 3 days (RAINAA inputs at longer delays were set to zero). Also shown are the actual data projected onto the surface. Note that the surface is relatively linear, and that the sensitivity of $DOD_{f,d}$ to RAINNAA @ τ = 1 is consistently less than @ τ = 3 days.

The overall impact of rainfall can be estimated from figure 5 as follows. The total increase in $DOD_{f,d} \approx 2$ milligram per liter (mg/L). This occurs when the RAINAA @ τ = 1 and 3 days are both ≈ 2 inches of rain. (Approximately 0.8 mg/L for τ =1 and 1.2 mg/L for τ =3 days.) Because RAINAA is a 2-day moving average, a value of RAINAA = 2.0 is equivalent to 4 inches of rainfall over 2 days, or 8 inches over 4 days. The sensitivity of $DOD_{f,d}$ to rainfall can be characterized on average as:

$DOD_{f,d} \approx$ 2 mg/L / 8 inches of rainfall over 2 days ≈ 0.25 mg/L per inch of rainfall. It should be noted that South Carolina's water-quality standard for the maximum impact of all point sources on the Cooper River is only 0.1 mg/L.

A second ANN model was created to examine how interactions among WL, XWL, SC, and WT affect DOD. The model used filtered and decorrelated versions of these variables, plus their 1-day derivatives, plus RAINAA as inputs to predict $DOD_{f,d}$. Figure 6 shows the model prediction of $DOD_{f,d}$ versus $XWL_{f,d}$ and $SC_{f,d}$ as a response surface. (To generate the response surface shown in figure 6, RAINAA and derivative inputs were set to 0, and "unseen" WL and WT inputs were set to the midpoints of their ranges.) A number of behavioral modes are apparent. Mode 1 shows that DOD remains low and nearly constant at high XWL. A high XWL, the difference between high and low tide, is an indicator of high tidal fluxes that would transport non-point source oxygen-consuming matter from proximal wetlands downstream to the main channel of the river. Mode 2 shows that DOD also remains low and nearly constant at high SC. A high SC indicates a reverse flow of seawater upriver so that the characteristics of the seawater dominate behavior at the gage. Mode 3 shows that at low XWL, DOD behavior is highly dependent on SC. The highest DOD value occurs at low XWL and SC, indicating conditions that have little tidal exchange. Resident fresh water is being neither aerated nor transported away, allowing non-point organic matter to decay undisturbed. Increasing SC, an indicator of reverse flow upriver, drastically reduces DOD. This is likely caused by the diversion of high quality water flowing downstream from the Cooper River's West Branch into the East Branch where the gage resides. As SC continues to increase, DOD behavior becomes that of Mode 2. Mode 4 shows that DOD declines quickly with

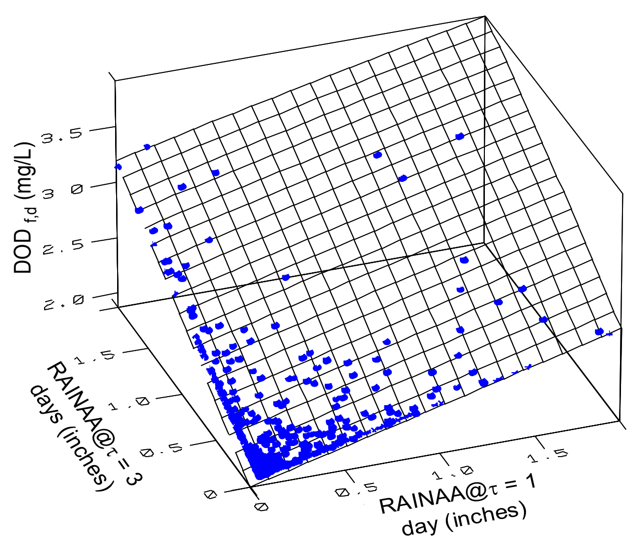

Figure 5: $DOD_{f,d}$ versus RAINAA @ $\boldsymbol{\tau}$ = 1 and 3

increasing XWL and tidal flux. Mode 5 shows that DOD increases at an intermediate but low SC. The surface about Mode 5 is likely a region of transitional behavior in which lower quality water from below the Tee is forced into the East Branch by higher tidal fluxes. In figure 6, the actual data are projected onto the response surface, showing that the data are densest in the region of greatest surface

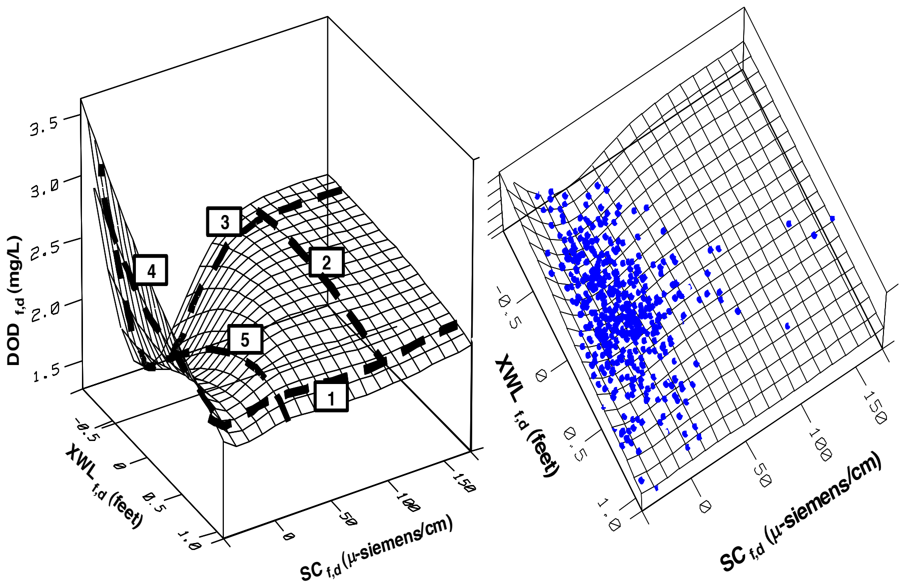

Figure 6: ANN response surface of $DOD_{f,d}$ versuss $XWL_{f,d}$ and $SC_{f,d}$. Different behavioral modes are marked at left. Actual values are projected onto the same surface at right, which has been tipped to show the distribution of the data.

complexity. This provides confirmation that the ANN has captured the very complicated behaviors of the natural system surrounding the gage. It also should be noted that the total computed range of $DOD_{f,d} \approx 2$ mg/L, indicating that non-point organic loading caused by tidal flooding has a dramatic impact on water quality. It was found that this range increased when $WL_{f,d}$ was decreased, theoretically increasing the organic concentrations in flood waters. The range also increased when $WT_{f,d}$ was increased, effectively providing for an increased level of microbial activity. The maximum computed range of $DOD_{f,d} \approx 4$ mg/L when both $WL_{f,d}$ was decreased and $WT_{f,d}$ was increased to the limits of their historical range.

CONCLUSIONS

Long term real-time gaging of hydrodynamic and water quality parameters, in combination with data-mining techniques, including signal processing and ANN models, can provide an excellent means to understand highly complex and interacting behaviors in an estuary. The location selected for this study, being largely unaffected by anthropogenic oxygen-consuming constituent sources, provided an excellent case for evaluating the effects of non-anthropogenic, non-point source loading using these tools. The sensitivity of DOD, the dissolved-oxygen deficit, to a 1-inch rainfall was estimated to ≈ 0.25 mg/L, and the overall sensitivity to tidally forced organic loading was 2.0 mg/L or more.

LITERATURE CITED

Hinton, G.E., 1992, How neural networks learn from experience, *Scientific American*, September 1992, 145-151.

Press, William H., S.A.Teukolsky, W.T. Vetterling, and B.P. Flannery, 1993, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press.

U.S. Geological Survey, 1981, Technical Memo. 81.11, Reston, VA.