ESTIMATING POINT-SOURCE IMPACTS ON THE BEAUFORT RIVER USING ARTIFICIAL NEURAL NETWORK MODELS

Paul A. Conrads[1], **Edwin A. Roehl Jr.**[2], **and William P. Martello**[3]

Abstract

The Beaufort River 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 32-kilometer section of the Intracoastal Waterway. The river is on the 303(d) list of impaired waters of South Carolina for low dissolved-oxygen concentrations. The Clean Water Act stipulates that a Total Maximum Daily Load must be determined for impaired waters.

Artificial neural network (ANN) models and other data mining techniques were applied in the Beaufort River system to quantify the relationships between the time series of four wastewater point-source discharges and the dissolved-oxygen concentrations recorded at seven real-time gages distributed about the system. The analysis included environmental factors such as water temperature, tides, and rainfall. This paper describes findings of the relationship between one of the point sources and a nearby gage. It was found that the effects of biochemical oxygen demand and ammonia loads on the dissolved-oxygen concentrations vary significantly with water temperature and tidal conditions. Depending on tidal conditions, calculations estimate that at a water temperature of 20° Celsius, a reduction of 100 lbs/day of 5-day biochemical oxygen demand from the point source will increase the dissolved-oxygen concentration at the nearby gage by 0.073 mg/L. The corresponding change in dissolved oxygen relative to 100 lbs/day of NH_3 is 0.16 mg/L.

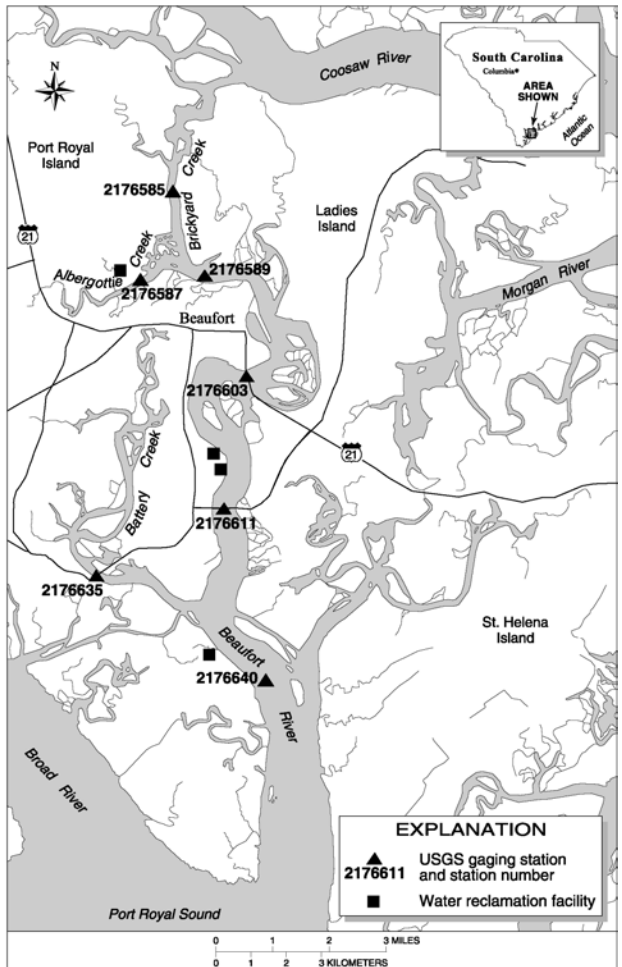

Figure 1: Beaufort River, S.C.

KEY TERMS: estuary, dissolved oxygen, point-source loading, neural network models

---

[1] U.S. Geological Survey, 720 Gracern Road, Suite 129, Columbia, SC, 29210, Phone: (803) 750-6140, Fax: (803) 750-6181; E-Mail: pconrads@usgs.gov

[2] Advanced Data Mining, 116 Sugar Mill Lane, Greer, SC, 29650, Phone: (864) 292 –1607, Fax: (503) 210-8015, E-Mail: earoehl@aol.com

[3] Jordan, Jones, and Goulding, 745 South Milledge Ave. Athens GA, 30605, Phone: (706) 353-2868, Fax: (706) 549-0423, E-Mail: bmartello@jjg.com

INTRODUCTION

A problem of great social importance is determining how to best use natural resources while preserving the quality of surrounding natural systems, such as surface water, groundwater, and air systems. Environmental regulatory agencies attempt to control exploitation using scientific means such as deterministic (physics-based) models that predict how a natural system will behave under scenarios of interest. In practice, however, the statistical accuracy of the models commonly is poor because natural systems tend to be too complex for state-of-the-art deterministic modeling methods. The poor accuracy of the models results in important decisions being made in the absence of unambiguous scientific findings.

The U.S. Geological Survey (USGS) cooperated in a study comparing artificial neural network models (ANNs) to deterministic finite-difference models of the Cooper River, a complex estuarial system (Conrads and Roehl, 1999). Both models were developed from 3 years of measurements of water level (WL), dissolved-oxygen concentration (DO), water temperature (WT), and specific conductivity (SC) that had been recorded by a network of gaging stations. The models predicted the river's hydrodynamic, mass transport, and water-quality behaviors. The comparison showed the ANNs to be significantly more accurate and quickly developed. The ANNs could also be deployed as compact programs that execute without iteration; a prototype control system was developed to investigate regulating wastewater discharges according to the river's assimilative capacity (Roehl and Conrads, 1999).

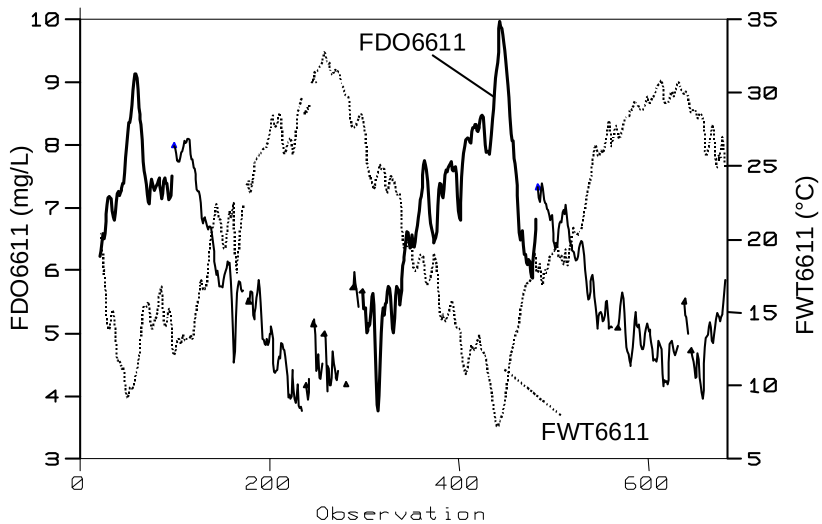

Figure 2: FDO and FWT for Beaufort River at Port Royal (station 02176611).

To gain a better understanding of the Beaufort River and its tributaries, Beaufort-Jasper Water and Sewer Authority, in cooperation with the USGS, established a network of seven gaging stations in the Beaufort River Basin in 1998. The gaging stations use satellite telemetry to transmit the data in "near" real-time (4-hour intervals) to the District Office in Columbia. This network consists of four stations on the Beaufort River, and one station each on Brickyard, Albergotti, and Battery Creeks (fig. 1). Each station records WL, WT, SC, and DO on a 15-minute interval. A precipitation gage is located at the Albergotti Creek gage. Three acoustic velocity meters (AVMs) were deployed in the spring of 2001 at two gages on the Beaufort River and at the Battery Creek gage to measure continuous (15-minute interval) tidal streamflow. In addition to the gaging network maintained by the USGS, the South Carolina Department of Health and Environmental Control gathers monthly water-quality monitoring data for the U.S. Environmental Protection Agency's STORET (STOrage and RETrieval) database. In addition, the National Weather Service has collected meteorological data since the early 1980's.

DESCRIPTION OF STUDY AREA

Beaufort County is a rapidly growing residential, retirement, and resort area along the southeast coast of South Carolina. The water resources of the area are crucial to the economic success of the region. The Beaufort River is a complex estuarine river system that connects the Coosaw River to the north and Port Royal Sound to the south (fig. 1). The river experiences semi-diurnal tides of approximately 3 meters at its confluence with Port Royal Sound. Continuous tidal streamflow measurements from the AVM at station 02176611 show large tidal oscillations of approximately 60,000 cubic feet per second (ft^3/s) and net streamflow towards the Coosaw River of approximately 6,000 ft^3/s. The Beaufort River and its tributaries support a variety of uses including shellfish grounds, fisheries nursery habitats, shipping access to Port Royal, receiving waters for wastewater effluent, and a 32-kilometer section of the Intracoastal Waterway. The river is on the 303(d) list of impaired waters of South Carolina for low dissolved-oxygen concentrations.

APPROACH

The variability of DO in the Beaufort River is a result of many factors including the quality of the water from Port Royal Sound to the south and the Coosa River to the north, the loading of oxygen-consuming constituents from the tidal marshes and other non-point sources, effluent from four permitted point sources and physical characteristics of streamflow, tidal range, salinity, and temperature. To evaluate whether an ANN could be used to determine the influence of point-source discharge loadings on DO, data from a gaging station near a permitted discharge was selected. Of the seven stations on the Beaufort River and its tributaries, the gage at Port Royal (02176611) is only 150 meters south of the largest of four point-sources in the system.

The data were comprised of 30-minute measurements for WL, SC, WT, and DO at station 02176611. The effect on DO of the oxygen-consuming constituents transpires on a time scale of several days. This effect can be difficult to discern when coupled to high frequency forces such as diurnal and semi-diurnal tidal water level, tidal flow, and ambient temperature. Therefore, the hourly time series were filtered using frequency domain filtering (Press and others, 1993) to remove diurnal and semi-diurnal periodic signal components (filtered variables are denoted by a “F” prefix, for example, FDO). 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 ANN models to overfit when correlated variables are used as inputs. Two years of point-source effluent data were obtained from four wastewater reclamation facilities that discharge effluent into the Beaufort River. Data from the facilities consisted of measurements of flow rates, biochemical oxygen demand (5-day) concentration (BOD_5), and ammonia concentration (NH_3). These measurements were typically taken once a week. The outfalls for the two facilities to the north of station 02176611 are located beside one another. The effluent data for these two facilities were combined and treated as one point-source in the analysis.

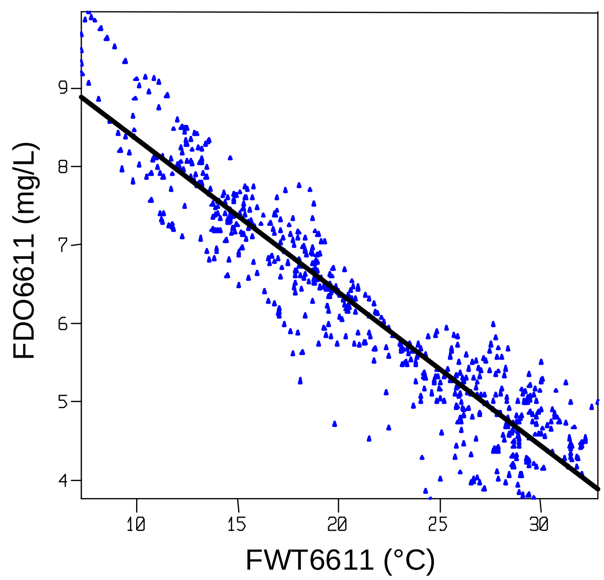

Figure 3: Scatter plot of FDO and FWT and least-squares regression line (R^2=0.88)

Rainfall data were collected from the Albergotti Creek station (02176587) and one of the water reclamation facilities near the Beaufort River at Port Royal gage. The dataset was augmented with calculated variables. The dissolved-oxygen deficit (DOD) was computed using an algorithm that assumes a constant barometric pressure (USGS, 1981). It is assumed that higher values of DOD connote higher levels of microbial activity. In addition, the difference between the high and low tide WL’s for each tidal cycle (XWL) were computed and then filtered as above.

Typically, the majority of the variability in DO is due to WT. Inspection of FDO and FWT shows their inverse relationship (Figure 2). Linear regression produces a coefficient of determination (R^2) of 0.88, indicating that approximately 88 percent of the variability of DO is explained by WT alone (Figure 3), and that only approximately 12 percent of the variability is caused by other factors. WT has two effects. One is that dissolved-oxygen saturation decreases with WT, and the other is that microbial activity that consumes DO also increases with WT (given sufficient DO and nutrients). The use of DOD rather than DO as the response variable of interest emphasizes the microbial activity effect in the DO signal.

The goal of this study was to quantify the effect that point-source discharges of oxygen-consuming constituents have on instream DO. Due to the limited number of data points of the effluent sampling concentrations data as compared to the gaging data (weekly values as compared to 30-minute data), a subset of the dataset was excised that included only the digitally filtered data of DO, WT, WL, and XWL for the day of the effluent. In addition, the 1-day derivatives of the DO and WT were computed and included in the dataset (1-day derivate of the filtered variables are denoted by an E prefix, for example, EDO or ESC). The sensitivity of the response variables, DO and DOD, to the explanatory variables of interest of BOD_5, NH_3, rainfall, and tidal range were determined using ANN models. The type of ANNs used were the multi-layer perceptrons described by Hinton (1992) that were trained using the back-propagation and conjugate gradient algorithms.

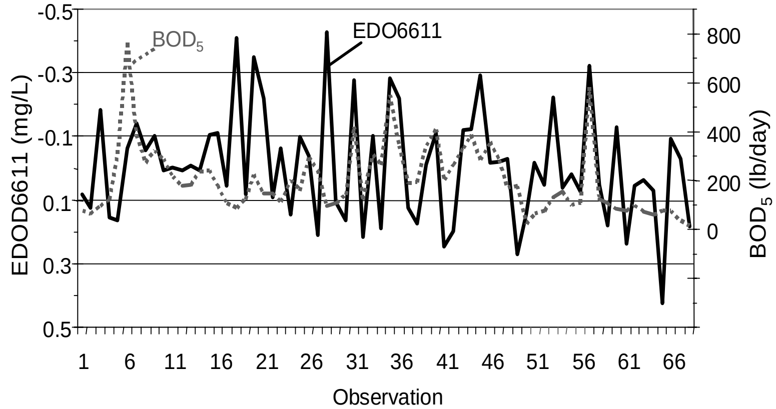

Figure 4: EDO and BOD_5 (at a 1 day time delay). Linear R^2 = 0.13.

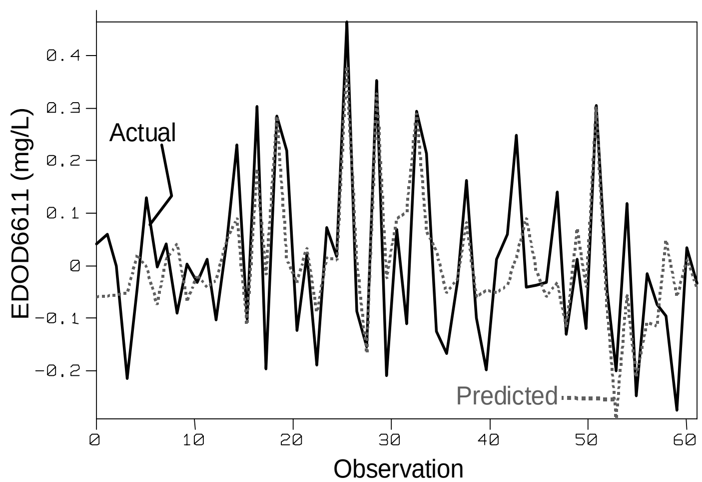

Figure 5: Measured and Predicted EDOD. ANN used BOD_5 as an input at a time delay of 1 day. R^2_{ANN} = 0.57.

RESULTS

Approximately 88 percent of the variability of DO is due to temperature, and all other factors account for only about 12 percent of the variability. Visual inspection of the BOD_5 loading from the water reclamation facility and the daily change in DO concentration at station 02176611 (Figure 4) shows a relation between the two variables (note that the EDO scale has been inverted so decreases in daily DO rise on the scale). The number of coincident peaks in the daily change in DO and BOD5 loading, for example observations 6, 31, 35, 39, and 58, indicate that the BOD_5 loading may account for a significant part of the remaining 12 percent of the variability in DO.

An ANN model of the EDOD, having BOD_5, rainfall, and decorrelated filtered WL, XWL, SC and WT as inputs, was generated to provide a more comprehensive assessment of the relationship between the BOD_5 and the DO. Figure 5 shows that the ANN fits most of the higher peaks in the EDOD. The R^2_{ANN} = 0.57, indicating that approximately 57% of the variability in the EDOD is accounted for by variability in the input variables. The functional form of the ANN's multivariate mapping of inputs to outputs can be understood by examining three dimensional response surfaces like that shown in Figure 6. The figure plots the ANN's prediction of EDOD versus the decorrelated ESC and the BOD_5. WT was set to 20° C. The actual data are projected onto the surface to show how the data are distributed in the (ESC, BOD_5) plane. Note the region where there is no actual data and the ANN's extrapolation through it. Looking from right to left, the EDOD increases with the BOD_5 for all values of decorrelated ESC, however, the sensitivity is much greater at negative values of decorrelated ESC than for positive values. A physical explanation of this observation is that the gage is on the ocean side of the discharge point and that a decreasing SC connotes freshwater flows from inland sources moving effluent towards the DO probe. The historical range of the BOD_5 is 15 to 765 lbs/day. The impact of BOD_5 on EDOD can be quantified as follows. Neglecting the region of no data, the EDOD range ≈ 0.50 - –0.05 = 0.55 mg/L. The sensitivity to 100 lbs/day of BOD_5 loading at 20° C can be estimated as 100 (0.55 mg/L) / (765 – 15 lbs/day) = 0.073 mg/L per 100 lbs/day of BOD_5.

The impact of the NH_3 discharge can be similarly evaluated. Figure 7 shows that predictions made by an ANN model of the EDOD, having NH_3, rainfall, and decorrelated filtered WL, XWL, SC and WT as inputs, generally runs through the middle of the actual data. The R^2_{ANN} = 0.31, indicating that approximately 31% of the variability in the EDOD is accounted for by variability in the input variables. It should be noted that the NH_3 input was delayed relative to the EDOD by 3 days,versus 1 day for the BOD_5, in the first model described above. The delays were chosen by testing different delay configurations and selecting those that produced the highest $R^2{}_{ANN}$'s. Figure 8 again plots the ANN's prediction of EDOD versus the decorrelated ESC and the NH_3, with WT set to 20º C, and the actual data projected onto the response surface. Again, the EDOD increases with the NH_3 for all values of decorrelated ESC; however, unlike the first model, the sensitivity is much greater at positive rather than negative values of decorrelated ESC. A possible physical explanation is complex. The Beaufort River is just a channel connecting the Coosaw River to the north with Port Royal Sound to the south. The channel geometry decreases greatly from south to north. Flow measurements by acoustic velocity meters indicate that the net flow direction is from the Port Royal Sound to the Coosaw River with tidal oscillation an order of magnitude greater than the daily mean streamflow. Reversing tidal streamflows would mix the slow reacting NH_3 southward (connoted by negative decorrelated ESC) and would allow NH_3 to be dispersed and diluted in the tidal exchange with Port Royal Sound. Conversely, flows to the north would trap a large portion of the NH_3 in the Beaufort River where it would reside long enough to have a discernable impact on the DO. Indeed, analyses at gages to the north indicate the presence of a "sag" where DO levels become severely depressed during sustained high point-source loading. Due to the large tidal prism, impacted water in the north is well mixed with water in the south so that the DO sensitivity to NH_3 is registered at the 2176611 gage.

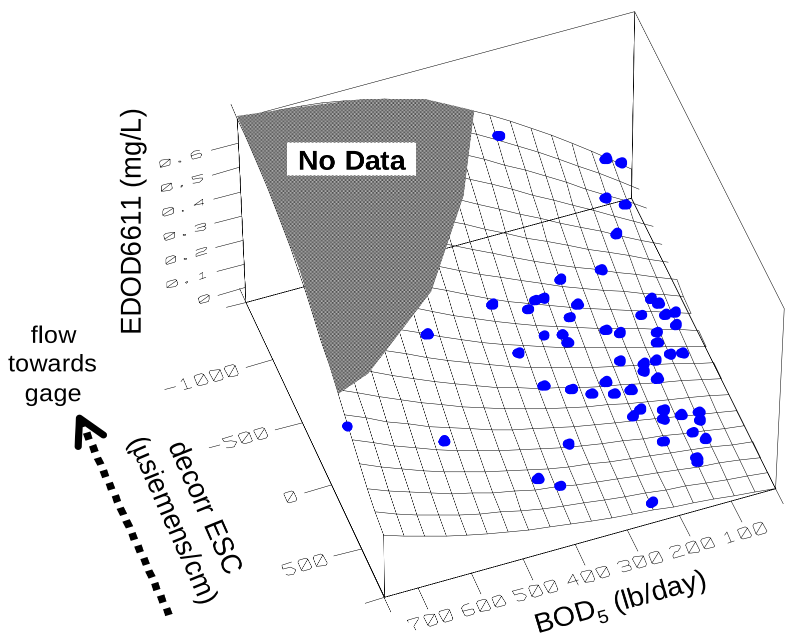

Figure 6: ANN Prediction of EDOD versus decorrelated ESC and BOD_5 at WT = 20C

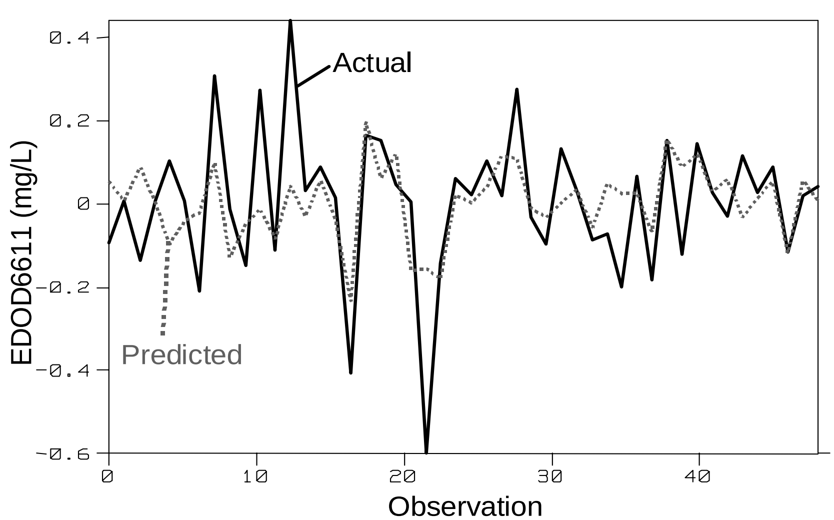

Figure 7: Measured and Predicted EDOD. ANN used NH_3 as an input at a time delay of 3 days. R^2_{ANN} = 0.31.

The historical range of the NH_3 is 15 to 263 lbs/day. Neglecting the region of no data, the EDOD range ≈ 0.10 - –0.3 = 0.40 mg/L. The sensitivity to 100 lbs/day of NH_3 loading at 20° C can be estimated as 100 (0.40 mg/L) / (263 – 15 lbs/day) = 0.16 mg/L per 100 lbs/day of NH_3.

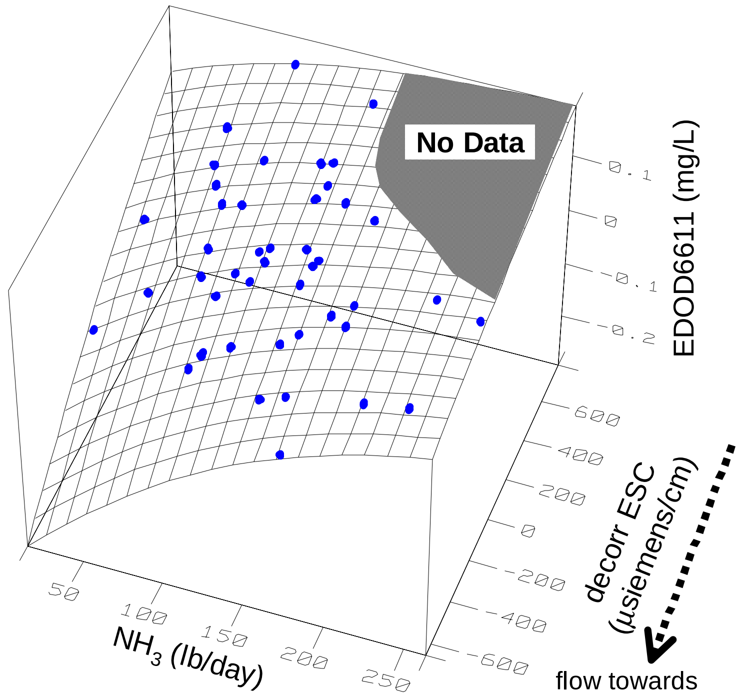

Figure 8: ANN Prediction of EDOD versus decorrelated ESC and NH_3 at WT = 20C.

CONCLUSIONS

In combination, long-term real-time gaging of water-quality parameters, signal processing, and ANN's can provide an excellent means to understand highly complex and interacting behaviors in an estuary. The location selected for this study provided an excellent case for evaluating the effects of point-source effluent loading on DO using these tools. The general findings are that a point-source's signals of BOD_5 and NH_3 can be correlated to DO at a nearby gage through ANN modeling; that sensitivities between these variables vary greatly with changing tidal and ambient conditions; and that a physical interpretation of the system's process physics can be readily made by examining ANN response surfaces. A final note is that South Carolina's water-quality standard for the maximum impact of all point sources on the Beaufort River is only 0.1 mg/L.

REFERENCES

Conrads, P.A., and Roehl, E.A. 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 15-16.

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

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

Roehl, E.A., and Conrads, P.A. 1999, "Real-Time Control for Matching Wastewater Discharges to the Assimilative Capacity of a Complex, Tidally Affected River Basin," South Carolina Environmental Conference, Myrtle Beach, March 15-16

U.S. Geological Survey, 1981, Technical Memorandum 81.11, Reston, Va.