Real Time Prediction of BOD for the CPW Plum Island Treatment Facility

Andrew W. Fairey, Charleston CPW, Charleston, SC
Joseph B. Busby, PhD, PE, OptiQuest Technologies, LLC Greenville, SC

ABSTRACT

Artificial neural networks (ANN's) were applied to develop soft sensors for the Plum Island WWT facility in Charleston, SC. The Plum Island WWT facility consists of two process trains that are both activated sludge processes. Both trains discharge to a common chlorine contact basin and outfall. Models have been developed for the treatment process that allow the BOD to be predicted in real time. The models for each process train have thirteen process inputs and have an R^2 of approximately 0.7. The outputs from each of the models for the treatment trains are used as inputs to the model for the final effluent BOD.

The models were developed using ANN's. ANN's are an effective tool for developing complex non-linear models. They are especially effective for including a large number of process variables in a predictive model.

The benefits to be gained are a reduced frequency of analysis of the wastewater to determine BOD concentrations and improved control of the treatment process since performance is known in "real time". Therefore as flow rates, BOD loading, and the ambient temperature change process parameters such as sludge age can be adjusted to keep the final effluent within permit limits. Also the process models can be queried to determine the best combination of operating conditions needed to achieve the desired results.

INTRODUCTION

The main wastewater treatment facility for the city of Charleston, South Carolina is the Plum Island WWT. The design flow rate for the facility is 36 MGD and it consists of two parallel activated sludge processes. The "A" process train is designed for 24 MGD and the "B" train for 12 MGD. The plant operates consistently within permit limits. The mean value of the effluent BOD over a five-year period was 4.75mg/l with a standard deviation of 2.4 mg/l.

The objective of the project reported in this paper was to develop an empirical model for the BOD in the final effluent using a combination of process operating conditions, influent characteristics, and ambient conditions as model inputs. The purpose of the model is to be able to demonstrate over a period of time that the frequency of BOD analyses throughout the plant can be reduced significantly without compromising reporting integrity. In addition it was desired to know in "real time" the operating effectiveness of the plant. This allows the operators to know when operating conditions are such that desired treatment levels may not be met. The model can then be used in a query configuration to assess the impact of changing such variables as sludge age, aeration rates, or flow ratios between the process trains in order to return treatment to desired operating levels.

MODEL DEVELOPMENT

ANN's are powerful analytical tools that can be used to solve a variety of problems. They provide a different approach to computing that involves developing mathematical structures that have the ability to learn. They are loosely inspired by academic investigations into modeling human learning behavior. ANN's use a set of processing elements (or nodes) analogous to neurons in the brain. These processing elements are interconnected in a network that can then identify patterns in data as it is exposed to the data. In a sense, the network learns from experience just as people do. This distinguishes neural networks from traditional computing programs that simply follow instructions in a fixed sequential order. The models presented in this paper were developed using ANN's since they lend themselves to modeling complex non-linear systems where there is sufficient data that describes the operating ranges of the process to be modeled.

The Plum Island facility has a SCADA system that collects hundreds of process variables. Data for a five-year period from 1996 through 2001 were retrieved and processed for developing the neural network models. The HRT for the plant, at average flow rates, is less than six hours and most of the data used for the model were collected on a daily basis. Some of the data are grab samples and some, such as the effluent BOD, are 24-hour composite samples. Because of the nature of the data and statistical inaccuracies of the BOD analysis process itself, eight day moving averages that were lead back to get in time sync were used for developing the models. The variables that were used for the models are shown in Table I.

The approach for developing the model for the final effluent was to first model the effluent for each of the treatment trains ("A" and "B") and have the outputs of these models feed a model for the final effluent that incorporates the impact of the effluent chlorination. This configuration is shown below in Figure 1.

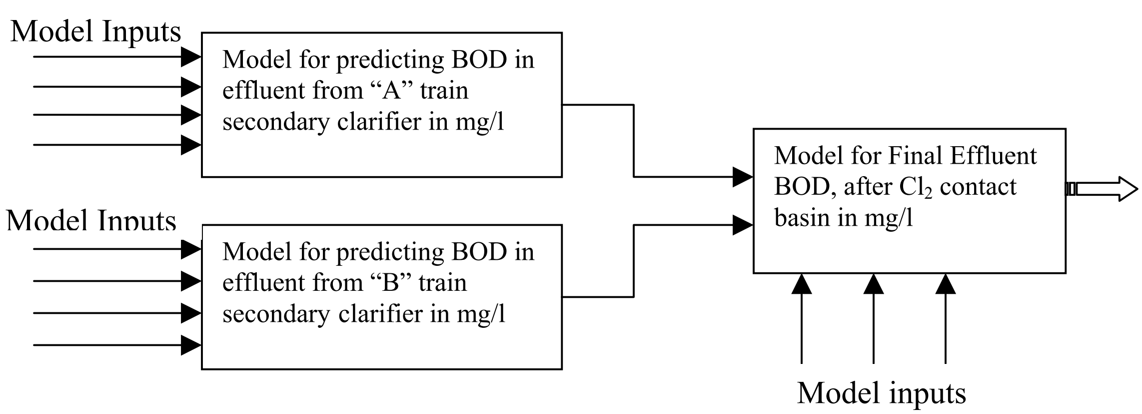


Figure 1 Model Configuration for Final Effluent BOD

The following graphs and tables present the results for developing each of the models. The tables show the inputs to each of the models and variables are listed in order of sensitivity with respect to their impact on the model output. The models are developed by establishing a part of the five-year database to be used for training and a part to use as test data. The curves presented show the training data compared to the output of the model.

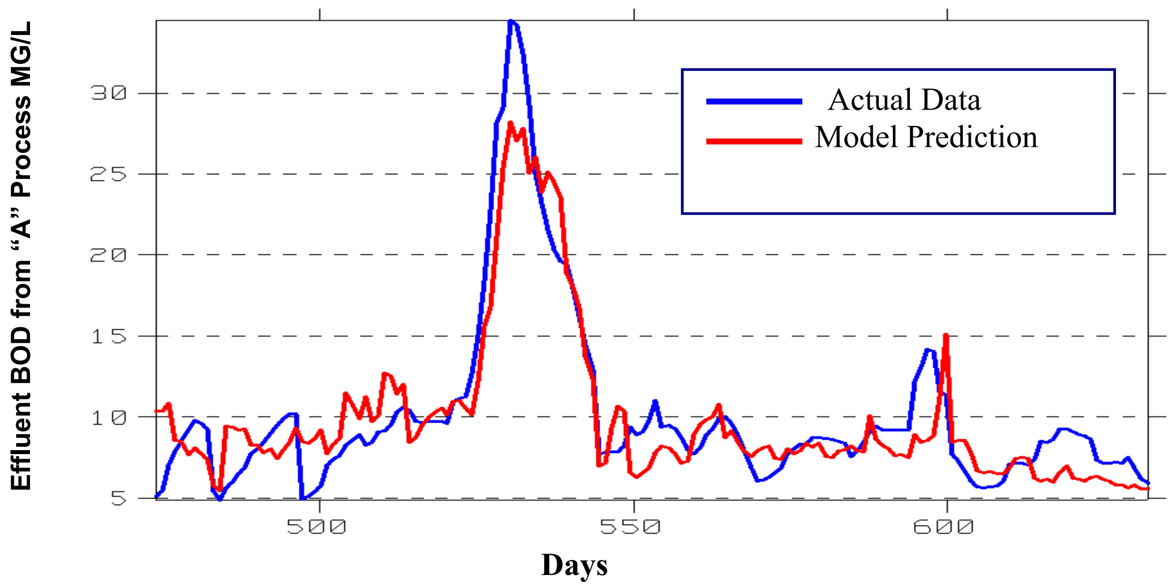


Figure 2 This is a plot of the model results for the "A" Process Train. The blue curve is the actual data for the effluent BOD and the red is the model results. This shows a section of the data that was used to develop the model. The R^2 for this model is 0.69.

Variable Name	**AveAbs Sens**	**Ave Sens**	**Variable Identification**
SCAETSSL	0.20341	+0.20341	Secondary Effluent "A" TSS in mg/l
MCRTA8L	0.14413	-0.14413	Mean Cell Residence Time "A" in days
INFNH3GL	0.12545	+0.12545	Influent NH3 Concentration in mg/l
ABATEML2	0.09648	-0.07810	Aeration Basin "A" Temp in degrees C
FENH3AL	0.08950	+0.08900	Final Effluent NH3 Concentration in mg/l
ABAO2DIF	0.06949	+0.00672	Difference between low and high O_2 mg/l
FLOWAL	0.06328	+0.05800	Flow rate through "A" Train
INSAL95	0.05218	-0.05218	Influent Salinity in ppt
THK1OFSS	0.04176	+0.04176	Thickener No. 1 Overflow SS in mg/l
WASALBAL	0.03896	+0.01894	Waste Flow from "A" Plant lb/day
ABAHIO2L	0.03660	-0.02258	Aeration Basin "A" High O2 in mg/l
FCCL2RES	0.03112	+0.03112	Cl2 Residual (Fecal Sample) mg/l
THK2OFSS	0.00764	+0.00213	Thickener No. 2 Overflow SS in mg/l

Table 1 The above is a lists the variables used as inputs for the model shown in Figure 2.

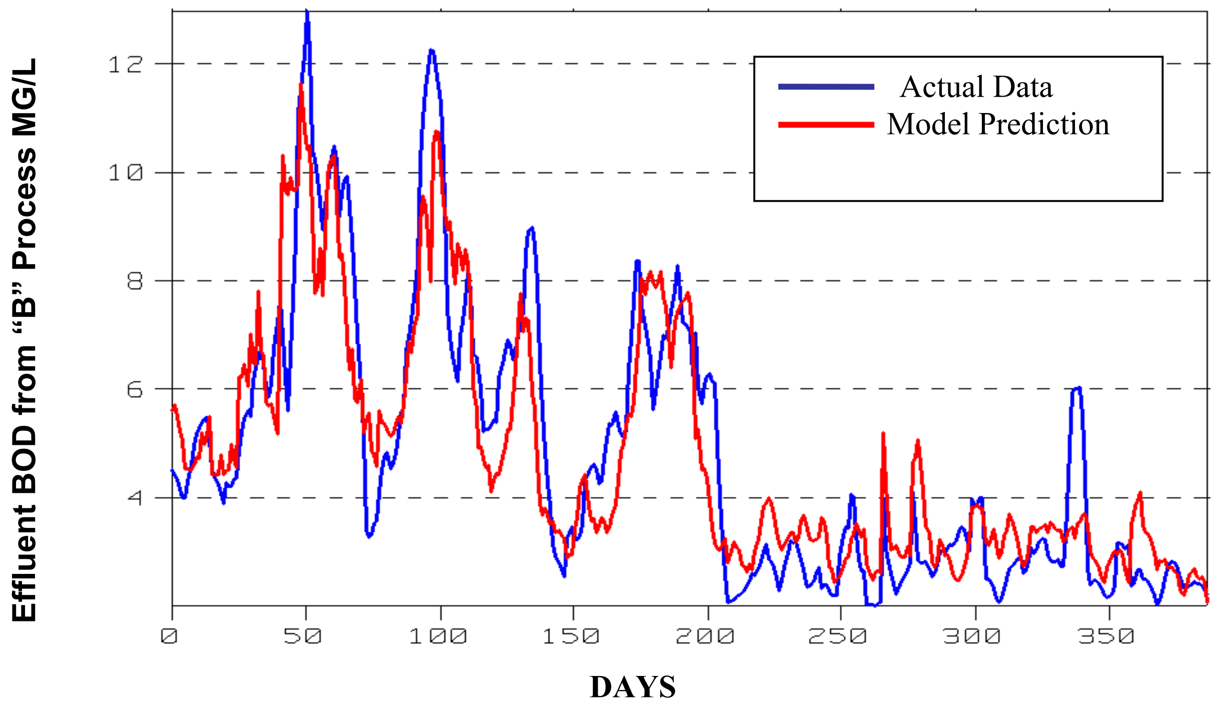


Figure 3 This is a plot of the model results for the "B" Process Train. The blue curve is the actual data for the effluent BOD and the red is the model results. This shows a section of the data that was used to develop the model. The R^2 for this model is 0.74.

Variable Name	AveAbs Sens	Ave Sens	Variable Identification
WASBQAL	0.20793	+0.20793	Waste Flow from "B" Plant MGD
FLOWBL	0.16360	+0.16360	Flow to "B" Plant MGD
SCBETSSD	0.10530	+0.10530	Secondary Effluent "B" TSS mg/l
ABBS30GL	0.10425	-0.04628	Aer. Basin "B" Settleable Sludge ml/l
ABBMLSSL	0.08996	-0.00694	Aer. Basin "B" MLSS mg/l
ABBTEMDL	0.06464	-0.01272	Aer. Basin "B" Temperature degrees C
ABBHIO2L	0.05652	-0.05409	Aer. Basin "B" High Dissolved O_2 mg/l
PREBBO5L	0.05381	+0.05137	BOD in Primary Clarifier Effluent mg/l
INSAL95L	0.04910	-0.04910	Influent Salinity ppt
ABBO2DIF	0.03582	+0.02781	Aer. Basin "B" Dissolved O2 Diff. mg/l
FENH3AL	0.01071	+0.01071	Ammonia in Final Effluent mg/l
FCCL2RES	0.00808	+0.00808	Cl2 Residual (Fecal Sample) mg/l
NH3DELTA	0.00521	+0.00517	Influent – Effluent Ammonia mg/l

Table 2 The above is a list the variables used as inputs for the model shown in Figure 3.

The model for the final effluent BOD actually consists of two separate models. The first model uses the inputs from the models presented in Figures 2 and 3 in addition to the flow rates through each process train. The residual from this model (the difference between the data and the model prediction) is then used to develop a second model for the final effluent. The inputs for the second model include the total suspended solids (TSS) in the final effluent in addition to the effluent water temperature. This approach is taken since there is a high level of correlation between the predictions by the models for the "A" and "B" trains and TSS in the final effluent.

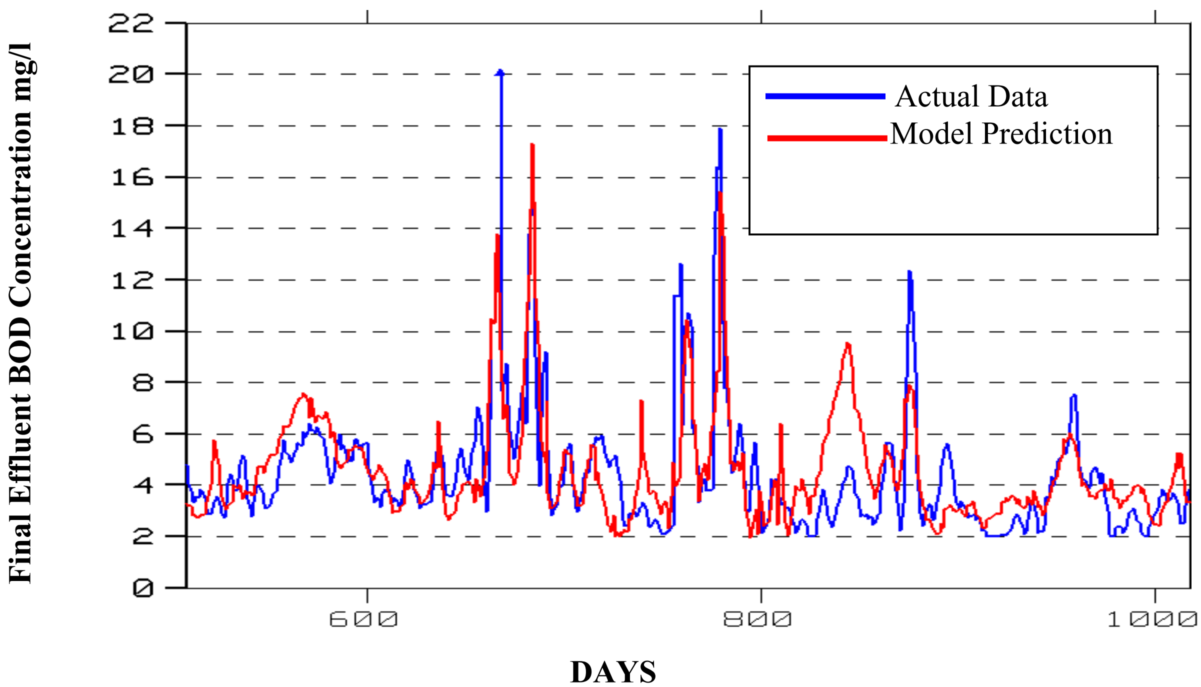


Figure 4 This graph shows the model results for the final effluent BOD. The red curve is the sum of two models and is the prediction and the blue curve is the actual data. The R^2 for the prediction is 0.66.

RESULTS

The model is now being implemented at the Plum Island WWT facility. Data was collected during November and December of 2002 and input to the model for validation. The results show that the correlation coefficient between the model prediction and the actual data is 0.886 that gives an R^2 of 0.78. The following graphs illustrate the model results for this period.

Models such as these are reflective of a given data period. As process conditions and influent characteristics change over time, new models need to be developed in order to capture the impact of these changes. Once a model has been developed, it is a relatively quick process to train and implement new models. As can be seen in the R^2 for the validation data, the models for this application continue to be valid even though the data used to train the models was through the summer of 2001.

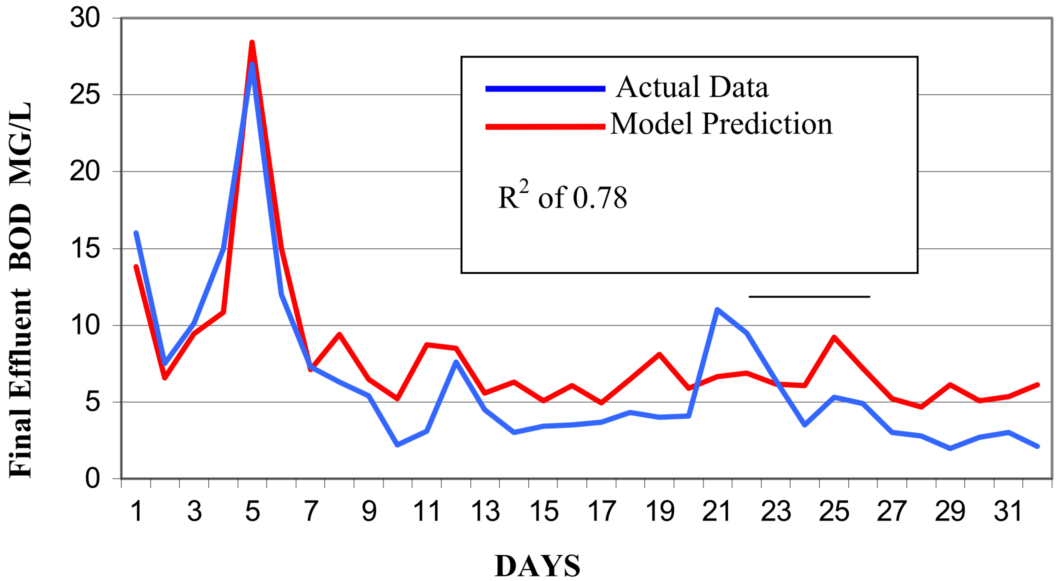


Figure 5 This graph shows the model results compared to the actual data for the BOD in the final effluent. The data was collected during November and December of 2002. The blue curve is the actual data and the red curve is the predicted BOD using the process models.

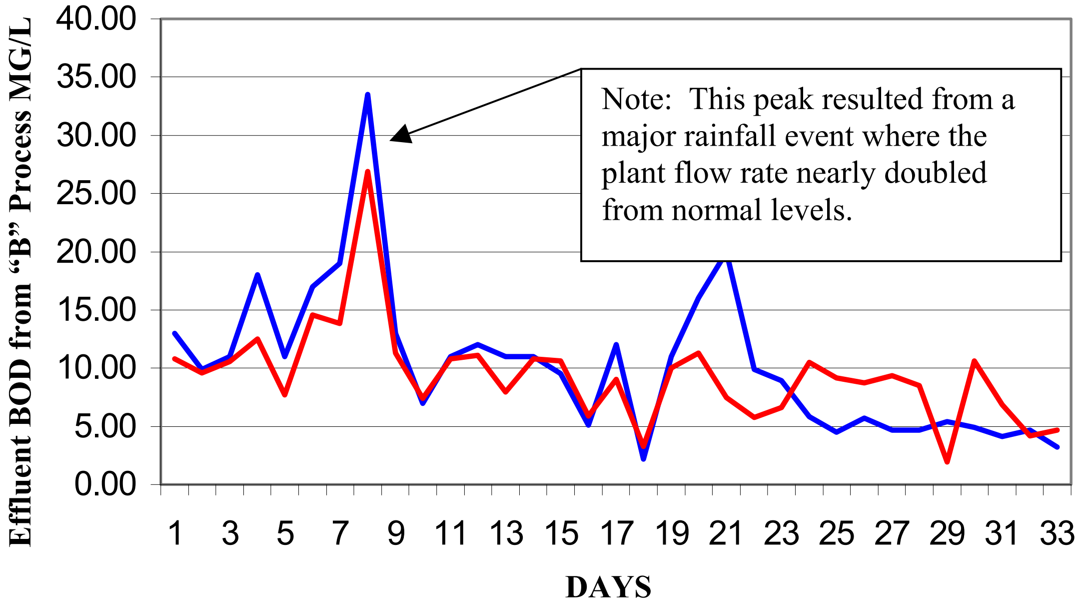
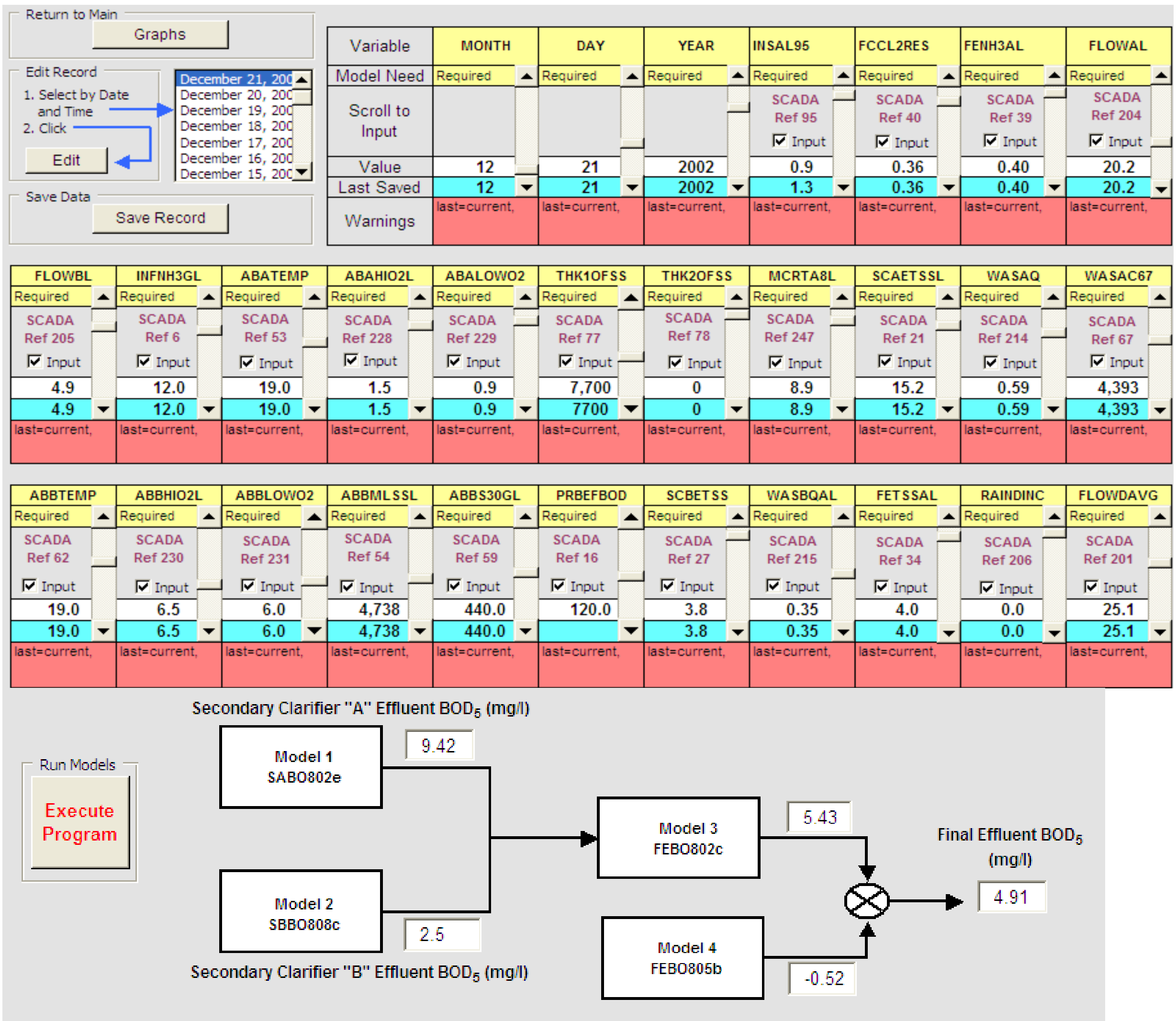


Figure 6 This graph shows the model results compared to the actual data for the BOD in the effluent from the "B" Process Train. The blue curve is the actual data and the red curve is the predicted BOD using the process models. The correlation coefficient between the data and model results is 0.75.

APPLICATION

The following graphic presents the interface that the operator uses to input data to the model. Initially this has been a manual operation with the plan being to tie the model directly to the SCADA system. This approach has been used for other applications similar to this where the prediction software was interfaced directly with a Proficy Data Historian so that data is downloaded on command of the operator.



As mentioned previously, one of the goals of the project was to provide a tool for the plant operators that would allow them to query the model to assess the impact of changing operating parameters. For example, what is the impact of changing the amount of sludge wasted from one of the operating trains. Using the input screen shown above the operator can move the slide bar for the sludge wastage variable, actuate the execute program button and immediately see the impact of this change on that train and on the final effluent.

Using the process models also allows an engineer to gain a better understanding of the complex interactions in the process. This is accomplished using the query techniques or graphical outputs from the models. For example Figure 7 below shows the impact that sludge wastage has on the effluent BOD from the final clarifier of the “A” process. An important consideration in using the model in this fashion is that changing the amount of sludge wasted also has an impact on the TSS in the effluent. This also impacts the BOD. Therefore queries like this have to carefully applied.

Figure 7 This XY plot shows the impact of changing the amount sludge wasted per day on effluent BOD from the final clarifier for the “A” treatment train. All other model variables are kept at their mean value.

FUTURE APPLICATIONS

The benefits of this modeling approach to reduce costs and improve operation are apparent. The application of ANN’s can be further expanded to include additional processes at the WWT facility. One opportunity that may offer considerable cost savings is a model that links the sludge handling processes with the activated sludge process. The model would allow determination of operating conditions that would minimize sludge production while still allowing the activated sludge process to meet desired treatment levels.