Use of Artificial Neural Networks to Predict Water Quality Variables at a Decentralized Wastewater Treatment Facility

Jacob Chicano[1], Kate Newhart[2], Tzahi Cath[2]
[1]University of Connecticut, [2]Colorado School of Mines

Introduction:

As the population across the United States continues to grow, so too does the need to increase our nation's wastewater treatment capacity and lower spending costs. While urban areas are sufficiently serviced by large-scale centralized wastewater treatment plants (WWTPs), decentralized facilities offer rural communities and expanding suburban areas an opportunity to service their wastewater close to the source, at lower cost due to reduced pumping distances. One drawback to these local WWTPs, however, is that their operators are not necessarily experts in the field of wastewater treatment[1]. Biochemical wastewater treatment processes, such as nitrification and denitrification, require contrasting operational conditions, naturally making treatment a complicated task. Careful monitoring and advanced understanding of these processes are necessary to ensure proper treatment. Consequently, decentralized WWTP operators require innovative operational monitoring techniques to ensure the plant is running properly.

Revolutionary advancements in artificial intelligence (AI) and machine learning make implementation of robust, data-driven systems, such as artificial neural networks (ANNs), at WWTPs as a form of system monitoring an appealing idea. ANNs offer potential to simplify data collection, analysis and system prognostics through the prediction of effluent quality, without having to run long, and sometimes hazardous, tests. Whereas sensors are prone to fouling and failure, ANNs are computational tools that are capable of capturing the obscure mathematical relationships between process variables, immune to technical failure so long as they are provided with reliable historical data. ANNs can be trained to learn a mathematical function relating a set of input data to a corresponding set of output data. The network may then be supplied with an input data set for which the outputs are not known, apply this learned function, and predict what the output is. Theoretically, this prediction can be thought of as ideal outputs under normal operating conditions (NOCs), and can be compared to sensor readings throughout the system. A large disparity between ideal and actual output values could alert operators to a potential fault in the system.

Objectives:

Task 1: Monitoring how much nitrate is released in wastewater effluent is a critical environmental issue. At all times, any monitoring scheme should be able to accurately feed back effluent nitrate concentration to ensure the plant is operating within federal law. Use an ANN to predict effluent nitrate concentration from easy-to-measure influent ammonium readings

Task 2: To substitute sensor readings with ANN predictions, the error between observed and predicted outputs should be minimized. One method of reducing this error is to find the ANN architecture that yields the lowest prediction error. Optimize the structure of the ANN so that forecasting error is minimized

Approach:

The complex relationships between the wastewater treatment process steps imply a highly nonlinear mathematical behavior of the system. As such, an ANN that can capture these nonlinear relationships between data sets is recommended, such as the nonlinear auto-regressive with exogenous inputs (NARX) network[2]. This network structure analyzes error between both past and current inputs and outputs, adjusting the weights between process variables accordingly.

While the network itself learns how to map a set of input patterns to a set of outputs, a trustworthy set of data is necessary to reliably estimate the mapping function. These data were collected at Mines Park in Golden, CO, from the sequencing batch-membrane bioreactor (SB-MBR) testbed, a demonstration-scale WWTP that models treatment at a decentralized WWTP. Between the days of April 15, 2018 and April 27, 2018, data were collected at one-minute intervals for a total of 16000 data points of both influent ammonium concentration and effluent nitrate concentration.

The coding language of choice to begin designing the ANN was MATLAB, due to its intuitive neural network toolbox that makes starting off easy for beginners. A time-series NARX network was developed in MATLAB, initially structured using an input layer with two neurons (for the reading of the two NH_4^+ sensors' data), one hidden layer with ten neurons, and an output layer with one neuron (NO_3^- effluent concentration). Time delay was initially set to two, meaning that estimation of the mapping function included the two most recent estimates of the output to minimize error. Network training was accomplished using the Levenberg-Marquardt backpropagation algorithm, with a tangent sigmoid activation function in the hidden layer, and linear activation in the output layer. 80% of all data points supplied to the network were used to train the network, so the network could learn as much about the variation in the data as possible. The remaining 20% was split equally between validation and testing to minimize prediction error and test the generalizability of the network when it is supplied with new information.

Following the completion of training, validation and testing, the finalized network was used to forecast up to two days' worth of effluent concentrations in one-minute intervals. The results were then averaged using an exponential weight moving average (EWMA) to relate the network's prediction accuracy over a longer period of time, and still allow the newly measured value to have a significant impact on the updated average.

Results and Discussion:

Minimization of prediction error is necessary to achieve the best forecasting potential possible from the network. An optimization loop was thus created to test the effect of increasing time delay. This loop created a network with time delay values ranging from 2 to 80, and showed that increasing time delay increased multi-step prediction performance until time delay was equal to 60. Increased accuracy is likely attributed to more data being used to approximate the mapping function. A similar loop was run to find the optimal number of hidden nodes in the hidden layer, with the number of nodes ranging from two to ten with twenty replicates of each. On average, six hidden nodes yielded the best multi-step performance.

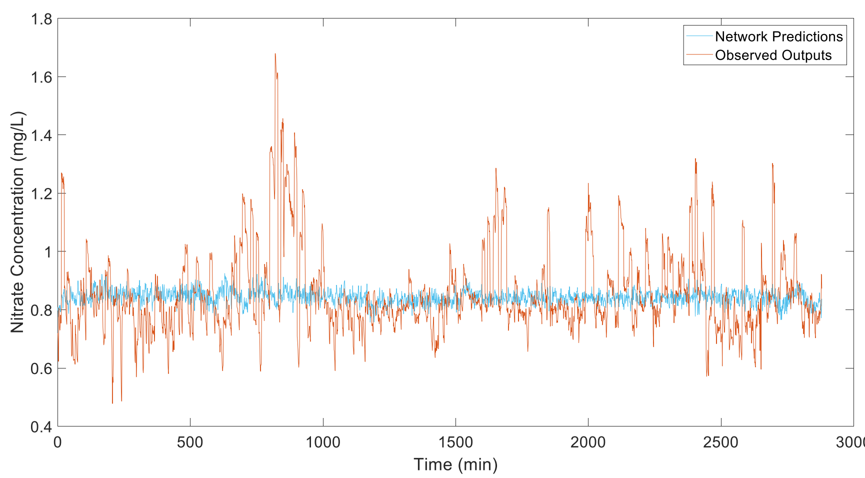

Figure 1. Evolution of ANN prediction accuracy with 2880 minutes (2 days) of SB-MBR effluent nitrate concentration at the Mines Park testbed. MSE = 0.0704

Figure 1 compares the optimized network's NO_3^- prediction to a set of known targets associated with the same NH_4^+ inputs. Visual inspection of the graph shows that the network has a difficult time predicting the magnitude of shifts in nitrate data, but captures a general average of the outputs.

Figure 2 depicts the EWMA taken using the data from Figure 1. Using an average that spans 240 minutes, prediction error, measured as mean square error, decreases an order of magnitude to 0.0037.

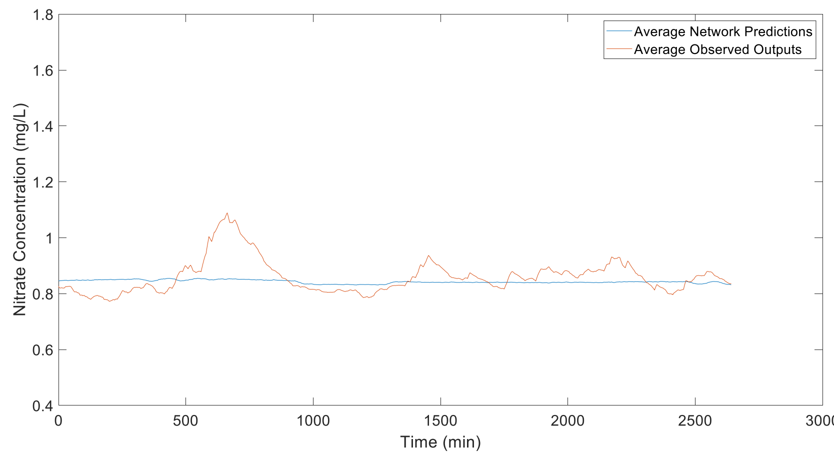

Figure 2. Exponential weighted moving average of both target and predicted effluent nitrate concentrations. Each point is the average of the current data point and 239 previous data points (with time = 0 representative of the average of the first 240 data points). MSE = 0.0037

Minute-by-minute predictions of effluent characteristics are seemingly difficult to obtain through the usage of ANNs. Poor prediction performance is likely attributed to the complex interactions that occur during wastewater treatment. While ammonium influent concentrations might be a generally good predictor of nitrate effluent concentration, a handful of other environmental, biological, and mechanical factors influence how much nitrate is produced and how much leaves the system at once. Averaging can be seen as a tool that can be used to account for and cancel out these influencing factors.

Conclusion:

ANN technology has difficulty developing accurate minute-by-minute predictions of effluent nitrate concentration given only influent ammonium concentration data, potentially due to some biochemical or mechanical influence that the network could not account for. However, a weighted moving average of the predicted and observed data shows that averaging can negate the effects of these influences.

It is possible that an ANN can be used to predict average effluent quality. However, it is prudent to note that unless these results can be improved by means of finding a way to precisely approximate effluent quality at each minute, then ANN monitoring technology provides no practical advantage over modern sensor technology. Further research should investigate incorporation of other process variables into an ANN, and their effect on forecasting accuracy. Additionally, an ANN's ability to predict a wastewater treatment system NOCs should be tested to ascertain if potential system faults can be identified before a system alarm is triggered.

References:

[1]Work for water. Water and Wastewater Treatment Plant Operators. Retrieved August 2, 2018, from http://www.workforwater.org/resourceforwaterprofessionals/page_int.aspx?id=271

[2]Lee, M., Hong, S., Choi, H., Kim, J., Lee, D., Park, J. (2008). "Real-Time Remote Monitoring of Small-Scaled Biological Wastewater Treatment Plants by a Multivariate Statistical Process Control and Neural Network-Based Software Sensor". *Process Biochemistry*. 43. 1107-1113.