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

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

Background:

- Decentralized wastewater treatment plants offer the potential to *increase the United States' water treatment capacity* while simultaneously *decreasing conveyance costs*.
- *Operators at smaller, decentralized facilities are typically not wastewater treatment engineers*. To ensure wastewater is being treated properly, operators require a greater degree of operational supervision and guidance. To ensure operators are well-informed about current and future wastewater treatment states, accurate and timely predictions of water quality parameters must be made.
- In this work, an artificial neural network (ANN) is used to predict effluent water quality of a decentralized wastewater treatment facility. The ANN is *capable of "reading" wastewater treatment influent quality and maps what the nitrate concentration should be under normal operating conditions (NOC) without the need for first principals modeling.*

Barriers to Reinvention:

- Reliance on sensor technology that is **variable and difficult to maintain** makes collection of reliable data difficult for water and wastewater treatment facilities
- **Model-based prediction techniques** are notoriously computationally and time expensive, prone to oversimplification, and lack robustness that would allow it to capture the dynamic nature of wastewater treatment and influent water quality
- The **structuring and training of an ANN is largely a trial-and-error process**, thus development of the network is both time consuming and typically suboptimal
- No single answer exists to the question of how much data is necessary to avoid **underfitting** and **overfitting**

Objectives:

Objective 1: Determine the optimal structure of an ANN for which influent ammonium data is used to predict effluent nitrate concentration

Objective 2: Test multi-step prediction accuracy of an ANN used to predict effluent nitrate concentration while system is under NOC

Approach:

Data collection:
Ammonium sensor: Vernier Ammonium Ion-Selective Electrode
Nitrate sensor: YSI IQ SensorNet NitraVis®

Influent ammonium and effluent nitrate data collected in one-minute intervals for a total of 11 days from SB-MBR at Mines Park testbed. All data sets gathered were normalized to zero mean and unit variance.

Approach (continued):

Network setup:

Optimal neural network set-up is a trial-and-error process of altering a variety of variables. Thus, methodology explained below depicts the final results, not the process.

Platform for network development: MATLAB
Network architecture: NARX network
Number of hidden layers/nodes: 1 layer / 6 nodes
Transfer functions: tangent sigmoid (hidden layer), linear (output layer)
Time delay: 60
% data used training, validation, testing: 80,10,10

Network learning:

Figure 1 depicts the general neural network learning process train. Following this step, the network is ready to be used to predict effluent nitrate concentration.

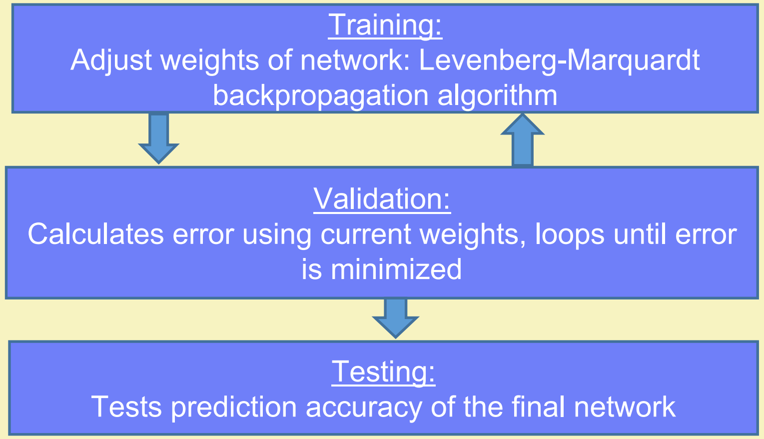

Figure 1. Neural network general learning algorithm flowchart

Results:

Optimization:

The optimal structure of the ANN is described in the "Network setup" section of "Approach". Multiple loops were run while varying the network properties listed. The network with the structure described was chosen because it, on average, consistently yielded the lowest multi-step prediction error.

Prediction Accuracy:

Figure 2 shows that the network has *difficultly learning how magnitude of effluent NO_3^- is affected by perturbations* in influent NH_4^+ concentration, but *seems to capture a long-term average of total NO_3^- outputted*.

An exponential weighted moving average (EWMA, Figure 3). was created. Averaging of data results in mean error decreasing an order of magnitude, possibly due to negation of the effects of short-term mechanical or biochemical anomalies.

Future Work:

- Incorporation of additional process variables into network to improve prediction accuracy
- Investigate use of ANN soft-sensor technology as fault identification system

Conclusion:

The ANN developed in this work requires that minute-by-minute prediction accuracy be improved. Averaging of data yields a more accurate, but unrealistic, prediction. Unless prediction accuracy can be improved, there is little reason to use ANNs in place of classical sensor technology.

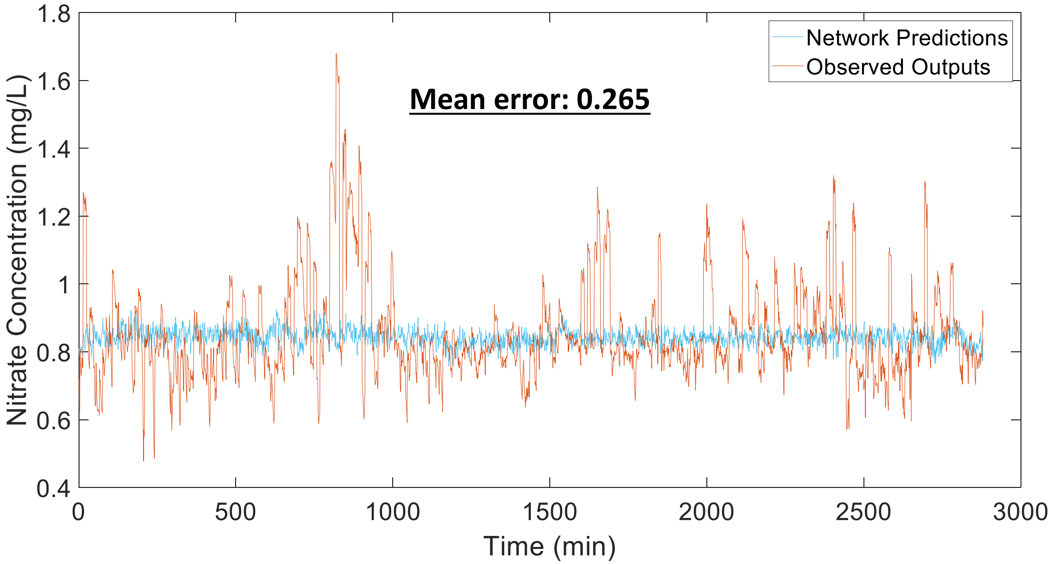

Figure 2. Neural network prediction of 2880 minutes (2 days) of SB-MBR effluent nitrate concentration at the MP testbed

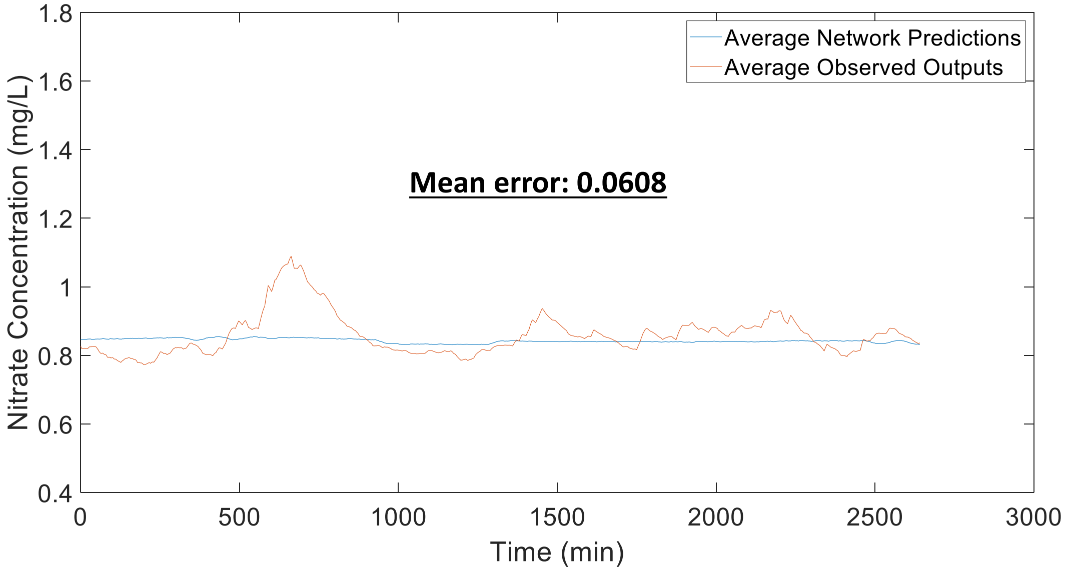

Figure 3. Exponential weighted moving average of both target and predicted effluent nitrate concentrations. Each point is the average of the current data point and 239 previous data points (with time = 0 representative of the average of the first 240 data points). While error decreases significantly, a constant average effluent quality prediction is unrealistic due to the transient nature of wastewater