PREDICTION OF CONTAMINANTS IN WATER DISTRIBUTION SYSTEMS USING ARTIFICIAL NEURAL NETWORKS

P. Romero-Gomez, R. G. Austin, and C. Y. Choi

Department of Agricultural and Biosystems Engineering
The University of Arizona, Tucson, Arizona
(Corresponding Author: C. Y. Choi, cchoi@arizona.edu)

ABSTRACT

Transport phenomena of contaminants in water distribution systems are examined using Artificial Neural Networks (ANNs). First, a contaminant is introduced in the reservoir, and its concentration throughout the system is predicted as a function of water demands at regional levels. For this purpose, a back-propagation ANN is trained, tested and validated with data obtained from both experiments in an actual water system and EPANET. Next, the most likely location of the chemical's intrusion point is tracked based on readings collected from several sensors placed in the water network. In order to minimize intrinsic errors, several parameters of the architecture and functions of these ANNs were thoroughly tested: a number of processing units in hidden layers, transfer functions, and learning rules, to name a few. The present study provides relevant information for alertness and preparedness in potential intentional and accidental contamination events.

Introduction

The fate of chemical components in drinking water networks can be described using physical models based on the underlying principles that govern the transport phenomena and/or with empirical methodologies based on observations and experimentation. For these physical models to perform adequately, it is essential to have a thorough understanding of all the variables involved in this process and to employ intensive computational tools for finding solutions. However, the use of physical models is often limited in field applications because a real water network is very complex and transient that the computing power needed becomes prohibitive. In addition, we do not yet fully understand the underlying functional relationships with regard to the entire water network. As an alternative method, Artificial Neural Networks (ANN) have been widely applied in different fields of engineering and science since their introduction in the 1960s. ANN is a modeling approach based on how biological neural systems are believed to work, and it is increasingly used because it offers a higher capacity to manage the intrinsic variability and uncertainty of predictions. Due to rising concerns regarding water security, utility managers need tools for efficient monitoring and decision-making, support of which must come from Water Quality (WQ) model outcomes.

Because water quality depends primarily on quality at its sources and disinfection treatments, numerous applications of ANNs are available for microbial quality of groundwater and other sources. Brion and Lingireddy (2003) collected water samples at the inlet of a drinking water plant on the Delaware River, USA, to determine *Crystosporidium* and *Giardia* concentrations. These were correlated to input measurements (*Clostridia perfringens*, *E. coli*, pH, coliforms, precipitation, turbidity and river flow) with ANNs. Results showed that ANNs can predict peak concentrations based on input parameters with high accuracy (88 %). In the same fashion, long-term and short-term forecasts of chlorine concentrations and hydraulic head of an aquifer with ANN were performed by Gümrah et al. (2000). Input data for the initial two years were generated with a physical model, and included time, six concentrations (from six previous time steps), and hydraulic heads from the studied site and a neighboring well. The average

absolute error for short-term prediction (6 months) was 3.5 %, whereas long-term prediction (8 year) absolute error averaged 18.5 %.

Milot et al. (2002) predicted the presence of trihalomethanes (THM, a by-product of chlorination) as a function of Dissolved Oxygen Demand (DOC), reaction time, pH, temperature and chlorine dose. Three prediction tools were compared in this study: ANNs, a logistic regression model, and multivariable regression. ANNs proved both effective and computationally efficient; given the size and nature of the database used, it was concluded that ANNs may perform similarly when analyzing other chemical components. Baxter et al. (2001) successfully applied ANNs to model the color in a river and evaluate the effect of water softening with multiple inputs: river flow rate, pH, water temperature, alum and lime dose, to name a few. Demand forecasting of the entire municipal system was also performed using summer index, weekday and weekend index, and peak water demand, among others. Lingireddy and Ormsbee (1998) conducted the calibration of water distribution systems with a back-propagation ANN. The input data were pipe roughness associated to a group of pipes and water demands, and the output was the observed pressure at several locations of the water system. It is remarkable that while numerous applications of ANNs on source water quality are available, few studies have examined water quality in a pressurized distribution system.

The purpose of this work is to present three applications of ANNs to predict water quality in distribution systems. Because water quality encompasses microbial and chemical contents in drinking water, we focus our applications primarily on modeling salt concentrations throughout the network. Furthermore, this study focuses on mass transport processes that occur from the reservoir to the delivery points, rather than the water sources. First, contaminant concentrations throughout an exemplary network are correlated to the local (grouped nodes) demands. The second application identifies the contamination point from a group of potential intrusion locations based on readings from four sensors placed in a highly interconnected network. Third, the axial mixing in a single pipe is examined in order to determine the parameters of the distribution curve downstream. ANNs seem desirable for analyzing these problems because solutions need to be developed without solving the micro-scale interactions that actually occur and are poorly understood. Also, the use of ANNs can be evaluated based on physical model outcomes and experimental/field data can be further integrated in order to enhance their performance.

Overview of Artificial Neural Networks

Artificial Neural Networks are the topic of numerous studies focusing on fundamental research areas as well as applications to different fields. Detailed descriptions of the approach on the subjects of civil engineering and water supply engineering are provided by Flood and Kartman (1994) and Neelakantan (2005). The general structure of a feed-forward Artificial Neural Network is depicted in Figure 1. It is composed of Processing Units (PE) organized in layers, and each PE is connected to all PEs in the previous and next layer. Such connections are not equal, but they have different *strengths* or *weights* that define the ultimate features of the network. The “feed-forward” characteristic of the entire network indicates that the information is received at the input layer, flows through the succeeding layers, and ultimately the processed outcomes are produced in the output layer. Individually, each PE contains a transfer function that produces an output based on the input received from the previous layer. The input is composed of all the connection weights multiplied ($w_{x,n}$) by their respective input signal plus the signal coming from a bias unit. Therefore, the output can be expressed as follows:

$$O_j^n = f\left(\sum_{i=1}^{I}\left(w_{ij}\cdot O_i^{n-1}\right)+b_n\right) \tag{1}$$

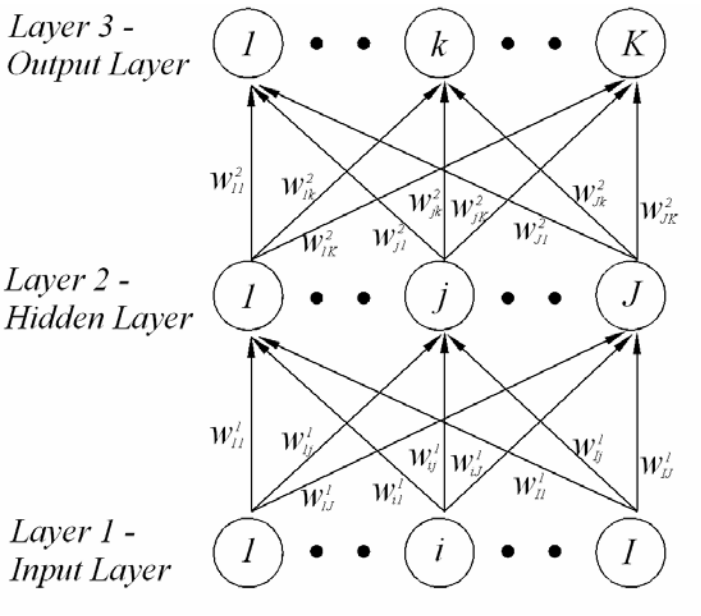

Figure 1. Feed-forward Artificial Neural Network

where O_j^n is output of the *j-th* PE that belongs to layer *n*, w_{ij} is the connection weight between the *j-th* PE in layer *n* and the *i-th* PE in the previous layer (*n-1*), O_i^{n-1} is the output of the *i-th* PE in the previous layer (*n-1*), and b_n is the signal from the bias unit (typically equal to 1). The type of transfer function *f(x)* is a major feature to be defined because it basically establishes how the information is processed at each PE and through the network. The logistic sigmoid function (Eq. 2) is commonly used as the transfer function and produces an output range from 0 to 1, making it a suitable choice when managing binary outputs. However, other transfer functions can also be used according to the modeled process and its input data. Once all the outputs from the layer *n* have been calculated, the series of outputs for the succeeding layer can be computed, and so forth. As the PEs pertaining to the input layer do not have the transfer function imbedded, they therefore produce no output, instead passing the input information onto the first hidden layer.

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

An ANN must be trained after its architecture, transfer functions, and other features have been defined. The training process consists in systematically presenting a set of inputs and their corresponding outputs to the network in order to establish the connection weights with which the network will further perform predictions. Feed-forward networks are typically trained using the back-propagation algorithm, which involves several steps for minimizing the difference between the calculated and presented output vectors. For this purpose, the tolerance, the weight-updating rule, and the initial connection weights are set. Next, an input vector is presented to the network and its resulting output vector is calculated. The calculated and correct output vectors are compared in order to find the error magnitude. This error magnitude is later used to calculate the degree of weight change and its direction (either increasing or decreasing) so that the error is reduced at the next iteration. The weights are updated and the process continues iteratively until the error is less than the tolerance or until the number of iterations is reached. The training process is the most time-consuming step in ANN modeling because it assumes that enough information of the phenomena or process to be modeled is available. Also, there is no general

guideline on the number of data points that are sufficient for the training process; experience is often the only tool to rely upon. However, ANNs' ability to be flexibly applied to a wide range of problems has made them a successful modeling tool.

Methodology for ANN application to WQ prediction

As mentioned previously, modeling with Artificial Neural Networks requires training and testing datasets. These datasets may be generated by experimentation or existing physical models. In this work, the EPANET hydraulic and WQ models were used to generate the required datasets. EPANET is a freeware program developed by the Water Supply and Water Resources Division of the U.S. EPA's National Risk Management Research Laboratory (Rossman, 2000), and is widely used for research on and management of drinking water systems. The ANNs employed in both applications presented here sought to reduce the amount of input information necessary when solving two fundamental problems regarding water quality in distribution networks: concentration prediction and contamination source identification.

Modeling of Contaminant Concentrations

A water distribution system usually has more than one water source powered by a pump, elevation pressure or a combination of both. It may be branched (Figure 2) or highly interconnected (Figure 4), or contain sections with both types of configurations. The water distribution system used in this study was set up as depicted in Figure 1 in order to work with a system easier to manipulate during the generation of training and testing datasets. The system was assumed to operate under steady-state conditions; no elevation change was defined for any point and the chemical injection point was simulated by implementing two pumps with the same curve. Two flows were of different chemical concentrations and converged at a point before water spread throughout the network. This was the only chemical intrusion point defined in the system. EPANET manages specific as well as generic chemicals whose diffusivity and reaction properties are set relative to chlorine in water at 20 °C. For the chemical considered here, the ratio relative to chlorine was set equal to 1. The dimensions of the pipe diameters and the total length in the distribution system were 8 in and 695 ft, 4 in and 2215 ft and 2 in and 2450 ft. Hazen-Williams formula are used for head loss calculations with a roughness coefficient of 100. Both pumps were defined with the same curve (one-point pump curve, 300 GPM – 30 ft). Bulk and wall reaction orders were equal to 1 and global bulk and wall reaction coefficients (K_b and K_w, respectively) were equal to -24.

The system was divided into six regions and each region was composed of eight demand points. Therefore, boundaries were delimited in order to define each region and its corresponding nodes, as depicted in Figure 1. Because managing information at the level of each demand point was impractical and time-consuming, the information was managed at a regional level instead. In actuality, water demand is a variable periodically monitored by water companies. Thus, the relationship of the regional water demand to the regional average chemical concentration was modeled with an ANN. Numerous scenarios that led to the training and testing datasets were implemented into EPANET. Regional water demands and their resulting contaminant concentrations were obtained from the EPANET hydraulic and WQ models, respectively. Chemicals injected into any water network undergo concentration decreases with respect to both distance and time. This degradation is due to chemical reactions along with a flawed mixing of the chemical in water. Consequently the activated Water Quality mode was "Chemical", which indicated that there were chemical reactions occurring within the fluid (bulk reaction), and in the interaction of the fluid and wall (wall reaction).The bulk and wall reactions created the spatial variation of contaminant concentration throughout the network.

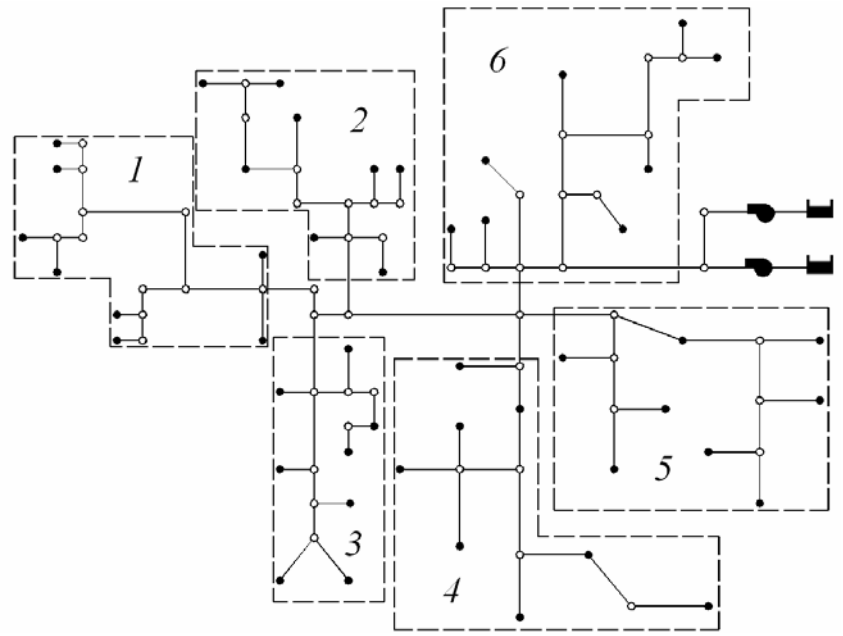

Figure 2. Water distribution system used for generation of training and test datasets. Region No. 1 through No. 6 are presented.

Water networks simulated with EPANET generate network files that contain a listing of junctions (nodes) and links (pipes) properties. Each scenario's corresponding network file was modified to set the water demands at each node. Random water demands for the regions were generated by utilizing a random number generator tool; these demands ranged from 0 to 80 GPM (regional water demand) and therefore tested the flexibility of the neural network to process any range and combination of water demands. The eight nodes pertaining to one region had the same water demand. Next, the scenario was run and the average chemical concentration at each region was calculated, resulting in a pattern composed of six values for each regional demand (input) and six values for the expected average chemical concentration (output). For the training set, a hundred patterns were obtained, where as the testing dataset contained twenty-five patterns.

A multi-layer feed-forward ANN was implemented with NeuralWorks Professional II, and was trained with the back-propagation learning algorithm. Initially, the network consisted of the input layer, the hidden layer, and one output layer with six units each. The Generalized Delta Rule (GDR) was used to update the connection weights, and the hyperbolic tangent (*tanh*) function was applied as the transfer function in each PE. The hyperbolic tangent function belongs to the sigmoid-type family of functions. Because bipolar input units were set, the input values were scaled to the range of -1 to 1, which corresponded to the minimum and maximum input water demands, respectively. The epoch, which is the number of training patterns presented to the network between weight updates, was set equal to 100. After the aforementioned network configuration was tested, two parameters were modified in order to reduce the error between the calculated and actual output values. First, the number of iterations was set from 10^4 up to $5x10^5$. Second, the hidden layer size was changed to five, seven, eight and ten PEs. The training dataset was used to record the Root-Mean Square (RMS), which was the indicator of the network performance. Bias units connected to all the units in the hidden and output layer were also implemented.

Contaminant Source Identification

Several techniques have been used to identify the intrusion point of contaminants in water distribution networks, mainly the simulation-optimization approaches. The main purpose of this task is to locate the point of introduction for the contaminant based on the information collected at the water quality monitoring units, should a deliberate or accidental contamination event occur. This would allow the cessation of the event before further chemical spread occurs and to propose control and decontamination strategies. Source identification is regarded as an inverse problem because the driving cause must be traced back based on information obtained from resulting effects. As a consequence, this ill-posed task depends strongly on the number and spatial locations of the monitoring units and may have endless solutions, especially when dealing with actual water distribution networks. Artificial Neural Networks, however, can elucidate the relationship between observed chemical concentrations in the water network (field data) and the origins of such chemicals (intrusion point), for small water networks. In this application, the chemical concentration readings at four locations were used as inputs for an Artificial Neural Networks. The outputs noted the most probable location out of four potential locations, as seen in Figure 3.

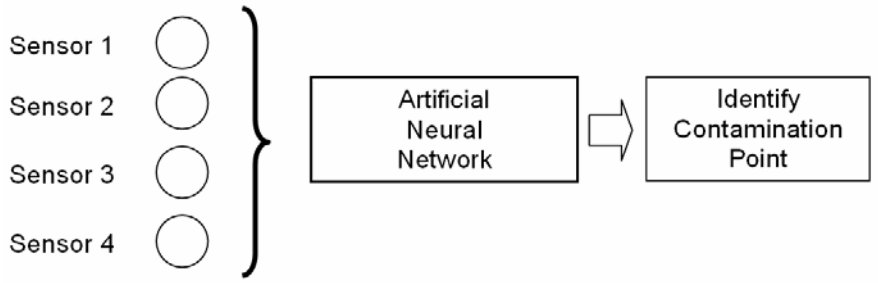

Figure 3. Diagram for application of ANNs to identify a contamination point

The 10x10-node water network depicted in Figure 4 was created for EPANET simulations. The pipe intervals between nodes were of the same length (100 ft, 2 in diameter, roughness coefficient for Hagen-Williams head loss formula equal to 100). Water was supplied by three pumps with one-point flow head curve (80 GPM – 50 ft). There was no elevation change among the nodes. Bulk and wall reactions were not considered; thus dispersion was the only mass transport mode affecting water quality. In order to simulate the sensor locations, four nodes were randomly placed throughout the network, as seen in Figure 4 (blue circles). Also, four potential contamination locations were randomly set (red circles), only one of which had a negative demand in order to simulate the contaminant intrusion for each scenario (10 GPM @ 1000 mg/L). A random number generator tool was used to create the 100 water demands (ranging from 0 to 5 GPM) that were to be set at the 100 nodes for each scenario. The nodes representing sensor units did not have water demand. The EPANET hydraulic and Water Quality models were run on each scenario in order to obtain the values of chemical concentration at the monitoring units. Thus, each pattern was composed of four inputs (four sensor readings) and four binary outputs (0 for the non-activated intrusion points, 1 for the activated intrusion point).

A feed-forward ANN for modeling this problem was trained with the back-propagation learning algorithm. After testing several parameters of the network (reduction of RMS by modifying number of iterations and PEs, and transfer functions), the final network contained an input layer with four scaled bipolar inputs (-1 to 1), two hidden layers with six processing units, and an output layer with four PEs. The Generalized Delta Rule (GDR) was used to update the connection weights, and the hyperbolic tangent (*tanh*) function was applied as the transfer function in each PE. The epoch was set equal to 200. The training dataset contained 200 patterns, whereas the testing dataset contained 64 patterns.

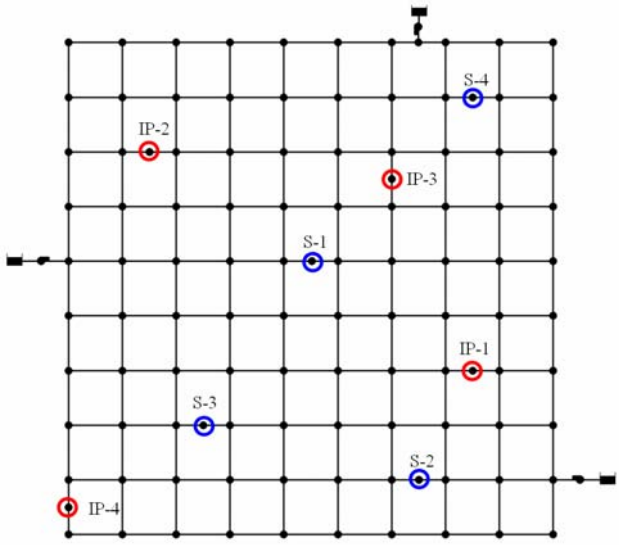

Figure 4. Water network for intrusion point identification

Results and Discussion

Modeling of Contaminant Concentrations

The generation of training and testing datasets was the most time-consuming task in both applications presented here, and is one of the major drawbacks of using Artificial Neural Networks in any application. Figure 5 shows the solute concentration trends with respect to water demands (at a regional level) obtained from the training dataset for the water distribution network presented in Figure 2. It is noteworthy that lower demands induce lower average concentrations, increasing monotonically as water demands increase. The concentration trends level off at water demands higher than approximately 40 GPM for all sections. Also, higher concentrations are observed at Section 1 when compared to the rest of the sections. Overall, concentrations at the sections nearer the injection point (water reservoirs) exhibit higher values than at those sections further from the water reservoirs. This may be explained by analyzing the mass transfer principles that drive the dispersion of reacting chemicals within the water network. The flow paths to regions closer to the injection point are shorter, which results in shorter residence time of the water. Such shorter residence times do not allow for bulk and wall chemical reactions to take place for long periods of time. Therefore, water chemical contents are not depleted as much as at sections further from injection points, for which residence times are longer.

A key feature of the Artificial Neural Network used in this application is the robustness and effectiveness achieved with it. As mentioned previously, water demands at all regions for each scenario were randomly defined, and their combinations were set without following any specific rule. Nevertheless, the ANN still provides consistent patterns of chemical concentrations with respect to water demands. This implies that the fundamental physical principles, though poorly understood and described in a macro-scale, are driving the chemical interactions that govern the transport phenomena occurring in a micro-scale. The prediction of chemical concentrations at regional level by using ANN modeling is useful because the installation of monitoring units at several locations of the water network is economically prohibitive. Therefore, reliable prediction tools may aid in providing the information necessary for managing local water quality. These tools, however, must be validated with field data and should be updated periodically in order to incorporate new scenarios that were not previously present. ANN models enable us to integrate recently measured field data that will potentially enhance the model performance. Furthermore, the ANN in this study used water demands as input data, making this application practical because water demand is constantly monitored at a household level.

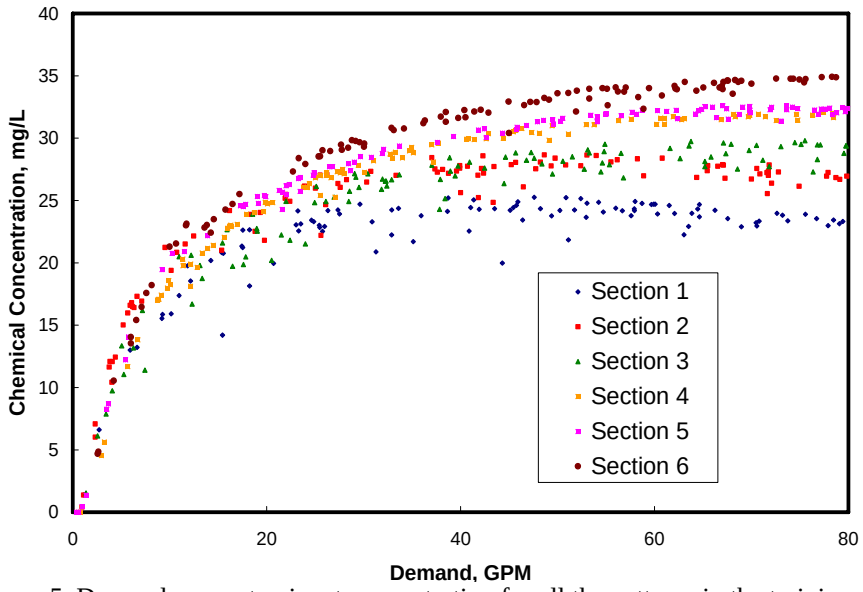

Figure 5. Demand vs. contaminant concentration for all the patterns in the training set

The direct comparison between EPANET and ANN chemical concentrations is depicted in Figure 6. It is remarkable that almost all points are clustered along the bisecting line, which indicates that the ANN results reproduce EPANET average concentrations with high correlation. However, a few outliers are found, mostly at low chemical concentrations (< 6 mg L^{-1}). There is no significant difference in correlation among the regions; however, regions further from the reservoirs (Sections 1, 2 and 3) exhibit correlations slightly lower than those regions closer to the reservoir (Sections 4, 5 and 6). This comparison is drawn from results obtained with the training dataset, for which the high correlations observed were expected. Initially, only twenty-five patterns were employed in training the neural network, which provided poor correlations between both ANN and EPANET outcomes. This suggested the need for increasing the training data patterns, which were accordingly set to one hundred patterns. This drawback is often found in modeling with ANNs, and although several training rules have been established for some ANN applications to actual problems, there is not yet a universal rule that can be applied to all modeling cases. Therefore, experience and a trial-and-error approach may help define the number of patterns required for training the network. This issue turns out to be of the utmost importance if ANN modeling is to be applied to a real-world water network. Because water quality data acquisition is expensive and time-consuming, critical decisions should be made with care regarding sampling frequency and locations for training.

After the training was completed, the ANN was run on the testing dataset, which contains patterns that the network has not "observed". Figure 7 shows the comparison of EPANET versus ANN solute concentrations for all regions of the water network. Even though the correlations (*R*-square) were slightly lower than those observed from the training dataset, the correlation values remain high for all regions ($R^2 > 0.93$). This helps confirm the robustness of the ANN proposed in this work because water demands at the testing patterns were also produced with a random generator tool. Similar to the results plotted in Figure 6, higher predictions were observed in near regions (Section 4 and 5) in comparison to those further from the reservoirs (Section 1, 2 and 3). Section 6 (the closest to the reservoir) was the exception in this observation, because its R-square value was the lowest from all regions and some outliers were visually noticed at low chemical concentration values (< 4 mg L^{-1}). The next step to successfully applying ANN modeling is to carry out a field study to validate the ANN's performance in real-world scenarios.

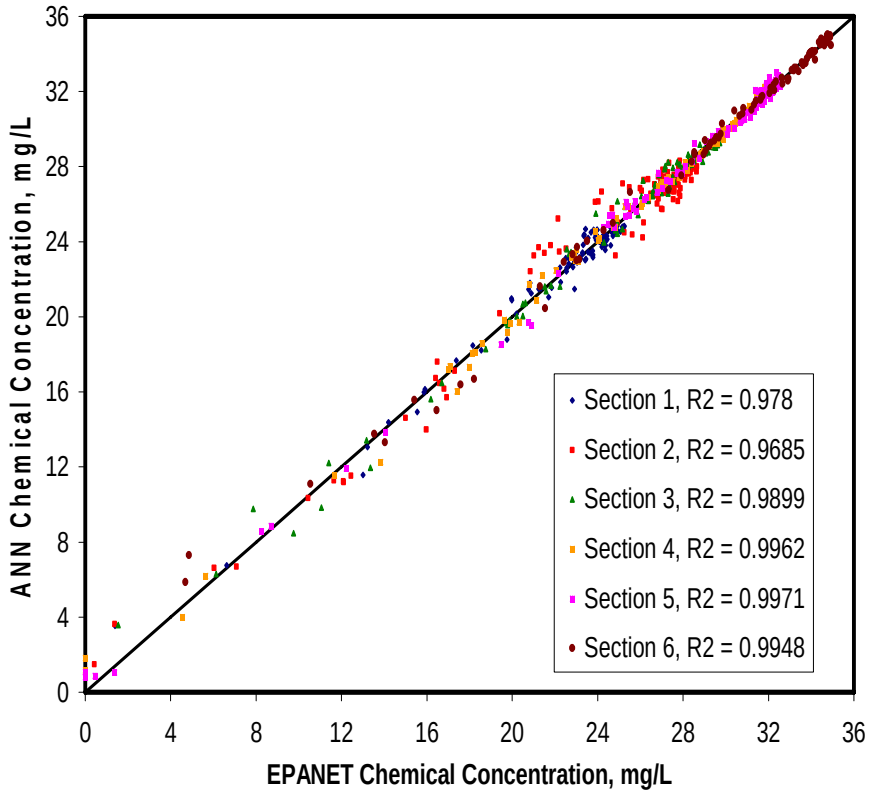

Figure 6 Calculated vs. expected contaminant concentration in different regions of the network, with training dataset

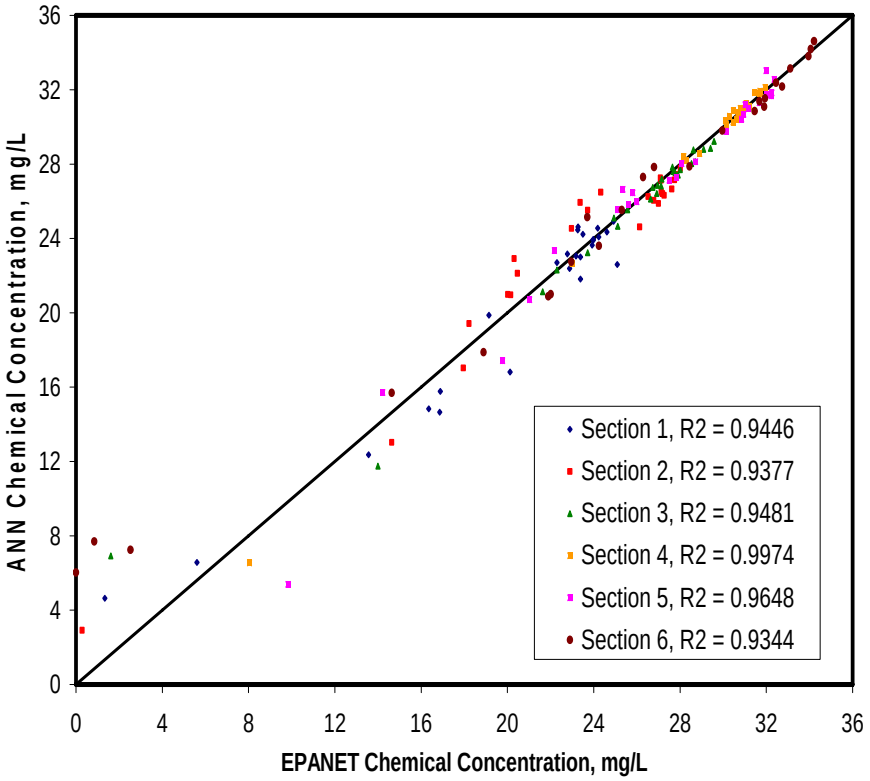

Figure 7 Calculated vs. expected contaminant concentration in different regions of the network, with testing dataset

Contaminant source identification

The second application based on 10x10 network (Figure 4) provided significant insight over the potential use of ANNs for modeling water quality in distribution systems. The generation of training and testing datasets was again the most laborious task along with the tests on several parameters of the neural network. Unlike the ANN used in the preceding application, the lowest RMS was achieved with two hidden layers of six PEs each, which demonstrates that there are several features of the ANN to be specifically defined before performing any calculations. Table 1 contains the information obtained when the ANN was run on the training dataset. In general, the injection points were traced back correctly for all the cases presented to the network. Injection Point 1 was the most difficult to identify based on the four sensor readings, as it is manifest for the lower percentage of properly classified cases (76 %). On the other hand, Injection Points 3 and 4 show the highest identification percentage, which was explained with a detailed analysis of the results. Whereas Injection Points 3 and 4 were mainly identified based on readings at Sensors 4 and 3, respectively, Injection Points 1 and 2 were traced back with combinations of all the sensor readings, which led to a larger number of possibilities.

The outcomes from testing patterns are presented in Table 2. They reproduce trends similar to training dataset results, as Injection Points 1 and 2 were most likely to be misidentified, whereas Injection Points 3 and 4 show high higher identification percentages. It should be noted that Sensors 2 and 4 did not produce signals as often as the others, which may reduce the actual number of monitoring units to 2. This is because Sensors 2 and 4 are located near the pumps, and to a certain degree are located upstream of the water flow. Consequently, the chemical injected cannot reach those locations and produce a signal. Thus, further studies on this application should be accompanied with an optimization of sensor placement in order to increase the confidence to effectively detect possible contamination events. However, the results show that there neural networks can potentially be used to identify a contamination source with the adequate sensor placement. This approach may also be used in combination with existing techniques in order to improve the outcomes as well as to supply information that cannot be readily obtained from existing techniques only. In addition, experimentation must verify the findings shown in this study and provide data that may be utilized for training the ANN. The construction and testing of a 5x5-node model network is currently underway at the Water Distribution Systems Laboratory of The University of Arizona in order to move ahead to the experimental phase of this research.

Table 1. EPANET (actual) versus ANN (modeled) Injection Points. Classification was carried out with training dataset

Injection Point (EPANET)	Injection Point (ANN)			
	1	2	3	4
1	38 (76 %)	8 (16 %)	0 (0 %)	4 (8 %)
2	1 (2 %)	39 (78 %)	0 (0 %)	10 (20 %)
3	0 (0 %)	6 (12 %)	44 (88 %)	0 (0 %)
4	0 (0 %)	6 (12 %)	0 (0 %)	44 (88 %)

Table 2. EPANET (actual) versus ANN (modeled) Injection Points. Classification was carried out with testing dataset

Injection Point (EPANET)	Injection Point (ANN)			
	1	2	3	4
1	11 (68.75 %)	4 (25 %)	0 (0 %)	1 (6.25 %)
2	1 (6.25 %)	13 (81.25 %)	0 (0 %)	2 (12.5 %)
3	0 (0 %)	2 (12.5 %)	14 (87.5 %)	0 (0 %)
4	0 (0 %)	1 (6.25 %)	0 (0 %)	15 (93.75 %)

Conclusions

The issue of assessment, monitoring and detection of contaminants in drinking water has become important to ensure alertness and preparedness in case of an accidental or deliberate contamination event. Although ANNs have been long applied to multiple studies regarding water quality, few applications exist for the analysis of water quality within a pressurized water network. ANNs are advantageous in that they provide a data-driven solution to the problem; they are not based on physical principles of the phenomenon. This allows us to manage water quality on a macro- rather than micro-scale, which is suitable for municipal water distribution systems that are highly complex and transient. This introductory work supports two potential uses of feed-forward ANNs trained with the back propagation algorithm to determine the contaminant concentration at multiple locations of the network as well as to identify the point in the water network where a contaminant was introduced. The input data came from existing physical models for hydraulic and water quality analysis in pressurized water systems, and thus, further improvements of our findings must include experimental data for training and validation of the ANN. The ultimate goal of this type of research is to provide understanding of the chemical and microbial transport processes that are relevant from both the engineering and public health standpoints.

Acknowledgements

This work is supported by the Environmental Protection Agency/Department of Homeland Security (under Grant No. 613383D).

Reference

Baxter, C.W., Q. Zhang, S.J. Stanley, R. Shariff, R-R.T. Tupas and H.L. Stark (2001). “Drinking water quality and treatment: the use of artificial neural networks.” *Canadian Journal of Civil Engineering* 28: 26-35.

Brion, G.M. and S. Lingireddy (2003). “Artificial neural network modelling: a summary of successful applications relative to microbial water quality.” *Water Science and Technology* 47 (3): 235-240.

Flood, I. and N. Kartman (1994). “Neural networks in Civil Engineering. I: principles and understanding,” *Journal of Computing in Civil Engineering* 8 (2): 131-148.

Gümrah, F. B. Öz, B. Gueler, and S. Evin (2000). “The application of artificial neural networks for the prediction of water quality of polluted aquifer.” *Water, Air, and Soil Pollution* 119: 275 – 294.

Kim, Minyoung. (2005). Surveillance System for Biological Agents in Water Systems. *PhD Thesis*. The University of Arizona.

Lingireddy, S. and L.E. Ormsbee (1998). “Neural networks in optical calibration of water distribution systems.” Artificial Neural Networks for Civil Engineers: Advanced Features and Applications. Ed. I. Flood and N. Kartam. American Society of Civil Engineers.

Neelakantan, T.R. (2005). “Artificial Neural Networks: An Overview,” Artificial Neural Networks in Water Supply Engineering. Ed. S. Lingireddy and G.M. Brion. American Society of Civil Engineers.

Milot, J. and M.J. Rodriguez, and J.B. Sérodes (2002). “Controbution of neural networks for modeling trihalomethanes occurrence in drinking water.” *Journal of Water Resources Planning and Management* 128 (5): 370-376.