# Deep Leaning Neural Networks for Determining Replacement Timing of Steel Water Transmission Pipes

Hong Jin Jun
Civil & Environmental Engineering
Univ. of Wisconsin-Madison
Madison, USA
jun9@wisc.edu

Jae K. Park
Civil & Environmental Engineering
Univ. of Wisconsin-Madison
Madison, USA
jkpark@wisc.edu

Cheol Ho Bae
Water Network Research Center
Korea Water Resources Corporation
Daejeon, South Korea
baech@kwater.or.kr

*Abstract*—Water main pipe breaks are an ongoing concern worldwide. Large-diameter steel water transmission mains (WTMs) transport a much larger volume of water and their failure leads to even greater damages than those seen in water networks with small diameter iron or PVC pipe lines. However, there is no predictive model for large-diameter steel WTMs, leaving retroactive maintenance as the sole means of prevention. The objective of this study was to predict the optimal replacement timing for large-diameter steel WTMs based on physical and environmental factors, using Deep Learning algorithms. The model was developed in four steps: (1) determine major factors, (2) determine the best model by comparing performances of three neural networks (NNs) (a shallow artificial NN, multiple hidden layered NN, Stacked autoencoder NN), (3) classify the data into homogeneous groups by an ANN-based clustering technique, and (4) perform the developed model for each group. The multiple hidden layered NN was found to be the best deep neural NN in forecasting a replacement timing of aging WTMs. Additionally, it is recommended that such ANN-based clustering methods be used in predicting a more accurate replacement timing of water networks and making a quantitative decision on replacement.

*Keywords—deep neural network, prediction, replacement*

## I. Introduction

### A. Background

WTMs are large-diameter pipes (more than 300 mm in diameter) that transport large volumes of water, including both raw water from natural water sources, as well as treated water to storage reservoirs or to smaller-diameter distribution networks connected to customers. As existing water networks get older, water main breaks become far more likely, resulting in water service interruption, large operational costs and considerable inconvenience to end users. It is reported that the infrastructure grade of the United States (U.S.) in the drinking water category is "D (Poor: At Risk)", implicating that the infrastructure is approaching the end of its service life with a high risk of failure [1]. Similarly, the multi-regional WTMs (about 5,000 km and 50 water supply facilities) installed in the 1960s in South Korea are experiencing increased water main breaks, as shown in Figure 1. If water main breaks could be accurately anticipated, water service outage, traffic congestion, operational costs and end user inconvenience could all be minimized. To this end, a reliable prediction model is needed in order to make the process of water main replacement as economical as possible.

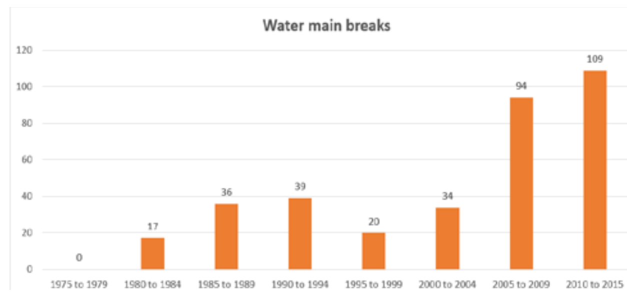


Figure 1. South Korea's multi-regional water main breaks over time (Source: K-water's Operation and Management Service System)

### B. The Ideal Predictive Model for Water Main Breaks

The ideal predictive model should address historical break data and other factors, both internal and external to the pipes themselves. Specifically, the factors affecting a service life of pipes can be categorized into three groups as follows:

- Physical factors: Pipe material, thickness, age, diameter, length, type of joints, coating, and traffic load.
- Environmental factors: Soil type, soil moisture, soil load, microbes, water quality (pH, residual chlorine, etc.).
- Operational factors: Internal water pressure, flow velocity, and operational and maintenance practices.

### C. Shortcomings of Previous Predictive Models

*a) Lack of Comprehensive Factor Analysis*: Previous models have addressed physical factors such as pipe thickness and/or environmental factors such as soil characteristics. However, these studies did not address water quality factors leading internal corrosion as contributors to the failure of pipes.

*b) Limited to Small Diameter Iron and PVC Pipes:* Furthermore, previous studies have been based on small diameter water distribution networks composed of iron or PVC pipelines. No studies have so far been conducted in developing a predictive model for steel WTMs. As WTMs transport a much larger volume of water and their failure leads to even greater damages than those seen in small diameter water main breaks, such studies are much needed. The need for studies focused on steel water mains is further underscored by the fact that extrapolation of the results of previous studies focusing on iron pipes may not be valid. This is due to the fact that iron and steel pipes are different in their composition, properties, lining and welding systems, the chemical reactions associated with their corrosion, and their failure characteristics [2].

### D. *Artificial Intelligence-Based Predictive Models*

Traditional studies have used statistical models that utilize historical break data to predict patterns. Shamir & Howard (1979) were the first to develop a regression model in which a pipe breakage is exponentially associated with its age. A comprehensive review of research conducted using the traditional statistical models can be found elsewhere [4].

The complex, multifactorial nature of WTM breaks requires a highly-sophisticated system in order to develop a reliable predictive model. Artificial NNs are an attractive modality for this purpose due to their impressive capacity to solve complex, real world problems. Jafar et al. (2010) employed ANNs to predict the number of failures of water networks (composed of asbestos cement, PE and iron) and others have used ANNs to predict failures of aging pipe lines (iron or PVC) [6][7][8]. However, studies using deep neural networks (DNNs) have not so far been used to predict WTM breaks.

A DNN is an artificial neural network (ANN) with multiple hidden layers of units between the input and output layers. Similar to ANNs with one hidden layer (Shallow ANNs), DNNs can model complex non-linear relationships. Additional layers allow DNNs to learn high-level features in data by using structures composed of multiple non-linear transformations [9]. DNNs include multiple hidden layered NNs (MLNNs), stacked auto-encoders NNs (ANNs), deep belief networks (DBNs) and convolutional NNs (CNNs), among others.

DNNs are thought to be an ideal modality for predicting large-diameter steel WTM breaks. DNNs have already been used successfully to solve numerous complex tasks, such as character recognition and stock market predictions. Their previous success has led to the recognition of DNNs as powerful tools that can be used to accurately predict outcomes in complex systems involving multiple parameters.

### E. *Study Objective*

In summary, the objective of the study is to develop a predictive model for large diameter steel WTM breaks, which includes both physical factors as well as water quality factors, with the aid of DNNs.

## II. Methodology

Figure 2 describes the methodology of model development including data pretreatment, determination of major factors affecting breaks of WTMs, comparison of DNNs, and potential improvements when employing a clustering method using a NN.

### A. *Data Collection*

Data were obtained from the database of Korea Water Resources Corporation (K-water), which maintains 32 drinking water treatment plants and supply facilities (50% of South Korea's drinking water) as a government-owned public utility.

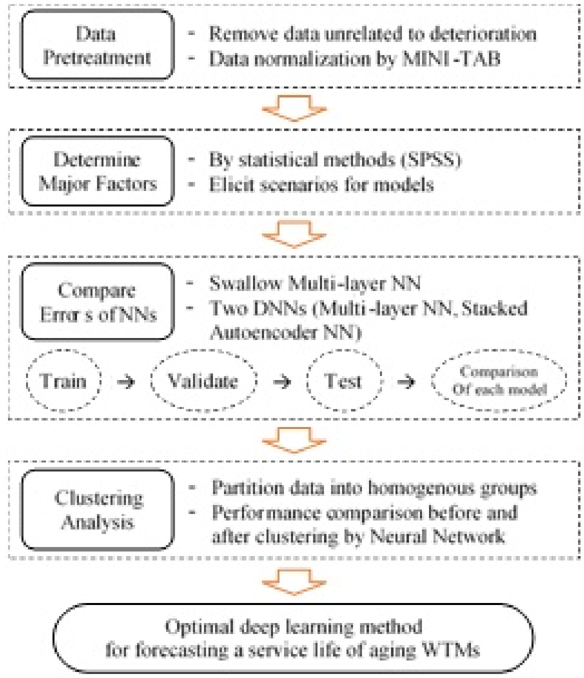


Figure 2. Methodology of model development

Data consisted of 855 historical breaks of steel WTMs over the observation period (1969 – 2015) with a total length of 2,783 km. The collected data include pipe diameter, thickness, age, coating, service type, length of segments, surface area, depth of cover, traffic loading, water quality (pH, alkalinity, dissolved oxygen, residual chorine, temperature, electoral conductivity) for inputs, and time to first break for target as shown in Table I. It is assumed that the water quality data measured at the monitoring stations in the DWTPs reflects the water quality throughout the WTMs. The break time of WTMs refers to the time to the first break of WTMs and start considering the replacement based on cost-benefit analysis before serious water service disruption occurs.

TABLE I. DESCRIPTIONS OF DATA ON WATER TRANSMISSION MAINS

| Type | Factors | Descriptions |
|---|---|---|
| Inputs (15) | Pipe diameter | Internal diameter of pipes |
| | Pipe thickness | Pipe manufacturing specifications |
| | Pipe age | Age of laid pipe |
| | Pipe Coating | No coating, Either internal or external coatings, Both coatings |
| | Service type | Raw, Settled, Purified water |
| | Pipe length | Length of a segment of laid pipe |
| | Pipe surface area | Area adjacent to environment (internal water quality and external soil) |
| | Depth of cover | Earth load pressure |
| | Traffic loading | Traffic road, Non traffic road |
| | Six water quality parameters | pH, Alkalinity, DO, Residual chlorine, Temperature, Conductivity |
| Outputs (1) | Time to first break | Pipe failure leading to water leakage |

### *B. Data Pretreatment and Statistical Analysis*

The historical break data (116 incidents) not caused by deterioration, such as breaks resulted from poor installation, faulty materials, were excluded amongst 855 break incidents.

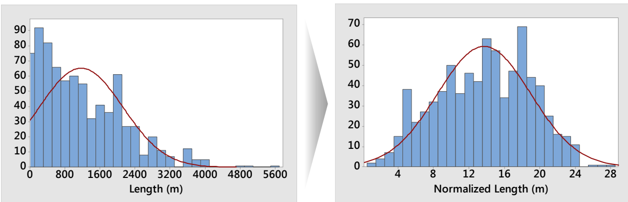


Figure 3. Data transformation for pipe length by Box-Cox Transformation

Real data are not at normal distribution while many statistical analyses including NNs assume normality. However, Box-Cox Power transformation can improve normality [10]. Achim et al. (2007) applied Box-Cox Power transformation to improve NN model's representation of the collected data when forecasting water pipe asset life [11]. Thus, Box-Cox Power Transformations were applied to the input factors to improve normality using MINITAB 2017. The sample of Box-Cox transformation applied to the input variable (Pipe length) is shown in Figure 3. And, influencing factors affecting a break time of steel WTMs were identified through the Pearson's correlation analysis using SPSS 18 (A statistical software).

### *C. Construction of DNN Models*

Two DNN methods (MLNNs and stacked ANNs) were used and compared with Shallow ANN to determine the optimal model. The NN Toolbox of MATLAB (Mathworks 2016) was used in all programing. Input data were randomly divided into three data sets: training (70%), validation (15%), and test data set (15%) to build and test the model.

#### *1) Shallow ANN and MLNN*

A typical MLNN is shown in Figure 4. Mathematically, a MLNN with *p*, $S^1$, $S^2$, and $S^3$ as the number of input, 1st hidden, 2nd hidden, and output nodes, respectively, is based on the following equation:

$$a^3_{s3} = f^3\left[\sum_{s2=1}^{S^2} w^3_{s3s2} \times f^2\left\{\sum_{s1=1}^{S^1} w^2_{s2s1} \times f^1\left(\sum_{R=1}^{R} w^1_{s1R} \times P_R + b^1_{s1}\right) + b^2_{s2}\right\} + b^3_{s3}\right] \quad (1)$$

where $a^3_{s3}$ is the output values; $P_R$ is the input values; $w^1_{s1R}$, $w^2_{s2s1}$, and $w^3_{s3s2}$ are the weights of connections of the input layer and the 1st hidden layer, of the 1st hidden layer and the 2nd hidden layer, and of the 2nd hidden layer and output layer, respectively; $b^1_{s1}$, $b^2_{s2}$, and $b^3_{s3}$ are the biases at the 1st hidden layer, the 2nd hidden layer, and output layer, respectively; $f^1$ is a sigmoid activation function; and $f^2$ and $f^3$ are the linear activation functions [5]. The used sigmoid activation function is as follows:

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

The learning process of MLNN is based on a series of connection weight adjustments to minimize errors between the predicted and target values. Inputs are first propagated forward through each layer of the network. The process of training a NN involves tuning the values of the weights and biases of the network to minimize mean square error between the predicted values ($a_i$) and the target values ($t_i$). It is defined as follows:

$$Error = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \quad (3)$$

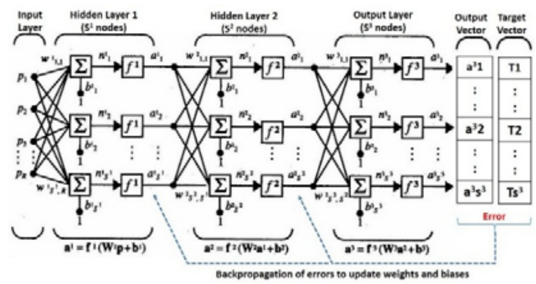


Figure 4. Typical multi-hidden layered NN (SOURCE from [12])

A first-order iterative optimization algorithm aimed at finding a minimum in the error surface was used to perform gradient-descents in weights and biases. This process was iterated for each epoch until convergence within a given tolerance [13]. The calculation procedures of Shallow ANN are the same as MLNN except for the number of hidden layers.

#### *2) Stacked Autoencoder Neural Network (Stacked ANN)*

Stacked ANN uses an unsupervised learning method for pretraining to help initializing network with good parameters. An autoencoder is pretrained to reconstruct the input as its output. After the pretraining, stacked ANN can run backpropagation on the entire network to finetune weights for the supervised task [14]. Because this backpropagation starts with good weights, its credit assignment is better and the learned model is likely to be better if backpropagation is run initially as shown in Figure 5 [15]. The following calculation method and procedures after the pretraining are identical to MLNN.

#### *3) Overfitting Avoidance*

DNNs with multiple hidden layers are powerful tools to learn complicated relationships between inputs and outputs. However, overfitting can be detrimental to such networks [16]. Cross-validation technique, namely early stopping, known as empirically better approach, was used to improve generalization of DNNs and avoid overfitting [15]. The error on the validation set is monitored during the training process. However, when the error on the validation set begins to rise, indicating they are overfitting as shown in Figure 6, the network returns the weights and biases, resulting in the least validation-set error [12].

### *D. DNNs with Clustering Analysis*

A better prediction was achieved by clustering water main data into homogenous groups [17]. The Self-Organizing Map (SOM) is a clustering tool. Data can be automatically organized into a meaningful two-dimensional order in which similar models are closer to each other in the grid than the more dissimilar ones [18]. For clustering analysis, the Neural Network Toolbox of MATLAB (Mathworks 2016) was used, which is based on Kohonen learning algorithm [12].

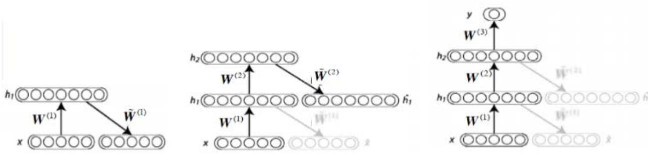

Figure 5. Typical Stacked ANN (Source from [15])

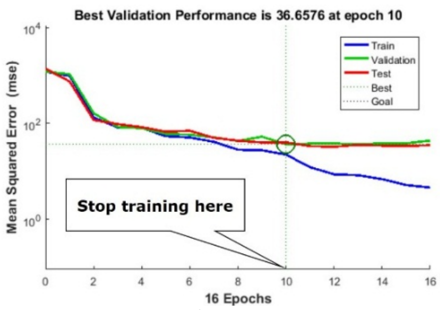


Figure 6. Examplary application of Cross-validation technique

The SOM is trained iteratively. For each input (X), the Euclidean distance between input and all the synaptic weights is calculated as follows [19]:

$$d_j(X) = \sum_{i=1}^{n} (x_i - w_{ji})^2 \tag{4}$$

where $X$ is input vectors with units($x_i$); $w_{ji}$ is the synaptic weights between input units $i$ and the neurons $j$ in the output layer; and $d_j(X)$ is the squared Euclidean distance between $X$ and $w_i$ for each neuron $j$.

The neuron closest to $X$ is declared the best matching neuron $I(X)$. After $I(X)$ is recognized, the weights of $I(X)$ as well as its topological neighbors are updated so that they are moved closer to the input vector in the input space. The vectors are updated following the Kohonen learning rule as follows [18]:

$$w_j(t+1) = w_j + a(t)T_{j,I(x)}((x_i - w_j(t)) \tag{5}$$

$$T_{j,I(X)} = \exp\left(-S^2_{j,I(X)}/2\sigma^2\right) \tag{6}$$

where $T_{j,\ I(X)}$ denotes the neighborhood function that decays with distance $S_{j,\ I(X)}$ between the $I(X)$ and the neurons $j$ in the output layer; $t$ is a time (epoch); σ denotes the size of the topological neighborhood; α is the learning rate ($0 < a < 1$). This learning process continues until the two-dimensional output map stops changing. Eventually, the inputs are grouped into clusters [20].

## III. Results

### A. Main Factors Affecting a Service Life of Steel WTMs

Ten factors (pipe age, length, water temperature, electoral conductivity, surface area, service type, residual chlorine, pH, coating, DO) had P-values less than 0.05.

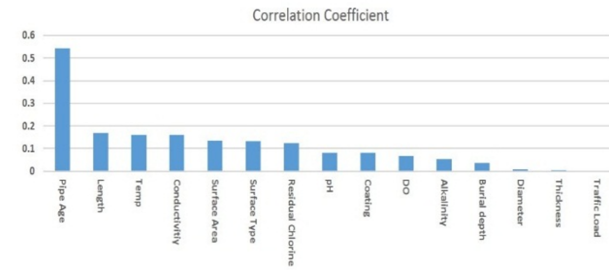


Figure 7. Correlation coef. of factors affecting a break time of steel WTMs

TABLE II. Scenarios for DNN models

| Category | Factors |
|---|---|
| Scenario 1 | Age, Length, Temp., Conductivity, Surface area, Service type, Chlorine, pH, Coating, DO, Alkalinity, Burial depth, Diameter, Thickness, Traffic load |
| Scenario 2 | Age, Length, Temp., Conductivity, Surface area, Service type, Residual chlorine, pH, Coating, DO |
| Scenario 3 | Age, Length, Temp., Conductivity, Surface area, Service type, Residual chlorine, pH, Coating |
| Scenario 4 | Age, Length, Temp., Conductivity, Surface area, Service type, Residual chlorine, pH |
| Scenario 5 | Age, Length, Temp., Conductivity, Surface area, Service type, Residual chlorine |
| Scenario 6 | Age, Length, Temp., Conductivity, Surface area, Service type |
| Scenario 7 | Age, Length, Temp., Conductivity, Surface area |

Their correlations with a break time of steel WTMs were significant at the 0.05 level, while other factors (burial depth, diameter, thickness, traffic load) appeared to contribute relatively little to steel WTM deterioration, unlikely considered important for predicting break of iron or plastic pipe in water distribution networks in existing researches.

Figure 7 shows that pipe age has the highest correlation with WTMs deterioration and water quality parameters such as temperature, residual chlorine, conductivity significantly affect WTP failure. Iron corrosion is controlled by water quality factors that may exert their influence during the time when the metal corrodes [21]. This indicates that water quality factors can be a good indicator for the prediction of steel WTMs break.

### B. Application of DNNs

*a) Setting-up Seven Scenarios for Models*: The scenarios established by the factors obtained through the correlation analysis are presented in Table II. It consists of seven scenarios. Scenario 1 includes all 15 factors. Scenario 2 includes pipe age, length, water temperature, electoral conductivity, surface area, service type, residual chlorine, pH, coating. The input data was sequentially excluded one by one in descending order of its relativity for determining the input data for other following five scenarios.

*b) Performance Comparison of Models by Scenarios:* Shallow ANN, MLNN, and Stacked Autoencoder NN were performed for each scenario. Table III shows the performances of the models by scenarios including correlations (R) and rooted mean squared errors (RMSE).

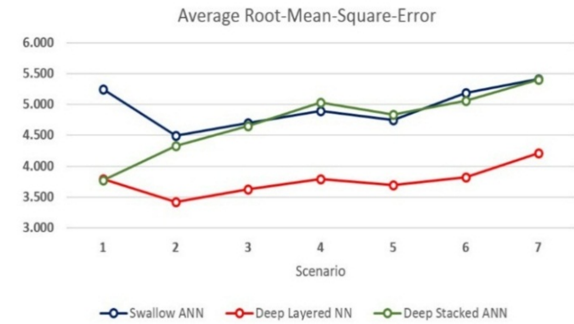


Figure 8. Performance comparison of models by scenarios

TABLE III. PERFORMANCES OF MODELS BY SCENARIOS

| Table Head | | Shallow ANN | Deep NN | |
|---|---|---|---|---|
| | | | *MLNN* | *Stacked Autoencoder NN* |
| Scenario 1 | R value | 0.74 | 0.83 | 0.85 |
| | RMSE | 5.24 | 3.80 | 3.77 |
| Scenario 2 | R value | 0.78 | 0.86 | 0.82 |
| | RMSE | 4.49 | 3.42 | 4.33 |
| Scenario 3 | R value | 0.79 | 0.84 | 0.82 |
| | RMSE | 4.70 | 3.62 | 4.65 |
| Scenario 4 | R value | 0.80 | 0.83 | 0.79 |
| | RMSE | 4.89 | 3.8 | 5.03 |
| Scenario 5 | R value | 0.82 | 0.84 | 0.81 |
| | RMSE | 4.74 | 3.69 | 4.84 |
| Scenario 6 | R value | 0.79 | 0.84 | 0.80 |
| | RMSE | 5.19 | 3.82 | 5.06 |
| Scenario 7 | R value | 0.76 | 0.78 | 0.76 |
| | RMSE | 5.41 | 4.21 | 5.40 |
| Overall | R value | 0.78 | 0.83 | 0.81 |
| | RMSE | 4.95 | 3.76 | 4.73 |

Figure 8 shows that MLNN has better accuracy than any other two algorithms (Shallow ANN and Stacked Autoencoder NN) across the scenarios. MLNN in Scenario 2 showed the highest forecasting performance in a break time of steel WTMs with the values of R and RMSE as 0.86 and 3.42 years, respectively, amongst the seven scenarios. The Stacked Autoencoder NN has the best performance in Scenario 1 and appears to be the better predictive model with many input variables. In contrast, for other scenarios (2 – 7), the Stacked Autoencoder NN appears to have similar performances to the shallow NN, unlikely that DNNs are better performances than shallow NNs, as well as implicating that it is very important to determine an optimal condition for each DNN because prediction performances of DNNs may vary based upon both modeling conditions such as number of input variables and applicable fields such as regression or image recognition. DNNs such as Stacked Autoencoder NN appear to have a better performance in many input factors than Shallow ANNs, whereas Stacked Autoencoder NN and shallow ANN have similar performances as input variables decrease.

### *C. Clustering WTM data by ANN*

Dataset was partitioned by similarity using the SOM algorithm.

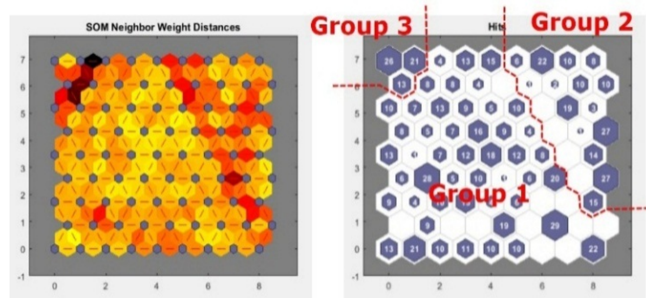


Figure 9. Clustering data into similar groups by ANN

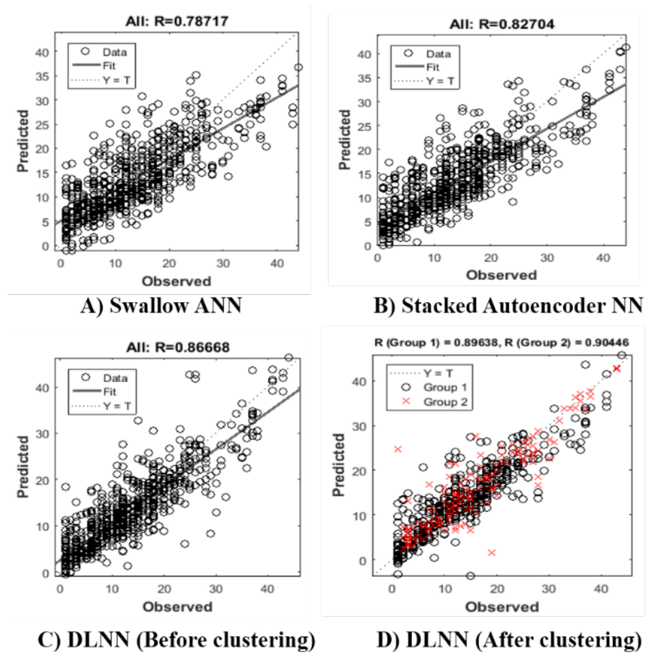


Figure 10. Representative scatter plots and R values of the predicted and the observed values for Shallow ANN, Stacked Autoencoder NN, MLNN before clustering, and MLNN after clustering in Scenario 2 in forecasting a replacement timing for multi-regional steel WTMs

Figure 9 shows that dataset was classified with 2D topology with 81 neurons, the red lines indicate the distances between neurons, and the darker colors represent larger distances on the SOM Neighbor Weight Distance map. The figures in the hexagonal neuron represents the number of homogeneous data on the Hits map of Figure 9. The SOM map shows data are clustered into three homogeneous groups.

### *D. Performance comparison before and after clustering*

The MLNN in Scenario 2 was performed for each of two groups because the third group should be discarded due to its small number of data. Table IV shows the performance improvements before and after clustering by ANNs. The performance for each group after partitioning into homogeneous group was improved by 4.4% and 4.9% in R values and by 42.4% and 58.8% in RMSE, respectively. Figure 10 illustrates the representative scatter plots and R values of the predicted and the observed values for shallow ANN, Stacked Autoencoder NN, MLNN before clustering, and MLNN after clustering in Scenario 2 in forecasting a replacement timing for multi-regional steel WTMs in South Korea. When all data points in a scatter plot lie exactly along a straight dotted line, the R value becomes 1. In the figure, the data lie more closely to the straight dotted line with higher R values from shallow ANN, to Stacked Autoencoder NN, to MLNN before clustering, and to MLNN after clustering.

### E. *Usefulness of the Developed MLNN Model*

From the results in Table III, Table IV and Figure 10, it can be concluded that the developed model was both reliable and robust in predicting the first break time for aging large-diameter steel WTMs. It is thought that this model can be supplemented to the "Man Entry and Visual Inspection" methodology currently in use. Specifically, if water supply engineers enter data (age, length, temp., conductivity, surface area, service type, residual chlorine, pH, coating, DO) for each segment of the operating WTMs, they can estimate when each of WTMs would experience their first break. Additionally, the predictive model may be used to create better operating conditions to extend the break time of WTMs by adjusting factors (e.g., residual chlorine) that affect WTMs deterioration.

## IV. CONCLUSIONS AND RECOMMENDATIONS

The Shallow ANN and DNNs (MLNN and Stacked Autoencoder NN) were performed with the various physical and water quality data collected during 1969–2015 and used to forecast a replacement timing of large-diameter steel WTMs using statistical methods and clustering technique, as compared with each NNs. The following conclusions can be drawn:

- Significant factors affecting steel WTM deterioration in the study area were determined by a statistical method: pipe age, length, water temp., conductivity, surface area, service type, residual chlorine, pH, coating, and DO. Thus, it is suggested that these data be monitored and collected to improve the accuracy of forecasting replacement timing of WTMs.
- It is elicited that MLNN is the best Deep Neural NN for large diameter steel WTMs, while the Stacked Autoencoder NN fitted predictive models with many input variables. MLNN had the better accuracy than the Stacked Autoencoder NN in a regression field like forecasting a first break time of aging WTMs, although the Stacked autoencoder NN has been known as one of the efficient Deep Learning algorithms in the areas of mainly classification and image processing.
- The optimal prediction model was established to estimate a replacement timing for large-diameter steel WTMs on the basis of historical water main break records and available physical and environmental data, using DNNs. The model was developed in four steps: (1) determine major factors and set up scenarios for models, (2) determine the best scenario and DNN by comparing performances, (3) classify data into similar groups by ANN-based clustering technique, and (4) perform the developed model for each group.
- The SOM clustering method using ANN was first applied to WTM management area. The method improved prediction accuracy significantly. It is thought to be a good tool for clustering data. Also, such ANN-based clustering method is recommended to be used in predicting a replacement timing of water networks.

TABLE IV. PERFORMANCE COMPARISON BEFORE AND AFTER CLUSTERING

| **Groups** | **MLNN for Scenario 2** | | **Improvements** |
|---|---|---|---|
| | ***Before clustering*** | ***After Clustering*** | |
| Group 1 | R value: 0.861<br>RMSE: 3.421 | R value: 0.899<br>RMSE: 1.97 | 4.4% ↑<br>42.4% ↑ |
| Group 2 | | R value: 0.903<br>RMSE: 1.41 | 4.9% ↑<br>58.8% ↑ |

The DNN-based predictive model proposed in this study is robust and can be used to reliably predict the first break time for aging large-diameter steel WTMs, reducing the amount of time and money spent on the direct inspection of existing WTMs and preventing unwanted water service interruption, operational costs and end user inconvenience resulting from WTM breaks. This model is expected to benefit both drinking water consumers and water supply engineers and consultants.

## V. REFERENCES

[1] ASCE, “2013 Report Card (ASCE),” 2014.

[2] AWWA, *Corrosion of Iron and Steel*, vol. 75190, no. 1. 1996.

[3] Shamir et al, “An analytic approach to scheduling pipe replacement.,” *Am. Water Work. Assoc. J.*, vol. 71, no. 5, May 1979, pp. 248–258, 1979.

[4] Kleiner et al, “Comprehensive review of structural deterioration of water mains: statistical models,” *Urban Water*, no. 3, pp. 131–150, 2001.

[5] Jafar et al, “Application of Artificial Neural Networks (ANN) to model the failure of urban water mains,” *Math. Comput. Model.*, vol. 51, no. 9–10, pp. 1170–1180, 2010.

[6] Ahn et al, “Predicting water pipe breaks using neural network,” *Water Sci. Technol. Water Supply*, vol. 5, no. 3–4, pp. 159–172, 2005.

[7] Al-Barqawi et al, “Condition Rating Model for Underground Infrastructure Sustainable Water Mains,” *J. Perform. Constr. Facil.*, vol. 20, no. 2, pp. 126–135, 2006.

[8] Nishiyama et al, “Forecasting breaks in cast iron water mains in the city of Kingston with an artificial neural network model,” vol. 923, no. August, pp. 918–923, 2014.

[9] Bengio, *Learning Deep Architectures for AI*, vol. 2, no. 1. 2009.

[10] R. S. Society, “The Box-Cox Transformation Technique,” vol. 2, no. May 2016, pp. 129–136, 2016.

[11] Achim et al, “Prediction of water pipe asset life using neural networks,” *J. Infrastruct. Syst.*, vol. 13, no. 1, pp. 26–30, 2007.

[12] MathWorks, *Neural Network Toolbox User ’ s Guide How to Contact MathWorks*, no. June. 2015.

[13] Hagan et al, *Neural Network Design*, vol. 2. 1995.

[14] P. Baldi, “Autoencoders, Unsupervised Learning, and Deep Architectures,” *ICML Unsupervised Transf. Learn.*, pp. 37–50, 2012.

[15] M. Craven, “Neural Networks,” 2016.

[16] Caruana et al, “Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping,” *Adv. Neural Inf. Process. Syst.*, pp. 402–408, 2001.

[17] Male et al, “Analyzing water main replacement policies,” *J. water Resour. Plan. Manag.*, vol. 116, no. 3, p. 362, 1990.

[18] Kohonen, “The self-organizing map,” *Neurocomputing*, vol. 21, no. 1–3, pp. 1–6, 1998.

[19] John A. Bullinaria, “Self Organizing Maps : Fundamentals What is a Self Organizing Map,” *Neural Networks*, pp. 1–15, 2004.

[20] Huang et al, “Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods,” *Mech. Syst. Signal Process.*, vol. 21, no. 1, pp. 193–207, 2007.

[21] AWWA, *Internal Corrosion of Water Distribution Systems*. 1996.