Memoirs of the Faculty of Engineering, Kyushu University, Vol. 65, No. 1, March 2005

Applications of Artificial Neural Networks for Optimal Pressure Regulation in Supervisory Water Distribution Networks

by

Haytham AWAD*, Akira KAWAMURA** and Kenji JINNO***

(Received December 15, 2004)

Abstract

This paper presents a simple and efficient technique for improving the existing optimal pressure regulation and leakage minimization algorithms for supervisory water distribution networks. With the assistance of Supervisory Control and Data Acquisition (SCADA) we have trained a Self-Organized Map (SOM), an unsupervised artificial neural network (ANN), to classify well regulated pressure cases for the water distribution network based on its actual values of flow meter readings which reflect the real network water demands or consumption. After training the SOM, a simulation step is used to classify the unregulated pressure cases into the different model classes. Based on these classifications the appropriate electrical motor valves setting of the well pressure regulation events are used for the unregulated ones. Regarding that all the available algorithms deal directly with the pressure regulation problem from an optimization point of view which required a computational time depending on the water network size and the used optimization method and in most cases requires also a network simplification method which is considered as another optimization problem. Using SOM as a pre-optimization method could prevent all errors resulting from applying optimization models, save its computational time and provides us with an on-line pressure regulation method. Computational results for Block 12 of Fukuoka City water distribution network using a short-term data set demonstrate the effectiveness of using SOM as a pre-optimization tool for regulating 74% of events within the target pressure range.

Keywords: Water distribution networks, Motor valve control, Pressure regulation, Leakage minimization, Self-Organizing Maps (SOM), Artificial Neural Networks (ANNs)

---

*Graduate Student, Department of Urban and Environmental Engineering
**Professor, Department of Civil Engineering, Tokyo Metropolitan University
***Professor, Institute of Environmental Systems

1. Introduction

Water loss occurs in all distribution systems - only the volume of loss varies, depending on the characteristics of the pipe network and other local factors, the water company's operational practice, and the level of technology and expertise applied to control it. These losses can be severe, and may go undetected for months or even years. The larger losses are usually from burst pipes, or from the sudden rupture of a joint, while smaller losses are from leaking joints, fittings, service pipes, and connections. The volume lost will depend largely on the characteristics of the pipe network and the leak detection and repair policy practiced by the company, such as: (i) The pressure in the network, (ii) Whether the soil type allows water to be visible at the surface or not, (iii) The "awareness" time (how quickly the loss is noticed), and (iv) The repair time (how quickly the loss is corrected). Examples of water loss as a percent of water supplied were summarized in an international survey by the International Water Services Association (IWSA) in 1991[1] as: (i) Developed countries 8-24%, (ii) Newly-industrialized countries 15-24%, and (iii) Developing countries 25-45%. In Japan, the average amount of water loss in the biggest 12 Japanese cities was 7.4% in 1997[2]. The lowest value occurred at Fukuoka City, Western Japan; 3.5% in 1997 which is considered at that time the minimum rate in the world according to the UN-habitat program[3].

Volume of water leakage from water distribution networks represents a significant amount from the total domestic water use; domestic water use in Japan was 16.4 billion m^3 in 1999 which represents 19% of the total water usage[4]. Therefore, in order to counteract water leakage a lot of water companies and waterworks bureaus around the world have established a "Water Leakage Prevention" sections which are responsible for: (i) detecting leaks from water supply network at an early stage followed by rapid repair, (ii) Improvement of water pipes status by replacing the deteriorated pipes and performing routine programs for pipes maintenance, and (iii) adjusting the water pressure.

It is well known that water leakage from a supply-and-distribution network is directly related to the system service pressure. The leakage volume increases proportionally with the increase in the average system pressure[5]. Therefore, in order to improve the existing system operation and management some waterworks bureaus in the developed countries around the world has started to establish the supervisory system of water distribution network in which sensor information from pressure gauges, flow meters, and electric valves connected to important points of the network are continuously sent to a control center where valve openings are adjusted depending on the situation in the network given by the sensor information. One of the purposes of the control is to minimize leakage and to maintain appropriate hydraulic pressures for the consumers. By controlling the distribution of hydraulic pressures in the network, pipe breaking could be lessened and water could be conserved. However to achieve this kind of control, it is necessary to accurately estimate the distribution of hydraulic pressures through the network and to properly control the valves.

Available mathematical models that explore pressure regulation or leakage minimization problem through optimal control valve settings could be divided into types. In the first type, the following mathematical statement for network pressure regulation is used as an objective function to be minimized[6),7)].

$$J=\sum_{j=1}^{nn}(H_j-H_j^T)^2\rightarrow \min \tag{1}$$

where H_j is the head at node j, H_j^T is the required target head at the same node and n_n

is number of junction nodes. The foregoing function is to be minimized under the following constraints: (i) mass continuity should be satisfied for each node of the water supply network, (ii) sum of the head losses around a closed loop must be equal to zero, and (iii) Minimum and maximum head constraint for each node in the network.

In the second type, the total amount of leaked water from the network will be used as an objective function to be minimized as the following equation[5].

$$J=\sum_{i}^{n_p} K_i L_i p_i^{1.18} \rightarrow \min \tag{2}$$

where n_p number of links, K_i is an unknown experimental coefficient depends on the value of service pressure, age of the pipe, deterioration of the pipe and the soil properties, L_i pipe length and p_i average service pressure of the studied pipe. Eq. (2) is to be minimized under the same constraints mentioned before.

For the previous two methodologies the required optimal valves settings are embedded in the constraints represented by the sum of the head losses around a closed loop, in which the hydraulic analysis of the network is performed using the Hazen-Williams empirical equation or Darcy-Weisbach in conjunction with Colebrook-White formula. The search-analysis frameworks of this optimization problem (see **Fig. 1**) require several repetitive analysis of pipe network.

Obtaining an optimal control of the distribution of system service pressures in a municipal water distribution networks has always faced combinatorial problems due to its complexity, scale of the problem, number of hydraulic variables to be optimized, variation of water demand and the difficulty in estimating the roughness coefficient of old pipes. Regarding that all the available algorithms deal directly with the pressure regulation problem from an optimization point of view which requires a computational time depending on the water network size and the used optimization method and in most cases required also a network simplification method which is considered as another optimization problem. By other words, the total computational could be divided into three main parts; optimization algorithm time which depends upon the used optimization method, the time required to solve the different components of the water distribution network in order to compute the value of the objective function and the time required to deal with the real water supply network if no simplification technique has been used. In addition to the previous three difficulties, the estimation of the uncertainty values of water supply networks presented by nodal network demands and pipes roughness could lead to significant errors in the application of the developed model.

In this paper, we suggest to improve the available control schemes by using a pre-

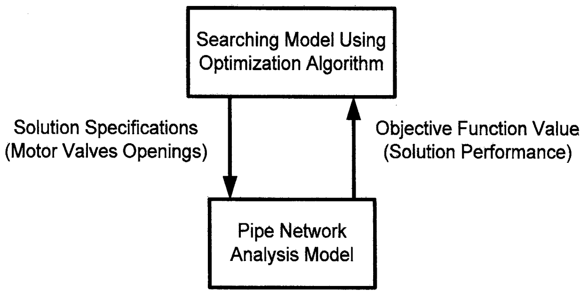

Fig. 1 Search-analysis frameworks.

optimization method based on learning from historically cases of valves operation. For the proposed algorithm, two different Artificial Neural Networks (ANNs), unsupervised and supervised learning algorithms, were applied to find the optimal electric motor valves setting to regulate pressure and minimize water leakage in supervisory water supply networks. First, a Feed-Forward Back-Propagation algorithm (FFBP), a supervised ANN, was applied to predict the values of hydraulic pressures at the locations of different pressure gauges. Second, a Self-Organizing Maps (SOM), an unsupervised ANN, was applied to regulate hydraulic pressure. Using the proposed model as a pre-optimization method could prevent all errors resulting from applying optimization models, save its computational time and provides us with an on-line pressure regulation method.

2. Materials and Methods

2.1 Feed-Forward Back-Propagation (FFBP) algorithm

The FFBP algorithm is well-known approach for the prediction applications. In this type, the network usually consists of an input layer, some hidden layers and an output layer. In its simple form, each single neuron is connected to other neurons of a previous layer through adaptable synaptic weights. Knowledge is usually stored as a set of connection weights and training is the process of modifying the connection weights. The network uses a learning mode, in which an input is presented to the network along with the desired output and the weights are adjusted so that the network attempts to produce the desired output. The weights, after training, contain meaningful information whereas before training they are random and have no meaning. The success of applying such supervised neural networks on any problem depends on training the net with sufficient range of data that spans a broad range of conditions. In this study we used the typical FFBP algorithm, which was presented by Rumelhart and McClelland[8]. For more details about ANN models and learning procedures, readers may refer to the literature[9),10),11]. General properties of neural networks, as well as their application for the prediction and forecasting of water resources variables, have been thoroughly covered in a number of publications[12),13]. For background information the reader could refer to this literature; only specific properties of the neural networks employed are given here. FFBP algorithm has been used in the present study to estimate pressure values at different pressure gauges of the application example of the supervisory water supply network.

2.2 Self-Organizing Maps (SOM)

The SOM is relatively a simple unsupervised neural network used for the categorization of input patterns into a finite number of classes. SOM consists of two layers units, the input units which are a one-dimensional array that provides simulation to a usually two-dimensional array of map space units (output units) and all units in the input layer are fully connected with the units in the output layer (**Fig. 2**). The neurons of the output layer which is preferably arranged into two dimensional grids for better visualization are connected to adjacent neurons by a neighborhood relation dictating the structure of the map. The arrangement of the output layer neurons are usually distributed in rectangular or hexagonal arrangement. Generally it is preferable to use the hexagonal lattice, because it does not favor horizontal and vertical directions as much as rectangular array[14].

When an input vector x is sent through the network, each neuron k of the output network, which is also called competitive layer computes the distance between the weight

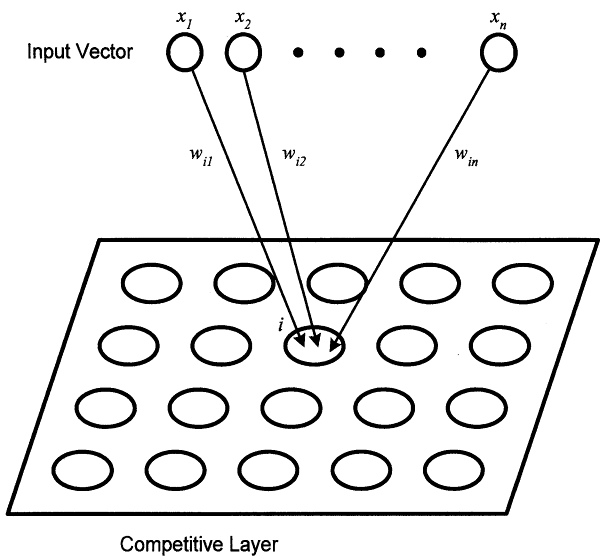

Fig. 2 Structure of SOM network.

vector w and the input vector x. Among all the output neurons, the so-called winning unit or Best-Matching Unit (BMU) is determined by the similarity between the weight vector w on that unit and the input vector x. For an input vector x, the BMU is determined by

$$\|x-w_c\|=\min_i\{\|x-w_i\|\} \tag{3}$$

in which the subscript c refers to the winning unit (BMU), $\|...\|$ is the distance measure and i refers to all units in the competition layer, in Eq. (3) each unit in the two-dimensional output layer is identified by a single subscript for simplicity. Accordingly, a second winning unit will be determined with respect to the second input vector, and so forth. At the end of competition only one unit in the competitive layer wins in corresponding to one input vector.

For the BMU and its neighborhood neurons, the weight vectors w are updated by the SOM learning rule.

$$w_i(t+1)=\begin{cases} w_i(t)+\alpha(t)h_{ci}(t)(x(t)-w_i(t)) \rightarrow if & i\in N_c(t) \\ w_i(t) & \rightarrow else \end{cases} \tag{4}$$

where α is the learning rate at time t; h_{ci} so-called neighborhood function that is valid for the neighborhood N_c.

The value of α varies from 0.0 to 1.0, and it controls the rate of learning. An α of 1.0 means it learns a new example as soon as it is presented. However, it forgets all previous examples of that class. Similarly, an α of 0.0 means that the network does not learn at all, but classifies new examples based on previous experiences only. The neighborhood function $h_{ci}(t)$ is a time-variable and a decreasing function $[h_{ci}(t)\rightarrow 0$ when $t\rightarrow\infty]$. It is often represented by a Gaussian function as follow

$$h_{ci}(t)=e^{-d_{ci}^2/2\sigma(t)^2} \tag{5}$$

where σ is the neighborhood radius at time t and $d_{ci}=\|r_c-r_i\|$ is the distance between map units c and i on the map grid. The training is usually performed in two phases. In the first phase, relatively large initial learning rate and neighborhood radius are used. In the second phase both learning rate and neighborhood radius are small right from the beginning. This procedure corresponds to first tuning the SOM approximately to the same space as the input data and then fine-tuning the map.

There are two different styles of training strategies. In sequential training the weights are updated each time when an input vector is presented. In batch training the weights are only updated after the presentation of all input vectors. In many applications, batch training type is the preferred option, as it forces the search to move in the direction of the true gradient at each weight update. However, several researchers suggest using the sequential type, as it requires less storage and "...makes the search path in the weight space stochastic... which allows for a wider exploration of the search space and, potentially, leads to better quality solutions"[13),15)].

After some training steps, the SOM will arrange high-dimensional input data along its two-dimensional output space such that similar inputs are mapped onto neighboring regions of the map which means that the similarity of the input data is preserved within the representation space of the SOM. Usually, in the SOM application, in order to ensure that all variables of any input vector x receive equal attention during the training process, it is important to normalize the input vectors to unit length before the training steps.

To measure the ability of SOM in arranging the different input vectors through its two-dimension grid, usually two evaluation criteria could be applied to measure the quality of SOM; resolution and topology preservation. For identifying and measuring the resolution of the SOM, we compute the quantization error[14)] which is the average distance between each data vector and its winning unit (BMU). The topographic error which used to present the accuracy of the training map in the preserving topology is also calculated. This error represents the proportion of all input data vectors for which first and second BMUs are not adjacent for the measurement of topology preservation. The topographic error can be calculated as follows[16)]:

$$\epsilon_t = \frac{1}{N}\sum_{k=1}^{N} u(x_k) \tag{6}$$

where N is the number of input vectors; $u(x_k)$ is 0.0 if the first and second BMU's of x_k are next to each other, other wise $u(x_k)$ is 1.0

2.3 Case study and data used

In order to illustrate the capability of the proposed pressure regulation model, Block 12 of the supervisory Fukuoka City water supply network is selected as a case study. In this Block (**Fig. 3**), there are 54 nodes, 74 pipes, and 9 inflows from outside the network at nodes 1, 3, 10, 17, 20, 41, 50, 51 and 54. For the telemeters attached to the network, there are 7 flow meters (*M1*, ..., *M7*), 20 electric motor valves (*V1*, ..., *V20*) and 11 pressure gauges (*P1*, ..., *P11*). It is noticed from **Fig. 3** that flow meters are connected to the main inlets and outlets and a valve is connected adjacent to each flow meter in order to control the flow entering or leaving the block. Motor valves are operated by remote control while pressure gauges and flow meters fitted to distribution pipes are monitored. One of the main objectives of the supervisory control of the water network of Block 12 is to regulate the pressure in all the network nodes between an upper target value (32 m) and a lower value (24 m). The values of flow rate passing each flow meter, the opening percentage of each motor valve and the pressure intensity at each pressure gauge are recorded every minute. The analyzed data of this study are based on one minute data for all flow meters, pressure gauges, and motor valves for a randomly selected two days (Saturday and Sunday, 9th and 10th of November

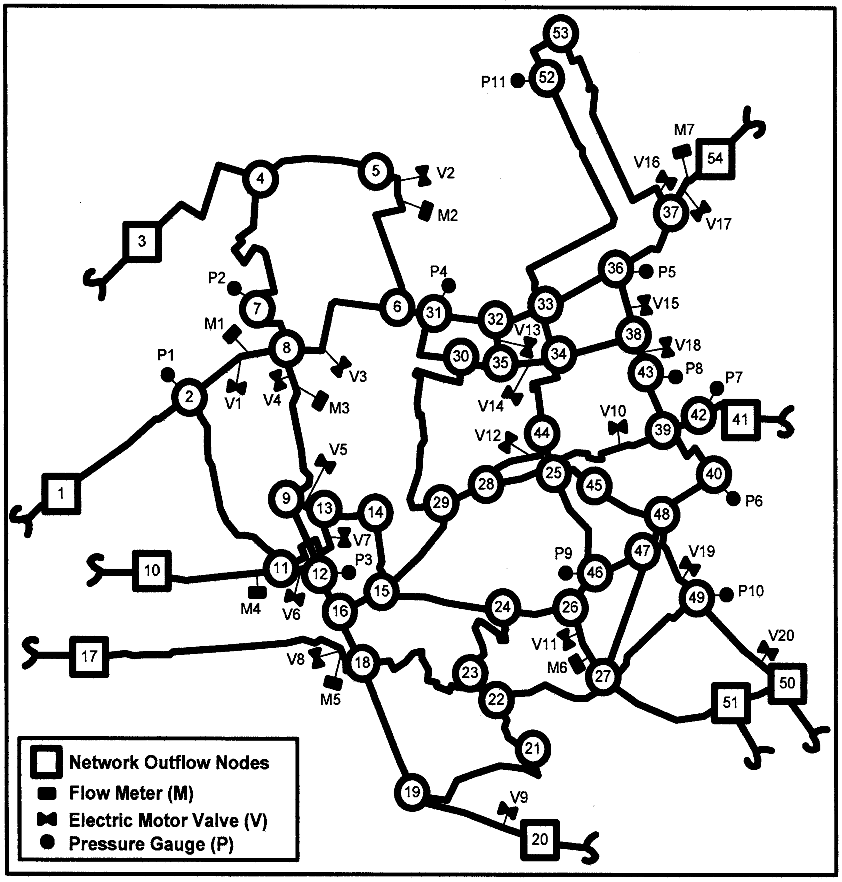

Fig. 3 Skeletonized water distribution network of Block 12.

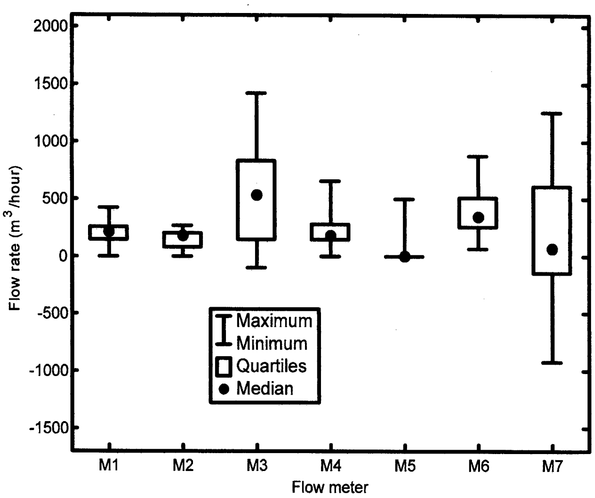

Fig. 4 Box-whisker plots for the 7 flow meters of Block 12.

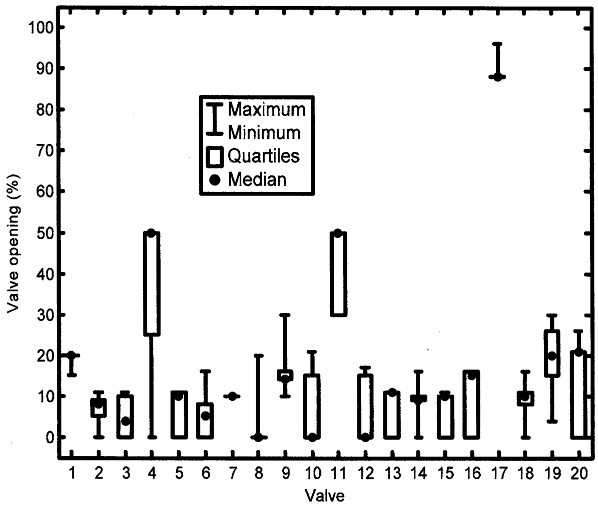

Fig. 5 Box-whisker plots for the 20 electric motor valves of Block 12.

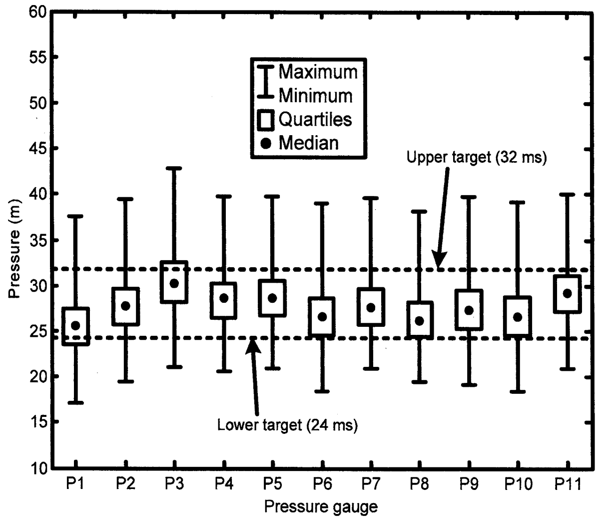

Fig. 6 Box-whisker plots for the 11 pressure gauges of Block 12.

2002). The total number of data for each telemeter equal to 2880 (total minutes during those two days).

Figures 4, 5 and **6** show the box-whisker plots for all the 38 telemeters of Block 12 for the data set used in this study. The box-whisker plots show the median, upper and lower quartiles, and also the maximum and minimum recorded values for each telemeter.

3. Estimating Hydraulic Pressures

One of the main steps of the proposed method which will be presented in the next section is to estimate the pressure values at the location of the 11 pressure gauges of the application example. Therefore, the FFBP model presented in this section will be used to perform this task.

The number of input and output nodes in the FFBP model is determined according to the nature of the studied problem, in the proposed model the number of input nodes are set to the total number of flow meters and electric motor valves (27 nodes) while the output nodes number are set to the total number of pressure gauges (11 nodes). Regarding that the dimension of the input vector is large; it is useful in this situation to reduce the dimension of the input vectors. Using one of the most effective procedures for performing this operation (principal component analysis) the total number of input nodes has been reduced from 27 to 7 nodes. The number of hidden layers and hidden nodes which depends on the complexity of the mathematical nature of the problem is determined by trial and error. One hidden layer with 40 nodes is found to be suitable to describe the relationship between the input and output variables. All transfer functions in the hidden and output layer are hyperbolic tangent functions.

Other additional information used in the model formulation is as follows: the mean squared error is used as an error function, batch mode of training is used in which all weights and biases are updated after presentations of all training vectors, the maximum number of learning counts is 2000, the initial weights and biases are randomly selected between -1 and 1, learning rate during training processes is 0.01 while the momentum constant is 0.9.

In developing the FFBP model, a cross-validation technique is used in which the data set is divided into three subsets; a training set, validation set and testing set. The odd minutes data are used for training while the even minutes data are divided to two subsets one for validation and the other for testing. A data pre-processing has been used because it may have a significant effect on model performance, all original data of input and output vectors of the three previous sets are scaled separately in the range of the used hyperbolic tangent functions (-0.9 to 0.9).

Figures 7 and 8 show scatter plots of the model estimated data versus observed data for pressure gauges P1 (lowest mean value) and P3 (highest mean value), respectively. The plotted results for those two pressure gauges are an example of the results obtained while all

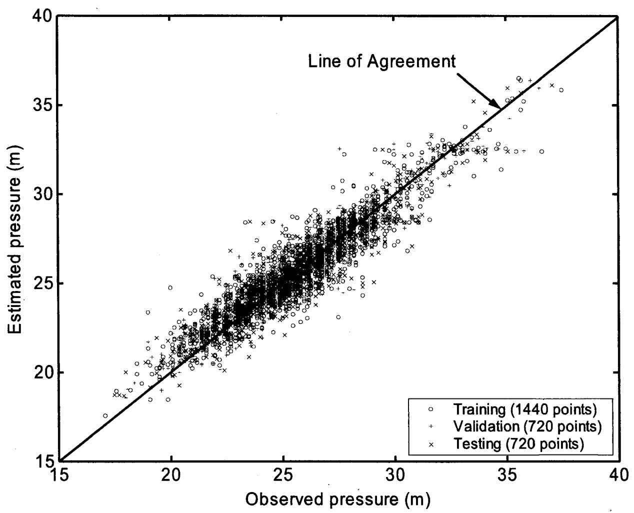

Fig. 7 Observed and estimated pressure values for pressure gauge P1 (lowest mean value).

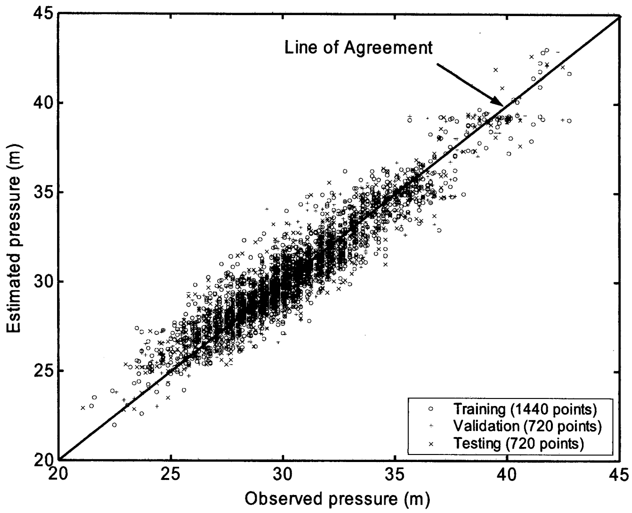

Fig. 8 Observed and estimated pressure values for pressure gauge P3 (Highest mean value).

remaining pressure gauges show same trend. In both figures there is good agreement between both estimated and observed data for all model sets, the training set (1440 points), the validation set (720 points) and the testing set (720 points). The Root Mean Square Error (RMSE) for the estimated pressure values is relatively acceptable for all pressure gauges; it varies between 1.066 m at pressure gauge P9 and 1.148 m at pressure gauge P3.

It is important to notice that the applicability of the FFBP model presented in this section is limited to the water distribution network of Block 12 City and also to the upper and lower values of telemetry data used in the training phase.

4. Proposed Method

Hydraulic pressures in water supply networks depend on several factors and values related to the system operation. The values that could affect the pressure at any node of the network could be divided into passive and active values. Passive variables are constant in any loading condition. Examples of passive values are pipes diameters and lengths.

For active variables, their values are changeable over the different loading conditions. In a water supply network operated without pumps (similar to that of Block 12), active variables are different nodal demands, electric motor valve openings, outflows from the network and hydraulic pressure at fixed grade nodes. The system responses due to all those variables (passive and active) are different pipes flow and the hydraulic pressure at all network nodes.

System response variables which should be included in the formulation of an optimal pressure regulation model in a water supply network could be simplified by selecting the hydraulic pressure at the location of all pressure gauges of the studied network. By this assumption, we have neglected all the pipes discharge as they are not related to the hydraulic pressure; and the hydraulic pressure in the majority of water supply network nodes because selecting the location of pressure gauges is considered as a good indication of hydraulic pressure in the different water supply network sub-areas.

All active variables could be determined from the reading of flow meters and electric motor valves attached to the pipes of the water network. Nodal demands are represented indirectly in the flow meter readings. Total nodal demands equal to the summation of flow meter readings connected to the water network main entrances excluding any internal flow meter reading. For pipes lengths and diameters which are considered as passive variables they could be considered in the system as embedded values because they have always fixed values. By other word, the water supply network system could be efficiently represented by: (i) flow meters readings, (ii) pressure gauges values; and (iii) electric motor valves openings. This assumption is correct when the system is represented by sufficient number of telemeters and more accurate results are obtained with the increase of the total number of observation points.

The model presented in this paper used the Self-Organized Map (SOM) to classify flow meter readings of well regulated pressure cases. After that a simulation step of flow meter readings of unregulated vectors is performed. Each vector of flow meter used for training or simulation has corresponding electric motor valve vector and hydraulic pressure vector. When the simulation step is performed, the electric motor valve vector of the unregulated case will be replaced by that of the regulated one and the resultant pressure will be tested in that case using the FFBP model presented in the previous section.

4.1 Model formulation

The application of SOM method as a pre-optimization tool for regulating pressure in water supply networks between an upper and lower target values is presented below and is illustrated in **Fig. 9**:

1. Input the simulation data which contains p vectors; each has p_1 readings from flow meters, p_2 readings from pressure gauges and p_3 readings from electric motor valves. The input matrix (I) has the dimension of $p \times (p_1 + p_2 + p_3)$.

2. Partition the input matrix (I) into two matrices (two groups of data) according to the values of pressure gauges. Group A presented by the matrix (I_{reg}) in which all p_2 components of its pressure gauges vectors are well regulated within the required target range ($>=$24m and $<=$32m). Group B presented by the matrix (I_{unreg}) in which any p_2 components of its pressure gauges vectors fall outside the target range ($<$24m or $>$32m). The size of I_{reg} is $p_{reg} \times (p_1 + p_2 + p_3)$ and the size of I_{unreg} is $p_{unreg} \times (p_1 + p_2 + p_3)$ in which $p_{reg} + p_{unreg} = p$.

3. With the data of group A presented by the matrix I_{reg} we will find a suitable size of SOM. This SOM will be constructed with the assistance of flow meters readings only, which represent the actual water demand of the network when the p_2 pressure points are well regulated and also the openings of the p_3 electric motor valves give a good operation case of the water supply network. The matrix used to construct a suitable SOM named FM_{reg} and it is a sub-matrix of I_{reg} and has the dimension of $p_{reg} \times p_1$.

4. Normalize the values of FM_{reg} between 0 and 1 so that each component of FM_{reg} will receive equal attention during the training process.

5. Assume a hexagonal arrangement of the output layer of the SOM which is preferable over the rectangular arrangement because it does not favor horizontal and vertical directions[14]. The minimum arrangement dimension is 2×2 while the maximum size is related to the problem size. In our application example which will be presented in the next section, the SOM size doesn't exceed 30×30.

6. Initialize and train the selected SOM subjected to Eqs. (3), (4) and (5).

7. Evaluate the selected SOM by computing both topographic and quantization errors.

8. Check convergence based on the values of topographic and quantization errors. If the

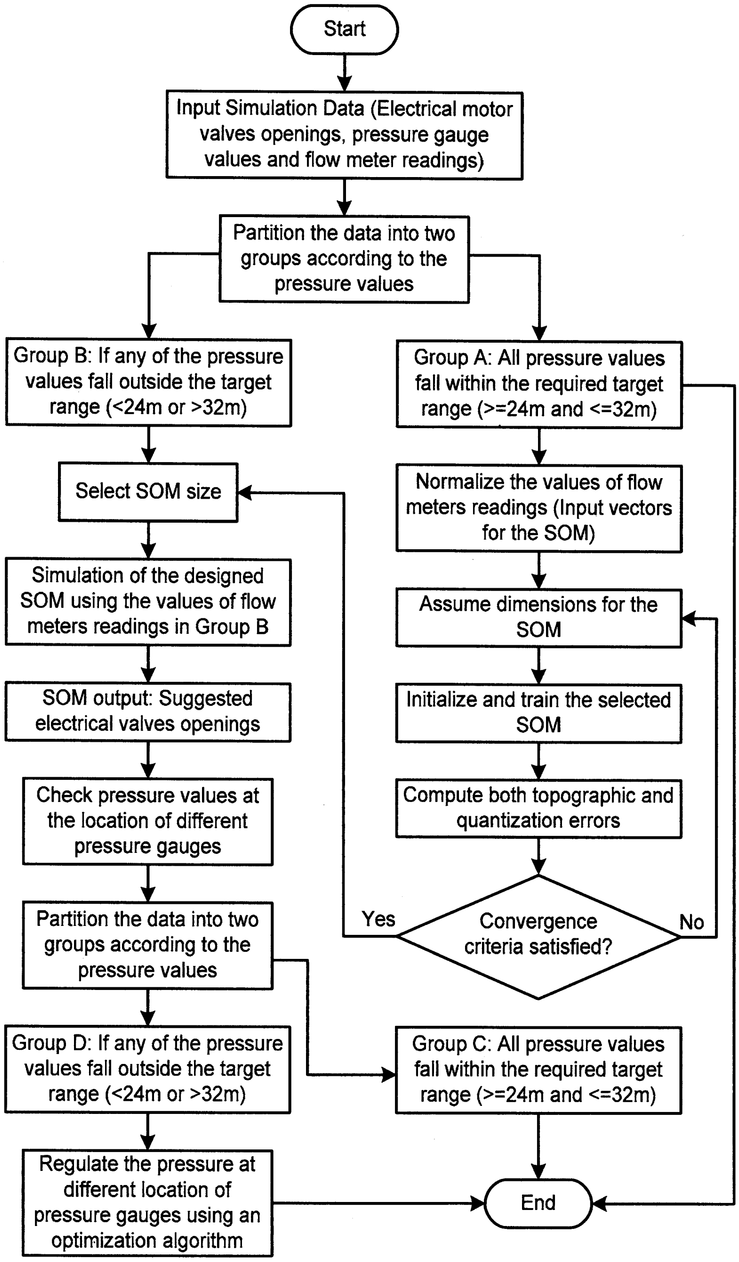

Fig. 9 Flow chart of the SOM model used for regulating pressure in water supply networks.

convergence criteria are satisfied, continue; otherwise return to step 5 by selecting another hexagonal arrangement.

9. Now we have an SOM of X neurons in x-direction and Y neurons in y-direction trained with the matrix FM_{reg} which represents the readings of flow meters of the different well regulated pressure cases. The SOM has $X \times Y$ neurons. In some units of the trained SOM there is a possibility that they didn't receive any flow meter vectors after the training processes. For the remaining units of SOM, they have one or more electric motor valve

vector of p_3 components which corresponds to a well regulated pressure case of flow meter readings and pressure gauges values. In this step we will make a simulation of the designed SOM using the values of flow meter readings in group B presented by a matrix FM_{unreg} which is a sub-matrix of I_{unreg} and has the dimension of $p_{unreg} \times p_1$.

10. Put $N=1$ in which N represents the unit number of the trained SOM and the maximum value of N is $X \times Y$.

11. For unit number N on the trained SOM. If there is a presence of vectors from both FM_{unreg} and FM_{reg}, step 12; otherwise step 13.

12. Replace all electric motor valve vectors of FM_{unreg} by that of FM_{reg} associated with unit number N. go to step 14.

13. This step is related to the units of SOM in which they have a minimum of one vector from FM_{unreg} and no vectors from FM_{reg}. In that case we will search in the neighborhood of unit N which varies between 2 units for the four units located at the corners of SOM and 6 units for any middle unit of the SOM. Replace all electric motor valve vectors of FM_{unreg} by that of FM_{reg} associated with the neighborhood units of unit number N.

14. Check pressure values at the p_2 locations of different pressure gauges using the modified matrix FM_{unreg}. Store the results according to the values of pressure gauges. Store in group C when pressure gauges vectors are well regulated within the required target range ($>=$24m and $<=$32m) while store in group D when any component of pressure gauges vectors fall outside the target range ($<$24m or $>$32m). Predicted values of hydraulic pressures are evaluated using the FFBP model presented in the previous section.

15. $N=N+1$; if $N<=X \times Y$ then step 11. Otherwise continue (step 16).

16. Evaluate the SOM efficiency to regulate the pressure in the location of pressure gauges between an upper and lower target values. The (I_{unreg}) matrix is divided now into two matrices; (O_{reg}) which is the summation of group C cases and have a dimension of $t_{reg} \times (p_1 + p_2 + p_3)$ and (O_{unreg}) which is the summation of group D cases and have a dimension of $t_{unreg} \times (p_1 + p_2 + p_3)$; $t_{reg} + t_{unreg} = p_{unreg}$. The SOM efficiency could be defined as the percentage of vectors that their hydraulic pressure has been regulated as:

$$\text{SOM efficiency} = \frac{t_{reg}}{p_{unreg}} \times 100 \tag{7}$$

For group D cases presented by (O_{unreg}) matrix, an evolutionary computing technique to regulate the pressure in all the network nodes within the required target range could be used[17]. The objective function of this model could be presented in the form of network pressure regulation; Eq. (1) or as the total amount of leaked water from the network; Eq (2).

5. Application

This section presents an application example for the proposed algorithm of applying SOM for optimal pressure regulation in water supply networks. The selected data set is for two days with one minute interval of telemetry data recorded from the 7 flow meters, 11 pressure gauges and 20 electric motor valves attached to the different nodes and pipes of Block 12 of Fukuoka City water supply network. Therefore, the total number of vectors (p $=2880$), $p_1=7$, $p_2=11$ and $p_3=20$.

The input matrix (I) has the dimension of 2880×38 representing all the data used for this application. As a first step, we have divided the input matrix (I) into two matrices according to the actual hydraulic pressure recorded at the 11 pressure gauges. The first matrix (I_{reg}) which will be used to construct the SOM has a 949 well regulated vectors in which all pressure gauges vectors are well regulated within the required target range (>= 24m and <=32m). The second matrix (I_{unreg}) has a dimension of 1931×38. For the pressure gauge vectors of this matrix at least one component of the 11 components falls outside the desired target range (<24m or >32m).

5.1 Map size

Input vectors to the SOM are all sets of flow meter readings for (I_{reg}) matrix known as FM_{reg} matrix with dimension of 949×7. The total numbers of training vectors are 949. Each vector contains 7 components representing the associated 7 flow meter readings. Normalization range of input vectors is from 0 to 1, batch mode of training is used, initial SOM weights are set randomly between 0 and 1, adaptive learning rate is used which varies between 0.1 to 0.4 and maximum allowed number of epochs is 200.

Different map size has been evaluated by calculating both topographic and quantization errors. All possible two-dimensional map sizes which vary from 2 to 30 neurons have been tested. In general, increasing the map size will increase the topographic error which is calculated using Eq. (6) while brings more resolution into mapping when the quantization error decreases. The map size selected to present the different classification of flow meters is hexagonal lattice with middle size of 11×16. At that size, the topographic and quantization errors equal 2.0021 and 1.4561, respectively. That's mean that there is only 19 vectors in which the first and second BMU aren't adjacent. For that selected map size the required number of epochs for convergence is 88.

Figure 10 shows the trained SOM units using a hexagonal lattice of size 11×16. The number written in the upper area of each unit indicates the total number of hits associated with those units (number of BMU) and representing the total number of flow meter vectors of well regulated cases (949 vectors). The minimum and maximum number of hits recorded for any neuron in the trained SOM is 0 and 26, respectively. The number of units in which they haven't any hits representing a regulated case in **Fig. 10** is 52 out of the total 176 units.

5.2 Regulating the hydraulic pressure

After training the SOM with the FM_{reg} matrix, a simulation step is done using the FM_{unreg} matrix which is a sub-matrix of I_{unreg} and has the dimension of 1931×7. The matrix FM_{unreg} represents all cases of flow meters values in which at least one component of the pressure gauge vectors fall outside the desired target range. The number written in the lower area of each unit in **Fig. 10** indicates the number of flow meter vectors in which the pressure is unregulated. The minimum and maximum number of hits recorded for any neuron in the simulated SOM is 0 and 165, respectively. The number of units in which they haven't any hits representing unregulated vectors in **Fig. 10** is 50 out of the total 176 units.

For the 176 units presented in **Fig. 10**, four types of units could be determined according to the existence of regulated and unregulated vectors:

1. Units that have both regulated and unregulated vectors. For those units, at least one hit is recorded in both the upper and lower areas of any unit. Total number of units for that type is 108 units. To regulate the pressure for those units, we have replaced all electric motor valve vectors of FM_{unreg} by that of FM_{reg} associated with each unit. After replacement, we have predicted the values of hydraulic pressures at the location of the 11 pressure gauges using the FFBP model.

2. Units that have only unregulated vectors. For those units, at least one hit is recorded in the lower area of the unit and there isn't any hit recorded in the upper area. Total number of units representing that type is 18 units. In that case, we have replaced all electric motor

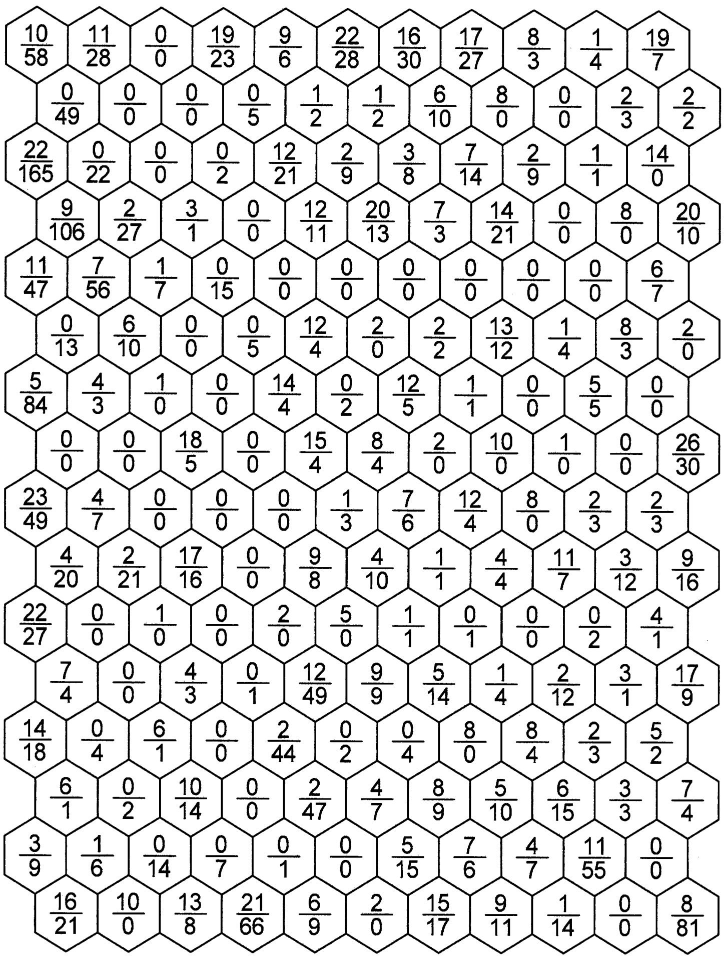

Fig. 10 SOM map of size 11×16 trained with the flow meters vectors of well regulated cases (upper number in each unit). The lower number in each unit indicates the number of vectors in which the pressure is unregulated.

valve vectors of FM_{unreg} by that of FM_{reg} associated with the neighborhood units. The number of the neighborhood units varies between 2 units for the four units located at the corners of SOM and 6 units for any middle unit of the SOM. After replacement, we have predicted the values of hydraulic pressures at the location of the 11 pressure gauges using the FFBP model.

3. Units that have only regulated vectors. For those units, at least one hit is recorded in the upper area of the unit and there isn't any hit recorded in the lower area. Total number of units representing that type is 16 units. Those units could be useful if the SOM is supplied with additional future data. The electric motor valve vectors associated with those units could replace any unregulated vectors fall in those units and by applying the methodology of *Type 1* the hydraulic pressure could be regulated.

4. Units which lack any number of hits in both upper and lower areas. In **Fig. 10** there are 34 units of that type. Those units could be useful if the SOM is supplied with additional future data. The electric motor valve vectors associated with the neighborhood units could replace any unregulated vectors fall in those units and by applying the methodology of *Type*

2 the hydraulic pressure could be regulated.

5.3 Results analysis

Comparing the situation before and after applying the proposed algorithm for the application example, **Table 1** shows a comparison of regulated and unregulated cases. The distribution of vectors is based on two-hour interval for the studied two-days. As example, first row in **Table 1** shows that there are 100 regulated vectors from 9:01 to 11:00 and 140 vectors at the same time interval in which hydraulic pressure exceeds the desired target range. After applying the SOM model, the number of the regulated cases is increased by 83 (improved using SOM model) and for the remaining 57 cases the SOM model failed to improve them. The SOM efficiency for that time interval is 59.29% calculated using Eq. (7). Maximum number of unregulated vectors after applying the SOM model is 73 recorded 5:01 to 7:00. **Table 1** shows that there are small number of regulated cases during night time; 15 cases from 1:01 to 3:00 and 19 cases from 3:01 to 5:00. After applying the proposed model, 170 and 171 cases has been improved for those two time intervals showing the effectiveness of the proposed algorithm during the night time in regulating the pressure when water demand is minimum. The maximum efficiency of the proposed model is 95.14% recorded for the time interval between 21:01 and 23:00. For all the application cases, 1437 cases have been improved out of the 1931 case with a percentage of 74.42%.

Table 2 compares between the pressure status at the location of each pressure gauge before and after applying the proposed algorithm. This comparison is based on the 1931 cases in which the pressure was unregulated. For example, pressure gauge (P1) has 863 cases in which the pressure is less than 24 m but after applying the SOM model those cases have been reduced to 242 cases only. For the same pressure gauge, the 87 cases in which the pressure is bigger than 32 m have been reduced to 22 cases after applying the proposed algorithm. The highest value of events in which the pressure exceeds the upper target limit (32 m) is recorded at the location of pressure gauge (P3) with 857 cases that have been reduced to 222 cases after applying the SOM model. All details of the pressure status for the 11 pressure gauges are presented in **Table 2**. When considering all the pressure gauges as one group, results show that: (i) the 996 cases in which the pressure is below than 24 m have been reduced to 256 cases, (ii) the 908 cases in which the pressure is bigger than 32 m have been

Table 1 Hourly distribution of pressure vectors.

Time		No. of well Regulated vectors	No. of unregulated Vectors	No. of vectors Improved using SOM	No. of unregulated vectors after applying SOM	SOM Efficiency (%)
From	To					
9:01	11:00	100	140	83	57	59.29
11:01	13:00	123	117	89	28	76.07
13:01	15:00	82	158	134	24	84.81
15:01	17:00	102	138	109	29	78.99
17:01	19:00	137	103	67	36	65.05
19:01	21:00	137	103	79	24	76.70
21:01	23:00	96	144	137	7	95.14
23:01	1:00	41	199	149	50	74.87
1:01	3:00	15	225	170	55	75.56
3:01	5:00	19	221	171	50	77.38
5:01	7:00	44	196	123	73	62.76
7:01	9:00	53	187	126	61	67.38
Sum		949	1931	1437	494	74.42

Table 2 Comparison between pressure vectors before and after applying the proposed SOM model.

Pressure Gauge	Before applying SOM				After applying SOM			
	P<24m	P>32m	P<24m and P>32m	P>=24 and P<=32m	P<24m	P>32m	P<24m and P>32m	P>=24 and P<=32m
P1	863	87	0	981	242	22	0	1667
P2	245	226	0	1460	83	139	0	1709
P3	33	857	0	1041	18	222	0	1691
P4	117	312	0	1502	50	157	0	1724
P5	95	325	0	1511	54	161	0	1716
P6	481	166	0	1284	134	105	0	1692
P7	213	223	0	1495	93	139	0	1699
P8	567	116	0	1248	200	82	0	1649
P9	246	215	0	1470	93	144	0	1694
P10	459	178	0	1294	116	91	0	1724
P11	52	434	0	1445	36	174	0	1721
All gauges	996	908	27	0	256	238	0	1437

reduced to 238 cases, (iii) there are 27 cases in which some points are below than 24 m and one or more gauges read more than 32 m, all those cases have vanished after applying the SOM model, and (iv) In general, 1437 cases have been improved out of the 1931 cases with a percentage of 74.42%.

Table 3 compares between the pressure status at the location of pressure gauges before and after applying the SOM model taking into consideration the number of unregulated components. For example, there are 88 cases out of the 1931 in which all the pressure gauges exceed the target range. Those cases are divided into 84 cases in which all the pressure gauges read more than 32 m and 4 cases that all pressure gauges read less than 24 m. after applying the SOM model those 88 cases have been reduced to 40 cases indicating an improvement of 48 cases with a 55 percent of improvement. Considering this comparison the highest number of unregulated vectors are 814 cases when there is one pressure gauge read a value exceed the target range (481 cases bigger than 32 m and 333 cases less than 24 m); those 814 cases has been reduced to 204 cases only (percentage of improvement is 75%). From the last column in **Table 3** we could deduce that the percentage of improvement increase with the decrease of the number of unregulated components in pressure vectors (55% when 11 components are unregulated and 75% in case of one unregulated component).

In **Table 4**, distribution of pressure vectors components with the different zones of water pressure are shown for all the 11 pressure gauges. There are nine different zones starting from "Zone I-A" to "Zone IV". Those nine zones are determined according to the hydraulic pressure before and after applying the proposed model. For example, "Zone I-A" presents any case in which the pressure is below the lower target value of 24 m before and after applying the model, in "Zone I-B" the pressure is below the 24 m before applying the SOM model and bigger than 32m after applying the model and "Zone IV" presents the situation in which the pressure is well regulated in both cases. Boundaries which determined the different zones are mentioned in the second and third rows in **Table 4**.

According to the different nine zones, three main classes could be determined based on the efficiency of the proposed model to regulate the pressure at the 11 tested points; (i) for "Zone I" and "Zone II" the pressure is unregulated after applying the SOM taking into consideration that "Zone II" presents the cases in which the original pressure is within the

Table 3 Comparison between pressure vectors components before and after applying the proposed SOM model.

No. of unregulated points in each vector	No. of regulated points in each vector	No. of unregulated points in each vector (P<24m)	No. of unregulated points in each vector (P>32m)	Number of vectors	Total number of vectors	No. of vectors Improved using SOM	No. of unregulated vectors after applying SOM	Percentage of improving (%)
11	0	0	11	84	88	48	40	55
		11	0	4				
10	1	0	10	24	62	42	20	68
		9	1	1				
		10	0	37				
9	2	0	9	29	53	35	18	66
		8	1	1				
		9	0	23				
8	3	0	8	16	36	28	8	78
		7	1	1				
		8	0	19				
7	4	0	7	19	65	49	16	75
		6	1	2				
		7	0	44				
6	5	0	6	35	101	82	19	81
		6	0	66				
5	6	0	5	54	138	108	30	78
		4	1	2				
		5	0	82				
4	7	0	4	45	145	114	31	79
		3	1	3				
		4	0	97				
3	8	0	3	43	154	116	38	75
		2	1	12				
		3	0	99				
2	9	0	2	78	275	205	70	75
		1	1	5				
		2	0	192				
1	10	0	1	481	814	610	204	75
		1	0	333				
0	11	0	0	949	949	------	------	------
		Sum		2880	2880	1437	494	

target range, (ii) "Zone III" presents the cases in which the SOM model has succeeded to regulate the pressure, and (iii) "Zone IV" presents the cases in which the pressure status has not changed; in both situations the pressure is within the target range. Total number of cases associated with all the nine zones for the 11 pressure gauges are presented in details in **Table 4** in which the last two rows indicate the number of cases recorded for the minimum and maximum value of each pressure vector. The summation of any row shown in **Table 4** presents the total number of unregulated vectors before applying the proposed model (1931 vectors).

Figures 11 and **12** show the relation between the initial and estimated pressure recorded at the location of pressure gauges (P1) and (P3), respectively. Boundaries of the different nine zones are plotted in both figures. Pressure gauges (P1) and (P3) are selected as representative of all pressure gauges because the majority of low pressure values are recorded at the location of (P1) while the majority of high pressure values are recorded at the location of (P3). "Zone III" in **Figs. 11** and **12** show the number of cases improved using

Table 4 Distribution of pressure vectors components with the different zones of water pressure.

Zone		I				II		III		IV
		A	B	C	D	A	B	A	B	
Pressure before applying SOM		<24m	<24m	>32m	>32m	>=24m and <=32m	>=24m and <=32m	<24m	>32m	>=24m and <=32m
Pressure after applying SOM		<24m	>32m	<24m	>32m	<24m	>32m	>=24m and <=32m	>=24m and <=32m	>=24m and <=32m
Pressure gauge	P1	110	13	17	0	119	5	736	74	857
	P2	9	8	21	29	66	89	215	189	1305
	P3	0	13	3	132	5	87	30	712	949
	P4	2	8	10	47	40	100	105	257	1362
	P5	4	7	9	51	43	101	82	267	1367
	P6	32	7	6	31	95	68	443	128	1121
	P7	13	8	7	42	72	90	193	173	1333
	P8	54	9	15	12	137	55	498	95	1056
	P9	11	9	3	44	73	97	232	162	1300
	P10	24	7	4	34	85	53	431	137	1156
	P11	1	7	4	70	28	100	47	357	1317
Min. pressure		139	13	0	0	104	0	884	71	720
Max. pressure		0	13	0	149	4	89	4	773	899

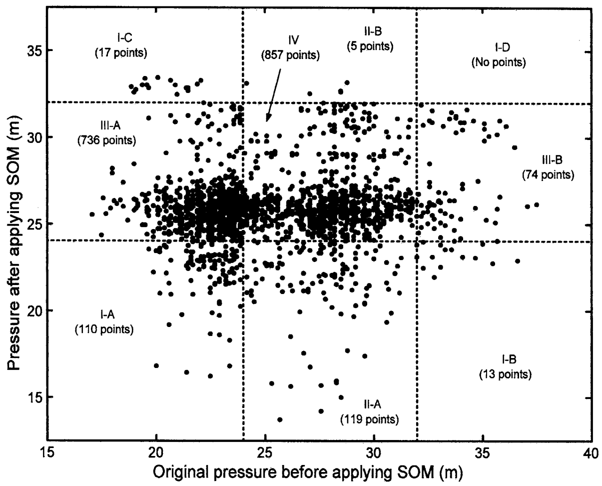

Fig. 11 Relation between initial and estimated pressure recorded at (P1).

the SOM model. For pressure gauge (P1), zones III-A and III-B show the improved 736 and 74 cases, respectively. For pressure gauge (P3), zones III-A and III-B show the improved 30 and 712 cases, respectively.

Points in zones II-A and II-B represent the cases that are regulated before applying the SOM model for the related pressure gauge. For those cases there is an existing of one or more pressure gauges in which their pressure is unregulated. After applying the proposed method, the pressure becomes unregulated for the studied pressure gauges and there is a probability that the pressure status is improved in other pressure gauges. In general all cases

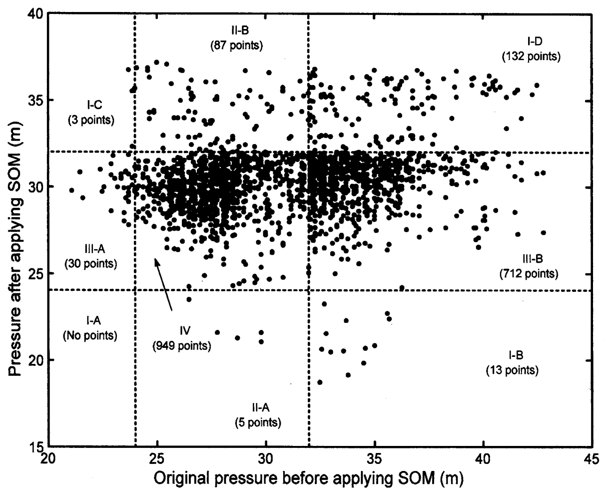

Fig. 12 Relation between initial and estimated pressure recorded at (P3).

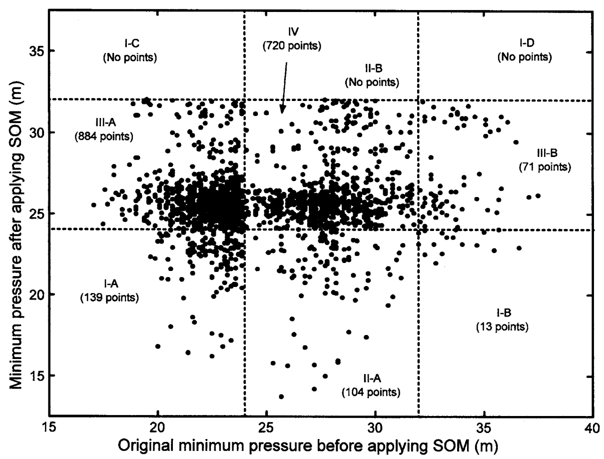

Fig. 13 Relation between minimum initial and estimated pressure for all pressure gauges.

presented in zone II are classified under the same category of the cases of zone I in which the SOM model fails to regulate the pressure within the required target range. Highest number recorded in zone II-A is 137 cases existing at the location of pressure gauge (P8) while the highest number recorded in zone II-B is 101 cases at the location of pressure gauge (P5); see **Table 4** for more details.

In **Fig. 13**, the relation between the minimum initial and estimated hydraulic pressure in all pressure vectors is plotted. Zones III-A and III-B show the improved 884 and 71 cases, respectively. The same procedure is repeated for the maximum initial and estimated hydraulic pressure in all pressure vectors (see **Fig. 14**). The improved 4 and 773 cases are shown in zones III-A and III-B, respectively.

The proposed model has successfully improved 1437 cases out of the 1931 unregulated cases for the application example. For the remaining 494 cases, results of other calculation

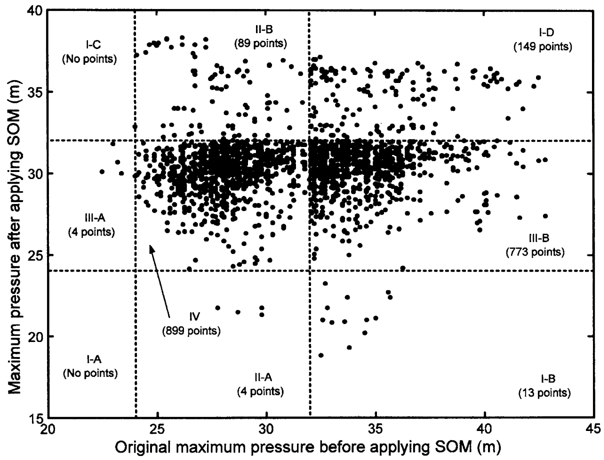

Fig. 14 Relation between maximum initial and estimated pressure for all pressure gauges.

show an improvement in pressure status using the proposed model in 298 cases.

6. Conclusions

The study performed in this paper demonstrates the potential of applying Self-Organizing Maps (SOM) for finding optimal electric motor valve settings to regulate the pressure in water supply networks between upper and lower target values. The presented application of Block 12 of Fukuoka City water distribution network achieves good performance for the short term data set used. The SOM has successfully regulated 1437 cases out of the 1931 unregulated cases in the application example. For the remaining 494 cases, an improvement in pressure status using the proposed model are recorded in 298 cases comparing to the situation before applying the SOM model.

The model considered in this paper is classified as an "expert system model" as it is based on learning from the past historical data. This model offers the opportunity of being used directly for the on-line optimal pressure regulation in water supply networks without any need to deal with the skeletonized network presented or knowing the real nodal demand. The presented model uses a FFBP model to predict the hydraulic pressure at the location of different pressure gauges.

The proposed method shows high efficiency in case of regulating pressure during daily operation with high performance for night time when there is an increase of hydraulic pressure in the network (from 1:00 a.m. to 5:00 a.m.). Efficiency of SOM algorithm depends on the number of points to be regulated. For the application example when all points are unregulated the percentage of improvement is 55% while in case of one unregulated point the efficiency increases to 75%.

This paper presents analysis and comparison for the situation of pressure before and after applying the proposed method. This comparison is done for all the location of the 11 pressure gauges. Significant improvement was recorded at the location of pressure gauges (P1) and (P3). Those two pressure gauges are the most critical points as there pressures for the majority of cases exceed the required target range before applying the proposed method.

For demonstrating the efficiency of the proposed method, we have used a short-term data set. The same procedure could be repeated for a long-term data set taking into consideration that the size of SOM should be big enough to represent the majority of operational cases.

This paper evaluates also the potential of applying FFBP algorithm for pipe network analysis, the model presented for the short-term data set of telemetry data has been applied successfully for the prediction of hydraulic pressures at the location of the 11 pressure gauges. The RMSE varies between a minimum value of 1.066 m at pressure gauge (P9) and a maximum value of 1.148 m at pressure gauge (P3).

Principal component analysis has been implemented successfully to reduce the big number of neurons in the input layer of the FFBP model. The number of neurons has been reduced from 27 to 7. By applying the principal component analysis the total number of weights and biases to be determined by the FFBP models is reduced.

FFBP model parameters has a significant effect on the model results, all those parameters have been selected by trial and error based on the most recommended values in the literature. A sensitivity analysis of different model parameters should be considered in the future improvement of presented models.

References

1) World Health Organization (2001). "Leakage management and control – a best practice manual." Publication of World Health Organization (WHO), http://www.who.int/
2) Fukuoka City Waterworks Bureau (1997). "Water leakage prevention measures." Publication of the Fukuoka City Waterworks Bureau.
3) UN-habitat (1998). "Best practices database in improving the living environment." United Nations Human Settlements Programme Database, http://www.bestpractices.org
4) MLIT (2002). "Water Resources in Japan." Publication of Land and Water Bureau, Water Resources Department, Japanese Ministry of Land, Infrastructure and Transport. http://www.mlit.go.jp/
5) Jowitt, P.W. and Xu, C. (1990). "Optimal valve control in water distribution networks." Journal of Water Resources Planning and Management, ASCE, 116(4), 455–472.
6) Miyaoka, S. and Funabashi, M. (1984). "Optimal control of water distribution system by network flow theory." IEEE Transactions on Automatic Control, AC-29(4), 303–311.
7) Savić, D.A. and Walters, G.A. (1995). "An evolution program for optimal pressure regulation in water distribution networks." Engineering Optimization, 24(3), 197–219.
8) Rumelhart, D.E., and McClelland, J.L. (1986). "Parallel distributed processing." MIT Press, Cambridge, Mass.
9) Caudhill, M., and Butler, C. (1992). "Understanding neural networks: Computer explorations." Vols. I and II, MIT Press, Cambridge, Mass.
10) Freeman, J.A., and Skapura, D.M. (1991). "Neural networks, Algorithms, applications, and programming techniques." Addison-Wesley, Reading, Mass.
11) Haykin, S. (1999). "Neural networks comprehensive foundation." Prentice Hall International, Second Edition.
12) Smith, M. (1993). "Neural networks for statistical modeling." Van Nostrand Reinhold, New York.
13) Maier, H.R., and Dandy, G., C. (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications." Environmental Modelling & Software, 15, 101–124.
14) Kohonen, T. (2001). "Self-Organizing Maps." third edition, Springer-Verlag, Berlin.

15) Hassoun, M.H. (1995). "Fundamentals of Artificial Neural Networks." MIT Press, Cambridge.
16) Kiviluoto, K. (1996). "Topology Preservation in Self-Organizing Maps." Proceedings of the International Conference on Neural Networks (ICNN'96), 294-299.
17) Awad, H., Kawamura, A., Jinno, K., and Kuno, Y. (2003). "Evolutionary computing techniques for optimal pressure regulation in water distribution networks", Annual Journal of Japan society of Civil Engineers, JSCE, Vol.47, pp.865-870.