# A Method for Estimating the Diameter of Water Pipes Using Artificial Neural Networks of the Multilayer Perceptron Type

Jacek Dawidowicz Białystok University of Technology, ul. Wiejska 45A, 15-351, Białystok, Poland

Abstract-Designing a water distribution system is inherently associated with hydraulic calculations, the primary purpose of which is to select the diameters of water pipelines. Computer programs may choose diameters, but most often, this task is up to the designer. It is necessary to control flow velocity; at the same time, however, it is also necessary to minimise pressure losses in the water mains network. In order to improve the above process, an artificial neural network was designed, which, after hydraulic calculations, evaluates the accuracy of the diameters selected on a classification basis. The output layer of the neural network consists of ten neurons corresponding to the nominal diameters of the water pipes. Classification of the correct pipeline diameter is done using "one-of-N" encoding, that is, only one neuron from the output layer is activated, thus selecting the correct diameter. Based on data from hydraulic calculations, the neural network diagnoses the diameter of the pipelines along individual sections of the water supply network and proposes appropriate values or accepts the existing ones.

# Keywords—water distribution system; hydraulic calculations; diameter selection; artificial neural networks; classification

# I. INTRODUCTION

Designing a water distribution system always involves hydraulic calculations, the primary purpose of which is to select the diameters of pipelines to carry the water. In this paper, steady-state analysis was applied for peak hour water demand (PHD). Flow rates for these conditions are the basis for ascertaining the diameters of the water mains [1]. Computer programs may choose diameters, but most often, this task is up to the designer. An artificial neural network was designed, which undertook classification having performed the hydraulic calculations and having estimated the accuracy of the selected pipeline diameters.

## II. REVIEW OF THE USE OF ARTIFICIAL NEURAL NETWORKS IN THE CALCULATIONS OF WATER DISTRIBUTION SYSTEMS

A review of artificial intelligence methods including artificial neural networks for the calculation of water distribution systems is provided in the article [2]. In the hydraulic computation of water distribution systems, artificial neural networks are used to assist in the taring of the simulation models [3, 4]. In their article [5], proposed recursive neural networks for calculating flows and pressure losses in looped water distribution systems. During simulations of hydraulic water distribution systems, pressure losses in individual water pipes are calculated using, among other things the Darcy-Weisbach formula. This requires calculations of the linear resistance coefficient, most often using the iterative method. Numerous studies proposed artificial neural networks for calculating coefficients of linear resistance [6,7], which allow calculation time to be shortened. Calculation modules, based on artificial neural networks, were also introduced into simulation methods used in the real-time control of water supply networks. The task of neural computing, in this case, is to simplify the computational model and accelerate calculations [8,9]. In the study [10], the purpose of simulation calculations, using neural networks, was the optimisation of pressure in the individual nodes of the network taking into account the predicted water demand. In the article [11], a neural network was described for the calculation of pressure losses in water pipes. The method of estimation of pressure levels and the pattern of pressure zones, using artificial neural networks, is described in the article [12].

#### III. AN INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS

In this paper, an artificial neural network of the multilayer perceptron type was used for classification [13,14]. The number of neurons in the output layer corresponds to the number of calculated parameters. The number of neurons in the hidden layer should be determined in the process of training the neural network. Initially, the number of neurons based on Kolmogorov's Theory can be assumed:

$$K = 2 \cdot N + 1 \tag{1}$$

where: K – is the initial number of neurons in the hidden layer, N- is dimension of the input vector  $X = [x_1, ..., x_n]^T$ . In the output layer of the network, the Softmax activation function was used according to the formula:

$$y = \frac{e^{S}}{\sum_{m=1}^{M} e^{S_{m}}}$$
  $y = (0...+1)$  (2)

where M - is the number of neurons of the output layer, S - is the value of the function of the post-synaptic potential.

This is an exponential function, the value of which is additionally normalised so that the sum of the activation of all M neurons of the output layer of the network is equal to 1. In addition to the fact that network signals are the basis for recognising the appropriate class, the output values of individual neurons can be interpreted, as an estimation of the probabilities that they belong to a given class [15]:

In the training of the network, the backpropagation error method was applied at first and then the quasi-Newton method. The training was done with validation, hence the set of all training examples was subdivided into training, validation and test subsets. In the network training process, a function of the Entropy Multiple Error (EME) was adopted. Entropy where each class corresponds to one neuron in the output layer (M> 1) [14]:

$$E_{EME} = -\sum_{t=1}^{T} \sum_{m=1}^{M} d_m^{(t)} \log y_m^{(t)}$$
(3)

where M - is the number of neurons of the output layer, T - is the number of examples in the training set, T = 1, ..., T - is the number of the training example, d - is the reference value of the neural output signal, y - is the calculated value of the neuron in the output layer.

The basis for classifier rating is an overall classification error defined as:

$$\varepsilon = \frac{n_{mis}}{n_{all}} \tag{4}$$

where  $n_{mis}$  - is the number of misclassified training examples,  $n_{all}$  - is the number of all training examples subjected to classification.

Another quality measure is relevance of the classification  $\eta$  defined as a complement to the value of one of the overall classification errors, namely:

$$\eta = 1 - \varepsilon = 1 - \frac{n_{mis}}{n_{all}} = \frac{n_{cor}}{n_{all}}$$
(5)

where  $n_{cor}$  - is the number of correctly classified training examples.

## IV. A REVIEW OF ARTIFICIAL NEURAL NETWORK Structures in the Evaluation of the Diameters of Water Supply Pipes

The use of artificial neural networks requires the preparation of data with an appropriate number of training examples. To this end, information on 36 existing medium and small sized water distribution systems was gathered; hydraulic calculations were then carried out for them. The results of the calculations were used to compile training data. For hydraulic calculation, the methodology took into account with

withdrawals occur along the pipe  $q_{sec}$  and the flow rate at the beginning  $Q_s$  and at the end  $Q_e$  of the calculated section of the water pipe, with the length L with the absolute roughness coefficient k, being assumed. Searching for a suitable, multilayer perceptron structure was started from a network with one hidden layer and with a set of 5 input variables: L,  $Q_s$ ,  $Q_e$ ,  $q_{sec}$ , k.

The individual neurons of the output layer correspond to the nominal diameters of the water pipes that were adopted as follows:

- DN90, DN110, DN160, DN225 for PE100 polyethylene pipes of the SDR17 series (EN 12201-2:2011),
- DN250, DN300, DN350, DN400, DN450, DN500 for ductile iron pipes (EN 545: 2010).

Sensitivity analysis of individual input variables, that is, their effect on the training error of the neural network, indicated that the variable  $q_{sec}$  can be omitted because its absence does not cause a significant deterioration of the neural network. Searching for a proper neural network structure with one hidden layer was started from 9 neurons in the hidden layer, with their number being then increased by the initial value in subsequent attempts to teach larger networks. A larger network was constructed when the neural network did not improve or where only insignificant improvement was observed after long learning cycles, suggesting an insufficient number of neurons, processing layers, or learning algorithms becoming obstructed by the local minimum. Networks with the same structure were trained several times to prevent them from becoming obstructed by the local minimum (Hornik 1991).

As a result of the training process of neural networks as a final solution has been selected neural network included in Table 1, pos. 4. The above neural network has the lowest number of the training error  $E_{EME}$  for the validation subset and a very high accuracy of classification. In the table 1, K – is the number of neurons in the hidden layer of the multilayer perceptron,  $E_{EME}\left(L\right)$  – is an error in the training subset,  $E_{EME}(V)$  – is an error in the validation subset,  $E_{EME}(T)$  – is an error in the test subset,  $\eta(L)$  – is the relevance of classification for the training subset,  $\eta(V)$  – is the relevance of the classification for a validation subset,  $\eta(T)$  – is the relevance of the classification for a test subset.

Adopted artificial neural network consists of the following elements:

- the input layer with neurons for 4 input variables: L, Q<sub>s</sub>, Q<sub>e</sub>, k.
- one hidden layer constructed of 36 neurons with a logistic activation function,
- an output layer made up of 10 neurons with the Softmax activation function, corresponding to the diameters of the pipes.

Table 2 shows the results of the classification in the form of an confusion matrix for the neural network from Table 1, pos. 4 for the test subset.

No.	К	E <sub>EME</sub> (L)	E <sub>EME</sub> (V)	E <sub>EME</sub> (T)	η(L)	$\eta(V)$	η(Τ)
1	9	0.293010	0.339211	0.409987	0.889465	0.883746	0.891195
2	18	0.244072	0.256846	0.330387	0.906092	0.897047	0.909018
3	27	0.234540	0.252878	0.296792	0.904229	0.895717	0.903166
4	36	0.205843	0.241232	0.300815	0.913009	0.899175	0.905560
5	45	0.221163	0.271595	0.323957	0.918994	0.908220	0.913541
6	54	0.226723	0.237513	0.296843	0.907688	0.895451	0.904762

TABLE I. NEURAL NETWORKS FOR THE ASSESSMENT OF THE DIAMETER OF WATER PIPES

	DN 90	DN 110	DN 160	DN 225	DN 250	DN 300	DN 350	DN 400	DN 450	DN 500
Total	615	656	5190	403	401	333	297	216	198	121
Correct	597	584	471	392	387	330	264	183	97	109
Incorrect	1	0	2	0	3	0	0	0	0	1
Undetermined	17	72	46	11	11	3	33	33	111	11
DN90	597	0	0	0	0	0	0	0	0	0
DN110	1	584	1	0	0	0	0	0	0	0
DN160	0	0	471	0	0	0	0	0	0	0
DN225	0	0	1	392	3	0	0	0	0	0
DN250	0	0	0	0	387	0	0	0	0	0
DN300	0	0	0	0	0	330	0	0	0	0
DN350	0	0	0	0	0	0	264	0	0	0
DN400	0	0	0	0	0	0	0	183	0	0
DN450	0	0	0	0	0	0	0	0	87	1
DN500	0	0	0	0	0	0	0	0	0	109

Figure 1 shows a diagram of network for the classification of pipeline diameters showing the activation of the neuron in the output layer and indicating selection of the diameter assigned to it.



PIPELINES

Table 3 shows the values of neuron activation with the Softmax of the output layer of the neural network for one of the training examples. Determination of the affiliation to one of the

classes consists in the selection of the neuron of the output layer, in which a value close to 1 appears; with the other neurons, the values should be close to 0, however this is practically unobtainable. For this reason, two threshold values are introduced, viz., the acceptance threshold and the rejection threshold, to which the activation level of the neurons of the output layer is compared. The activation level above the acceptance threshold results in the object being accepted into the class, while the activation value below the reject threshold indicates that the object is not affiliated to any class. In this task, the acceptance threshold is set at 0.95, while the rejection threshold is set at 0.05. If this condition is not met, the case is described as indefinite, that is, the network is unable to classify the object into any class. In the example in Table 5, the activation level is 0.8281449, which causes the example to remain unclassified, but it can be assumed- and with very high probability - that the diameter is DN110. As a consequence, the neuron activation values for the unclassified pipelines should be analysed during calculations also.

FABLE III. ACTIVATION VALUES OF NEURONS IN THE OUTPUT
LAYER OF THE NEURAL NETWORK FOR THE EVALUATION OF
PIPELINE DIAMETERS

Pipeline diameter assigned to the output layer neuron	Activation of the output layer neuron
DN90	0.00000000000000000
DN110	0.82814490000000000
DN160	0.1718551000000000
DN225	0.00000000000000000
DN250	0.00000000000000003
DN300	0.00000000000000000
DN350	0.00000000000000000
DN400	0.0000000000000100
DN450	0.00000000000000001
DN500	0.00000000000000002
Activation sum:	1.00000000000000000

# V. SUMMARY

The above methodology and neural network was developed having in mind programmes for the hydraulic calculation of water distribution systems in which it will evaluate the results obtained from the viewpoint of the pipeline diameters adopted. Based on data from hydraulic calculations, the neural network diagnoses the diameter of the pipelines on the individual sections of the water supply network and proposes appropriate values or accepts the existing ones. The neural network obtained is characterised by a high index of classification accuracy, that is, in evaluating the diameter. However, it should be borne in mind that there may be cases of non-classification of the diameter or of misclassification, albeit this is somewhat unlikely. In the results of the calculations, the neuron activation value from the output layer of the neural network should appear for each diameter. The neural network was designed solely to assist the process of selecting diameters. The final decision on the selection of a diameter belongs to the person undertaking the calculations.

#### REFERENCES

- K. Lansey and L.W. Mays, "Hydraulics of water distribution systems", in: Water distribution systems handbook, L.W. Mays, Ed. New York: McGraw-Hill, 2000, pp.4.1-4.29.
- [2] A. Czapczuk, J. Dawidowicz, and J. Piekarski, "Metody sztucznej inteligencji w projektowaniu i eksploatacji systemów zaopatrzenia w wodę", Rocznik Ochrona Srodowiska (ISSN 1506-218X), vol. 17, no. 2, 2015, pp. 1527-1544 (in Polish).
- [3] S. Lingireddy and L.E. Ormsbee, "Neural networks in optimal calibration of water distribution systems", in: Artificial neural networks for civil engineers: advanced features and applications, I. Flood and N. Kartam, Eds., Cincinnati: ASCE, 1998, pp 53–76.
- [4] J. Saldarriaga, R. Gómez, and D. Salas, "Artificial intelligence methods applicability on water distribution networks calibration", in: Critical transitions in water and environmental resources management, G. Sehlke, D.F. Hayes, and D.K. Stevens, Eds., Reston: ASCE, 2004, pp. 1-11, doi:10.1061/40737(2004)248.
- [5] C. Xu, F. Bouchart, and I.C. Goulter, "Neural networks for hydraulic analysis of water distribution systems, Innovation in Computer Methods for Civil and Structural Engineering", Cambridge: Civl-Comp Press, 1997, pp. 129-136.
- [6] S. M. Besarati, P. D. Myers, D. C. Covey, and A. Jamali, "Modeling friction factor in pipeline flow using a GMDH-type neural network",

Cogent Engineering, vol.2, iss. 1, 2015, https://doi.org/10.1080/23311916.2015.1056929.

- [7] W. H. Shayya and S. S. Sablani, "An artificial neural network for noniterative calculation of the friction factor in pipeline flow", Computers and Electronics in Agriculture, vol.21, iss. 3, pp. 219–228, 1998.
- [8] L. Yongchao and L. Wending, "Water Supply System of Telemeter and Remote Control Based on Neural Fuzzy Control Technique", Proceedings of the 5th International Symposium on Test and Measurement Conference ITSM, 2003, pp.1269-1272.
- [9] M. Damas, M. Salmeròn and J. Ortega, "ANNs and Gas for predictive controlling of water supply networks", Proceedings of the IEEE-INNSENNS international joint conference on neural networks (IJCNN2000), vol. 4, Como" IEEE, 2000, pp 365–368, doi: 10.1109/IJCNN.2000.860799.
- [10] A. Bargiela, "High performance neural optimization for real time pressure control", Proceedings of High Performance Computing Conference, HPC Asia'95. Chap. AL34, Taipei, 1995, pp. 1-8.
- [11] A. Czapczuk, J. Dawidowicz and J. Piekarski, "Application of Multilayer Perceptron for the Calculation of Pressure Losses in Water Supply Lines", Rocznik Ochrony Środowiska, vol. 19, 2017, pp. 200-210.
- [12] J. Dawidowicz, "Evaluation of a pressure head and pressure zones in water distribution systems by artificial neural networks", Neural Computing & Application, 2017, doi:10.1007/s00521-017-2844-8.
- [13] S. Haykin, "Neural networks: a comprehensive foundation", New Jersey: Prentice Hall Int., 1999.
- [14] C.M. Bishop, "Neural networks for pattern recognition", Oxford: University Press, 1996.
- [15] J. S. Bridle, "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition", in: Neurocomputing: Algorithms, Architectures and Applications, F. Fogelman-Soulie, J. Herault, Eds., New York: Springer-Verlag, 1990, pp. 227-236.