Rocznik Ochrona Środowiska

Volume 23	Year 2021	ISSN 1506-218X	pp. 835-844

https://doi.org/10.54740/ros.2021.057 open access

Received: 09 November 2021 Accepted: 14 November 2021 Published: 06 December 2021

Kohonen Artificial Networks for the Verification of the Diameters of Water-pipes

*Jacek Dawidowicz**
Białystok Technical University, Poland
https://orcid.org/0000-0003-1742-9463

Andrzej Czapczuk
F.B.I. TASBUD Capital Group, Warsaw, Poland
https://orcid.org/0000-0002-9293-1173

Wojciech Kruszyński
Białystok Technical University, Poland
https://orcid.org/0000-0002-4326-0705

**corresponding author's e-mail: j.dawidowicz@pb.edu.pl*

Abstract: The design of the water distribution system is inherently linked to the execution of calculations, which aim, among other things, to determine the flow rate through individual pipes and the selection of diameters at the appropriate speed. Each step in the calculations is followed by an evaluation of the results and, if necessary, a correction of the data and further calculations. It is up to the designer to analyse the accuracy of the calculation results and is time-consuming for large systems. In this article, a diagnostic method for the results of hydraulic calculations, based on Kohonen Network, which classifies nominal diameters [DN] on the basis of data, in the form of flows, has been proposed. After calculating the new variant of the water distribution system, the individual calculation sections are assigned to the neurons of the topological map of Kohonen Network drawn up for nominal diameters. By comparing the diameter used for the calculation, with the diameter obtained on the topological map, the accuracy of the chosen diameter can be assessed. The topological map, created as a result of labelling the neurons of the output layer of the Kohonen Network, graphically shows the position of the classified diameter, relative to those diameters with similar input values. The position of a given diameter, relative to other diameters, may suggest the need to change the diameter of the pipe.

Keywords: water distribution system, hydraulic calculations, diameter selection, Kohonen Network, evaluation of calculation results

1. Introduction

The design of the water distribution system is inherently linked to the execution of calculations, the purpose of which is, among other things, determination of the flow rate through the individual pipes, the selection of diameters for maintaining appropriate speeds, the calculation of pressure losses and pressure levels at the nodes. It seems that classical algorithms with a formalised course, can be supplemented with much more advanced computational techniques derived from the field known as artificial intelligence (Konar 2005, Bishop 1995). The scope of this approach includes such methods as artificial, neural networks, expert systems and genetic algorithms. In this paper, an attempt is made to supplement classical methodology by calculating water distribution systems using elements of artificial neural networks. A uni-directional artificial neural network, the so-called "neural network" was used in this paper, the Kohonen Network (Kohonen 2001, Kangas & Kohonen 1996).

An important challenge in the operation of water supply systems is the effective detection of leaks. In the article, by (Aksela et al. 2009) a method based on a self-organising map for the detection of leaks, in the water supply network, was presented. The data used for network training and validation consists of flowmeter readings and reported leak locations. The most important factor facilitating the self-organising, map-based modelling of leakages is the developed leakage function. The experimental results, presented, show that a model, trained on flow data, can detect leaks in a specific area of the water supply network. In the article by (Brentan et al. 2018) presented a grouping method based on self-organising maps, in combination with k-mean algorithms, in order to obtain groups that can be easily identified and used to make decisions supporting the design, operation and management of water distribution systems. In the article by (Blokker et al. 2016) the application of the self-organising, map technique of the SOM was analysed, in order to determine how this method could be used in the numerical analysis of water quality, in water distribution systems. An overview of the methods of artificial intelligence, including SOM, for water supply issues, is given in the work by (Czapczuk et al. 2015). The problem of assessing the accuracy of the selection of diameters of water supply pipes was also addressed in the work by (Dawidowicz et al. 2018) using the K-Nearest Neighbours method, and in the work also by (Dawidowicz 2012), where the method of inducing the rules of the expert system was used. A comparison of two artificial intelligence methods for predicting water supply failure is included in the paper (Kutylowska 2016).

2. Introduction to Kohonen Network

In the 1950's, the idea of a self-organising system, *i.e.,* one that changes its structure on the basis of information coming to it from the environment, the so-called

SOM - Self Organizing Map, was used for the first time. Kohonen used the concept of self-organisation for artificial neural networks and proposed a network called 'Self-organising mapping'.

This, today, but with various modifications, is the most popular type of self-organising network and is named after its inventor- Kohonen.

2.1. The Kohonen Network structure

Kohonen nets are used for a non-model classification. Their aim is to select from a certain population, described by a multi-dimensional data vector $\boldsymbol{X} = [x_1,x_2,\ldots,x_i,\ldots,x_N]^T$, possibly homogeneous groups (clusters) in terms of considered features (input variables). They consist of two layers: input and output. Figure 1 shows a two-dimensional network, while Figure 2 shows a two-dimensional network. The neurons of the input layer (i = 1,..., N), are only used for entering data into the network, without performing any processing. In the output layer of the network, there are radial neurons, hence it is called a *radial layer*. Individual radial neurons are connected to all inputs and a weight is assigned to each connection. The collection of all connection weights, for each radial neuron, creates a vector of weights $\mathbf{W}=[w_1, w_2,\ldots,w_i,\ldots, w_N]^T$, the so-called *prototype or codebook vector*. The number of neurons in the output layer is determined by the network designer. Neurons in the output layer are not connected to each other and do not transmit information to each other but are connected by a neighbourhood relationship that affects the way neurons learn.

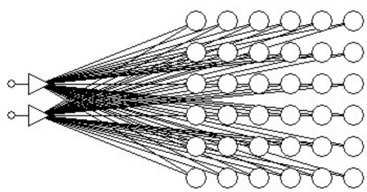

Fig. 1. Diagram of an example of the two-dimensional Kohonen Network for N = 2 (source: own study)

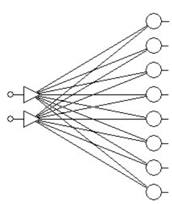

Fig. 2. Diagram of an example of the one-dimensional Kohonen Network for N = 2 (source: own study)

2.2. Learning the Kohonen network

Kohonen Networks are taught using an unsupervised learning algorithm (*self-organising learning*, *unsupervised learning*) where no *'d output values'* of the teaching data are used. In the case of Kohonen Network, we are dealing with *competitive learning,* so-called. Network learning is done by repeatedly showing examples of learning data, in the vector form X**,** along with modifications of the output neuron scales 'W**'**. The network is presented with additional input data, without information on the output signal which the network is to generate for a particular learning vector. The input signal is assumed to belong to one of several classes, but the classes sought are not known and the network seeks to detect them on its own. Similar input signals should be recognised as belonging to the same class. In this way, Kohonen Network tries to determine the structure of the data and the clusters of learning examples present in them.

After training the Kohonen Network, individual neurons should be assigned appropriate class labels, if known. Only then can the radial neurons act as classifiers. Each input signal is assumed to belong to one of several classes and the network output value identifies the class to which the signal belongs. After the learning process, each radial neuron of the output layer, or more precisely the vector of its weights (the so-called master /pattern vector), becomes the pattern or *"centre"* of a group of closely related input signals. After assigning the corresponding labels (names) of the individual classes to the individual neurons of the output layer, a so-called topological map is created. Class assignment is performed using the K-L Nearest Neighbours algorithm, in which a given neuron is assigned a label, based on the labels of the K-Nearest teaching cases. However, the condition is that at least L of K Neighbours has the same class, otherwise the label of the neuron will be "unknown".

The topological map graphically determines the position in the output layer of neurons, describing individual classes, their neighbourhood and the presence of clusters. In the case of a trained network, it is expected that similar input signals should elicit similar network responses, hence the arrangement of neurons, representing similar classes, should be similar on a topological map, forming certain groups.

3. Kohonen Network in the assessment of the diameters of water pipes

Numerical experiments were carried out to test the applicability of Kohonen's Network in assessing the diameters of the water distribution system. Sequential learning is used, i.e., learning examples are repeatedly presented to the network.

In order to compile a data set for the teaching neural networks, information on 33 existing medium *-and small-sized*- water supply systems was collected. Hydraulic calculations were performed for the above water distribution systems for different variants of water uptake from nodes, so as to obtain the widest possible range of data for teaching neural networks. Due to the large amount of data, a procedure was developed to convert the calculations' results for individual sections of the calculation wires to the appropriate format and save them in a set of training examples. Calculations were performed for different values of the absolute roughness coefficient. Based on the results of the hydraulic calculations, for the maximum water intake hour Q_{hmax}, 13,923 teaching examples were obtained. The calculation uses a methodology that takes into account nodal and sectional expenditure (Mielcarzewicz 2000). In this case, the teaching dataset was divided into two subsets: teaching and testing, covering 70% and 30% of the examples, respectively.

3.1. An overview of Kohonen Network solutions tested, in the assessment of the diameters of water pipes

Firstly, the network was trained in the form of a chain, consisting of 10 neurons in the output layer. In this case, at the learning stage, it is not possible to assign appropriate diameters to individual neurons, since the 'without a teacher' method was used. The purpose of this training was to verify whether the network automatically assigns input vectors corresponding to individual pipe diameters.

The set of input variables L, Q_p, q_{odc}., Q_k, k was assumed. The above network is shown schematically in Fig. 3.

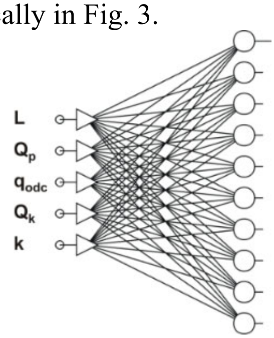

Fig. 3. Diagram of Kohonen one-dimensional network for assessing cable diameters (source: own study)

Kohonen Networks, with a square grid. The same set of input variables was used as for a one-dimensional network. The learning outcomes for these networks are presented in Table 1. The sensitivity analysis of the input variables for the Kohonen Network is presented in Table 2. The results indicate that two variables are relevant for this type of network Q_p and Q_k.

Network learning with two input variables: Q_p and Q_k described in Table 3. Sensitivity analysis for the above variables, showed that they are very important in the functioning of the network.

Table 1. Kohonen neural networks, for assessing wire diameters (5 input variables) (source: own study)

Basic data of neural networks:

Number of entries: 5
Input variables: L, Q_p, q_{odc}., Q_k, k
Output variable: none
Output layer function: Euclidean measure
Error function: Kohonen

Net no.	Network type	Number of neurons in the input layer	Number of neurons in the output layer	Error in the training set	Error in the test set	Accuracy of classification in the learning set	Accuracy of classification in the test set
1	Kohonen	5	10 (10x1)	0.210393	0.2090614	0.001436	0.001675
2	Kohonen	5	100 (10x10)	0.0857	0.08404	0.1953846	0.1809478
3	Kohonen	5	225 (15x15)	0.06057	0.05871	0.3802051	0.362853
4	Kohonen	5	400 (20x20)	0.04815	0.04718	0.4781538	0.4511728
5	Kohonen	5	900 (30 x 30)	0.06177	0.0641	0.4339118	0.4079371
6	Kohonen	5	1600 (40 x 40)	0.04854	0.05161	0.6235863	0.5745367
7	Kohonen	5	2500 (50 x 50)	0.03795	0.04232	0.7567988	0.6865344

Table 2. Sensitivity analysis of neural network input variables from Table 1 (source: own study)

Net no. from Table 7.19	Type of data subset	Assessment parameter of variable sensitivity	Kohonen's input variable				
			LD	Q_p	q_{odc}	Q_k	K
1	Learning set	Rank	5	2	3	1	4
		Error E_i	0.1910124	0.3235362	0.2023955	0.3239577	0.2002289
		Quotient	0.9078833	1.53777	0.9619873	1.539773	0.9516893
2	Learning set	Rank	4	2	5	1	3
		Error E_i	0.07702	0.2705157	0.06907	0.2708085	0.08325
		Quotient	0.8987833	3.156711	0.8059887	3.160128	0.9715154
3	Learning set	Rank	3	2	5	1	4
		Error E_i	0.05226	0.264514	0.04887	0.2646739	0.04952
		Quotient	0.8628372	4.36696	0.8067703	4.3696	0.8175303
4	Learning set	Rank	3	2	4	1	5
		Error E_i	0.04266	0.2625269	0.04186	0.262776	0.04051
		Quotient	0.8860484	5.452492	0.8694931	5.457665	0.8413751
5	Learning set	Rank	4	5	2	1	3
		Error E_i	0,06524	0,05822	0,188856	0,18910	0,06774
		Quotient	1,056091	0,9424292	3,057326	3,061276	1,096599
6	Learning set	Rank	4	5	2	1	3
		Error E_i	0,05317	0,04498	0,1844805	0,1848024	0,05361
		Quotient	1,095457	0,9268248	3,80085	3,807482	1,104571
7	Learning set	Rank	3	5	2	1	4
		Error E_i	0,0474278	0,03446	0,1831879	0,1834781	0,04701
		Quotient	1,249718	0,9081402	4,826982	4,834629	1,238826

A diagram of Kohonen net, in the form of a 10x10 rectangular grid is shown in Fig. 4.

Table 3. Kohonen neural networks for assessing wire diameters (2 input variables) (source: own elaboration)

Net no.	Network type	Neurons in the input layer	Neurons in the output layer	Error in the training set	Error in the test set	Accuracy of classification in the learning set	Accuracy of classification in the test set
Basic data of neural networks: Number of entries: 2 Input variables: Q_p, Q_k Output variable: none Output layer function: Euclidean measure Error function: Kohonen							
1	Kohonen	2	100 (10x10)	0.01997	0.0222	0.8285128	0.831738
2	Kohonen	2	225 (15x15)	0.01161	0.01303	0.8953846	0.898277
3	Kohonen	2	400 (20x20)	0.004524	0.005345	0.9368205	0.94136
4	Kohonen	2	625 (25x25)	0.003046	0.003473	0.9365128	0.935376
5	Kohonen	2	900 (30x30)	0.002189	0.002556	0.9758974	0.971757
6	Kohonen	2	1225(35x35)	0.001694	0.001897	0.9775385	0.969603

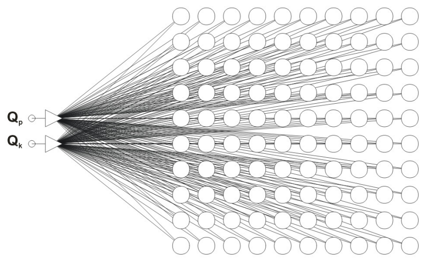

Fig. 4. Kohonen Network diagram in the assessment of cable diameters (source: own elaboration)

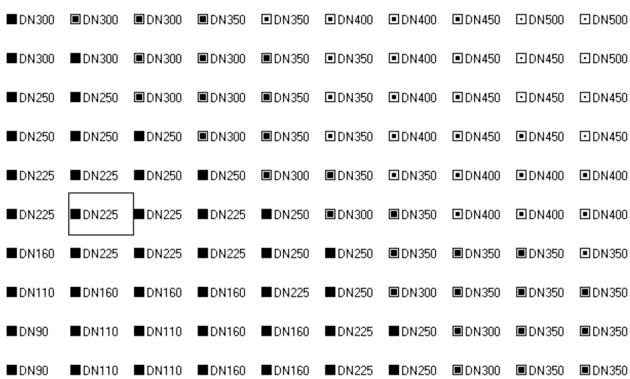

Fig. 5. An example of a topological map for the classification of pipe diameters (source: own elaboration)

Labels of individual pipe diameters were assigned to the neurons of the output layer of the Kohonen Network, thus determining their significance. In Kohonen Network, there is a regularity that says that the proximity of signals in the input space of the network equals the similarity of neighbouring objects on the topological map, described by these signals. Therefore, neurons describing the same diameters lie side by side in groups and gradually pass into diameters adjacent to the typeset.

3.2. A detailed description of *Kohonen Network* in the assessment of the diameters of water pipes

As a result of Kohonen neural networks learning, a structure with 2 neurons in the input layer and 900 neurons in the output layer was selected, arranged on a square grid with a width of 30 by 30 neurons (Table 3, item 5). The initial flow values $_p$ [l/s] in the Q branch are given at the network input together with the final flow Q_k [l/s]. In the case of the Kohonen Network, the effect of the remaining variables on the learning error was so insignificant that it was decided to abandon it.

Activation of Kohonen Network for the entire data set, after assigning labels describing the diameters of the wires to individual neurons, made it possible to evaluate the accuracy of the classifications obtained. The results presented in Table 4 indicate a certain number of incorrect classifications, which, however, represent a small percentage of all data. The accuracy of the classification for the learning set is 0.9758974, while, for the test set, it is 0.9717568.

Table 4. Results of the classification of water pipe diameters, using Kohonen Network for the test subset (source: own study)

	DN 90	DN 110	DN 160	DN 225	DN 250	DN 300	DN 350	DN 400	DN 450	DN 500
Total	568	598	535	457	299	428	462	419	248	164
Correct	558	555	510	448	293	424	457	409	247	159
Incorrect	10	43	25	9	4	3	2	9	0	3
Unspecified	0	0	0	0	2	1	3	1	1	2
DN90	558	39	0	0	0	0	0	0	0	0
DN110	10	555	22	0	0	0	0	0	0	0
DN160	0	4	510	9	0	0	0	0	0	0
DN225	0	0	3	448	1	0	0	0	0	0
DN250	0	0	0	0	293	2	0	0	0	0
DN300	0	0	0	0	3	424	2	0	0	0
DN350	0	0	0	0	0	1	457	9	0	0
DN400	0	0	0	0	0	0	0	409	0	0
DN450	0	0	0	0	0	0	0	0	247	3
DN500	0	0	0	0	0	0	0	0	0	159

3.3. Application of *Kohonen Network* in assessing the selected water pipe diameters of a water distribution system

Each calculation procedure is followed by an evaluation of the results and, if necessary, correction of the data and subsequent calculations.

The diagnostic method of Kohonen Network**,** classifies nominal diameters DN based on input data in the form of Q_p and Q_k. After the calculation of the new variant of the water distribution system, the individual sections of the calculation are assigned to the neurons of the topological map, drawn up for the nominal diameters. By comparing the diameter used for the calculation, with the diameter obtained on the topological map, the accuracy of the chosen diameter can be assessed.

A topological map, created as a result of the labelling neurons of the output layer, graphically shows the position of the classified diameter, relative to those diameters with similar input values. The position of a given diameter relative to other diameters may, for example, suggest the need to change the diameter of the duct, when the neuron describing the diameter, on a given section, is surrounded by neurons corresponding to other diameters.

4. Summary and conclusions

Various Kohonen Network structures, *viz*., the number and type of inputs and the size of the output layer were analysed. A sensitivity analysis was carried out to determine the impact of individual inputs on the way the network operates. A series of numerical experiments allowed a set of neural networks with different structures to be created and networks with the best parameters to then be selected.

Kohonen Networks can be used to assess the diameters of water supply lines. The advantage of this solution is the topological map, which graphically shows the position of a given diameter, relative to other diameters, depending on the parameters describing the computational section.

References

Aksela, K., Aksela, M., & Vahala, R. (2009). Leakage detection in a real distribution network using a SOM. *Urban Water Journal, 6*(4), 279-289.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition.* Oxford: Oxford Univ. Press.

Blokker, E. J. M., Furnass, W. R., Machell, J., Mounce, S. R., Schaap, P. G., & Boxall, J. B. (2016). Relating Water Quality and Age in Drinking Water Distribution Systems Using Self-Organising Maps. *Environments, 3*(2).

Brentan, B., Meirelles, G., Luvizotto, E., & Izquierdo, J. (2018). Hybrid SOM plus k-Means clustering to improve planning, operation and management in water distribution systems. *Environmental Modelling & Software, 106*, 77-88.

Czapczuk, A., Dawidowicz, J., & Piekarski, J. (2015). Artificial Intelligence Methods in the Design and Operation of Water Supply Systems. *Rocznik Ochrona Srodowiska, 17*, 1527-1544.

Dawidowicz, J. (2012). Expert System for Evaluation of Water Distribution System Created with an Inductive Inference. *Rocznik Ochrona Srodowiska, 14*, 650-659.

Dawidowicz, J., Kruszynski, W., Andraka, D., & Czapczuk, A. (2018). Assessing the Diameters of Water Pipes Using the k-Nearest Neighbours Method in the Calculations of Water Distribution Systems. *Rocznik Ochrona Srodowiska, 20*, 528-537.

Kangas, J., Kohonen, T. (1996). Developments and applications of the self-organizing map and related algorithms. *Mathematics and Computers in Simulation*, *41*, 3-12.

Kohonen, T. (2001). Self-Organizing Maps. *Springer Series in Information Sciences*, *30*, Springer-Verlag.

Konar, A. (2005). *Computational Intelligence: Principles, Techniques and Applications.* Springer-Verlag.

Kutylowska, M. (2016). Prediction of Water Conduits Failure Rate – Comparison of Support Vector machine and Neural Network. *Ecological Chemistry and Engineering a-Chemia i Inzynieria Ekologiczna A, 23*(2), 147-160.

Mielcarzewicz, W.E. (2000). *Calculation of water supply systems.* Warszawa: Arkady (in Polish).