Periodica Polytechnica
Civil Engineering

61 (4), pp. 873–881, 2017
https://doi.org/10.3311/PPci.10429

RESEARCH ARTICLE

Neural Network Approach for Availability Indicator Prediction

Małgorzata Kutyłowska[1*]

Received 20 December 2016; Revised 05 April 2017; Accepted 26 April 2017

Abstract
The principal aim of this research was to find out if artificial neural networks could be employed to predict the availability factor for water mains, distribution pipes and house connections. Modelling by means of artificial neural networks (ANNs) was carried out using the Statistica 10.0 software package. Operating data from the years 1999–2005 were used to train the ANNs while data from the next seven years of operation were used to verify the model. The optimal model (characterized by the lowest mean-square error) contained 11 hidden neurons activated by the exponential function. The linear function was used to activate the 3 output neurons. 185 training epochs sufficed to train the ANN, using the quasi-Newton method. The correlation between the availability indicator experimental values and the modelling results would remain high, amounting during model verification to $R^2 = 0.740$, $R^2 = 0.823$, $R^2 = 0.992$ for respectively water mains, distribution pipes and house connections. As the availability indicator prediction example shows, the artificial neural networks are a promising tool enabling quick and easy analysis of failure frequency. It is possible to train the ANN further and change the number of training epochs and hidden neurons as well as the activation functions and training methods.

Keywords
artificial intelligence, availability indicator, failure analysis, reliability, water-pipe network

[1]Faculty of Environmental Engineering
Wrocław University of Science and Technology
Wybrzeże St. Wyspiańskiego 27, 50-370 Wrocław, Poland
*Corresponding author, email: malgorzata.kutylowska@pwr.edu.pl

1 Introduction

Research on the frequency of failure of water supply networks has been conducted in Poland and in the world for many years [1–4]. Still, besides determining the failure frequency level of water conduits on the basis of operating data, it is also necessary to supplement such studies with the results of the modelling, using typical models or artificial intelligence and Monte Carlo simulations: probability of the failure occurrence and values of failure rate [5–8] or the amount of water losses through pipe breaks [9]. Therefore one of the aims of this research was to show that mathematical modelling could be used to determine the availability indicator and reliability level of water conduits.

1.1 Reliability and failure analysis

An analysis of the failure frequency of civil and environmental engineering structures is one of the elements in the assessment of their operational reliability level. The deep failure and reliability analysis of water supply as well as sewerage systems in Poland were carried out by M. Kwietniewski and J. Rak [10]. Authors indicated and described the results of previous studies as well as pointed out that new methods, e.g. mathematical modelling should be used for the assessment of technical condition of buried infrastructure in the future. This subsection presents basic information on broadly understood reliability. In the next subsections artificial intelligence would be used for assessing one chosen reliability indicator.

The operational reliability of an engineering structure is a property (expressed by reliability indicators) consisting in the ability of the structure to perform, over a specified time and in given operating conditions, the functions imposed on it [11]. One of the principle functions of an engineering structure is its operability. When damage occurs, the structure becomes more unreliable than before the failure. Although damage does not result in the total loss of reliability, it brings down its level. The latter can be estimated using selected reliability indicators and an analysis of failure frequency in the given operating conditions [11].

The water distribution network (guaranteeing treated water transmission, under proper pressure and in the required amount, to individual consumers and industry) is one of the major components of the whole water supply system and is called as critical infrastructure [12]. It should be characterized by high operational reliability and a low failure frequency. Therefore the risk of damage and other adverse factors which lower the operational reliability of the system should be properly assessed and reduced to an acceptable level by means of the available mathematical methods, e.g. Bayesian model [12] or methods of good management [13, 14]. This means that the failure frequency of the particular water supply network components (pumping stations, clean water reservoirs, fittings and water conduits) should be thoroughly analyzed. Line components, i.e. pipelines (water mains, distribution pipes and house connections) can be quite easily analyzed with regard to failure frequency since water companies should, as a rule, have the operating data (such as material, diameter, length, year of construction and type and date of failure) needed to determine the principal reliability indicators. In this paper, a reliability analysis is carried out on the basis of water distribution system line components. Water conduits are considered to be renewable elements, i.e. they can be repaired and upgraded during their operation, whereby they can be made fully operational [11]. The failure frequency of renewable components can be analyzed on the basis of, among other things, the following reliability indicators [11]: average time between failures, average renewal time, renewal intensity μ, failures stream parameter ω and availability indicator *AI*. The failure frequency analysis and the assessment of the reliable operation of water conduits in a selected water distribution network carried out in this paper are based on one of the above reliability indicators, i.e. the availability indicator.

Availability function *AI(t)* defines the probability of object operability maintenance at any time t. It is worth saying that probabilistic models are also used for other purposes, e.g. flood protection [15] and proper dimensioning of sewerage system [16]. This function (relation (1)) is also called a nonstationary availability indicator describing the sum of probabilities of occurrence of individual object operational event B_i [10, 11].

$$AI = P\left(\bigcup_{i=0}^{\infty} B_i\right) = \sum_{i=0}^{\infty} P\left(B_i\right) \tag{1}$$

In practical engineering considerations, for a selected (sufficiently long) time interval and water conduit operating data (length L, [km], failure intensity λ [fail./(km·year)] and average renewal time T_o, [year]) one can use the following stationary availability indicator [10, 11, 17]:

$$AI = \lim_{t\to\infty} AI\left(t\right) = \frac{\mu}{\mu+\omega} = \frac{1}{\left(1+\lambda\cdot L\cdot T_o\right)} \tag{2}$$

Thus the stationary availability indicator defines the probability that a given object will be operational at any instant t sufficiently distant from the time when the object was put into service [11].

One of the main aims of research on the analysis of failure frequency and operational reliability of underground infrastructure (e.g. water conduits) is to identify the causes of the failures in order to adopt proper methods of preventing them. Moreover, it is vital to assess the effect of a given component which keeps failing on the reliable operation of the whole system. It should be noted that all the investigations are conducted during normal water supply system operation. Considering that the whole system operates in real time, the above fact is important from the point of view of the analysis of the failures which have occurred. Therefore one should bear in mind that the data used in the analysis of failure frequency and to determine the reliability level are operating data acquired from water and sewerage companies. The latter should keep reliable records on the whole technical infrastructure under their supervision. Appropriate protocols and reports should contain survey information (e.g. material, diameter, age, water conduit laying depth, specifications of local pumping stations and water reservoirs), information on the failures which have occurred, the repairs and rehabilitations done and the planned upgrades. Each water distribution network is different (has its peculiarities), whereby the records and the ways of keeping archives differ between water companies. This means that one should adopt an individual approach to the analysis of the failure frequency of each water supply network, which sometimes makes it difficult to make generalizations.

Mathematical modelling based on operating data is used to determine the failure frequency and reliability of water conduits. Kleiner and Rajani [18, 19] reviewed the main statistical models and models based on the physical properties of water conduits. They indicated the advantages and disadvantages of the models used to determine the failure frequency of water supply networks. A few years earlier Khomsi et al. [20] had used computer simulation (a program written in TURBO PASCAL) to estimate the operational reliability level of a water distribution network. According to the authors [20], the model developed by them can be used not only to determine the reliability level of the existing water supply networks, but also to design new water distribution networks. Authors [20, 21] proposed following relationship for availability indicator calculation:

$$AI = 1 - \frac{\lambda\cdot L}{365} \tag{3}$$

Tabesh et al. [21] used fuzzy logic and artificial neural networks (ANNs) to determine the failure frequency and operational reliability levels of water conduits. They also used a relation (equation (3)) proposed by Khomsi et al.

The present paper continues and further elaborates the subject discussed in the previous paper [22] by the author. Since stationary indicators λ and *AI* can be relatively easily determined on the basis of operating data they are most often used to estimate the failure frequency and operational reliability levels. Therefore as part of the present research indicator *AI* (further referred to as the experimental availability indicator) was determined on the basis of the operating data (using relation (2)) made available by the water company in a medium-sized Polish town. The experimental availability indicator was then compared with the one predicted by means of artificial neural networks. In addition, it was checked whether the values calculated from the formula proposed by Khomsi et al. [20] were similar to the ones obtained from the modelling by means of ANNs.

1.2 Artificial neural networks

Artificial neural networks have been used for forecasting, modelling and optimization purposes for many years. They are an alternative to the mathematical models used for predicting the failure frequency and operational reliability of water supply networks. ANNs constitute a tool which can be used to quickly and accurately predict and forecast many dynamically changing parameters, such as the failure intensity (failure rate) indicator and the availability indicator of water conduits. The structure of artificial neural networks is created on the basis of existing data sets and fitted to the vectors of the training images in order to obtain the best possible verification results. Thanks to the fully automatic estimation of output vectors one can employ ANNs to solve highly complex problems in the case of which it is impossible to unequivocally determine (experimentally or using typical mathematical models) the correlations between the phenomena occurring in the failure frequency analysis.

The variability, and also very often the randomness, of the above indicators is due to, e.g., the climatic conditions, the location of the conduits (type of soil, mining damage areas, etc.), their age and methods of their rehabilitation, which poses difficulties for typical mathematical models. Owing to the properties of artificial neural networks it is possible to analyze nonlinear phenomena and approximate the functions of many variables. As mentioned above, the analysis of the failure frequency of water conduits is precisely an analysis of the functions of many dynamically changing parameters, such as age, pipeline laying depth, temperature, soil and wheeled vehicle load. It is not possible to analyze all the factors having a bearing on the failure frequency level of water conduits, since the data are not recorded because of the lack of time and sufficient funds or are virtually impossible to verify. Therefore ANNs can be employed in such situations.

The theory of ANNs is described in detail in the literature on the subject [23, 24]. Currently artificial neural networks are successfully used to predict hourly water demand distributions [25, 26], determine water pressure and quality in the distribution network [27, 28], predict the dam water level [29] and so on.

As regards the prediction of reliability indicators and the general analysis of failure frequency, ANNs have been used mainly to determine the number of failures [30], predict the failure intensity indicator [31] and estimate pipeline construction costs [32]. The literature reports published so far indicate that the prediction of the availability indicator by means of ANNs is not extremely advanced, which encouraged the author to undertake this subject.

2 Material and methods

Operating data received from the water and sewerage company in a medium-sized town in Poland were used to predict the availability indicator of water mains, distribution pipes and house connections by means of artificial neural networks. City has about 70,000 inhabitants almost 100% of whom are connected to the system. The water pipelines are laid almost in every street, so the compaction of the pipes is very high. The daily average supply is about 8,300 m^3/d. Because of the topography and the height of the buildings, there are two water supply zones with pressures of 3.6 bar and 5.6 bar, respectively. The network includes: (1) main conduits of both grey cast iron and steel transit, with diameters of DN 350 to 700 mm, (2) distribution pipes with diameters of DN 80 to 250 mm, made variously of grey cast iron, steel, PE, PVC or AC, (3) house connections with diameters of DN 25 to 200 mm, made of steel or PE.

As mentioned above the experimental availability indicator values were compared with the results of prediction by means of ANNs and with the *AI* values calculated from relation (3).

Modelling by means of ANNs was done in Statistica ver. 10.0. Data from the years 1999–2005 were used to train artificial neural networks while data from the next 7 years of service (2006–2012) were used to verify the model (the availability indicator was forecasted on the basis of the latter data). 50%, 25% and 25% of the data were used to respectively learn the ANNs, test the created models and validate them. The procedure was as described below. Twenty artificial neural network models were built and trained on the basis of the operating data for the years 1999-2005. Five models characterized by the lowest root mean-squared error and the relatively good agreement between predicted and experimental (calculated from relation (2)) availability indicator values were selected. The five models (No. 2, 3, 6, 7 and 14) were verified by inputting data for the next years (2006–2012) and determining their root mean-squared errors. The main parameters of these five models are listed in the table 1. It seems reasonable to choose optimal model on the basis of results from verification step because the data from years 2006–2012 were not known for the model previously. Finally, one model characterized by the lowest error and the best fit of the values generated by the ANN to the experimental results was selected. Further in this paper this model is referred to as the optimal model. One of the simplest and most popular artificial neural network structures, i.e. a multilayer perceptron with input neurons W, hidden layer neurons U and output neurons O (MLP W-U-O), was used to predict the availability indicator. When creating ANN models respectively minimum and maximum 1 and 12 hidden neurons were used in a single hidden layer. The ANNs were trained using the quasi-Newton method, the steepest descent method and the conjugate gradient method. The best results would be obtained when the first method was used. The number of training epochs would range from 1 to 672. The linear function, the logistic function, the exponential function and the hyperbolic tangent were used to activate the hidden and output layer neurons. A simplified (in comparison with the previous research [22]) approach to the selection of input variables for the ANN was adopted in this study. Similar approach as here was proposed in another author's work [5] but in relation to another reliability indicator prediction – failure rate. The aim was to find out whether by inputting the most basic information, which can be easily recorded by any water company, into the ANN model one would obtain results convergent with the experimental data. The following parameters were treated as input signals (6 neurons): the number of water mains failures (N_m), the number of distribution pipes failures (N_r), the number of house connections failures (N_p), the length of water mains (L_m), the length of distribution pipes (L_r) and the length of house connections (L_p). Three neurons being the indicators of availability of water mains AI_m, distribution pipes AI_r and house connections AI_p were output signals. An exemplary ANN architecture is shown in fig. 1.

Table 1 Main parameters of five ANN models

No.	Model	Number of learning epochs	Activation function-hidden layer	Activation function-output layer
2.	MLP 6-12-3	243	logistic	linear
3.	MLP 6-8-3	118	logistic	hyperbolic tangent
6.	MLP 6-6-3	235	logistic	hyperbolic tangent
7.	MLP6-10-3	12	exponential	linear
14.	**MLP 6-11-3**	**185**	**exponential**	**linear**

The range of variation of the input parameters and that of the experimental availability indicator during training and verification step are shown in table 2.

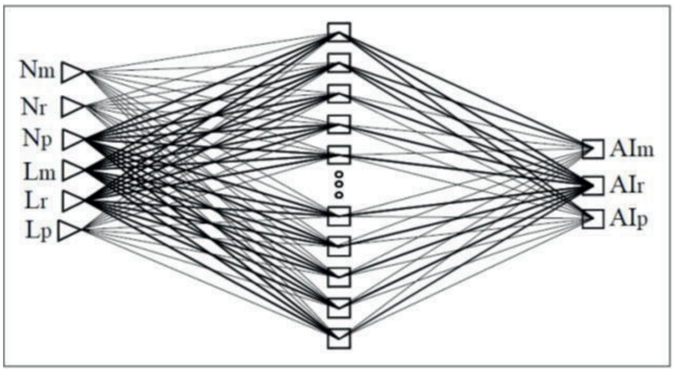

Fig. 1 Exemplary architecture of ANN model for availability indicator prediction

Table 2 Range of input and output parameters

Learning step (1999–2005)				
N_m	N_r	N_p	L_m, km	L_r, km
5–16	37–66	20–26	28.0–28.3	99.2–107.5
L_p, km	AI_m	AI_r	AI_p	
35.4–39.2	0.985533–0.995386	0.942776–0.967080	0.975858–0.982218	
Verification step (2006–2012)				
N_m	N_r	N_p	L_m, km	L_r, km
2–11	27–48	15–32	28.3–29.6	108.9–115.1
L_p, km	AI_m	AI_r	AI_p	
39.8–45.0	0.990056–0.998201	0.957719–0.975581	0.971744–0.986613	

3 Results and discussion

The optimal model selected from 20 ANN models showed the best agreement between the predicted results and the experimental ones. The figure 2 shows the correlation between experimental and predicted values of availability indicator of water mains, distribution pipes and house connections in verification step concerning one chosen optimal model No. 14 with the highest determination. The model (MLP 6-11-3) had 11 hidden neurons activated by the exponential function. The linear function was used to activate the 3 output neurons (AI_m, AI_r, AI_p). 185 training epochs were enough to train the ANN. The training quality of 0.99 was achieved in the training process while the testing quality during verification was at the level of 0.86. In the approach adopted in this study the ANN model contained 3 output neurons representing the three indicators of availability (of water mains, distribution pipes and house connections). This means that the training and testing quality mentioned above does not refer to the prediction of the particular indicators of availability of water mains, distribution pipes and house connections, but to the whole model. The approach (with 3 neurons at the output) adopted here differs radically from the ways of predicting the failure intensity (failure rate) indicator proposed in the earlier paper by the author [5] where separate ANN models would be created to predict the indicators for distribution pipes and house connections.

Figures 3–5 present the results of the prediction (by means of the adopted optimal ANN model), only for training step (the results of verification phase are shown in the figure 2) of the indicators of availability of water mains (fig. 3), distribution pipes (fig. 4) and house connections (fig. 5). Also the experimental values of this indicator and the ones calculated from relation (3) are displayed in the figures 3–5.

Figure 3 shows that the prediction of the availability indicator was practically ideal at the ANN training stage. The largest relative error between the experimental results and the predicted ones amounted to 0.07% (the year 2005). The prediction of AI, based on the data from the next years (fig. 2) (i.e. somewhat different from the training vector), carried a maximum error of 0.42%, which shows that this model is highly suitable for determining the availability indicator of water mains.

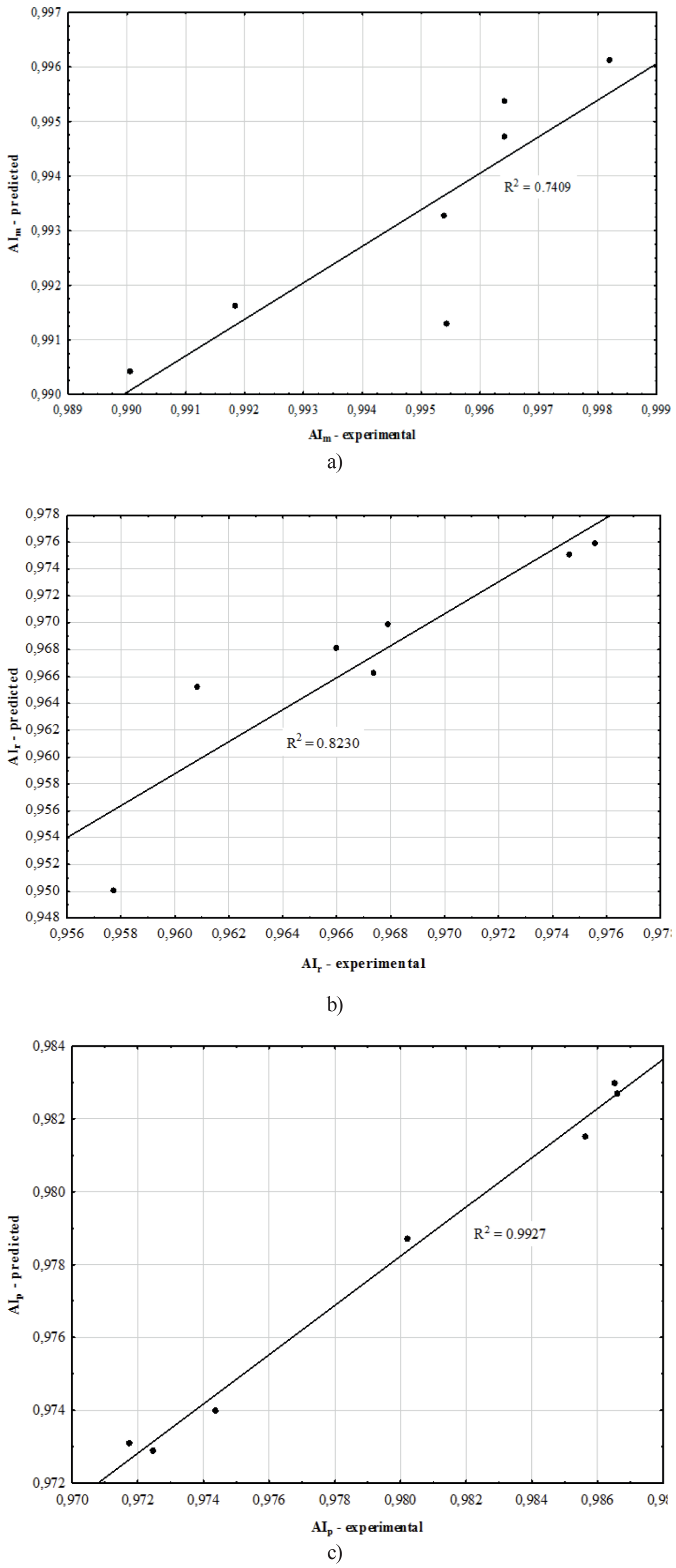

Fig. 2 Model No. 14 – correlation between experimental and predicted values of availability indicator a) water mains, b) distribution pipes, c) house connections

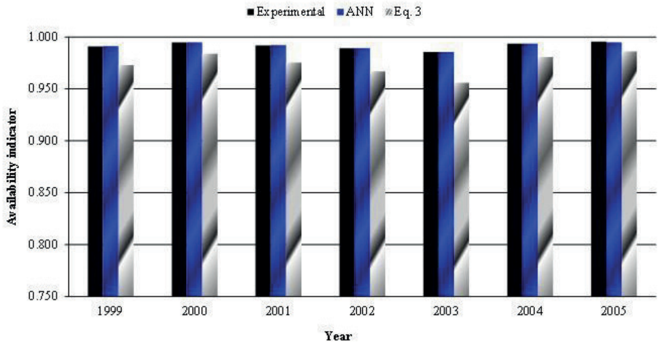

Fig. 3 Experimental and predicted availability indicator of water mains

Relatively large discrepancies between the experimental results and the *AI* values calculated from relation (3) proposed by other authors [20, 21] are noticeable. Also Tabesh et al. [21] compared the results of modelling by means of ANN with the ones calculated from the relation proposed by Khomsi et al. [20]. There is no agreement between the availability indicator values calculated from relation (3), the experimentally determined values of this indicator and the results of prediction by means of ANN. It is apparent (figs 3–5) that the availability indicator values calculated from formula (3) are lower than the experimental values. This means that this indicator is underestimated. Therefore in the particular case of the water distribution network in the medium-sized Polish town better results are obtained using the ANN model instead of relation (3). However, each water supply system is different and the distribution of failures over time (having a bearing on indicator *AI*) in another water supply network will differ from the one observed in the particular case. This means that the relations proposed by other authors cannot be outright rejected since they can be applicable to other municipal water systems.

High agreement between the experimental results and the ones obtained from the ANN model (model MLP 6-11-3) at the training stage is also observed for the prediction of the availability indicator of distribution pipes. The maximum error amounted to 0.21%. Also the model verification process was characterized by high correlation ($R^2 = 0.823$, fig. 2) and the maximum error did not exceed 0.80%. Similarly as in the case of water mains, relation (3) would yield underrated values of indicator *AI*. The underestimation of the availability indicator means that also the operational reliability level would be incorrectly calculated.

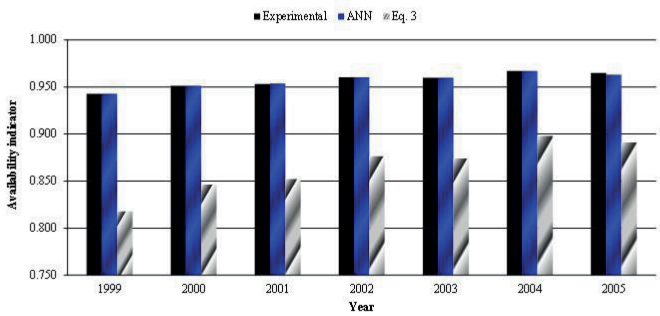

Fig. 4 Experimental and predicted availability indicator of distribution pipes

Moreover, one should bear in mind that failures of water mains and distribution pipes result in undesirable events which affect the functioning of the whole water supply system, the town and the population. Such events include: a pressure drop or no water in the water supply network, local secondary pollution of water in the pipeline, water outflows to the terrain surface and a necessary closure of the roadway section under which the damaged water conduit runs. In order to maximally eliminate the effects of failures one should strive not only to immediately repair any damage, but also to estimate the future level of reliability of the pipeline. The ANN model (MLP 6-11-3) proposed in this paper represents an attempt to create a tool for predicting reliability indicators (for a selected water supply system) in a relatively simple way. The verification of the model, based on data previously unknown to the ANN, yielded satisfactory results being in good agreement with the experimental ones (fig. 5). Such basic parameters as the number of failures and the length of conduits were the inputs to the ANN. Since each year the water company records such data the presented model (trained on data from the years 1999–2005 and verified using data from the years 2006–2012) can be used to determine the availability indicator (as one of the indicators showing the level of operational reliability) in the subsequent years. However, one should note that the proposed model (MLP 6-11-3) is not a universal model which can be used to predict indicator *AI* in every water supply network since it has been verified for only the particular municipal water system. Of course, not only this model's architecture (the number and type of input signals and the size of the hidden vector), but also the activation function and the training algorithm can be changed for predicting the availability indicator in another water distribution network in another town. Moreover, it is always possible to additionally train the ANN in order to achieve even higher agreement with the experimental results.

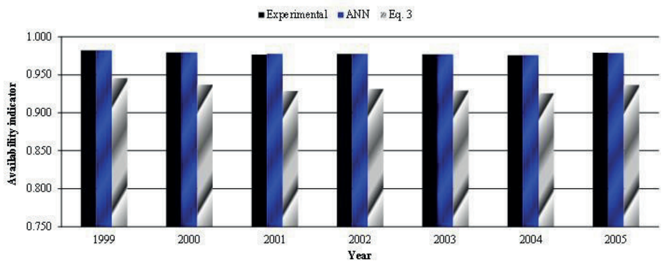

Fig. 5 Experimental and predicted availability indicator of house connections

As regards house connections, the maximum error of predicting indicator *AI* amounted to 0.08% and 0.41% for respectively ANN training (fig. 5) and model verification (fig. 2). Also the correlation between the experimental values and the ones obtained from the MLP 6-11-3 model was high ($R^2 = 0.992$). Therefore the ANN model can be used to predict the indicator of availability of house connections. Similarly as for the other types of pipelines, *AI* values calculated from relation (3) were found to considerably depart from the values yielded by the ANN. In the case of house connections, operational reliability and the possible consequences of a failure have a more local dimension (water supply will be reduced for a few hours only to the inhabitants connected to a given service pipe) than in the case of distribution pipes or water mains. The consequences of damage and unreliable operation have no significant effect on the operational reliability of the whole water distribution network. Of course, this does not relieve the operators of the duty of caring for each type of underground infrastructure, including house connections. The results of modelling related to water mains, distribution pipes and house connections in verification process (fig. 2) are characterized by a little bit lower convergence (in comparison with learning step) between experimental and predicted values of availability indicator. This is quite normal situation because in verification step the data for modelling were unknown to ANN model previously. The model was trained using different data set and that is the reason why the correlation is not ideal when we put to the model information from the next 7 years of operation.

It is also worth indicating that Spearman rank correlation (SR) in verification step (optimal model MLP 6-11-3) is relatively high and equals to 0.8829, 0.9643 and 0.9286 for water mains, distribution pipes and house connections, respectively. Spearman rank correlation was conducted to check whether the results of ANN model follow the monotonic increase or decrease of real values of *AI*. According to the results mentioned above one can concluded that the experimental and predicted values of availability indicator are strongly correlated and are characterized by increasing function because the values of coefficients 0.8829, 0.9643 and 0.9286 are near value of +1. The Spearman rank correlation is also proposed by other authors to check the importance of model parameters [33, 34].

The ANN models proposed in the paper contain relatively basic input variables as length of pipelines and the number on registered damages. Moreover, it was checked whether creating new additional models including different input parameters as material, diameter and length of the pipes could be also suitable approach for availability indicator prediction. The models for *AI* prediction of distribution pipes and house connections were created separately. It means that approach was a little bit different than proposed above. In the output layer only one neuron characterized by one indicator *AI* existed. The procedure of choosing the optimal model was the same as described in the section "material and methods". The optimal model for distribution pipes has 7 input neurons (material – was treated as quality parameter and one neuron correspond to one type of material: cast iron, steel, PE, PVC, asbestos cement, diameter and length) and 13 hidden neurons activated by exponential and hyperbolic tangent functions, respectively. The optimal model for house connections has 6 input neurons (material: cast iron, steel, PE, galvanized steel, diameter and length) and 12 hidden neurons activated by hyperbolic tangent and logistic functions, respectively. The results of prediction using additional models related to verification step are displayed in the table 3.

Table 3 Results of additional ANN modelling

Year	Distribution pipes, R2=0.5486; SR=0.7143		House connections, R2=0.5522; SR=0.6071	
	Experimental	**ANN**	**Experimental**	**ANN**
2006	0.960822	0.965463	0.971744	0.978499
2007	0.974636	0.966296	0.972462	0.979762
2008	0.965975	0.965964	0.980216	0.980328
2009	0.967349	0.966631	0.986535	0.981351
2010	0.967869	0.966828	0.986613	0.981356
2011	0.975581	0.966403	0.985613	0.981572
2012	0.957719	0.965081	0.974373	0.981390

New mentioned above models are a little bit more complicated (more input and hidden neurons) and the correlation between experimental and predicted values of indicator *AI* is even lower (see table 3) than for model MLP 6-11-3 which was earlier described in this paper. The results mean that there is no relationship between more complicated data (e.g. material, diameter) included to the input layer and good convergence related to experimental and modelled values of *AI*.

Today, an analysis of failure frequency on the basis of operating data and modelling results is one of the basic measures which must be adopted when determining the level of operational reliability and planning rehabilitations of water conduits. The experimental availability indicator results (and also the results of *AI* modelling by means of the ANN) show that the level of reliability of the water mains, the distribution pipes and the house connections in the analyzed water distribution system is quite high. However, the operational reliability of a given water supply network should be assessed on the basis of more than one reliability indicator. Therefore the conclusions emerging from the above analysis should be considered jointly with failure intensity indicator modelling results. The ANN model (MLP 6-11-3) for predicting the availability factor should be used together with the model for forecasting the failure rate indicator. Owing to this comprehensive approach one will be able to determine the failure frequency and operational reliability levels of a given water supply system. Moreover, it is necessary to remember that assessing the failure frequency of water pipes should be carried out together with e.g. impact of water leakages on the soil (suffusion processes) [35] because it is also very important problem related to the environmental engineering.

Artificial neural network approach is like "black box" modelling. To achieve reliable solutions it sometimes is required to apply "trial and error method" during creation of the network model. Such approach does not allow to penetrate deeply inside the way of forming the network structure, especially in the hidden layer (e.g. weight adaptation, changes related to activation functions). It is needed to create the artificial network for each problem separately. The solved problem extorts the number of neurons, the kind of activation function and the training methods. ANN approach could be used for prediction of some reliability indicators in other water systems, but models should be created separately for each system due to e.g. different operational conditions, different diameters and materials of water pipes and different number of registered damages. We should distinguish several limitations of using ANN. It is necessary to collect huge data base. This requirement is sometimes difficult to realize by water utilities. Some important, from forecasting point of view, variables are not registered. The situation is still improving because more and more water utilities attach importance to connect information about damages with GIS. It is also necessary to analyse deeply the data which are used as the training parameters in the process of neural network modelling. In some cases it is necessary to eliminate the outliers. If there is too much eliminated outlier data we should answer the question whether the modelling is rational. Unfortunately, real operational data are burden sometimes by huge mistakes due to e.g. improper or incomplete registration or lack of collecting all information in GIS database. In such cases it is necessary to eliminate or remain outliers reasonably.

4 Conclusions

The results of predicting the availability indicator of water mains, distribution pipes and house connections by means of an ANN model have been presented. The results of this research complete the author's earlier studies on the applicability of ANNs to the prediction of reliability indicators. The following final conclusions can be drawn from the analysis of the modelling results:

- artificial neural networks constitute a promising tool for predicting the availability indicator (AI_m, AI_r, AI_p) in a relatively simple and quick way;
- the adopted optimal model MLP 6-11-3 would accurately predict the indicator of availability of water mains, distribution pipes and house connections. The model has only 6 input signals (the number of failures and the length of water mains, distribution pipes and house connections), which undoubtedly is an advantage since such data, recorded by every water company, are easily available;
- the exponential function and the linear function in the MLP 6-11-3 model were used to activate respectively the hidden and the output neurons. The ANN was trained using the quasi-Newton algorithm and 185 training epochs;
- the ANN was trained on the operating data from the years 1999–2005 while the model was verified using the data for the next 7 years of operation of the water supply network in the medium-sized Polish town;

- the correlation between the experimental availability indicator values and the ones yielded by the MLP 6-11-3 model was high during both ANN training and verification (forecasting). In the case of training, this correlation for respectively water mains, distribution pipes and house connections amounted to: $R^2 = 0.993$, $R^2 = 0.991$ and $R^2 = 0.946$. In the case of verification, the coefficient R2 and Spearman rank correlation amounted to respectively: 0.740, 0.882, 0.823, 0.964 and 0.992, 0.821 for water mains, distribution pipes and house connections. The maximum relative error between the experimental results and the ones obtained from the ANN was below 1.0%, which shows that the ANN was properly trained and that the availability indicator value was correctly predicted;
- the ANN can always be additionally trained on another set of input data. Also the number of training epochs and hidden neurons as well as the activation functions and the training methods can be changed. Thus ANNs are relatively easy to implement and use when functions of many variables are to be approximated, owing to the fact the ANN's architecture can be relatively quickly adapted.

The work was realized within the allocation No. 0401/0006/17 awarded for Faculty of Environmental Engineering Wroclaw University of Science and Technology by Ministry of Science and Higher Education in years 2017–2018.

References

[1] Pietrucha-Urbanik, K. "Failure analysis and assessment on the exemplarywater supply network.". *Engineering Failure Analysis*, *57*, pp. 137–142. 2015. https://doi.org/10.1016/j.engfailanal.2015.07.036

[2] Kutyłowska, M., Orłowska-Szostak, M. "Comparative analysis of water–pipe network deterioration–case study.". *Water Practice and Technology*, *11*(1), pp. 148–156. 2016. 10.2166/wpt.2016.018

[3] Boxall, J.B., O'Hagan, A., Pooladsaz, S., Saul, A. J., Unwin, D. M. "Estimation of burst rates in water distribution mains.". *Water Management*, *160*(2), pp. 73–82. 2007. https://doi.org/10.1680/wama.2007.160.2.73

[4] Martinez-Codina, A., Castillo, M., Gonzales-Zeas, D., Garrote, L. "Pressure as a predictor of occurrence of pipe breaks in water distribution networks.". *Urban Water Journal*, *13*(7), pp. 676–686. 2016. https://doi.org/10.1080/1573062X.2015.1024687

[5] Kutyłowska, M. "Comparison of two types of artificial neural networks for predicting failure frequency of water conduits.". *Periodica Polytechnica Civil Engineering*, *61*(1), pp. 1–6. 2017. https://doi.org/10.3311/PPci.8737

[6] Wang, Y., Zayed, T., Moselhi, O. "Prediction models for annual break rates of water mains.". *Journal of Performance of Constructed Facilities*, *23*(1), pp. 47–54. 2009. https://doi.org/10.1061/(ASCE)0887-3828(2009)23:1(47)

[7] Moglia, M., Davis, P., Burn S. "Strong exploration of a cast iron pipe failure model.". *Reliability Engineering and System Safety*, *93*(6), pp. 863–874. 2008. https://doi.org/10.1016/j.ress.2007.03.033

[8] Wilson, D., Filion, Y., Moore I. "State-of-the-art review of water pipe failure prediction models and applicability to large-diameter mains.". *Urban Water Journal*, *14*(2), pp. 173–184. 2017. https://doi.org/10.1080/1573062X.2015.1080848

[9] Suchorab, P., Kowalska, B., Kowalski D. "Numerical investigations of water outflow after the water pipe breakage.". *Annual Set The Environment Protection*, *18*(2), pp. 461–427. 2016.

[10] Kwietniewski, M., Rak, J. „Niezawodność infrastruktury wodociągowej i kanalizacyjnej w Polsce.". "Reliability of water supply and wastewater infrastructure in Poland.". Polish Academy of Science, Warszawa, Poland. 2010. (in Polish)

[11] Wieczysty, A. "Niezawodność systemów wodociągowych i kanalizacyjnych.". (Reliability of water and sewer systems: reliability theory and its applications. Part I and Part II) Kraków University of Technology PublishingHouse, Kraków, Poland, 1990. (in Polish)

[12] Tchórzewska-Cieślak, B. "Bayesian model of urban water safety management.". *Global NEST Journal*, *16*(4), pp. 667–675. 2014. http://journal.gnest.org/sites/default/files/Submissions/gnest_01344/gnest_01344_published.pdf

[13] Boryczko, K., Tchórzewska-Cieślak, B. "Analysis of risk of failure in water main pipe network and of delivering poor quality water.". *Environment Protection Engineering*, *40*(4), pp. 77–92. 2014. https://doi.org/10.5277/epe140407

[14] Mamo, T. G. "Risk-based approach to manage aging urban water main infrastructure.". *Journal of Water Supply: Research and Technology – AQU*, *64*(3), pp. 260–269. 2015. https://doi.org/10.2166/aqua.2014.052

[15] Kádár, I., Nagy, L. "Comparison of different standards based on computing the probability of failure of flood protection dikes.". *Periodica Polytechnica Civil Engineering*, *61*(1), pp. 146–153. https://doi.org/10.3311/PPci.9501

[16] Kaźmierczak, B., Wdowikowski, M. "Maximum rainfall model based on archival pluviographic records – case study for Legnica (Poland).". *Periodica Polytechnica Civil Engineering*, *60*(2), pp. 305–312. 2016. https://doi.org/10.3311/PPci.8341

[17] Hotloś, H. "Ilościowa ocena wpływu wybranych czynników na parametry i koszty eksploatacyjne sieci wodociągowych.". (Quantitative assessment of the effect of some factors on the parameters and operating costs of water-pipe networks.) Wrocław University of Technology Publishing House, Wrocław, Poland, 2007. (in Polish)

[18] Kleiner, Y., Rajani, B. "Comprehensive review of structural deterioration of water mains: statistical models.". *Urban Water*, *3*(3), pp. 131–150. 2001. https://doi.org/10.1016/S1462-0758(01)00033-4

[19] Rajani, B., Kleiner, Y. "Comprehensive review of structural deterioration of water mains: physically based models.". *Urban Water*, *3*(3), pp. 151–164. 2001. https://doi.org/10.1016/S1462-0758(01)00032-2

[20] Khomsi, B. D., Walters, G. A., Thorley, A. R. D., Ouazar, D. "Reliability tester for water distribution networks.". *Journal of Computing in Civil Engineering*, *10*(1), pp. 10–19. 1996. https://doi.org/10.1061/(ASCE)0887-3801(1996)10:1(10)

[21] Tabesh, M., Soltani, J., Farmani, R., Savic, D. "Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling.". *Journal of Hydroinformatics*, *11*(1), pp. 1–17. 2009. https://doi.org/10.2166/hydro.2009.008

[22] Kutyłowska, M. "Neural network approach for failure rate prediction.". *Engineering Failure Analysis*, *47*, pp. 41–48. 2015. https://doi.org/10.1016/j.engfailanal.2014.10.007

[23] Suzuki, K. "Artificial neural networks – architectures and applications.". InTech, Chicago, USA, 2013.

[24] Hertz, J., Krogh, A., Palmer, R. "Introduction to the theory of neural computation.". Addison-Wesley Publishing Company, Redwood City, 1991.

[25] Cieżak, W., Cieżak, J. "Routine forecasting of the daily profiles of hourly water distribution in cities. An effectiveness analysis.". *Environment Protection Engineering*, *41*(2), pp. 179–186. 2015. https://doi.org/10.5277/epe150215

[26] Behboudian, S., Tabesh, M., Falahnezhad, M., Ghavanini, F.A. "A long-term prediction of domestic water demand using preprocessing in artificial neural network.". *Journal of Water Supply: Research and Technology – AQUA*, *63*(1), pp. 31–42. 2014. https://doi.org/10.2166/aqua.2013.085

[27] Ridolfi, E., Servili, F., Magini, R., Napolitano, F., Russo, F, Alfonso, L. "Artificial neural networks and entropy-based methods to determine pressure distribution in water distribution systems.". *Procedia Engineering*, *89*, pp. 648–655. 2014. https://doi.org/10.1016/j.proeng.2014. 11.490

[28] Cuesta Cordoba, G.A., Tuhovčák, L., Tauš, M. "Using artificial neural network models to assess water quality in water distribution networks.". *Procedia Engineering*, *8709*, pp. 399–408. 2014. https://doi.org/10.1016/j.proeng.2014.02.045

[29] Ünes, F., Demirci, M., Kisi, Ö. "Prediction of Millers Ferry Dam reservoir level in USA using artificial neural network.". *Periodica Polytechnica Civil Engineering*, *59*(3), pp. 309–318. 2015. https://doi.org/10.3311/PPci.7379

[30] Jafar, R., Shahrour, I., Juran, I. "Application of artificial neural networks (ANN) to model the failure of urban water mains.". *Mathematical and Computer Modelling*, *51*(9–10), pp. 1170–1180. 2010. https://doi.org/10.1016/j.mcm.2009.12.033

[31] Shirzad, A., Tabesh, M., Farmani, R. "A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks.". *KSCE Journal of Civil Engineering*, *18*(4), pp. 941–948. 2014. https://doi.org/10.1007/s12205-014-0537-8

[32] Dinu, P. A., Hussein, M., Bouferguene, A., Fernando S. "Artificial neural network model for cost estimation: city of Edmonton's water and sewer installation services.". *Journal of Construction Engineering and Management*, *136*(7), pp. 754–756. 2010. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000184

[33] Pasha, M.F.K., Lansey, K. "Effect of parameter uncertainty on water quality predictions in distribution systems-case study.". *Journal of Hydroinformatics*, *12*(1), pp. 1–12. 2010. 10.2166/hydro.2010.053

[34] Jung, D, Yoo, D., Kang, D., Kim, J. "Linear model for estimating water distribution system.". *Journal of Water Resources Planning and Management*, *142*(8), 2016. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000664

[35] Iwanek, M., Suchorab, P., Karpińska-Kiełbasa, M. "Suffosion holes as the results of a breakage of a buried water pipe.". *Periodica Polytechnica Civil Engineering*, Online first. https://doi.org/10.3311/PPci.9728