# CIVIL AND ENVIRONMENTAL ENGINEERING REPORTS

# ARTIFICIAL NEURAL NETWORK AS A VIRTUAL SENSOR OF NITRATE NITROGEN (V) CONCENTRATION IN AN ACTIVATED SLUDGE REACTOR

Lesław PŁONKA[1]
The Silesian University of Technology, Gliwice, Poland

A b s t r a c t

The paper discusses the use of an artificial neural network to control the operation of wastewater treatment plants with activated sludge. The task of the neural network in this case is to calculate (predict) the readings of the probe measuring the concentration of nitrate nitrogen (V) in one of the biological reactor tanks. Neural networks are known for their ability to universal approximation of virtually any relationship, including the function of many variables, but the process of "training" the network requires the presentation of many sets of input data and corresponding expected results. This is a difficulty in the case of wastewater treatment plants, because some key process parameters are usually not measured online (samples are taken and measurements are taken in the laboratory), and even if they are, the time intervals are large. Bearing in mind the aforementioned difficulty, this work uses a set of input data consisting only of information that can be measured with measuring probes.

As a result of the conducted experiments a high compliance of the probe's prediction with the expected values was obtained. The paper also presents data preparation and the network "training" process.

Keywords:     activated sludge process, artificial neural networks

---

[1] Corresponding author: The Silesian University of Technology, Faculty of Energy and Environmental Engineering, Department of Environmental Biotechnology, Akademicka 2 p.714c, 44-100 Gliwice, Poland,  leslaw.plonka@polsl.pl, +48 32 237 26 94

## 1. INTRODUCTION

Municipal wastewater treatment plants are installations operating under conditions of continuous variation of both substrate and hydraulic load. Both of these variable factors require continuous and dynamic control of the operation of the plants. The acquisition of knowledge about the operating conditions of wastewater treatment plants can be divided into two main ways: online measurements and laboratory tests, the former being obviously suitable for the purpose of dynamic plant control. The indications of the measuring probes directly influence the control of the plant operation and the final effect of wastewater treatment, i.e. the ecological effect of the whole plant. Therefore, it is important that these probes work correctly and any possible errors are detected and corrected. One of the techniques of detecting irregularities of the measuring probe indications or the occurrence of unusual conditions in the operation of a wastewater treatment plant is the use of artificial intelligence techniques to implement a virtual measuring sensor [4]. In this work an artificial neural network was used as a base to create a virtual sensor of nitrate nitrogen (V) concentration in the denitrification chamber of a biological reactor with activated sludge. The indication of this probe shows whether the effect of removing nitrogen compounds from wastewater is correct. This is a very important information given that nitrogen is a biogenic element, the excess of which is one of the causes of water eutrophication. Reducing the amount of nitrogen compounds in the wastewater leaving the plant is one of the main tasks of the whole plant [8]. It is also worth noting that the nitrate (V) concentration values measured at the end of the denitrification zone should usually be low, i.e. below 2 mg/dm$^3$ and even below 1 mg/dm$^3$[2]. In practice, this means that the control system must be resistant to accidental fluctuations in the results, while these fluctuations can represent a large percentage of the measured value. This creates an additional difficulty in detecting a situation in which the measuring probe for some reason loses the measurement precision or the measuring probe works properly but the working conditions of the treatment plant differ from the typical ones, e.g. as a result of industrial wastewater discharge disturbing the plant. The use of an additional, virtual sensor that calculates the predicted values of the probe indications may be helpful in detecting the above-mentioned situations.

The aim of the research described in this paper was to create a virtual nitrate nitrogen (V) sensor based on an artificial neural network, which as input data will use the indications of other probes located in different locations of the sewage treatment plant.

## 2. GENERAL INFORMATION ABOUT NEURAL NETWORKS

Artificial neural network is a kind of artificial intelligence technique that tries to describe a non-linear relationship between the input and output of a complex system using historical data (e.g. measurement, process data). Artificial neural network is an information processing structure that consists of units called neurons [11]. Neurons are usually organized in layers. Input signals are input into the "input layer" and then pass through the "hidden" layers to the output layer. The number of neurons in the first (input) layer must be equal to the number of input signals. Similarly, the number of neurons in the output layer is equal to the number of output signals. Each neuron can be connected to one or more units in subsequent layers. The figure 1 below illustrates a simple "perceptron" type network with three input units, three units in the hidden layer and one output neuron.
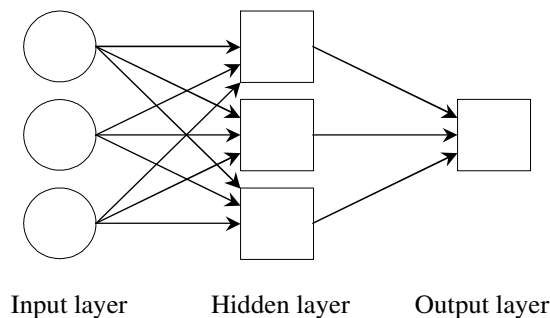


Input layer          Hidden layer          Output layer

Fig. 1. A simple "perceptron" type neural network diagram

Each "neuron" of the network has a defined activation function by which the values given at the input are converted into one output value. The argument of the activation function is the sum of the input values, which in turn are calculated as the products of the outputs of the neurons from the previous layers and the "weights" values. Weights are assigned to specific connections (arrows in the above diagram) during a process known as "training" the network. The network training process consists in presenting the input data multiple times and calculating the difference between the output provided by the network and the expected value. Based on this difference, special algorithms change the connection weights in such a way that the calculation error is slightly smaller. As a result of multiple presentation of the input data and correction of weights, it is possible to obtain a neural network capable of not only correctly calculating the result for the data presented during training, but also for data outside the "training set".

Multilayer perceptron networks (MLPs) are the neural networks most commonly trained using the "backpropagation" algorithm, and the method itself is not new

at all. MLP was presented for the first time in 1958 [7]. These are supervised networks and therefore require training to achieve the desired response. With one or two hidden layers, neural networks can map (always approximately) virtually any input-output relationship [1]. Many contemporary research and applications of neural networks use MLP techniques, eg [3, 7].

Nowadays, due to the constantly increasing efficiency of computers, "deep learning" techniques and networks are more and more often used, which are described as "deep neural networks". While perceptrons usually have one or at most two hidden layers, "deep" networks have multiple layers. Such solutions can be applied precisely because of the increased performance of computers, because their "training" is much more computationally demanding than that of "shallow" networks, i.e. with up to 2 hidden layers [12].

## 3. METHODOLOGY

There are many types of artificial neural networks that differ in their internal structure (the way neurons are linked, not just their number or number of layers) and training methods. In this study, it was decided to use a layered network, the structure of which is dynamically expanded during the training process (cascade training). During training, more neurons are added, each of which creates another layer of network. An exemplary diagram of such a network is shown in the figure 2 below.



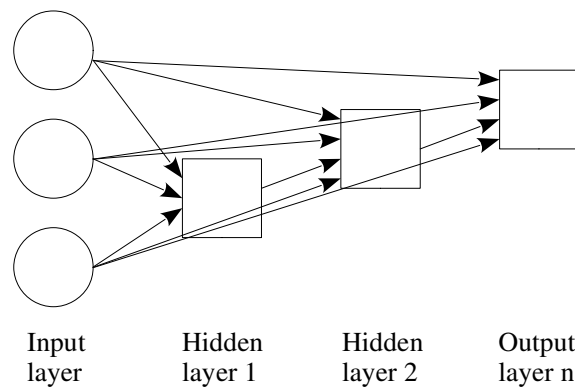| Input layer | Hidden layer 1 | Hidden layer 2 | Output layer n |

Fig. 2. Simplified diagram of a layered artificial neural network obtained as a result of cascade training

Cascade training, during which the network structure is built, has the advantage that it is not necessary to determine the size of the network before starting the training. This is important because until today it has not been possible to formulate

an exact recipe for selecting the size of the network. As a result, the number of layers and the number of neurons in each layer are selected by trial and error.

A neural network, whose task is to correctly reflect the complex and often unknown relationships between input and output data during training, usually requires the presentation of a large number of sample sets of inputs and outputs. In this work, the task of the neural network is to calculate the concentration of nitrate nitrogen (V) on the basis of the readings of the measuring probes measuring the concentrations and flows in other places of the installation. It was decided to use dynamic computer simulation as a data source for two reasons. The first reason is that by using the BSM1 mathematical model widely described in the literature, we obtain data commonly considered as valuable material for any analysis of the activated sludge process. The second reason is the possibility of obtaining a large amount of input data necessary in the process of training the neural network.

The figure 3 below shows the technological scheme of a biological wastewater treatment plant with activated sludge adopted in the BSM1 model. The segment of denitrification zone in which the neural network is to calculate the concentration of nitrate nitrogen (V) is marked in blue.
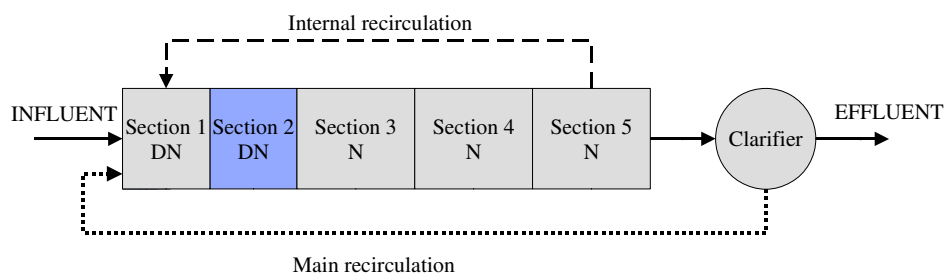


Fig. 3. Process diagram: Benchmark Simulation Model no. 1 (BSM1). Prepared by the IWA Taskgroup on Benchmarking of Control Stategies for WWTPs. "DN" - denitrification zone; "N" - nitrification zone

For the purpose of the computer simulation, the "STOAT" application was selected, whose compatibility with the BSM1 model is described in [1].

STOAT can be used to simulate wastewater and sludge treatment processes. A characteristic feature of the program is that it uses both models based on COD measurements (ASM1, ASM2d, ASM3) and BOD measurements (ASAL1...5). The models of the ASM family are widely used and considered to be the best for the purposes of mathematical modelling of the activated sludge process, however, their common disadvantage is the need to have knowledge of the wastewater flowing into the treatment plant which goes far beyond the set of typically conducted measurements ($BOD_5$, COD, suspended solids concentration and

nitrogen and phosphorus). The ASAL models simply use a typical set of measurements of concentrations and indicators of pollutants as input data, however, due to the success of the ASM models, work on them has not been continued for a long time and some aspects of calculations using ASAL models need to be improved especially when combining activated sludge processes with methane fermentation of sludge [10].

The BSM1 model is a set consisting of:
- mathematical model ASM1,
- input data describing the flow rate of wastewater,
- concentrations of contaminants in these wastewater with division into fractions according to the requirements of the ASM1 model,
- detailed description of the technological system of the plant: volume and purpose of objects, method of connections (flows and recirculations) and control.

Due to the fact that for the purposes of computer simulations, the formal writing of the ASM1 model requires software implementation and here the differences between the individual simulators arise - for the purposes of the BSM1 model STOAT offers a special version of ASM1 tested and agreed with the working group of BSM1 developers [1].

Neural network training was carried out using the "FannTool" tool - a graphical interface to the popular library of software that implements the process of neural network training - FANN [5, 6].

Computer simulation of the sewage treatment plant operation in accordance with the BSM1 model allows obtaining detailed information on the course of the process. In practice, only some of the data available through computer simulation can be measured online. For example, COD measurement, although it is possible with an analyzer, but the time intervals between measurements are much longer than in the case of measurement probes measuring e.g. oxygen or nitrogen concentration. In this work, it is assumed that all the information about the process is omitted, which is calculated by the simulator but in practice cannot be obtained online.

The instantaneous values of all state variables are the result of processes occurring in the near and slightly further past. Therefore, there is no point in expecting that the neural network will be able to correctly calculate the value of nitrate nitrogen (V) concentration in the reactor solely from the values of measurements taken at the same time. Not only the current measurements but also a number of measurements from the nearest past should be included in the input data set.

## 4.  RESEARCH RESULTS

### Selection of input data for training the neural network

As a result of the tests, it was finally decided to use the following set of data, on the basis of which the training of the neural network was carried out:

Table  1. Input data for training the artificial neural network

| Measured value | Location | Unit |
|---|---|---|
| Flow | Influent water | $m^3/h$ |
| Ammonia nitrogen | Influent water | $mg\ N/dm^3$ |
| Ammonia nitrogen | Segment 2 | $mg\ N/dm^3$ |
| Dissolved oxygen | Segment 5 | $mgO_2/dm^3$ |
| Nitrate nitrogen (V) | Segment 5 | $mg\ N/dm^3$ |
| Ammonia nitrogen | Segment 5 | $mg\ N/dm^3$ |
| KLa (aeration intensity) | Segment 5 | $[1/h]$ |
| Ammonia nitrogen | Effluent water | $mg\ N/dm^3$ |
| Nitrate nitrogen (V) | Effluent water | $mg\ N/dm^3$ |

The network training data set contains the measurement data listed in Table 1 measured at the moment and 9 previous measurements (measured at 72s intervals), ie. each measurement is represented by 10 values: the current and 9 previously measured. Thus, a table with 90 data columns and the last one, 91st, containing the expected value of nitrate nitrogen (V) concentration in the second denitrification segment was obtained. The expected value is needed to compare this value with the result of network calculations during the network training process and, on this basis, to determine the error and then change the weights of connections between the neurons of the network.

The computer simulation "lasted" 84 days and as a result 8065 lines of data were obtained. So the table containing the full set of information for the network training was 91 columns and 8065 rows.
The next stage of data preparation was the "normalization" process. If as an activation function inside the neuron is used a function whose values cannot be greater than 1, the neural network will not be able to give a result greater than 1. Therefore, the input data is scaled from the original values usually to a range (-1, 1) and this process is called normalization. In such a situation, the results of network calculations should be scaled back to obtain the correct values.
Then the data table was divided into two sets: a learning set and a test set. The former is used to train the network and the latter to check whether the neural network is able to properly calculate results for values that were not given as

examples during the training phase. The learning set contained 4833 lines and the test set 3223.

**Neural network training and calculation results**

After obtaining the data sets (training and testing), the network training procedure was launched. It is worth remembering that cascade training is also a change in the value of connection weights and the development of the structure.
The table 2 summarizes the network training process.

Table 2. FANN library parameter names and values (adopted automatically by software)

| FANN library parameter name | FANN library parameter values |
|---|---|
| num_layers | 11 |
| cascade_output_stagnation_epochs | 12 |
| cascade_candidate_change_fraction | 0,01 |
| cascade_candidate_stagnation_epochs | 12 |
| cascade_max_out_epochs | 150 |
| cascade_min_out_epochs | 50 |
| cascade_max_cand_epochs | 150 |
| cascade_min_cand_epochs | 50 |
| cascade_num_candidate_groups | 2 |
| bit_fail_limit | 3,50E-01 |
| cascade_candidate_limit | 1,00E+03 |
| cascade_weight_multiplier | 4,00E-01 |
| cascade_activation_functions_count | 10 |
| cascade_activation_functions | 3 5 7 8 10 11 14 15 16 17 |
| cascade_activation_steepnesses_count | 4 |
| cascade_activation_steepnesses | 0,25 ; 0,5 ; 0,75 ; 1 |
| layer_sizes | 91 1 1 1 1 1 1 1 1 1 1 |

In figures 4-5 sample results of neural network calculations are presented. The "Desired value", i.e. values from the computer simulation, was marked with a continuous line. The "Calculated values" should correspond as well as possible to the simulation results.
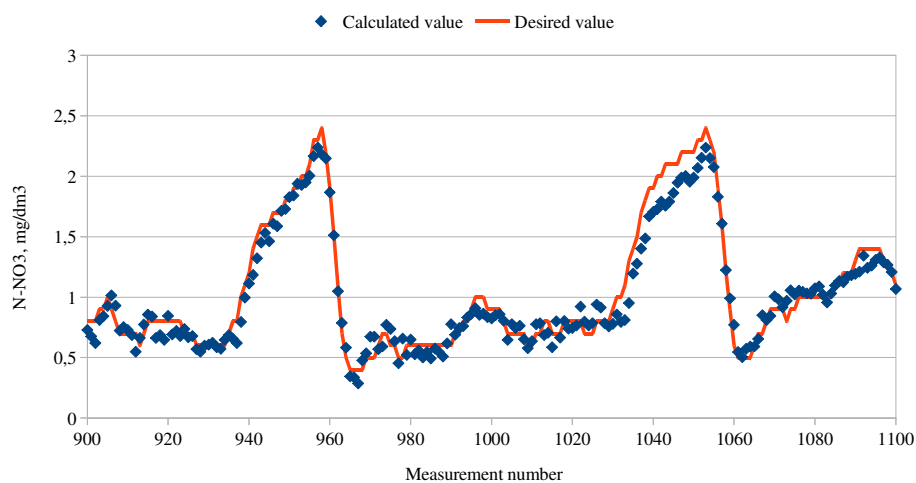
Fig. 4. Example 1: Results of neural network calculations - values of nitrate nitrogen
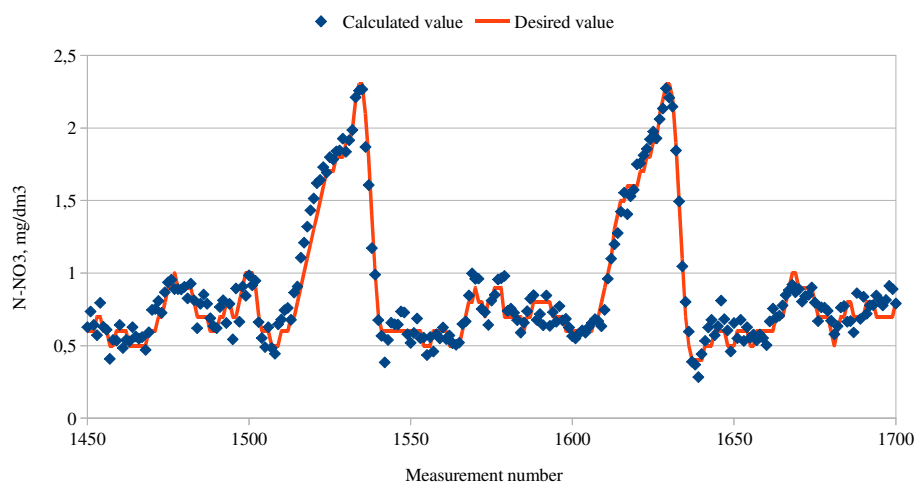(V) concentration in the denitrification chamber.



Fig. 5. Example 2: Results of neural network calculations - values of nitrate nitrogen
(V) concentration in the denitrification chamber.

The average calculation MAE (MAE - mean absolute error) for the entire test set
(3223 data rows) is 0.094 mg $N/dm^3$.

The next stage of research was to repeat the above actions but for artificially distorted data. An artificial, random noise of +- 2% of each individual value from the table was introduced to the table of data from the computer simulation.
Similarly as before, two sets of data were obtained: a teaching and a test set. The settings of the network training parameters listed in Table 2 remain the same.

The figures 6,7 show sample results of neural network calculations made for artificially noisy data.
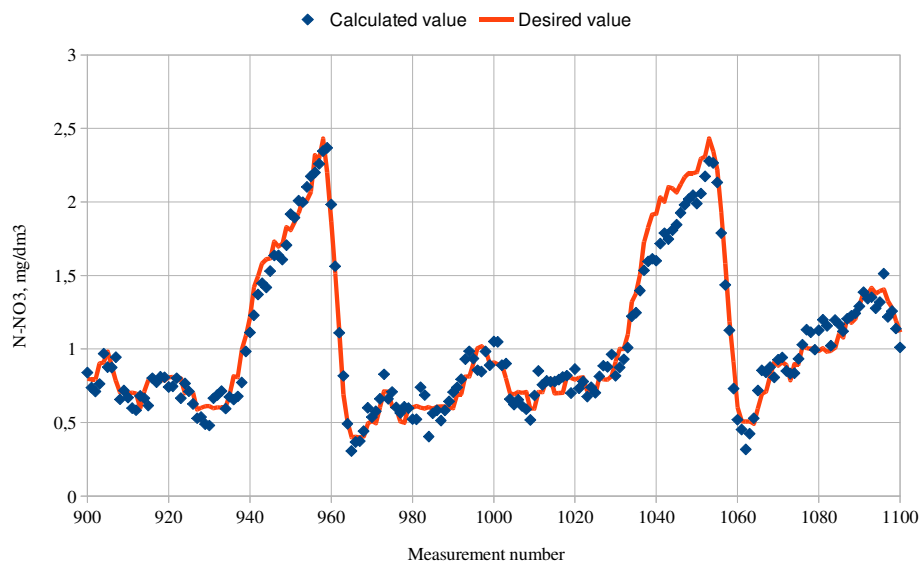


Fig. 6. Example 3: Results of neural network calculations made for artificially noisy data - values of nitrate nitrogen (V) concentration in the denitrification chamber
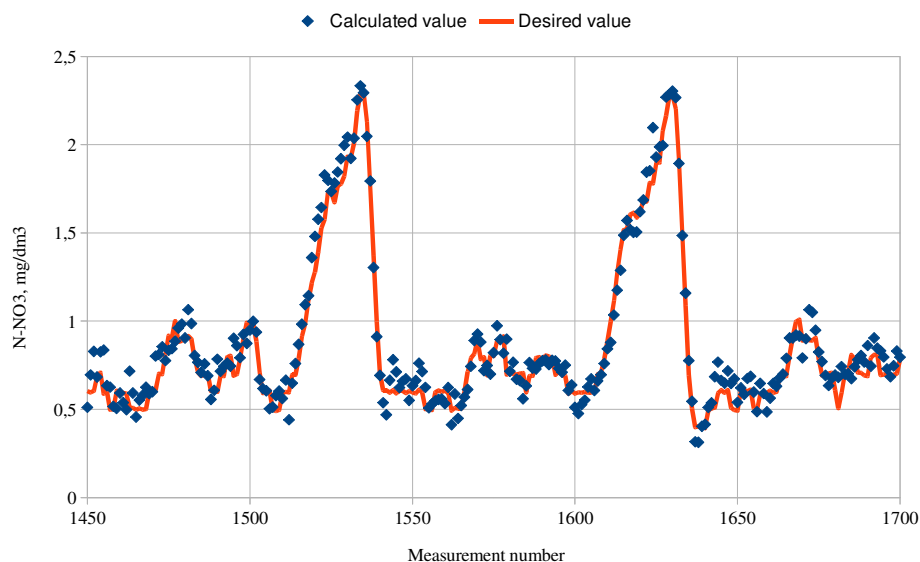
Fig. 7. Example 4: Results of neural network calculations made for artificially noisy
data - values of nitrate nitrogen (V) concentration in the denitrification chamber

The average calculation MAE for the entire test set (3223 data rows) is 0.097
mg N/dm$^3$.

It is worth noting at this point that both values of average error: 0.094 mg N/dm$^3$
for undistorted input data and 0.097 mg N/dm$^3$ for distorted data are below the
sensitivity of the currently produced probes. Depending on their design,
measuring range and purpose (activated sludge or wastewater), these probes have
different accuracies, e.g. ± 3 % of the measured value +0.5 mg/ dm$^3$ or ± 5 % of
the measured value +0.2 mg/ dm$^3$.

When trying to compare the quality of the obtained artificial neural networks with
the results of works available in the literature, it should be remembered that
differences, even only in the number of teaching examples, will affect the
possibility of obtaining good results.

For example, Wąsik et al. [13] obtained artificial neural networks acting as a
virtual sensor of nitrate nitrogen (V) concentration in wastewater treatment plant
effluent. Measurements of other forms of nitrogen in the same place were used as
input data. The reported calculation errors (MAE) are values 0.481 mg N/dm$^3$,
0.951 mg N/dm$^3$ and 2.199 mg N/dm$^3$ for three different neural network
architectures selected as the best of the 50 tested. These values are clearly higher
than in this study, but Wąsik et al. had 134 measurements of nitrogen content in

samples of wastewater flowing out of the plant over a period of 6 years. As a reminder: neural networks in this study were taught using more than 8000 data rows and each row contained 90 measurements (9 measurements from the present and another 81 from the past). Three networks described by Wąsik et al. had a number of inputs equal to 8, 4 and 6, not 90.

Finally, it is worth noting that the results of the neural network calculations are of similar quality for input data without distortion and artificially distorted data. This issue will be the subject of further research.

## 5. CONCLUSIONS

The example of the presented results shows clearly that the neural network is able to calculate the concentration of nitrate nitrogen (V) in the denitrification chamber with good accuracy. This work uses a set of input data consisting only of measurements that can easily be conducted online. Such selection of input data enables easier application in practice.

The key operation on the data is to construct a table containing data from the current and previous measurements. This information turns out to be necessary and sufficient for the neural network to be able to "learn" the complex relationships between the various process parameters.

The second stage of the work consisted in checking to what extent a slight distortion of the data will cause a deterioration in the quality of calculations. For this purpose, as described above, an artificial noise was introduced to the input data - both in the "teaching" and "testing" set. It turns out that two percent data distortion practically does not affect the quality of calculations. The average calculation error for undistorted data was 0.094 mg $N/dm^3$, while for noisy data this value was 0.097 mg $N/dm^3$.

## REFERENCES

1. Buscema M 1998. Back Propagation Neural Networks. *Substance Use & Misuse*, **33(2).**
2. Copp, J (Ed.) 2002. *The COST Simulation Benchmark: Description and Simulator Manual*. Luxembourg: Office for Official Publications of the European Community.
3. Dymaczewski, Z [red.] 2011. Sewage treatment plant operator's guide (*Poradnik eksploatatora oczyszczalni ścieków).* Poznań Polskie Zrzeszenie Inżynierów i Techników Sanitarnych Oddział Wielkopolski.
4. Kristianti, N et al. 2018. Prediction of Peat Forest Fires Using Wavelet and Backpropagation , *IJITEE*, **Vol. 2, No. 2.**

5.  Lichuan, Liu, Kuo, SM and Zhou, M 2009. *Virtual sensing techniques and their applications,* International Conference on Networking, Sensing and Control, Okayama, pp. 31-36.

6.  Nissen, S 2003. *Implementation of a Fast Artificial Neural Network Library (FANN).* Copenhagen: Department of Computer Science, University of Copenhagen.

7.  Nissen, S et al.. FANN - Fast Artificial Neural Network Library, <http://leenissen.dk/fann/wp/>, access 2020-11-30.

8.  Rosenblatt, F 1958. The Perceptron: A Probabilistic Model for information Storage and Organization in The Brain, *Psychological Review,* **65(6).**

9.  Sadecka, Z 2010. *Basics of biological wastewater treatment* (*Podstawy biologicznego oczyszczania ścieków)* Warszawa: Wydawnictwo Seidel-Przywecki.

10. Shao, S et al. 2019. Extended Kalman Filter Method based on Backpropagation Neural Network in Current Sensor Online Calibration , *IOP Conf. Series: Materials Science and Engineering* **631.**

11. Stokes, AJ, West, JR, Forster, CF and Davies, WJ 2000. Understanding some of the differences between the COD- and BOD- based models offered in STOAT. *Water Res.,* **34(4).**

12. Tadeusiewicz, R 2001. *Introduction to neural networks (Wprowadzenie do sieci neuronowych)* Kraków: StatSoft Polska.

13. Velandia, NS et al. 2017. Applications of Deep Neural Networks, Internetional *Journal of Systems Signal Control and Engineering Application* **10 (1-6)**, 61-76.

14. Wąsik, E, Chmielowski, K, Studziński, J and Szeląg, B 2018. Application of artificial neural networks to forecasting total nitrogen content in secondary effluent from treatment plants (Zastosowanie sztucznych sieci neuronowych do prognozowania zawartości azotu ogólnego w odpływie z oczyszczalni ścieków) *Ochrona Środowiska,* Vol. **40 No. 1.**