sciendo

CIVIL AND ENVIRONMENTAL ENGINEERING REPORTS

E-ISSN 2450-8594

CEER 2022; 32 (4): 0428-0439
DOI: 10.2478/ceer-2022-0066
Original Research Article

SHORT-TERM FORECAST OF OXYGEN CONCENTRATION IN NITRIFICATION CHAMBER USING ARTIFICIAL NEURAL NETWORK

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

Abstract

Due to the difficulties in implementing other methods of removing organic compounds and nitrogen from wastewater, municipal wastewater treatment plants use classical processes (nitrification and denitrification) that require large energy expenditure on aeration. The problem of high energy consumption concerns every treatment plant using aerobic activated sludge, hence the constant attempts to introduce possibly intelligent aeration control techniques. In this study, a short-term (hourly) forecast of oxygen concentration in the aeration chamber was calculated under the conditions of changing values of wastewater flow and pollutant concentrations as well as active aeration control according to an unchanging algorithm. Artificial neural networks were used to calculate the forecast. It is shown that an accurate prediction can be obtained by using different sets of input data but depending on what data we choose, the neural network required to obtain a good result has a more or less complex structure. The resulting prediction can be applied as part of a system for detecting abnormal situations and for preventing excessive energy consumption through unnecessary over-oxygenation of activated sludge.

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

One of the aspects of the idea of sustainable development is care for the natural environment. In practice, this means the necessity to build and operate sewage treatment plants wherever there are human clusters [2]. For example, in Poland, according to the latest available data of the Polish Central Statistical Office ("Statistics Poland"), in 2020 the number of all municipal wastewater treatment plants was 3,281 [4].

Due to the type of wastewater treatment methods and related processes used, wastewater treatment plants are divided into: mechanical, chemical, biological and biological with increased removal of nutrients. In 2019, the largest number of municipal wastewater treatment plants in Poland used biological methods of wastewater treatment (75%). Plants enabling increased removal of nutrients were 25% [4].

Biological wastewater treatment plants most often use the activated sludge process. This is an aerobic process, which means that oxygen must be supplied to the biological reactor, usually in the form of air pumped into the fine-bubble diffusers, or alternatively by means of surface aerators. Aeration of the contents of a biological reactor is usually a process that consumes a lot of electricity. The percentage contribution of aeration to the total energy consumption of a wastewater treatment plant varies depending on technological and operational conditions. It is often reported that the contribution is about 50% or more [12]. Oxygen supplied to the biological reactor is consumed by activated sludge in two processes: aerobic growth of microorganisms connected with biodegradation of organic pollutants present in sewage, and in the process of nitrification. The specific oxygen consumption of activated sludge in the process of growth and biodegradation of carbon compounds depends mainly on the temperature and sludge age, and ranges from about 1 kg O_2/kg BOD to about 1.3 kg O_2/kg BOD. Oxygen consumption in the nitrification process is assumed to be 4.3 kg O_2/kg N [0]. If one assumes that the ratio of BOD/N loads in the effluent flowing into the biological reactor is about 5 (this is a typical value), it turns out that nitrification consumes about 36% of the total oxygen. In practice, full removal of nitrogen is usually applied, i.e. in addition to nitrification, denitrification is used, which results in partial "recovery" of oxygen, and then the nitrogen removal process as a whole does not consume it as much. An alternative to nitrification / denitrification process for nitrogen removal that requires less oxygen is the Anammox process (anaerobic ammonium oxidation) which was envisioned as early as 1941, but it was not until the nineties that a group of bacteria capable of carrying out this process was identified at Delft University [11]. Unfortunately, due to the fact that Anammox bacteria are slow growing with a doubling time of 11-20 days [15] it is very difficult to conduct this process in the main activated sludge reactor but nitrogen removal from leachate after digested sludge dewatering is more commonly used [1, 10, 9].

Faced with the difficulty of implementing other means of removing organic compounds and nitrogen from wastewater, municipal wastewater treatment plants use classical processes that require high energy input for aeration. The problem of high energy consumption applies to essentially every plant using activated sludge, hence the constant attempts to introduce possibly intelligent aeration control techniques. There are different strategies to improve control: some are based on a mathematical model of the aeration process [12, 13], others (mostly the newer ones) use artificial intelligence techniques to build specific control algorithms [5]. The application of artificial neural networks in the calculation of complex phenomena, such as activated sludge process, requires a large amount of data at the stage of network training. Obtaining this data from the measurement devices of the treatment plant is not always easy. Hence, the mathematical model of wastewater treatment plant with activated sludge BSM1 created as a standardized platform for process studies and control strategies is eagerly used [1].

2. METHODOLOGY

There are many types of artificial neural networks, which differ in their internal structure - the way neurons are connected and their number or number of layers - and in their training methods. In this study, it was decided to use a layered network whose structure is dynamically extended during the training process (cascade training). During training, additional neurons are added, each of which forms another layer of the network, so that successive layers consist of only one neuron, but there are many layers. An example of such a network is shown in Figure 1 below.

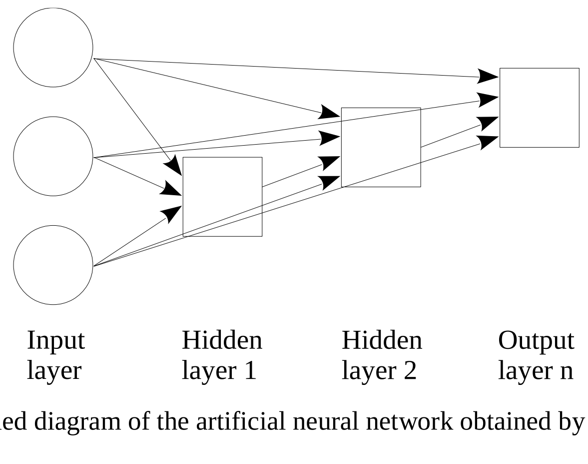


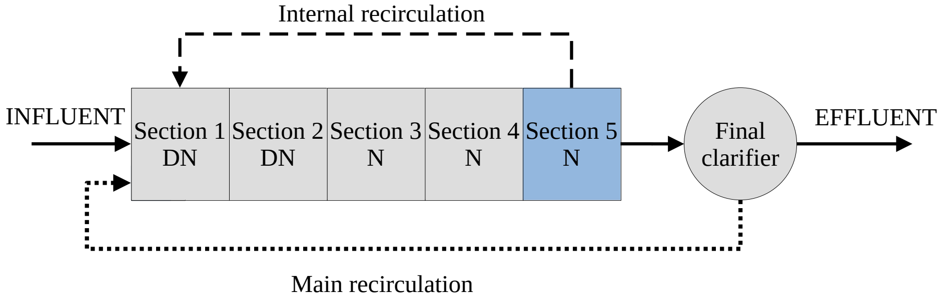
Fig. 1. Simplified diagram of the artificial neural network obtained by cascade training

Cascade training, during which a network structure is built, has the advantage that the size of the network does not need to be determined before training begins. This is important because, to date, it has not been possible to formulate an exact recipe for the selection of the network size. As a result, the number of layers and the number of neurons in each layer are selected on the basis of some fuzzy rules, one's own experience or directly by trial and error.

In the present work, cascade training was used not only to obtain neural networks that fulfil their purpose but also as a method to determine the required network size. The size of the network in this case should be considered to be not only the number of neurons but also the number of connections.

A neural network tasked with correctly reflecting 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 the present work, the task of the neural network is to calculate a prediction of the oxygen concentration in the aeration chamber based on readings of measurement probes measuring concentrations and flows elsewhere in the plant. It was decided to use a dynamic computer simulation as a data source for two reasons. The first reason is that, using the BSM1 mathematical model, which is widely recognised and described in the literature, we obtain data widely recognised as valuable material for any analysis of the activated sludge process. The second reason is that it is relatively easy to obtain the large amount of input data necessary for the neural network training process.

Figure 2 below shows the technological scheme of the biological wastewater treatment plant with activated sludge adopted in the BSM1 model. The blue colour indicates the segment of the nitrification zone where the neural network is to calculate the oxygen concentration.



DN = denitrification - non-aerated section
N = nitrification - aerated section

Fig. 2. Process diagram: Benchmark Simulation Model no. 1 (BSM1). Prepared by IWA Taskgroup on Benchmarking of Control Strategies for WWTPs

The "STOAT" application was selected for the computer simulation, the compatibility of which with the BSM1 model is described in [14]. A characteristic feature of the program is the possibility of using both models based on COD measurements (ASM1, ASM2d, ASM3) and BOD (ASAL1 ... 5). Models from the ASM family are widely used and considered the best for mathematical modeling of the activated sludge process, but their common disadvantage is the need to know the wastewater flowing into the treatment plant, which goes far beyond the set of typically performed measurements (BOD_5, COD, suspension concentration, nitrogen and phosphorus concentration). ASAL models use a typical set of measurements of pollutant concentrations and indicators as input data, but due to the success of ASM models, work on them has not been continued for a long time and some aspects of calculations using ASAL models require improvement, especially when combining activated sludge processes with methane fermentation of sludge [14].
The BSM1 model is a set consisting of:
- ASM1 mathematical model,
- input data describing the wastewater flow rate and the concentration of pollutants in these wastewater broken down into fractions in accordance with the requirements of the ASM1 model,
- a detailed description of the plant's technological layout: size and purpose of the facilities, connection method (flows and recirculation) and control.

The specific implementations of the ASM1 model from different software vendors may differ slightly and therefore STOAT offers a special version of ASM1 tested and agreed with the BSM1 programmers' working group [14].

Teaching neural networks was carried out using the "FannTool" tool - a graphical interface to the popular software library implementing the process of learning neural networks - FANN [0, 8].

Computer simulation of wastewater treatment plants using the BSM1 model allows obtaining even more detailed information about the process flow than is practically possible, because only some of the data available in the computer simulation can be easily measured online. For example, as a result of the simulation, we obtain information about COD in any place of the installation. Although there are automatic COD analyzers, they do not perform measurements as quickly and often as measurement probes that measure the concentration of, for example, oxygen or nitrogen. In this paper it is assumed that all information about the process, which is calculated by the simulator, but in practice cannot be obtained online or it is difficult - will be ignored - the neural network will not use this data for training.

The instantaneous values of all state variables are the result of processes occurring in the more and more distant past. Therefore, it makes no sense to expect that the neural network will be able to correctly calculate the value of the oxygen concentration in the reactor only from the values of the measurements carried out at the same time. The input data set should include not only current measurements, but also several measurements from the immediate past.

3. RESEARCH RESULTS

The study aimed to obtain a neural network capable of calculating a short-term prediction of the oxygen concentration in the aeration chamber labelled Section 5 in Fig. 2. In the BSM1 model, it was assumed that nitrification chambers 3 and 4 would have unchanged aeration and that air volume control would be only in Section 5. It is easy to guess that with a relatively constant activated sludge concentration, the oxygen concentration in Section 5 depends mainly on the instantaneous loads: BOD_5 and nitrogen flowing into the reactor and the intensity of aeration. Hence, the following set was chosen as input data for the tests:

Table 1. Set of measurements to be considered as possible training data for the neural network

No.	Place of measurement	Type of measurement
1	Influent	Flow (m^3/h)
2	Influent	Total COD (mgO_2/dm^3)
3	Influent	Ammonia (mg/dm^3)
4	Section 3	Dissolved oxygen (mgO_2/dm^3)
5	Section 4	Dissolved oxygen (mgO_2/dm^3)
6	Section 5	Dissolved oxygen (mgO_2/dm^3)
7	Section 5	Nitrate (mgN/dm^3)
8	Section 5	Ammonia (mgN/dm^3)
9	Section 5	KLa (1/h)

The KLa value represents the capacity of the aeration system in practice. Stating this capacity in such units is typical of ASM mathematical models. It can be considered that for a specific aeration system and weather conditions there is a strong relationship between the KLa value and the amount of oxygen delivered to the reactor. Providing KLa makes the units used independent of the quality and degree of wear of the diffusers and other aeration system components.

Using a computer simulation, 8065 rows of data were generated, each of which was staggered in relation to the next by 15 minutes (the simulation lasted 84d). Thus, data were obtained where, for each row, there was information about the current state, past states (previous rows) and future states (next rows). Hence, it was possible to calculate a 'forecast' of the oxygen concentration as the average value of the four measurements following the 'present' measurement. This particular forecast was the value that was fed to the neural network during learning along with the other data, but during the 'stand-alone' calculations the neural network was tasked with calculating this forecast without knowing its value.

The values listed in Table 1 were used as input data. These were the 'present' time data and a number of previous row data, i.e. historical data. The number of previous rows used as input to the neural network varied for the different tests and ranged from 4 to 14.

The different tests also differed in the subset of data - not all the data listed in Table 1 were always used as input to the network. This means that richer or less rich sets of information were tested, from which the forecast was calculated. Moreover, more or less historical data was used.

Each test consisted of training three neural networks based on the same data. Each of the three networks then performed forecast calculations on data not used during the training phase. The results of the calculations were averaged and MSE and MAE were calculated as measures of the quality of the resulting forecast.

As a result of the many tests carried out, it has become apparent that the best results are obtained using the data set shown in Table 2.

Table 2. The input data set for which the most accurate oxygen concentration forecast was obtained

No.	**Place of measurement**	**Type of measurement**
1	Influent	Flow (m^3/h)
2	Influent	Ammonia (mg/dm^3)
3	Section 5	Dissolved oxygen (mgO_2/dm^3)
4	Section 5	Nitrate (mgN/dm^3)
5	Section 5	Ammonia (mgN/dm^3)
6	Section 5	KLa (1/h)

As previously mentioned, the task of the neural network was to calculate the oxygen concentration forecast as an average value over the next hour (an average of 4 measurements taken every 15 minutes each). Each forecast was in this case calculated based on the current state and 9 previous states making a total of 60 inputs. This automatically means that the input layer size of the neural network was also 60 neurons.

The figures below show an example excerpt of the calculation results, indicating the correct forecast (solid line, 'Desired'), the forecast calculated by the neural network ('Calculated') and the calculation error.

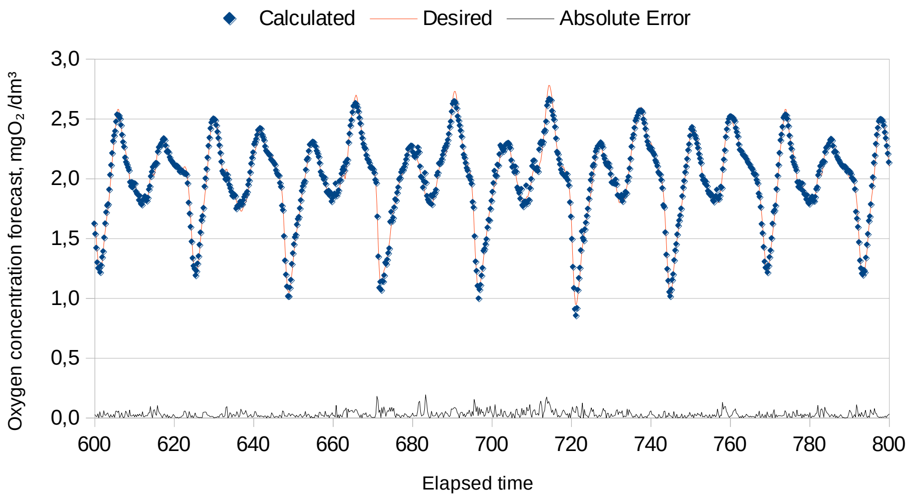


Fig. 3. Example 1: Results of neural network calculations - oxygen concentration forecast values

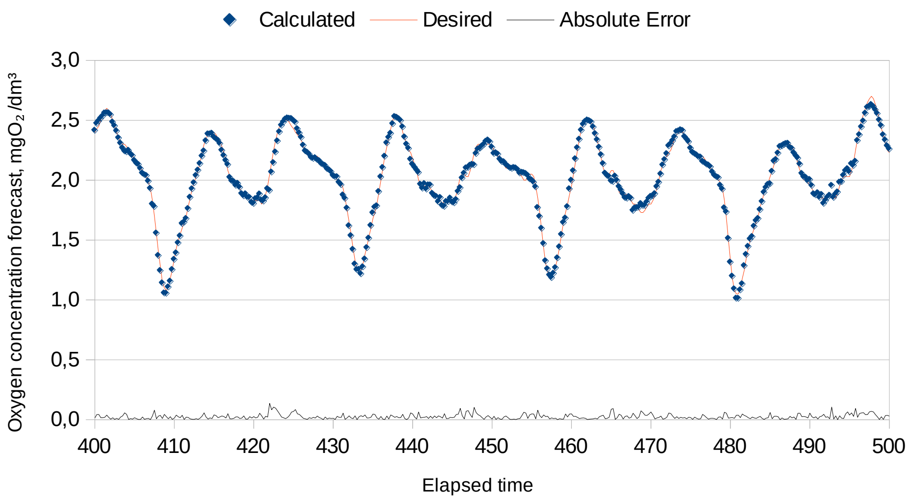


Fig. 4. Example 2: Results of neural network calculations - oxygen concentration forecast values

For the data set shown in Table 2, a number of tests were carried out to determine what amount of 'historical' data is needed to get the best possible result from the neural network calculations. A summary of the test results is shown in Table 3.

The MSE and MAE error values are calculated for each test separately and then also calculated for the averaged calculation results of the three neural networks. It can be seen that if we perform three tests for three different, separately trained neural networks and average the results of the calculations of these networks then the result obtained will have a smaller error than if the calculations were performed by a single network.

Due to the high accuracy of the results obtained, a limit on the number of hidden neurons was set at 10. The results presented in Table 3 show that there is no need to increase the number of historical data beyond 9 measurements (this corresponds to 135 minutes).

Table 3. Test results of neural networks calculating oxygen concentration prediction. The value 'H' indicates the number of 'historical' data rows, for example H=5 for six data rows: the current row and five previous measurements

'H' value	Network	Number of inputs	Number of layers	Total neurons and biases	Total connections	MSE	MAE
5	1	36	10	48	462	0,00291	0,03970
	2	36	10	48	462	0,00244	0,03669
	3	36	10	48	462	0,00298	0,04076
	avg(1,2,3)					0,00197	0,0328
7	1	48	10	60	594	0,00309	0,03985
	2	48	10	60	594	0,00245	0,03631
	3	48	10	60	594	0,00263	0,03726
	avg(1,2,3)					0,00192	0,03201
9 (the best)	1	60	10	72	726	0,00231	0,03542
	2	60	10	72	726	0,00260	0,03834
	3	60	10	72	726	0,00193	0,03239
	avg(1,2,3)					0,00149	0,02882
11	1	72	10	84	858	0,00229	0,03614
	2	72	10	84	858	0,00259	0,03778
	3	72	10	84	858	0,00212	0,03364
	avg(1,2,3)					0,00155	0,02924
13	1	84	10	96	990	0,00227	0,03420
	2	84	10	96	990	0,00325	0,04066
	3	84	10	96	990	0,00320	0,04061
	avg(1,2,3)					0,00174	0,03037
15	1	96	10	108	1122	0,00359	0,04287
	2	96	10	108	1122	0,00380	0,04476
	3	96	10	108	1122	0,00269	0,03710
	avg(1,2,3)					0,00185	0,03127

It is theoretically possible to increase the accuracy of the calculation by removing the size constraint of the neural network. Table 4 shows an example result of such a calculation for the most promising input data set.

Table 4. Test results of neural networks calculating oxygen concentration prediction. Automatic determination of the network size, without the limitation of 10 neurons in the hidden layers. The value 'H' indicates the number of 'historical' data rows, for example H=9 for 10 data rows: the current row and 9 previous measurements

'H' value	Network	Number of inputs	Number of layers	Total neurons and biases	Total connections	MSE	MAE
9 (the best)	1	60	15	77	1096	0,00053	0,01739
	2	60	20	82	1491	0,00044	0,01541
	3	60	16	78	1173	0,00043	0,01543
	avg(1,2,3)					0,00029	0,01259

4. CONCLUSIONS

A series of tests was performed to obtain artificial neural networks capable of calculating the forecast of the average value of oxygen concentration for the next hour (average value of four measurements performed every 15 minutes). The tests were performed with the use of a cascade algorithm, which selects the internal structure of the network itself (the number of neurons) in such a way as to obtain the lowest possible calculation error (MSE). Initially, the tests differed in the data set - various subsets were selected from Table 1. For the best set (Table 2), it was additionally determined how much historical data is needed to obtain the best possible result (Table 3). The obtained neural networks each time automatically took the size of 10 neurons in hidden layers, because due to the high accuracy of calculations, it was decided to limit the size of the network. Tests have also been conducted without this limitation. The results show the possibility of an additional increase in the accuracy of calculations, but the size of the network, despite the lack of limitation, did not increase above 82 neurons and biases.

From the research carried out, it appears that obtaining a good prediction of oxygen concentration should be a fairly easy task. This is important because the training of the neural network carrying out such a task should be repeated over and over again in such a way that the network is learned on the current data. It should be noted that the forecast results depend on the adopted aeration control algorithm (they adapt to this algorithm in a way). Thus, if the treatment plant operator finds that it is worth correcting the algorithm based on the forecast, it will also be necessary to update the neural network structure on an ongoing basis. Tests carried out show that different network configurations and different input data configurations produce decent computational results and there is a clear chance that reprogramming the neural network can take place without human supervision, automatically.

Concluding on the results obtained, it can be said that the use of artificial neural networks to improve aeration control and thus reduce the energy consumption of the plant - is possible.

ACKNOWLEDGEMENTS

This work was supported by the Faculty of Power and Environmental Engineering, Silesian University of Technology as part of statutory research.

REFERENCES

1. Al-Hazmi, H, Lu, X, Grubba, D, Majtacz, J, Kowal, P and Mąkinia, J 2021. Achieving Efficient and Stable Deammonification at LowTemperatures—Experimental and Modeling Studies. *Energies* **14(13)**, 3961. https://doi.org/10.3390/en14133961
2. Copp, JB (Ed.) 2002. The COST Simulation Benchmark: Description and Simulator Manual (a product of COST Action 624 & COST Action 682). Office for Official Publications of the European Community.
3. Dasgupta, P 2007. The idea of sustainable development. Sustain Sci. **2**,5–11. https://doi.org/10.1007/s11625-007-0024-y
4. DWA - German Association for Water, Wastewater and Waste. (2020). DWA -A 131 Wymiarowanie jednostopniowych oczyszczalni ścieków z osadem czynnym (Dimensioning of Single-Stage Activated Sludge Plants). Wydawnictwo Seidel-Przywecki Sp. z o.o.
5. Environment 2021. Statistics Poland, Spatial and Environmental Surveys Department. https://stat.gov.pl/en/topics/environment-energy/environment/environment-2021,1,13.html
6. Li, M, Hu, S, Xia, J, Wang, J, Song, X and Shen, H 2020. Dissolved Oxygen Model Predictive Control for Activated Sludge Process Model Based on the Fuzzy C-means Cluster Algorithm. *International Journal of Control, Automation and Systems* **18(9)**, 2435–2444. https://doi.org/10.1007/s12555-019-0438-1
7. Nissen, S 2003. Implementation of a Fast Artificial Neural Network Library (FANN). Copenhagen: Department of Computer Science, University of Copenhagen.
8. Nissen, S et al., FANN - Fast Artificial Neural Network Library, <http://leenissen.dk/fann/wp/>, access 2020-11-30.
9. Ochs, P, Martin, BD, Germain, E, Wu, Z, Lee, P-H, Stephenson, T, van Loosdrecht, M and Soares, A 2021. Evaluation of a Full-Scale Suspended

Sludge Deammonification Technology Coupled with an Hydrocyclone to Treat Thermal Hydrolysis Dewatering Liquors. *Processes* **9(2)**, 278. https://doi.org/10.3390/pr9020278

10. Pęciak-Foryś, G, Barbusiński, K and Filipek, K 2020. Analysis of the possible application of deammonification technology in the municipal wastewater treatment plant in Zabrze. *Architecture, Civil Engineering, Environment* **12(4)**, 115–123. https://doi.org/10.21307/ACEE-2019-057
11. Remy, M, Hendrickx, T and Haarhuis, R 2016. Over a Decade of Experience with the ANAMMOX Reactor Start-up and Long-Term Performance. *Proceedings of the Water Environment Federation* **2016(12)**, 4393–4405. https://doi.org/10.2175/193864716819706554
12. Rieger, L, Alex, J, Gujer, W and Siegrist, H 2006. Modelling of aeration systems at wastewater treatment plants. *Water Science and Technology*, **53(4–5)**, 439–447. https://doi.org/10.2166/wst.2006.100
13. Schraa, O, Rieger, L and Alex, J 2017. Development of a model for activated sludge aeration systems: Linking air supply, distribution, and demand. *Water Science and Technology* **75(3)**, 552–560. https://doi.org/10.2166/wst.2016.481
14. 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)**.
15. Strous, M, Kuenen, JG and Jetten, MSM 1999. Key physiology of anaerobic ammonium oxidation. *Applied and Environmental Microbiology* **Vol. 65 (7)**, pp. 3248–3250.

Editor received the manuscript: 06.10.2022