Science, Engineering & Education,8, (1), 2023, 73-76

Virtual Sensor for Measuring the Concentration of Cu(II) in River Waters

Kosta Boshnakov*, Bogdana Koumanova, Emil Mihailov

University of Chemical Technology and Metallurgy, Sofia, Bulgaria

Received 10 September 2023, Accepted 15 October 2023

ABSTRACT

The possibility of creating a virtual sensor for measuring the concentration of copper ions (Cu(II)) in water bodies was investigated. The creation of a virtual sensor is based on the possibility of using cheap and reliable sensors for easily measurable variables that are mathematically associated with the target variable. The values of the parameters characterizing the water quality have been collected for a long time and the surface water samples have been taken from the same points of the river water. Based on a correlation analysis, two input variables were selected for the virtual sensor - pH and electrical conductivity. At a certain stage of research, temperature was added as an input variable, but this did not lead to a significant increase in the accuracy of the virtual sensor. To create mathematical models for the purposes of the developed virtual sensor, a large number of artificial neural networks with feedforward propagation were trained and regression models were derived. In the article, the mathematical models that describe the measurements carried out with the highest accuracy are presented.

Keywords: *virtual sensor, copper ions, artificial neural networks, regression models.*

INTRODUCTION

Copper pollution of the world's water resources is becoming increasingly serious and poses a serious threat to human health and aquatic ecosystems [1]. Getting into the human body larger amounts of copper can lead to irritation in the digestive system, and excessive intake of copper increases the risk of Alzheimer's and other neurological diseases.

Continuous monitoring of Cu (II) in water bodies will provide an opportunity for timely response of the responsible services and thus reduce the danger to human health and aquatic ecosystems.

The most widely used methods for copper detection are traditional atomic spectrometry, UV-Vis spectrometry and electrochemical analysis. Studies have been conducted to measure

**Correspondence to: Kosta Boshnakov, University of Chemical Technology and Metallurgy, 8 Kliment Ohridski blvd., 1797 Sofia, Bulgaria, E-mail: kb@uctm.edu*

Cu (II) in drinking water with a fiber optic chemical sensor [2], but there is no information on studies related to the measurement of Cu (II) in waters containing high concentrations.

Virtual sensors are mathematical models that use measurements of cheap and reliable physical sensors as input variables, and as a result of the mathematical model, the values of a variable that is difficult and expensive to measure, or unmeasurable with available means, for which the virtual sensor is intended [3], are obtained. As mathematical models, static models from the class of intelligent systems, non-linear regression models, etc., are most often used.

The aim of the present work is to develop a virtual sensor for continuous measurement of the concentration of copper ions (Cu(II)) in river waters based on continuous measurements of easily measurable variables - pH, water temperature T, dissolved oxygen concentration (DO), and electrical conductivity (EC). A target variable is the concentration of copper ions, Cu(II).

PRELIMINARY DATA ANALYSIS

The study is based on the measurements of the above-mentioned variables over a period of several years. After an analysis of the available measurements, more than 250 sets of data with measurements were selected for complete completion with the necessary data for deriving mathematical dependencies. Table 1 shows a sample of 262 data sets.

For a preliminary assessment of the influence of the easily measured variables on the target variable, a correlation analysis was conducted. Table 2 presents the correlation coefficients between easily measured variables and copper ion concentration.

Regardless of the fact that from Table 2 it is obvious that significant correlation coefficients are r(pH, Cu(II) = -0.792 and r(EC, Cu(II)) = 0.895, their significance was also proven by a statistical method, which gives grounds for building a mathematical model initially a dependence on the type Cu (II) = f(pH, EC) was sought.

Table 1 Sample and statistics for the measurements carried out.

No	pH	T	DO	EC	Cu (II)
		^{0}C	mg/l	μS/cm	μg/l
1	6.53	4.0	8.00	370.0	646.0
2	6.47	8.6	5.96	160.0	66.0
…	…	…	…	…	…
131	8.86	9.7	6.84	261.0	52.0
…	…	…	…	…	…
262	9.15	24.1	14.30	400.0	19.5

Table 2. Correlation coefficients.

	pH	T	DO	EC
pH	1	0.14	0.22	-0.745
T	0.14	1	-0.098	-0.089
DO	0.22	-0.098	1	-0.086
EC	-0.745	-0.089	-0.086	1
Cu(II)	-0.792	-0.081	-0.080	0.895

DERIVING MATHEMATICAL DEPENDENCIES FOR SOFTWARE SENSOR TARGETS

Research was conducted on the possibility of mathematical description of the dependence of Cu(II) primarily on pH and EC, and then the temperature T was also included in the research process. For this purpose neural networks of the class with straight distribution and regression dependencies were applied.

Mathematical models with neural networks

To determine the initial structure of the neural network, the rule of thumb is used that the number of neurons n in the hidden layer is calculated by the expression $n = 2m+1$, where m is the number of input variables, following this rule for $m = 2$, $n = 5$.

Table 3 presents the neural networks with input variables pH and EC and output - the Cu(II) concentration. The first digit in the neural model notation is the number of input neurons, the second is the number of neurons in the hidden layer, and the third is the number of output neurons.

In the hidden layer, the activation functions are hyperbolic tangent, and for output neurons, the activation function is linear [4, 5]. R is the coefficient of multiple correlation, and Radj is the coefficient of multiple correlation taking into account the number of evaluated coefficients. The Table 3 also includes the number of coefficients to be evaluated in the neural network.

The same procedures were performed to derive neural models under three input variables pH, T and EC (Table 4).

From the results presented, it can be seen that neural mathematical models have high accuracy.

Mathematical models with regression dependencies

Based on the 262 data sets, regression mathematical models were derived for Cu (II) with two and three input variables, in the first case the input variables were pH and EC, and in the second case - pH, T and EC.

Linear mathematical models, incomplete quadratic models with and without a free term have been consistently derived [6, 7]. The mathematical models describing the available 262 measurements with the highest accuracy and the corresponding multiple correlation coefficients are presented in Table 5.

The very close values of R between the two models 2 and 3 and also of Radj give reason to consider that temperature has a minor influence.

Table 3. Trained neural networks for input variables pH and EC.

	Neural model	Number of coefficients	R	Radj
NN model 1	2-5-1	21	0.9832	0.9817
NN model 2	2-4-1	17	0.9830	0.9821
NN model 3	2-3-1	13	0.9793	0.9783

Table 4. Trained neural networks for input variables pH, T and EC.

	Neural model	Number of coefficients	R	Radj
NN model 4	3-7-1	36	0.9711	0.9660
NN model 5	3-5-1	21	0.9831	0.9816
NN model 6	3-3-1	19	0.9853	0.9842

Table 5. Regression models.

	Mathematical model	Number of coefficients	R	Radj
Reg. model 1	Cu (II) = -3463 + 29,980 EC - 3,635 pH* *EC + 50,07 pH^2	4	0.9769	0.9766
Reg. model 2	Cu(II)= -904pH+28.997 EC-3.498 pH* *EC+107.8 pH^2	4	0.9809	0.9860
Reg. model 3	Cu (II) = -1158 pH + 32,653 EC - - 3,904 pH*EC - 0,1339 T*EC + 148,5 pH^2	5	0.984	0.9839

CONCLUSIONS

Based on the study it can be concluded that:

- In order to implement a virtual sensor for measuring Cu (II) in the river basin studied, neural and regression mathematical models were derived.
- Input, easily measurable variables included as inputs to the virtual sensors are pH, temperature T and conductivity EC.
- Six neural and three regression mathematical models are presented, which describe all the available measurements with the highest accuracy.
- According to the derived mathematical models, obviously the inclusion of temperature as an input variable slightly increases the corresponding values of R and Radj.
- The large number of the derived mathematical models for the implementation of a virtual sensor for measuring Cu (II) in the river water under investigation allows choosing the most suitable one for implementation in the hardware construction of the virtual sensor.

Acknowledgements

This work has been carried out in the framework of the National Science Program "Environmental Protection and Reduction of Risks of Adverse Events and Natural Disasters", approved by the Resolution of the Council of Ministers No 577/17.08.2018 and supported by the Ministry of Education and Science (MES) of Bulgaria (Agreement No Д01-271/09.12.2022).

REFERENCES

1. Y.Liu, H.Wang, Y.Cui, Nan Chen, Removal of Copper Ions from Wastewater: A Review, Int. J. Environ. Res. Public Health, 2023, 20, (5), 3885; https://doi.org/10.3390/ijerph20053885
2. M. Pesavento, A. Profumo, D. Merli, L. Cucca, L. Zeni, N. Cennamo, An Optical Fiber Chemical Sensor for the Detection of Copper(II) in Drinking Water, Sensors, 2019, 19, (23), 2019, 5246. https://doi.org/10.3390/s19235246
3. D. Martin, N. Kuehl, G. Satzger, Virtual Sensors, Bus Inf Syst Eng., 63, (3), 2021, 315-323. https://doi.org/10.1007/s12599-021-00689-w
4. E. Bezerra, Overview of Artificial Neural Networks, 2019, https://eic.cefet-rj.br/~ebezerra/wp-content/uploads/2018/02/2019-UFPR-SMNE.pdf
5. M. Madhiarasan, M. Louzazni, Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Application, Journal of Electrical and Computer Engineering,2022, v. 2022, Article ID 5416722 https://doi.org/10.1155/2022/5416722
6. Multiple Linear Regression, https://www.colorado.edu/amath/sites/default/files/attached-files/lesson12_multregression.pdf
7. K. Boshnakov, B. Koumanova, Marian Varbanov, Emil Mihailov, Silviya Lavrova - Popova, Evaluation of Key Indicators Based on Automatically Measured Variables and Forecasting the Spread of Pollutants, Science, Engineering & Education, 6, (1), 2021, 55-64.