Ecological applications based on bacterial community abundance in reservoirs using an artificial neural network approach

Ivana Radojević[1*], Aleksandar Ostojić[1], Vesna Ranković[2]

[1] University of Kragujevac, Faculty of Science, Institute of Biology and Ecology, 34000 Kragujevac, Radoja Domanovića 12, Republic of Serbia; e-mail: ivana.radojevic@pmg.kg.ac.rs
[2] University of Kragujevac, Faculty of Engineering, Department for applied mechanics and automatic control, 34000 Kragujevac, Sestre Janjić 6, Republic of Serbia

* *Corresponding author*

DOI: 10.46793/ICCBI23.317R

Abstract: The objective of this study is to analyze the influence and predict abundance the heterotrophic bacteria (psychrophile; mesophile) and facultative oligotrophic bacteria as a reflection of ecological relationships in reservoirs and water quality. We used artificial neural networks (ANNs) to develop models based on input variables derived from two different reservoirs. The neural network models were developed using experimental data which is collected for ten years. Although reservoirs have a different position, different morphometric qualities, trophic state and dominant bacterial community there is a possibility of predicting these bacterial communities with the same input parameters. Comparing the modeled values by ANN with the experimental data indicates that neural network models provide accurate results. The important conclusion of this work is that ANNs can provide a flexible and applicable tool in monitoring water quality across bacterial communities in reservoirs.

Keywords: ecological application, feedforward neural network, reservoir, water quality

1. Introduction

The microbiological properties of water represent a significant and valid indicator of its quality. Heterotrophic bacteria show a high level of metabolic activity and contribute to a large extent to the process of decomposition of organic matter and its transformation. Contrary to heterotrophic bacteria the dominance of facultative oligotrophic bacteria in waters with a small amount of dissolved organic matter which favors their development as autochthonous community is confirmed. The basic problem in the case of water quality monitoring is the complexity associated with analyzing a large number of variables. The artificial neural networks (ANNs) have been successfully used as tools in the fields of water quality prediction and forecasting [1]. Much less common are the examples of the application of ANN networks as tools for predicting ecological status in aquatic systems based on the knowledge of heterotrophic and oligotrophic communities. LOBOVA et al. [2] studied the anthropogenic impact on the prediction of community heterotrophic bacteria, and RADOJEVIC et al. [3] applied a neural network for predicting the facultative oligotrophic bacteria in two reservoirs with different trophic states. Due to the significance of heterotrophic and facultative oligotrophic bacteria in analyzing complex microbiological relations, and as indicators of water quality and ecological conditions, the objective of this study is the development of different ANN models used for the prediction of their abundance.

2. Materials and methods

2.1 Study area and water quality data

The city of Kragujevac (in the central part of Serbia) is supplied with water from Gruža and Grošnica reservoirs. The data set used in this study was generated by monitoring of the water quality of Gruža and Grošnica reservoirs. The data set includes the data of the laboratory for water quality inspection of the public utility company for water supply in Kragujevac. Monthly sampling was carried out during the period of ten years (1998 to 2008). Three permanent sampling sites were selected for qualitative and quantitative sampling for Grošnica reservoir and five sampling sites for Gruža reservoir Samples were taken at every 5 m of depth. The data set for Gruža reservoir includes 455 data samples. The available set of data was divided into training, validation, and test subsets. In the training process of the FNN 273 samples were used. The ANN model was tested by using 91 randomly selected data. 91 data was taken for the validation set. The data set for Grošnica reservoir includes 199 data samples. In the training process of the FNN, 121 samples were used. 39 randomly selected data were taken for the validation set. The ANN model was tested by using 39 selected data. The variables of those data sets is: water temperature (T), dissolved oxygen (DO), pH, manganese (Mn), iron (Fe), electrical conductivity (EC), m-alkalinitet (m-alk.), total phosphate (TP), nitrites ($NO2^-$), ammonia ($NH4^-$), chlorophyll-a (Chl-a), chemical oxygen demand (COD), 5-day biochemical oxygen demand (BOD5), number of heterotroph (psychrophile) (Hp), heterotroph (mesophile) (Hm) and facultative oligotroph (FO). Basic statistics of the measured water quality variables in Gruža Reservoir and Grošnica Reservoir during this period are given in the paper RADOJEVIC et al. [3].

2.2 Performance criteria and sensitivity analysis

The sensitivity analysis is based on the correlation of change percentages between inputs and outputs, where the least sensitive input is identified and deleted [4]. Pearson’s correlation coefficient is one of the most used when choosing appropriate inputs for an artificial neural network. The correlation coefficient is defined as the degree of correlation between measured and modeled values:

$$R = \frac{\sum_{i=1}^{N_o}(y_i - \bar{y})(y_{mi} - \bar{y}_m)}{\sqrt{\sum_{i=1}^{N_o}(y_i - \bar{y})^2 \sum_{i=1}^{N_o}(y_{mi} - \bar{y}_m)^2}} \quad (1)$$

where and represent network output and the measured value of i-th element and represent their average and represent the number of observations.

2.3 Artificial neural network: model and training algorithm

To predict the condition of bacterial communities in reservoirs, there has been used non-recurrent neural network (feed-forward neural network, FNN) with different numbers of variables, depending on which bacterial community had been modeled. Levenberg-Marquardt backpropagation (LMA) was used in this study. MATLAB Neural Network Toolbox was used for the implementation of neural networks. ANN model, training algorithm, and performance criteria are given in the previous paper [5].

3. Results and discussion

3.1 Modeling and predicting H, Hm and FO using ANNs

The first phase of modeling was designed to determine suitable sets of input parameters for modeling Hp, Hm, and FO. Based on the correlation coefficient for the whole data set, the initial input parameters for a particular model were selected. The second phase of modeling used only the sets of input parameters yielding good results in the first phase and was targeted at finding the most suitable set of input parameters and the best-performing network structure. The number of hidden neurons increased stepwise to a maximum of 35. Table 1 shows the obtained models with the best performances. The reservoirs, models, selected inputs, outputs, network architecture, correlation coefficients for the whole data set, are shown. It is obvious from Table 1 that only modeling and predicting of Hp was not possible with the same input parameters in both reservoirs. In Gruža Reservoir model ANN-1-Hp shows the best qualities with the smallest number of inputs.

Table 1. Summary of artificial neural networks developed for predicting different bacteria community levels in reservoirs in Serbia.

Reservoir	ANN - output	Input variables selected	Architecture[1]	r (for all data)
Gruža	ANN-1-**Hp**	FO, Hm, COD	3-30-1	0.970
Grošnica	ANN-2-**Hp**	FO, T, TP, Fe	4-30-1	0.944
Grošnica	ANN-3-**Hp**	FO, T, pH, Chl-*a*	4-10-1	0.915
Gruža	ANN-4-**Hm**	Hp, FO, T, DO, COD, BOD5	6-30-1	0.881
Gruža	ANN-5-**Hm**	Hp, FO, T, DO, EC, Fe, Mn	7-30-1	0.944
Grošnica	ANN-6-**Hm**	FO, Hp, T, DO, EC, Fe, Mn	7-30-1	0.935
Gruža	ANN-7-**FO**	Hp, T, NO_2^-	3-30-1	0.947
Grošnica	ANN-8-**FO**	Hp, T, NO_2^-	3-30-1	0.945

[1]The architecture of an ANN model is the ANN structure that presents the number of neurons in each of three layers—input, hidden, and output; [2] r is the linear correlation between predicted values and measured values.

LOBOVA et al. [2] with help of artificial neural networks, and over a number of heterotrophic bacteria, successfully predict the condition of the ecosystem of a lake exposed to anthropogenic influence. Modeling and predicting Hm in two reservoirs were possible with the same input parameters. For Gruža Reservoir model ANN-5-Hm and Grošnica Reservoir model ANN-6-Hm, have the same input parameters, whereby the correlation coefficients (r) for all data are almost the same for both models. Predicting and modeling of Hm in Gruža Reservoir is possible with a smaller number of input parameters (ANN-4-Hm). RADOJEVIC et al. [3], showed that in both reservoirs modeling and predicting of FO with the same input parameters is possible (11). Table 1 shows that the number of input parameters was reduced from 11 to 3 (ANN-7-FO and ANN-8-FO). The parameters were selected according to sensitivity analysis, whereby, the top three parameters were selected. Correlation coefficients (r) for all data are better in new models (Table 1) than in the previous [3].

3. Conclusions

The models with the same inputs can predict, equally well, the number and the dynamics of bacterial communities in reservoirs with different positions, morphometric characteristics, trophic status, and dominant bacterial community. The gained knowledge contributes to a better understanding of relations, to ecological correlation and causality of the bacterial community, to more effective control, as well as to the reduction of expenses needed for monitoring the state and managing the water resources.

Acknowledgment

This research is funded by the Ministry of Education and Ministry of Science, Technological Development and Innovation, Republic of Serbia, Grants: No. 451-03-47/2023-01/ 200122.

References

[1] K. CHAU, *A review on integration of artificial intelligence into water quality modeling*, Marine Pollution Bulletin, 52 (2006): 726-733.

[2] T.I. Lobova, Yu.P. Lankin, L.Yu. Popova, *Assessing the anthropogenic impact on Lake Shira from antibiotic resistance of heterotrophic bacteria by neural networks methods*, Mikrobiologiya, 76 (2007) 263-270.

[3] Radojevic I., Lj. Comic, V. Rankovic, A. Ostojic, M. Topuzovic, *Applying Neural Networks for Predicting the Facultative Oligotrophic Bacteria in Two Reservoirs with Different Trophic State*, Journal of Ecological Protection and Ecology, 14 (2013): 55-63.

[4] Maier, H.R., Dandy, G.C. (1996): *The use of artificial neural networks for the prediction of water quality parameters*. Water Resources Research. 32 (4): 1013-1022.

[5] V. Ranković, J. Radulović, I. Radojević, A. Ostojić, Lj. Čomić, *Neural network modeling of dissolved oxygen in the Gruža reservoir, Srbija,* Ecological modeling, 221 (2010) 1239-1244.