Scholar Publishing Group
Water Pollution Prevention and Control Project
https://doi.org/10.38007/WPPCP.2023.040104
ISSN 2790-2730 Vol. 4, Issue 1: 29-38

A Neural Network-based Approach for Assessing the Energy Value of Regional Water Environment Pollution Losses and Its Application

Quan Xiao and Lu Zhao*

School of Economics, Management and Law, Hubei Normal University, Huangshi 435002, Hubei, China

sonelluu@163.com

**corresponding author*

Keywords: Neural Networks, Regional Water Environment, Pollution Loss, Energy Value Assessment

Abstract: With social development and population growth, people's demands for quality of life and quality of life are increasing, and higher requirements for drinking water safety have been put forward. Therefore, it is important to study and establish WE pollution loss assessment models for WE management work and reasonable economic policy formulation. However, most of the traditional models use a single factor as the input variable to assess the trend of pollutant concentration and the intensity of pollutant discharge; while the WE pollution loss assessment method based on energy value analysis theory uses a single factor as the input variable for energy value analysis to assess the intensity and concentration of pollutant discharge, but this method cannot handle multi-factor and multi-variable data, and can only use the fuzzy mathematics in However, this method cannot handle multi-factor and multi-variable data, and can only use fuzzy mathematics to evaluate and calculate the model; and it cannot deal with the non-linear relationship between multiple pollutants and single pollutant concentration at the same time. Therefore, in order to better solve the above problems, this paper proposes a neural network (NN) model based on the regional WE pollution loss energy assessment method; the method uses NNs to assess the energy value of a single factor; at the same time, the energy value of the regional pollution economic value is calculated to obtain the average energy value of the whole region; finally, the weighted average method is used to calculate the corresponding weighted average energy value of each unit and pollutant discharge intensity and other information. Finally, the weighted average method is used to calculate the weighted average energy value and the pollutant emission intensity of each unit. Compared with traditional methods, this method is not only better able to deal with multi-source data problems and non-linear relationships, but also more efficient in solving complex problems.

1. Introduction

Environmental pollution has become an important constraint on human survival and development, and has gradually become a global problem. According to the United Nations, by 2050, the world will produce 35 billion tonnes of toxic waste each year, including industrial, agricultural and municipal waste (including medical waste). In addition, more than 70 billion tonnes of wastewater will be produced each year, equivalent to 15% of the world's population. This is a serious water pollution problem caused by the untreated discharge of large amounts of domestic sewage. At the same time, environmental problems such as eutrophication of water bodies due to the massive discharge of pollutants by humans are becoming increasingly prominent. Therefore, water is one of the important resources on which human beings depend for survival and development [1-2].

In a related study, Angelika et al. used dynamic isotope equation modeling based on big data to study the two-way mechanism between water environmental pollution and the quality of economic growth [3]. It was argued that importance should be attached to sustainable economic development, elimination of factors of environmental pollution and timely adjustment of environmental policies to promote harmony between ecological environment and economic growth. victor et al. developed a model of diffuse waterborne pathogens with different rates of dispersal and spatial heterogeneity in order to explore the impact of environmental pollution on the spread of waterborne diseases [4]. The basic reproduction number R0 was defined and its threshold role was shown. Haider et al. investigated a cooperative approach to multi-player cooperative surface vehicles (USVs) for chemical contaminant source monitoring in dynamic water environments [5]. Based on the "Infotaxis" algorithm for multi-USVs, an improved shared probability was proposed to address the low success rate and inefficiency of multi-USVs in cooperative exploration due to cognitive differences. The cognitive differences between USVs are reconciled by introducing a confidence factor. The success rate and efficiency of exploration were improved.

In order to better understand the pollution loss situation of the regional WE and to plan [6-7], design and optimise future pollution control engineering projects, it is necessary to comprehensively evaluate the degree of damage that may be caused to the regional WE by each impact factor brought about by the project. Therefore, this paper establishes a regional WE energy value assessment model based on BP NN and energy value assessment methods, and validates the model.

2. Design Studies

2.1. Methods for Quantifying the Carrying Capacity of the Water Environment (WE)

(1) Indicator System Evaluation Methodology

The indicator system evaluation method is the most widely used [8]. The aim of this method is to assess the performance of the aquatic environment by creating a system of assessment indicators, processing the values of each indicator using statistical methods, assigning different weights to different indicators and finally calculating the final result using a weighted sum. Through much research and innovation by scientists, the indicator system approach is increasingly being used to quantitatively assess the performance of the aquatic environment. Currently, the main indicator system assessment methods are the vector function method, the fuzzy global assessment method and the hierarchical analysis method. These methods are relatively easy to choose for quantitative assessment and analysis, and are used to quantify and understand the degree of good and bad impacts on the aquatic environment, but the choice of assessment indicators is easily influenced by the subjectivity of the selector, making the assessment results inaccurate and unreliable [9-10].

(2) Principal Component Analysis

The aim of the principal component analysis method is to reduce the dimensionality of a multivariate system by replacing the original multivariate variables with a few component variables, while assigning objective weights to each component variable to avoid subjectivity. This method has the advantage of objectivity, analytical simplicity and a simple estimation process, but difficulties are encountered in determining the criteria for parameter estimation and in selecting principal components and control points.

(3) A multi-objective optimisation approach.

When using this approach to assess the carrying capacity of the aquatic environment, the constraints and the required optimisation objectives are first defined and then explored using a mathematical model. The advantage of this approach is that the transfer target can be defined in detail and the constraints on transfer capacity can be quantified, but the difficulty lies in the selection of the optimisation target and the solution of the model.

(4) System dynamics approach

The system dynamics approach, also known as the SD model, takes into account primarily social, economic and demographic factors. It focuses on factors that affect the system as a whole and uses system dynamics software to perform dynamic studies and analyses [11-12].

(5) Artificial NN Method

The artificial NN approach involves the manual construction of a NN to achieve a specific research objective. The method involves constructing an artificial NN to understand the indicator parameters affecting the carrying capacity of the aquatic environment at different times, and gradually developing the structure and function of a map-integrated system. In the application, the structure of the network is determined according to the study object, knowledge of the study object is obtained by studying typical sample data, the contribution of the network is estimated and the carrying capacity of the aquatic environment is analysed. Although artificial NN models are easy to implement, the accuracy of their training is elusive and therefore error is high [13-14].

2.2. WE Issues

(1) Water resources in the precipitation area are in short supply. The average annual precipitation, annual runoff and the amount of water entering the precipitation in a precipitation basin are all decreasing year by year, and the wide area of the waters of a precipitation and the large amount of reeds and other plants in the precipitation have led to a large amount of evaporation from the water surface; the subsurface conditions in the precipitation area are complex, and there is also a large amount of seepage from the bottom of the lake; at the same time, the construction of the upstream reservoir has intercepted a large amount of incoming water, ultimately leading to a shortage of water resources in a precipitation [15-16].

(2) Unbalanced water supply. A precipitation is affected by evaporation, seepage and other factors, while production and domestic water consumption around the precipitation area and within the precipitation is high, making the water resources available to a precipitation greatly reduced, and the over-extraction of groundwater is a serious problem, with an imbalance in the water supply relationship [17-18].

(3) The pollution of water bodies is serious. The upstream Fu River is the main pollution-receiving river of a city, and waste water discharged from industry, agriculture and domestic use flows into a certain precipitation through the Fu River. At the same time, the concentration of pollutants in a certain precipitation increases due to the large amount of sewage discharged from village life and livestock breeding in the precipitation, and the water pollution problem becomes increasingly prominent.

(4) Degradation of ecological functions. In recent years, the climate of a certain precipitation area has become warmer and drier, and changes in hydrological elements such as precipitation and surface runoff have affected the relationship between precipitation and surface runoff, while at the same time the ecological environment of a precipitation has been seriously damaged due to the impact of human and social development.

3. Experimental Research

3.1. Establishment of the "NN Method"

In order to construct a model that can quickly assess pollution losses, it is first necessary to screen monitoring data from various pollution sources, outfalls and wastewater treatment plants in the study area and eliminate some monitoring indicators that have a large but not serious impact, such as items such as permanganate index, ammonia nitrogen and chemical oxygen demand.

Secondly, in order to construct an energy value assessment model that can quickly assess pollutant concentrations at each source or outfall, data pre-processing, including data normalisation and normalisation, is required.

Thirdly, it is necessary to determine the relationship between the weighting coefficients of each indicator in the source and the amount of pollution loss in relation to the weighting coefficients and the amount of pollution loss in each environmental factor.

Fourthly, it is necessary to build an energy assessment model based on the interrelationships between the indicators in order to obtain the required energy assessment information.

Fifthly, a regional WE energy assessment model was developed to analyse the extent of pollution and pollution loss at each unit point and to obtain the required energy assessment information.

3.2. Early Warning Methods for WE Carrying Capacity Evaluation

(1) Control chart

The control chart method is a method of describing the overall probability distribution process based on the probability distribution of the sample size, based on the statistical laws followed by the data, and judging whether it is in a normal state through the analysis of the sample data. The control chart method is determined by two parameters: the mean value of the indicator μ and the standard deviation σ. The value of the indicator should fall within the $\mu \pm 3\sigma$ interval with 99.73% probability, with $\mu+3\sigma$ as the upper limit of the indicator and $\mu-3\sigma$ as the lower limit of the indicator. The formulae are as follows.

$$UCL_x = \bar{x} + A_2\bar{R} \tag{1}$$

$$CL_x = \bar{x} \tag{2}$$

$$UCL_x = \bar{x} - A_2\bar{R} \tag{3}$$

Where

UCLx - upper control limit.

CLx - control centre limit.

LCLx - lower control limit.

x - the sample mean.

A2 - the control threshold factor.

R - the mean value of the extreme differences.

(2) BP NN Model

BP NNs are one of the most widely used types of artificial NNs with proven algorithms and robust performance in pattern recognition and data classification and prediction. In this work, a BP NN model was created using MATLAB software and three layers were chosen as the topology of the model (Figure 1). the BP NN model was used to calculate the total WE intensity index for a specific WE event and to determine the warning level of the WE.

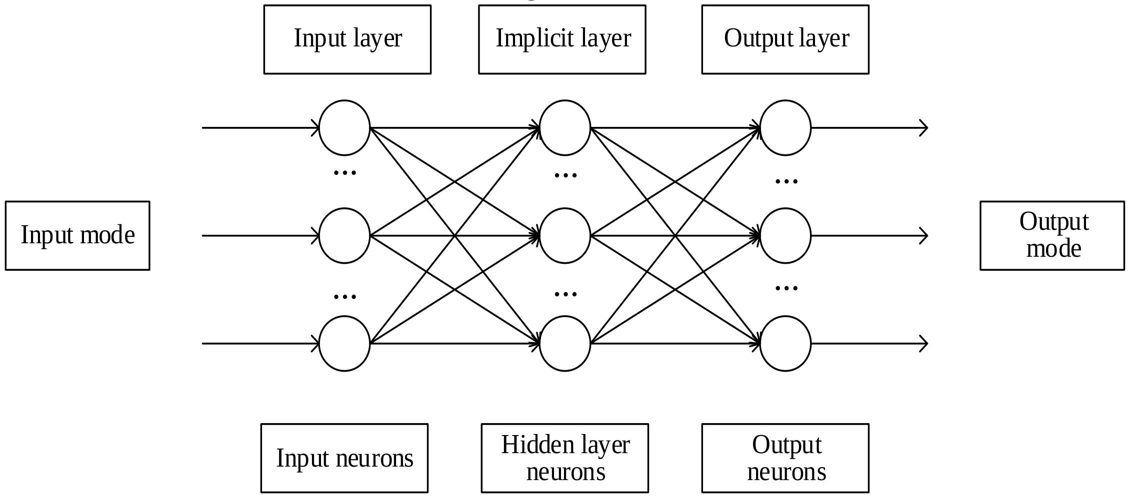

Figure 1. BP NN topology

3.3. Effectiveness of Feature Extraction Methods

In order to verify the effectiveness of the features extracted based on information entropy and gradient boosting tree algorithm, here the logistic regression algorithm is used as the base classifier, and the features extracted based on information entropy and gradient boosting tree algorithm are compared with the original water quality data without feature extraction and the features extracted based on principal component analysis, and these three types of features are used as the input to the logistic regression discriminant algorithm for the effectiveness of the feature extraction algorithm The evaluation process diagram is shown in Figure 2.

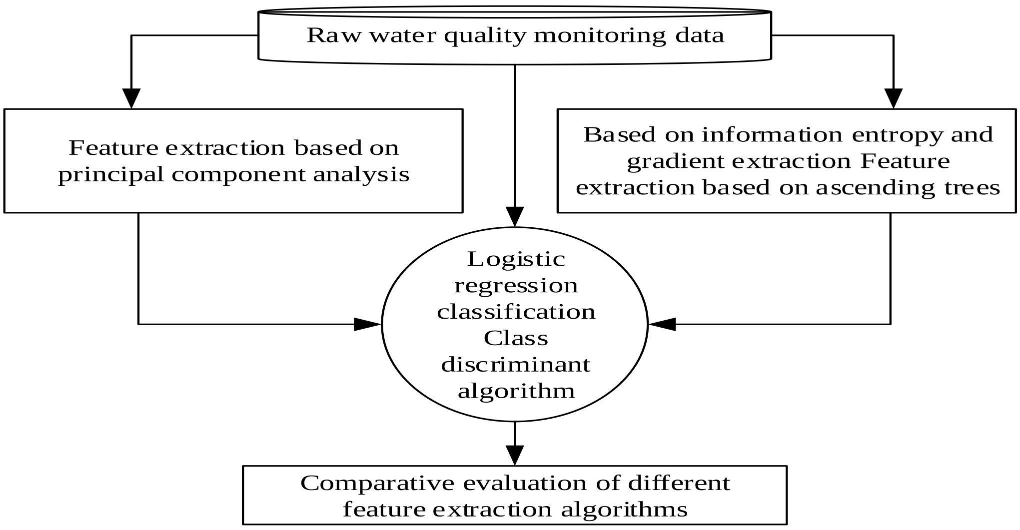

Figure 2. Flow of feature extraction algorithm effectiveness evaluation

4. Experimental Analysis

4.1. Discussion on the Effectiveness of Feature Extraction Methods

The pre-processed data set is divided into a training set and a test set, in which the training set accounts for 80% of the training of the feature extraction algorithm and the classification and discrimination algorithm, and the test set accounts for 20% to evaluate the effectiveness and reliability of the algorithm. The anomaly detection task is essentially a dichotomous algorithm, which is a judgement of whether the river water is polluted or not, and the water quality samples affected by different types of pollution sources are considered as one category, and the normal river water quality samples are classified as another category. In contrast, the pollution source category discrimination task is essentially a multi-classification algorithm, which not only makes a judgement on whether river water is affected by pollution sources, but also specifically discriminates between categories of pollution sources. The accuracy, recall and fl score metrics of the discrimination results for the three categories of features on the test set are shown in Table 1.

Table 1. Discriminant results of the three types of features on the test set

Tasks	Feature extraction algorithms	Accuracy rate	Recall rate	Flscore
Detection of water quality anomalies (dichotomous)	Raw water quality data	0.921	0.884	0.935
	Principal component analysis	0.931	0.896	0.943
	Algorithms used in this paper	0.970	0.960	0.977
Pollution source category identification (multiclassification)	Raw water quality data	0.722	0.680	0.693
	Principal component analysis	0.743	0.700	0.712
	Algorithm used in this paper	0.812	0.777	0.786

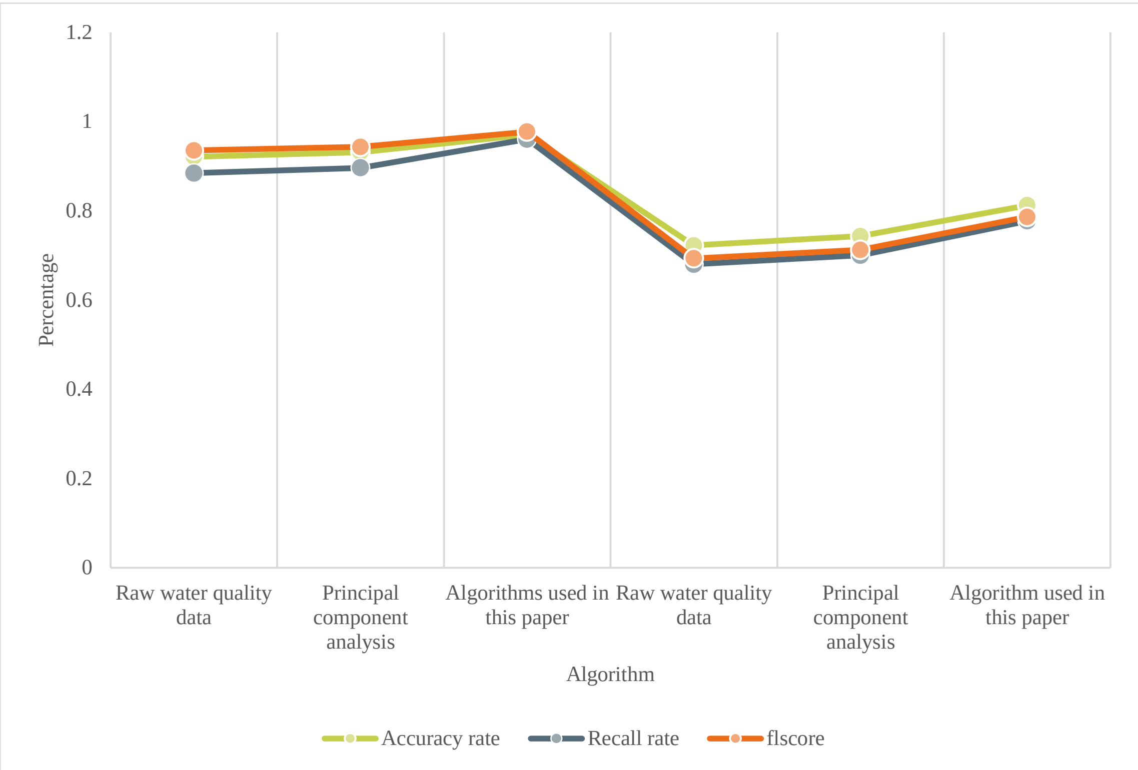

Figure 3. Analysis of the discriminatory results of the three types of features on the test set

As can be seen from Figure 3.

(1) The discrimination algorithm generally performed better on the water quality anomaly detection DH,0 task than on the pollution source category discrimination task.

(2) Using raw water quality data and features extracted by principal component analysis as input for water quality anomaly detection tasks, although with high accuracy and low recall.

(3) The feature extraction algorithm used in this paper achieves optimal results in all evaluation indexes regardless of the task of water quality anomaly detection DH,4 or the task of pollution source category discrimination, which shows the effectiveness of the feature extraction algorithm based on information entropy and gradient decision tree.

4.2. Analysis of Experimental Results

The models associated with the predictions performed in this paper have been experimentally validated to achieve optimal prediction results. Based on multiple model prediction evaluation methods, this work compares the graphical convolutional time series prediction model with other comparative models, using 2423160 water quality monitoring stations as the data set used for prediction, using the first 70% of the data as the training set, 10% of the data as the validation set, and the last 20% of the data as the test set, where the validation set is mainly used when the model accuracy on the data in the validation set cannot be The validation set is used to stop the model from learning on the training set early when the accuracy of the model cannot be improved on the validation set, to prevent overfitting from occurring. The performance of multiple models on the last 20% of the test data was evaluated and the final experimental comparison results for each model are shown in Table 2.

Table 2. Comparison of prediction accuracy of different models

Methods	RMSE	MAE	MAPE
ARIMA	0.219	0.154	1.796
BPNN	0.168	0.111	1.304
GCN	0.159	0.118	1.362
LSTM	0.174	0.115	1.365

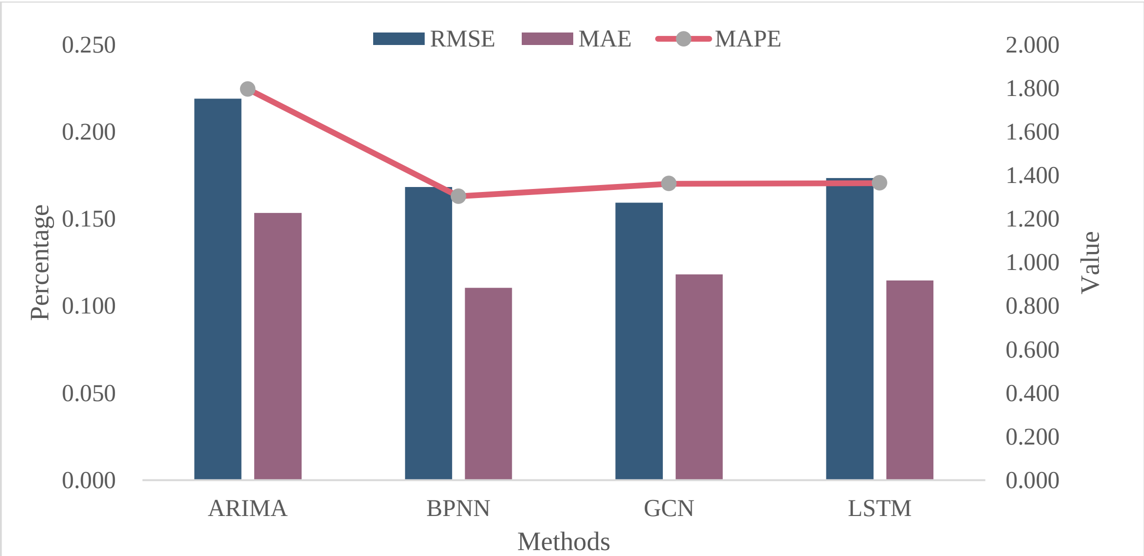

Figure 4. Comparative analysis of prediction accuracy of different models

In Figure 4, ARIMA indicates the use of the autoregressive difference moving average algorithm, BPNN represents the NN model, GCN indicates the graph convolutional NN prediction model, and LSTM indicates the use of the long and short-term memory NN. The three metrics of RMSE, MAE and MAPE show that GCN has the best predictive performance.

5. Conclusion

With rapid economic and social development, water resources are being over-consumed by humans, while the phenomenon of water pollution is becoming more and more serious. However, the total amount of water resources and the ability of water bodies themselves to withstand the pollution load are limited, and these problems can in turn become a major factor hindering social development. The assessment of pollution losses in the water environment is an important environmental and economic decision. This paper proposes a neural network-based approach to regional water environment pollution loss assessment, firstly by learning the hydrological theory through extensive research, and then by using artificial neural networks to analyse a large amount of data in order to obtain the required data, and to verify the accuracy of the model in the actual calculation of water pollution in a certain region.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Amal Agarwal, Lingzhou Xue. Model-Based Clustering of Nonparametric Weighted Networks With Application to Water Pollution Analysis. Technometrics. (2020) 62(2): 161-172. https://doi.org/10.1080/00401706.2019.1623076

[2] Swati Chopade, Hari Prabhat Gupta, Rahul Mishra, Preti Kumari, Tanima Dutta. An Energy-Efficient River Water Pollution Monitoring System in Internet of Things. IEEE Trans. Green Commun. Netw. (2021) 5(2): 693- 702. https://doi.org/10.1109/TGCN.2021.3062470

[3] Angelika Zube, Dominik Kleiser, Alexander Albrecht, Philipp Woock, Thomas Emter, Boitumelo Ruf, lgor Tchouchenkov, Aleksej Buller, Boris Wagner, Ganzorig Baatar, Janko Petereit. Autonomously mapping shallow water environments under and above the water surface. Autom. (2022) 70(5): 482-495. https://doi.org/10.1515/auto-2021-0145

[4] Victor Manuel Zezatti, Alberto Ochoa, Gustavo Urquiza, Miguel Basurto, Laura Castro, Juan Garcia. The Implementation of a Nickel-Electroless Coating in Heat Exchanger Pipes Considering the Problem of the Environmental Conditions ofthe Cooling Water Without Recirculation to Increase the Effectiveness Under Uncertainty. Int. J. Comb. Optim. Probl. Informatics. (2022) 13(4): 73-82.

[5] Haider A. H. Alobaidy, Rosdiadee Nordin, Mandeep jit Singh, Nor Fadzilah Audullah, Azril Haniz, Kentaro Ishizu, Takeshi Matsumura, Fumihide Kojima, Nordin Bin Ramli. Low-Altitude-Platform-Based Airborne IoT Network (LAP-AIN) for Water Quality Monitoring in Harsh Tropical Environment. IEEE Internet Things J. (2022) 9(20): 20034-20054. https://doi.org/10.1109/JIOT.2022.3171294

[6] Totan Garai, Harish Garg. Possibilistic multiattribute decision making for water resource management problem under single-valued bipolar neutrosophic environment. Int. J. Intell. Syst. (2022) 37(8): 5031-5058. https://doi.org/10.1002/int.22750

[7] Suresh Muthulingam, Suvrat S. Dhanorkar, Charles J. Corbett. Does Water Scarcity Affect Environmental Performance? Evidence from Manufacturing Facilities in Texas. Manag. Sci. (2022) 68(4): 2785-2805. https://doi.org/10.1287/mnsc.2021.4013

[8] Tristan Thielmann. Environmental conditioning: Mobile geomedia and their lines of becoming in the air, on land, and on water. New Media Soc. (2022) 24(11): 2438-2467. https://doi.org/10.1177/14614448221122190

[9] V. A. Miklush, I. A. Sikarev, Tatiana M. Tatarnikova. Organization of Environmental Monitoring of the Port Water Area by Processing an Anti-Interference Signal from a Vessel Traffic Control System. Autom. Control. Comput. Sci. (2021) 55(8): 999-1004. https://doi.org/10.3103/S0146411621080204

[10] Nikolaos Laoutaris, Costas Iordanou. What do information centric networks, trusted execution environments, and digital watermarking have to do with privacy, the data economy, and their future? Comput. Commun. Rev. (2021) 51(1): 32-38. https://doi.org/10.1145/3457175.3457181

[11] Ehsan Alvandi, Mojtaba Soleimani-Sardo, Sarita Gajbhiye Meshram, Behnam Farid Giglou, Mohammad Reza Dahmardeh Ghaleno. Using Improved TOPSIS and Best Worst Method in prioritizing management scenarios for the watershed management in arid and semi-arid environments. Soft Comput. (2021) 25(16): 11363-11375. https://doi.org/10.1007/s00500-021-05933-9

[12] Sivaraman Eswaran, Daniel Dominic, Jayapandian Natarajan, Prasad B. Honnavalli. Augmented intelligent water drops optimisation model for virtual machine placement in cloud environment. IET Networks. (2020) 9(5): 215-222. https://doi.org/10.1049/iet-net.2019.0165

[13] Uma Maheswari S, Vasanthanayaki C. Secure medical health care content protection system (SMCPS) with watermark detection for multi cloud computing environment. Multim. Tools Appl. (2020) 79(5-6): 4075-4097. https://doi.org/10.1007/s11042-019-7724-z

[14] Alexander S. Komarov, Alain Caya, Mark Buehner, Lynn Pogson. Assimilation of SAR Ice and Open Water Retrievals in Environment and Climate Change Canada Regional Ice-Ocean Prediction System. IEEE Trans. Geosci. Remote. Sens. (2020) 58(6): 4290-4303. https://doi.org/10.1109/TGRS.2019.2962656

[15] Mina Emami Khansari, Saeed Sharifian. A modified water cycle evolutionary game theory algorithm to utilize QoS for IoT services in cloud-assisted fog computing environments. J. Supercomput. (2020) 76(7): 5578-5608. https://doi.org/10.1007/s11227-019-03095-y

[16] Norimitsu Sakagami, Mizuki Suka, Yu Kimura, Eito Sato, Takahiro Wada. Haptic shared control applied for ROV operation support in flowing water. Artif. Life Robotics. (2022) 27(4): 867-875. https://doi.org/10.1007/s10015-022-00788-1

[17] Khoury Boutrous, Fatiha Nejjari, Viceng Puig. Reliability-Aware Zonotopic Tube-Based Model Predictive Control of a DW Network. Int. J. Appl. Math. Comput. Sci. (2022) 32(2): 197-211.

[18] M. S. Patel, A. A. Khalifa, M. S. Liew, Zahiraniza Mustaffa, Andrew Whyte. Downtime cost analysis of offloading operations due to influence of partially standing waves in Malaysian

waters and development of graphical user interface. Ann. Oper. Res. (2022) 315(2): 1263-1289. https://doi.org/10.1007/s10479-020-03845-9