IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 11 November 2022, accepted 26 December 2022, date of publication 29 December 2022, date of current version 5 January 2023.

Digital Object Identifier 10.1109/ACCESS.2022.3233298

RESEARCH ARTICLE

Monitoring and Forecasting of Water Pollution by Heavy Metals

MAKPAL ZHARTYBAYEVA[1], NURZHAN MUNTAYEV[1], SYMBAT TULEGENOVA[2], ZHANAR ORALBEKOVA[1], ZHANAR LAMASHEVA[1], AND KAZIZAT ISKAKOV[1]

[1]Department of Computer and Software Engineering, Faculty of Information Technologies, L.N.Gumilyov Eurasian National University, 010000 Astana, Kazakhstan
[2]Department of Botany, Faculty of Biology and Geography, Karaganda Buketov University, 100028 Karaganda, Kazakhstan

Corresponding author: Nurzhan Muntayev (muntayevnurzhan@gmail.com)

This work was supported by the Grant from the Ministry of Education and Science of the Republic of Kazakhstan within the framework under Project AP09058557.

ABSTRACT As a result of anthropogenic impact, significant changes occur in the morphometric, hydrological, chemical and other characteristics of water bodies, which in turn leads to changes in the structure, productivity and condition of aquatic ecosystems. Pollution of water bodies is the result of anthropogenic activities that lead to negative consequences - deterioration of water quality, threat to water bodies, deterioration of life and health of people. In order to monitor the water ecosystems of Akmola region, the research team took water samples from water bodies (Nura, Ishim Rivers of the Republic of Kazakhstan, Akmola region and Taldykol, Maybalyk and Zhaltyrkol Lakes), collected them in a database and conducted research.We constructed a mathematical model, after which the results were compared with full-scale data.According to water samples, the level of pollution of water sources was evaluated, the obtained data was entered into the database of the water ecosystem of Akmola region, and the collected data was predicted using the ARIMA (autoregressive integrated moving average) model, mathematical model and the R programming language.

INDEX TERMS ARIMA, forecasting, hydrochemical parameters, physical and chemical indicators, surface water, water quality assessment.

I. INTRODUCTION

The initial quality of water depends on the source of its origin. Surface water (lakes, reservoirs, streams, and rivers), which is a source of drinking water for the indigenous population, is usually of poor quality and requires extensive treatment. The ground water is of good quality. However, they can still be contaminated by agricultural runoff or surface and underground burial of liquid waste, including filters from landfills of solid household waste. Other sources, such as spring and rainwater, have different levels of quality [1], [2].

Thus, pollution of water sources and its impact on the environment is one of the most common environmental problems. Despite the fact that Kazakhstan occupies a large territory, the quality of water in all surface reservoirs of the Republic remains unsatisfactory. Together with wastewater, pollutants enter reservoirs, which affects the aquatic ecosystem, primarily hydrobionts [3], [4].

Water quality in all surface reservoirs of the country remains unsatisfactory. Along with wastewater, pollutants enter reservoirs and affect the aquatic ecosystem, primarily aquatic organisms [5], [6]. As a result of anthropogenic impact, significant changes occur in the morphometric, hydrological, chemical and other characteristics of water bodies, which in turn leads to changes in the structure, productivity and condition of aquatic ecosystems. Pollution of water bodies is the result of anthropogenic activities that lead to negative consequences - deterioration of water quality, threat to water bodies, deterioration of life and health of people [7]. Pollution contributes to an increase in the content of micro - and macronutrients in fresh and marine waters, bottom sediments, and living organisms above the natural background of a particular territory [8].

Various analytical models are used to predict water quality, which can be divided into traditional, based on statistical

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci.

models, and non-traditional, using artificial intelligence (AI) approaches. Methods of KNN, machine learning, artificial neural networks (ANN) can be attributed to non-traditional methods, and regression analysis, time series methods can be attributed to traditional methods [9]. A review of the literature has shown that artificial neural networks (ANN) are at the peak of popularity in modeling the prediction of water pollution [10], [11]. But this method requires a huge database and often a trial and error method is needed to determine the correct ANN architecture. Traditional methods of hydromonitoring, conceptual and numerical models of groundwater, were used to predict changing hydrogeological conditions and processes based on in-depth knowledge during monitoring. For the accuracy and reliability of the data, sufficient and more accurate data are needed for calibration and verification of the hydrological system. One of them is the need to introduce a large number of input parameters necessary for the modeling process. At the same time, many models are unable to cope with predictive hydrogeological uncertainties and non-linearity and give them a quantitative assessment. Failure to recognize and evaluate uncertainties and non-linearity can lead to an incorrect representation of the actual system, which contributes to poor performance of the groundwater model and reduces the accuracy of forecasting [12], [13], [14], [15], [16], [17], [18]. The work of Cheng et al. proves that the methods of autoregressive integrated moving average (ARIMA), SVN and AN have a more accurate result [19]. Due to the fact that the environmental quality assessment system based on maximum permissible concentrations has many disadvantages, for example, it does not take into account the interaction of various pollutants with each other, accumulation in organisms, interaction with bottom sediments, etc., therefore, an attempt was made to assess the state of aquatic ecological systems using the mathematical modeling method and the ARIMA model.

We collected water samples from the reservoirs of the Akmola region, laboratory analysis, classification of the analysis results, built a mathematical model, after which the results were compared with natural data, then a forecast was made based on ARIMA.

II. MATERIALS AND METHODS

In the following years, the method of induction plasma emission spectrometry has been widely used in the study of extremely low amounts of heavy metals (according to the Maximum allowable concentration (MAC) from environmental objects (including water). In the research areas, Induction Plasma Emission Spectrometer (IPES) was used, one of the most effective and highly sensitive methods for detecting heavy metals that may be caused by anthropogenic influences. The concentration of heavy metals taken for the study did not exceed the MPC. That is, the level of pollution of the region is satisfactory.

The form of periodic sampling approved by the International Organization for Standardization has been applied. From 2019 to 2022, in order to control the quality of surface water in the Akmola region, water samples were taken from 5 reservoirs (the Nura and Ishim rivers, as well as Maybalyk, Zharlykol, Taldykol lakes, etc.). Water samples were taken on rectilinear sections of drainage paths in turbulent, mixed flows. All samples were taken only after familiarization with the location of production workshops, water consumption and technology of wastewater treatment plants. More than 700 data were collected.During the study of surface water samples, physical and chemical indicators of water quality (water temperature, transparency, hydrogen Index, total water hardness, concentration of heavy metals) are determined.

The research team determined hydrochemical indicators of water to assess the quality of surface waters of Akmola region (Nura, Ishim Rivers and Maybalyk, Zharlykol, Taldykol Lakes). Water samples taken from reservoirs were analyzed for 13 physical and chemical parameters, including ion concentrations (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Fe total, Cl^{-}, SO_4^{2-}, HCO_3^{-},NO_3^{-},CO_3^{2-}). Water samples from surface water sources were also taken by atomic-emission spectroscopy (inductively bound plasma) in the Azimuth-Geology chemical-analytical laboratory of 6 heavy metals (cadmium, copper, manganese, nickel, lead and Mercury).

III. RESULTS

Based on the tasks of the research work – to study statistical data on the state of surface water in Akmola region, data from statistical agencies of this region were used.

The results obtained during the research work are described below.

According to the unified classification, water quality is evaluated as follows:

The state of bottom sediments in rivers in Akmola region in the second half of 2021.

According to samples taken from the bottom sediments of Lake Maybalyk, the average concentration of cadmium was 0.0001 mg/dm^3, nickel 0.005 mg/dm^3, manganese 0.009 mg/dm^3, copper 0.002 mg/dm^3, lead 0.001 mg/dm^3.

According to samples taken from the bottom sediments of lake Zharlykol, the average concentration of cadmium was 0.0001 mg/dm^3, nickel 0.001 mg/dm^3, manganese 0.003 mg/dm^3, copper 0.002 mg/dm^3, lead 0.001 mg/dm^3.

According to samples taken from the bottom sediments of Lake Taldykol, the average concentration of cadmium was 0.0001 mg/dm^3, nickel 0.01 mg/dm^3, manganese 0.181 mg/dm^3, copper 0.004 mg/dm^3, lead 0.001 mg/dm^3.

Data on samples taken from the Nura River:

On the Nura River: water quality at points 1.1, 1.2,1.3,1.4,1.5 belongs to class 4: magnesium 41 mg/l. Magnesium concentration does not exceed the background class. The average value of sulfates in water samples taken from each point is 263 mg/l.

According to samples taken from the section of Ch.Aitmatov street along Taldykol Lake, the average concentration of cadmium was 0.0001 mg/dm^3, copper 0.004 mg/dm^3, manganese 0.181 mg/dm^3, nickel 0.01 mg/dm^3, lead 0.001 mg/dm^3.

We can consider sulfate ions in water samples taken from River Lakes, which are the objects of our research, as indicators of anthropogenic pollution (table 1).

The chemical parameters of ground water play an important role in the classification and evaluation of water quality. It is noted that the best results can only be obtained when studying the ion complex in water, and not the concentration of individual ions [9]. Chemical classification also shows the concentration of various dominant cations, anions and their relationship.

Modern anthropogenic load and increasing man-made processes change the hydrochemical parameters of water to a deterioration in their quality.

Field studies conducted to analyse surface monitoring data and assess surface water have shown the presence of chemical pollutants in river lakes in the Akmola region.

IV. MATHEMATICAL MODEL

The research team considered the diffusion, convective and combined - diffusion-convective mechanisms of the propagation of heavy metals and other pollutants in water. In this part of the research, we used classical equations of mathematical physics [20]. The diffusion mechanism of the spread of pollutants is used to describe processes in reservoirs with stagnant or low-flowing water (for example, reservoirs, ponds, lakes, canals, etc.). This mathematical model is based on the diffusion equation.

The intensity of ingress and mixing of contaminated substances in reservoirs is influenced by various factors, such as climatic conditions, relief, landscape features, terrain soils, etc.

To begin with, let's consider a one-dimensional diffusion problem in relation to a narrow, extended body of water with standing water: a small river, stream, pipe or channel into which a limited amount of the pollutant enters. Then you can ignore the width and depth of the reservoir, since they are insignificant compared to the length, then we will consider diffusion only in one direction – along the length of the reservoir.

When constructing a mathematical model of diffusion processes for one-dimensional problems, differential equations of the parabolic type are used. Equations of this type will allow us to characterize processes by two independent variables: time and spatial coordinate.

One-dimensional diffusion task with initial and boundary conditions:

$$\frac{\partial K}{\partial t} = D\frac{\partial^2 K}{\partial x^2} K\left(0, t\right) = K_0; K\left(x, 0\right) = 0 \tag{1}$$

here K – the concentration of the contaminant; K_0 – the initial concentration; t – the propagation time, x-the length of the investigated area.

To further solve the task, we use the Poisson Integral

$$K\left(x, t\right) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\propto} \frac{1}{\sqrt{Dt}} e^{-\frac{(x-\xi)^2}{4Dt}} \alpha(\xi) d\xi \tag{2}$$

TABLE 1. **Information on surface water quality in Akmola region in the second half of 2022.**

Water form and connectedness	Pollution characteristics	
Nura River	The total hardness is 7.8-7.9, the water temperature is 0.1 °C, the hydrogen index is 8.3-8.38	
Location:1.0 km below the village of Kabanbai batyr	4th class	magnesium-41 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location:1.5 km below the village of Kabanbai batyr	4th class	magnesium-41 mg/dm^3, iron -0,24 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location:2.0 km below the village of Kabanbai batyr	4th class	magnesium-41 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location:2.5 km below the village of Kabanbai batyr	4th class	magnesium-40 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location:3.0 km below the village of Kabanbai batyr	4th class	magnesium-41 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Ishim River (Nur-Sultan)	The total hardness is 11.8-12.0, the water temperature is 0.1 °C, the hydrogen index is 8.35-8.43	
Location: Nur-Sultan, Almaty district, 0.6 km from Koksay street and the bypass road	4th class	magnesium – 70 mg/dm^3, chlorides - 457 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Ishim district, Syganak street, under the Arys bridge	4th class	magnesium – 73 mg/dm^3, chlorides - 479 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Ishim district, Sarayshyk street, under the Arkhar bridge	4th class	magnesium – 68 mg/dm^3, chlorides - 479 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Saryarka district, Kabanbai batyr avenue and ZH. Shanin street, under the Karaotkel bridge	4th class	magnesium – 70 mg/dm^3, chlorides - 479 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Saryarka district, 0.8 km from Sh. Beisekov street to 189 street	4th class	magnesium – 70 mg/dm^3, chlorides - 487 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Maybalyk Lake	The total hardness is 18.6-19.5, the water temperature is 0.1 °C, the hydrogen index is 7.93-8.43, the water sample is yellowish	
Location: Nur-Sultan, Ishim district, 10.0 km from the bypass road	4th class	chlorides – 1179 mg/dm^3, magnesium-141 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Ishim district, 6.3 km from the bypass road	4th class	chlorides – 1223 mg/dm^3, magnesium-140 mg/dm^3, the actual concentration of magnesium exceeds the background class.

TABLE 1. ***(Continued.)*** **Information on surface water quality in Akmola region in the second half of 2022.**

Location: Nur-Sultan, Ishim district, 4.7 km from N. Nazarbayev International Airport	4th class	chlorides – 1312 mg/dm , magnesium-146 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Ishim district, 3.0 km from the bypass road	4th class	chlorides – 1294 mg/dm^3, magnesium-147 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Nur-Sultan, Ishim district, 1.5 km from N. Nazarbayev International Airport	4th class	chlorides – 1294 mg/dm^3, magnesium-146 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Zharlykol Lake	The total hardness is 7.6-7.9, the water temperature is 0.1 °C, the hydrogen index is 8.66-8.7, the water sample contains sediment	
Location: 8.0 km from Zhalgyzkuduk village	4th class	chlorides – 496 mg/dm^3, magnesium – 66 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: 8.6 km south - east of Zhalgyzkuduk village	4th class	chlorides – 496 mg/dm^3, magnesium – 66 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: 10.0 km southwest of Zhalgyzkuduk village	4th class	chlorides – 496 mg/dm^3, magnesium – 68 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: 10.4 km from Zhalgyzkuduk village, 1.8 km from Lake Karasor	4th class	chlorides – 496 mg/dm^3, magnesium – 64 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Taldykol Lake	The total hardness is 41.7-42.2, the water temperature is 0.1°C, the hydrogen index is 8.04-8.06, the water sample contains sediment	
Location: 0.8 km between Ch. Aitmatov and Syganak streets	4th class	chlorides – 1773 mg/dm^3, magnesium -241 mg/dm^3, sulfates-1963 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: Urker district, 2.2 km from Korgalzhyn Highway	4th class	chlorides – 1773 mg/dm^3, sulfates – 1963 mg/dm^3, magnesium-235 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Ishim River (Nur-Sultan)	The total hardness is 6.5-6.7, the water temperature is 0.1°C, the hydrogen index is 7.56-8.42, the water sample contains sediment	
Location: along the Arnasay dam, 3.0 km from the village of Arnasay	> 5th class	chlorides – 234 mg/dm^3, sulfates – 192 mg/dm^3, magnesium-39 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: along the Arnasay dam, 2.3 km from the village of Arnasay	> 5th class	chlorides – 234 mg/dm^3, sulfates – 192 mg/dm^3, magnesium-39 mg/dm^3, the actual concentration of magnesium exceeds the background class.

TABLE 1. ***(Continued.)*** **Information on surface water quality in Akmola region in the second half of 2022.**

Location: along the Arnasay dam, 1.8 km from the village of Arnasay	> 5th class	chlorides – 234 mg/dm^3, sulfates – 177 mg/dm^3, magnesium-39 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: along the Arnasay dam, 1.2 km from the village of Arnasay	> 5th class	chlorides – 234 mg/dm^3, sulfates – 177 mg/dm^3, magnesium-39 mg/dm^3, the actual concentration of magnesium exceeds the background class.
Location: along the Arnasay dam, 0.7 km from the village of Arnasay	> 5th class	chlorides – 234 mg/dm^3, sulfates – 172 mg/dm^3, magnesium-40 mg/dm^3, the actual concentration of magnesium exceeds the background class.

where $\alpha(\xi)$ – the action of the initial concentration; $F(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\lambda^2} d\lambda$ – error function

$z = \frac{x}{2\sqrt{Dt}}$, as a result, we obtain the following equation:

$$K(x,t) = K_0 [1 - F(z)] \quad (3)$$

This equation will help determine the distribution of the concentration K(x,t) and estimate the time for which the concentration of the pollutant in the reservoir will become greater than the maximum permissible. $K_{ma} = \delta \cdot K_0$ so we calculate the value of the maximum permissible concentration at some point of the channel x, where $\delta < 1$. Using the ratio we get: $K_{ma} = K_0 \left[1 - F\left(\frac{x}{2\sqrt{Dt_{ma}}}\right)\right]$, $t_{ma} = \frac{x^2}{4D^2 r^2}$, r- the root of the equation, F(z) = 1δ.

Inorganic pollutants have diffusion coefficients from 0.4 to 3.0 cm^2/day, and organic pollutants have diffusion coefficients from 0.3 to 1.8 cm^2/day. In practice, when industrial wastewater pollutes reservoirs, the concentration of pollutants in the source is maintained at a constant level of K_0, at the beginning of the channel, the substance enters, which inevitably leads to some convection. To partially account for this factor, we can introduce a longitudinal mixing coefficient, or a convective diffusion coefficient, which is one or two orders of magnitude greater than the calculated values of the diffusion coefficient, which we determine experimentally. For a one-dimensional diffusion problem with a constantly active pollution source, the initial and boundary conditions take the following form: $K(0, t) = K_0(t)$, $K(x, 0) = 0$, $K(0, 0) = K_0$, $K(\infty, t) = 0$. We assume that $K_0(t)$-the known dilution dependence of the initial concentration of the source K with the time value t. We assume that $K_0(t) = K_0 - k \cdot t$.

Since we decided to simplify the issue, we will assume that the reservoirs of the Akmola region are flat-bottomed, then it will be convenient to consider the problem of the distribution of contaminated substances, heavy metals in a flat space, for this it is enough to solve the two-dimensional diffusion issue [21], [22].

We believe that the distribution of pollutants (heavy metals) over the depth of the reservoir is uniform, and the diffusion process takes place along the length and width.

The two-dimensional diffusion equation in a given region ω has the following form [20]:

$$\frac{\partial K}{\partial t} = D(\frac{\partial^2 K}{\partial x^2} + \frac{\partial^2 K}{\partial y^2}) \tag{4}$$

For the case when there is a constantly active source of pollution with a density $p(x, t)$ on the boundary section from $-i$ to $+ i$, and the initial concentration of heavy metals in space is 0, the boundary conditions are as follows:

$$K(x, y, 0) = 0 \left. \frac{dK}{dx} \right| = \begin{cases} p(x, t), & when - i \leq x \leq i \\ 0, & whenx \notin [-i, i] \end{cases} \tag{5}$$

where $p(x, t)$ – is the impurity flux density; x, y – are the corresponding coordinates along the length and width of the reservoir.

$K(x, 0) = K_0$;$-i \leq x \leq i$;

$$K(x, y) \left.\right|_{y=\infty} ; K(x, y) |_{x=\pm\infty} = 0$$

We can consider a similar problem for the three-dimensional nature of the distribution of heavy metals - diffusion in the form of a sphere. This usually happens when the source of pollution is located in an unlimited space, for example, at some depth of a reservoir, and in the environment the initial concentration is zero, that is, $K(R, 0) = 0$ at $R \neq 0$. The propagation of heavy metals in a homogeneous medium occurs symmetrically in all directions: along x, y, z. Hence, the diffusion equation:

$$\frac{dK}{dt} = D\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \tag{6}$$

The diffusion equation in a spherical coordinate system:

$$\frac{dK}{dt} = D\left(\frac{\partial^2 K}{\partial R^2} + \frac{2}{R^2}\frac{\partial K}{\partial R}\right) \tag{7}$$

Here R-the radius vector.

Based on these mathematical expressions, calculations were made to determine the concentration of heavy metal when a certain amount of it enters the reservoir in a certain period of time.

The time to reach the maximum permissible concentration K_{ma} with a change in the initial concentration of the substance has changed by 9 times: 1 – at $K_0 = 1$ M/l; 2 – at $K_0 = 9$ M/l.

To evaluate the mathematical model, we compared the experimental data with the data obtained by calculations. The comparison showed that the calculated data are in good agreement with the full-scale data. The relative error was 5%.

V. FORECAST

So far, various models have been developed, which generally fall into three categories: (1) hydrodynamic models based on the Saint-Venant equations (2) statistical models (e.g. autoregressive models). integrated moving average class (ARIMA) 2) and machine learning models (eg, artificial neural network (ANN) [18], [23], [24], [25].

All laboratory data on water collections were included in the database of the aquatic ecosystem of the Akmola region, based on the collected data, forecasts were made using the ARIMA model and the R programming language.

In the period 2019-2022. (twice a year, in autumn and spring) about 700 data collected from control points were selected. Samples were taken from the Nura, Ishim Rivers and Taldykol, Maybalyk and Zhaltyrkol Lakes of the Akmola region of the Republic of Kazakhstan.

Heavy metals are highly toxic and dangerous compounds. An increase in the concentration of heavy metals in water can lead to deterioration of human health, and in some cases even to death. One of the main causes of pollution of water sources with heavy metals is the impact of anthropogenic factors on the environment. Water hardness, mineralization, hydrogen index in determining water quality, etc. indicators play an important role. According to these indicators, it is possible to determine the level of water pollution and prevent water blooming. Thus, the classification of heavy metals and water quality indicators was carried out, and the measurements of the elements given in the following table were obtained (table 2).

TABLE 2. Units for magnetic properties.

first analysis	Second analysis
• mercury	• Carbonates
• cadmium	• bicorbonates
• copper	• chlorides
• manganese	• sulfates
• nickel	• nitrates
• lead	• anion sum
	• calcium
	• magnesium
	• sodium and potassium
	• iron
	• cation sum
	• minerals
	• dry residue

Data on concentrations of the following heavy metals in atmospheric waters are collected at monitoring points: mercury, cadmium, copper, manganese, nickel, lead.

The analysis revealed that the concentration of heavy metals copper, manganese and nickel in water is high (Fig.6). With this in mind, we predict models for them. The ARIMA model was chosen as the forecasting model. ARIMA uses three main parameters (p, d, q) expressed as integers. Therefore, the model is also written as ARIMA (p, d, q). Together, these three parameters take into account seasonality, trend, and noise in the dataset [24]:

The ARIMA seasonal model is used to track seasonality — ARIMA (p,d,q) (P,D,Q)s. Here (p, d, q) are the non—seasonal parameters described above, and (P, D, Q) follow the same definitions, but refer to the seasonal component of the time series. The parameter s defines the periodicity of the time series (4 - quarter periods, 12 - year periods, etc.). In our

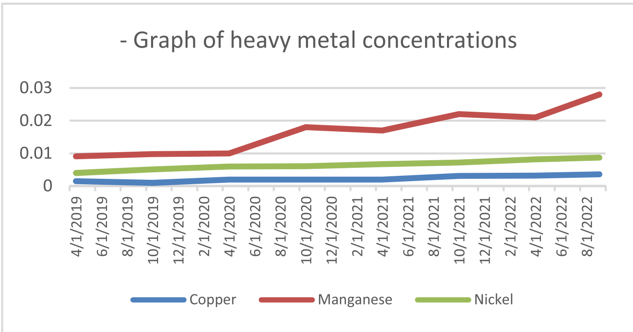

FIGURE 1. Concentration of heavy metals along the Ishim River, mg/dm^3.

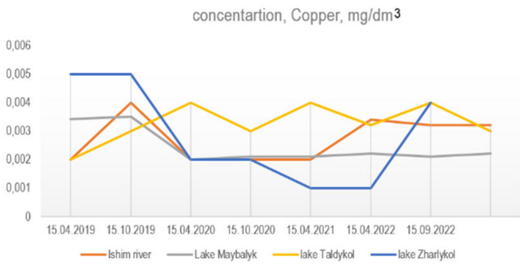

FIGURE 2. Copper concentration in the studied water bodies of Akmola region, mg/dm^3.

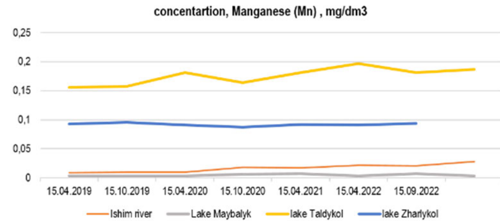

FIGURE 3. Manganese concentration in the studied water bodies of Akmola region, mg/dm^3.

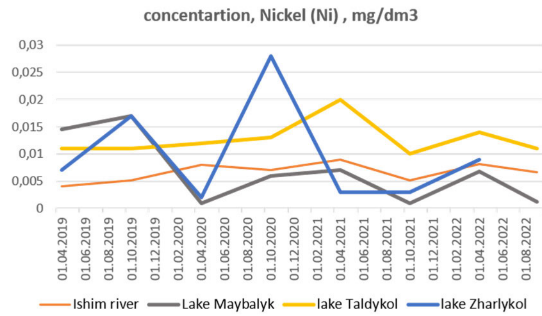

FIGURE 4. Nickel concentration in the studied water bodies of Akmola region, mg/dm^3.

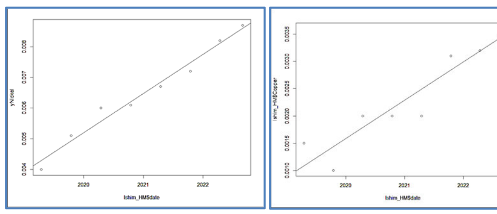

FIGURE 5. Graph of the approximating function for making forecasts based on statistical data for Manganese, Nickel, Copper.

case, 6-half-year periods were taken due to spring and autumn sampling.

The concentration of heavy metals along the Ishim River is shown in the following graph (Figure-1). In order to optimize data processing, it was decided to predict them by the type of heavy metal (Copper, Manganese, Nickel) and depending on the sampling location (Ishim, Taldykol, Zharlykol, Maybalyk).

According to the results of the conducted studies of the Ishim river, the limit of the concentration of manganese has been exceeded in it. Determination of the concentration of manganese in water was carried out after filtration of water through a membrane filter with a pore size of 0.45 microns, that is, dissolved forms of manganese were detected. The MPC in water are 0.1 mg /l for manganese. A mutagenic effect can be observed in organisms that are in their turn due to the high marginal concentration of manganese in the environment. The existence of this element leads to the fact that there is an overspending of livelihood in the body. During the analysis of the dynamics of manganese concentrations in the Ishim water in 2019-2022, it was found out that in the surface layer of water, the increase in the concentration of manganese is largely influenced by natural conditions created in winter and spring. Using the results of laboratory analysis, we graphically compared the values of the concentration of heavy metals of the Ishim River with the indicators of the lakes Zharlykol, Taldykol and Maybalyk (Figures 2-4).

According to the graphs of Manganese concentration, we can notice the stationarity of time series, according to the graphs of Copper and Nickel concentration, the non-stationarity of time series, according to the definition, a one-dimensional time series is called stationary if its probabilistic characteristics (parameters of a random variable) are constant.A time series is called non-stationary if at least one of the probabilistic characteristics is unstable. Such a graph is typical for a random walk, i.e. a process of approximately the form $y_t = y_{t-1} + u_t$ [24]. Next, we are faced with the task of obtaining a stationary series. We obtain stationarity using a logarithm, then we test the series for stationarity using the methods described above: visual analysis of the graph, visual analysis of ACF and PACF (Fig. 5,6)., tests for unit roots. If a stationary series is obtained, then we proceed to the next point, if not, then we apply the operator of taking a sequential difference and repeat the testing. In practice, the sequential difference is taken, as a rule, no more than twice. For stationary series, it is advisable to calculate auto-correlation function (ACF) and partial auto-correlation function (PACF), since there is a correlation between the values. Therefore, it makes sense to calculate ACF and PACF, since the correlation between random variables standing on the steps can vary over time.

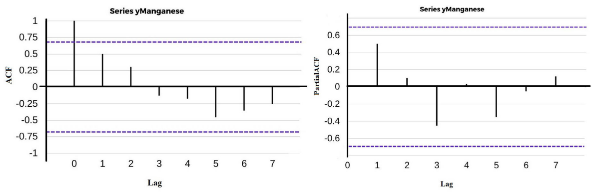

FIGURE 6. ACF and PACF for Manganese indicators in the Ishim River.

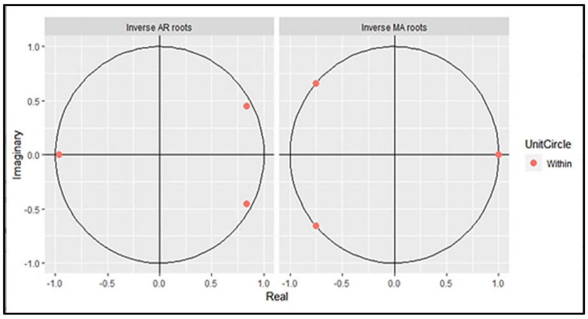

FIGURE 7. AR and MA roots.

TABLE 3. AIC and BIC coefficients.

Model	coefficients					
	Manganese		Copper		Nickel	
	AIC	BIC	AIC	BIC	AIC	BIC
Model1	-65.228	-65.644	-79.823	-79.931	-79.823	-79.931
Model2	-81.284	-81.609	-71.926	-72.967	-71.926	-72.967
model3	-77.747	-77.963	-66.632	-67.257	-66.633	-67.257

The autocorrelation diagram shows that the autocorrelation value basically does not exceed the boundary value, although the autocorrelation value of the 1st order exceeds the boundary, we will consider this an accident, and the autocorrelation value will not exceed a significant boundary in other cases. And we can expect that from 1 to 20 will sometimes exceed the 95% confidence interval.

For non—seasonal series, Rob Hyndman recommends taking lag, h=10, for seasonal h=2m, where m is the periodicity of seasonality, i.e. h=24 for monthly data.

For ARMA (p, q), can see where the roots of the AR and MA parts are located (Fig 7):

Our models for the values of the Manganese concentration in the Ishim River are as follows, from them we must choose the best one according to the Akaike information criterion (AIC) and Bayesian information criterion (BIC) coefficients, and then make a forecast based on it. We have selected 3 models with the following values

model1 <- arima(y, order = c(1, 1, 0))

model2 <- arima(y, order = c(3, 2, 1))

model3 <- arima(y, order = c(2, 3, 1))

Calculate the coefficients AIC and BIC (table -3).

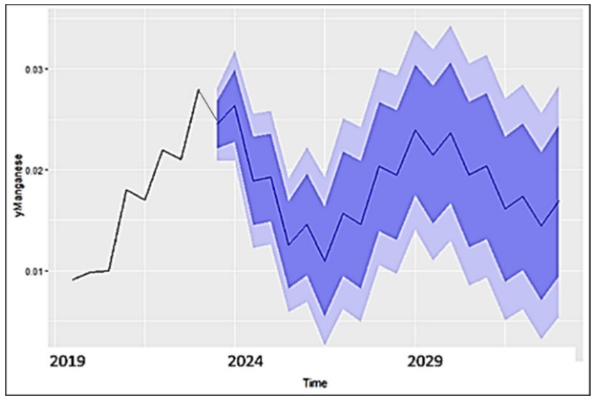

FIGURE 8. Forecast of Manganese concentration indicators in the Ishim River.

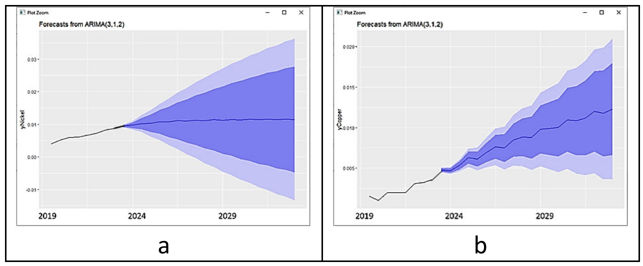

FIGURE 9. Forecast of Nickel and Copper concentration indicators in the Ishim River (a – Nickel, b-Copper).

Let's select the best model according to the AIC penalty criterion:

$$AIC = -2\ln l + 2k$$

where $\ln l$- the logarithm of the likelihood function, and k- the number of model parameters.

The more parameters, k, the more complex the model, the higher the AIC. The lower the likelihood function, l, that is, the lower the probability of obtaining the available data for a given model, the higher the AIC. And the BIC indicator is determined by:

$$BIC = 2\ln L + \ln n \cdot k.$$

The lower the value of one of these criteria for a number of models under study, the better the model will match the data. As we can see, out of the three models, the second ARIMA model turned out to be the best according to the AIC and BIC criteria (3,2,1) (figures 8-9, table 4).

We will do the same work for the Nickel and Copper concentration indicators in the Ishim River.

Next, a forecast was made using the Arima method for data obtained from the Ishim rivers, Maibalyk, Taldykol and Zharlykol lakes of the Akmola region of the Republic of Kazakhstan by laboratory studies.

Next, we calculated the coefficients according to the Akaike AIC, BIC criterion to determine the best models (Table 5).

TABLE 4. Quantitative indicators of predicting the concentration of heavy metals in the ishim river.

Date	Manganese, mg\dm^3					Copper, mg\dm^3	Nickel, mg\dm ^3
	Forecast	Low 80	High 80	Low 95	High 95	Forecast	Forecast
15.04.23	0,0273	0,0250	0,0296	0,0238	0,0308	0,0039	0,0039
15.10.23	0,0319	0,0297	0,0341	0,0285	0,0353	0,0043	0,0043
15.04.24	0,0325	0,0301	0,0349	0,0288	0,0362	0,0046	0,0046
15.10.24	0,0378	0,0352	0,0403	0,0338	0,0417	0,0050	0,0050
15.04.25	0,0379	0,0348	0,0410	0,0331	0,0426	0,0053	0,0053
15.10.25	0,0423	0,0392	0,0453	0,0376	0,0470	0,0056	0,0056
15.04.26	0,0436	0,0401	0,0470	0,0383	0,0488	0,0060	0,0060
15.10.26	0,0477	0,0441	0,0513	0,0422	0,0532	0,0063	0,0063
15.04.27	0,0487	0,0447	0,0526	0,0427	0,0547	0,0067	0,0067
15.10.27	0,0527	0,0488	0,0566	0,0467	0,0587	0,0070	0,0070
15.04.28	0,0543	0,0499	0,0586	0,0476	0,0609	0,0074	0,0074
15.10.28	0,0579	0,0534	0,0624	0,0510	0,0648	0,0077	0,0077
15.04.29	0,0594	0,0546	0,0642	0,0521	0,0667	0,0080	0,0080
15.10.29	0,0631	0,0582	0,0680	0,0556	0,0705	0,0084	0,0084
15.04.30	0,0649	0,0596	0,0701	0,0568	0,0729	0,0087	0,0087
15.10.30	0,0682	0,0628	0,0736	0,0599	0,0764	0,0091	0,0091
15.04.31	0,0701	0,0644	0,0758	0,0614	0,0788	0,0094	0,0094
15.10.31	0,0734	0,0676	0,0793	0,0645	0,0824	0,0098	0,0098
15.04.32	0,0754	0,0693	0,0816	0,0660	0,0849	0,0101	0,0101
15.10.32	0,0786	0,0722	0,0849	0,0689	0,0882	0,0104	0,0104

The following figures show a graphical visualization of the forecasting results (figure 10, 11)

We compared the prediction results obtained using ARIMA with the results of mathematical modeling. The comparison showed that the relative error was 5.2%, which indicates that the calculated data of the two models are in good agreement.

As we can see from the graphs obtained as a result of forecasting, the concentration of heavy metals will increase. Obviously, when summed up, the concentration of heavy metals is a great danger. To reduce the activity of some enzymes (beta-glucosidase, catalase), even a low concentration of heavy metals (up to 0 mg per 1 kg) is sufficient. As a rule, elevated concentrations of copper and manganese are recorded during the flood period, the waters of which are mostly formed on the surface of the catchment area. At the same time, there are differences in the seasonal dynamics of metals. The highest values of copper concentration were recorded in November. Then it can be assumed that the autumn concentration peaks are most likely associated with anthropogenic influence. Attention is drawn to the fact that if the "natural" course of the concentration of metals were determined solely by the runoff, then during August-December their accumulation should have been at the same level [25].

TABLE 5. Coefficients according to the akaike aic, bic criterion.

Model	coefficients					
	Manganese		Copper		Nickel	
	AIC	BIC	AIC	BIC	AIC	BIC
Model1_ Taldykol	-45,058	-45,166	-80,853	-80,961	-80,853	-80,961
Model2_ taldykol	-33,001	-34,043	-60,266	-61,307	-60,266	-61,307
Model3_ Taldykol	-25,237	-26,799	-63,839	-64,463	-63,839	-64,463
Model1_ Maybalyk	-71,055	-71,163	-59,881	-59,989	-80,853	-80,961
Model2_ maybalyk	-54,910	-55,951	-43,741	-44,782	-60,266	-61,307
Model3_ Maybalyk	-40,286	-41,848	-45,259	-45,884	-63,839	-64,463
Model1_ Zhyltyrkol	-27,830	-29,058	-49,018	-50,245	-49,018	-50,245
Model2_ zhyltyrkol	-43,281	-44,322	-71,926	-72,967	-60,266	-61,307
Model3_ Zhyltyrkol	-16,929	-19,633	-37,101	-39,805	-37,101	-39,805

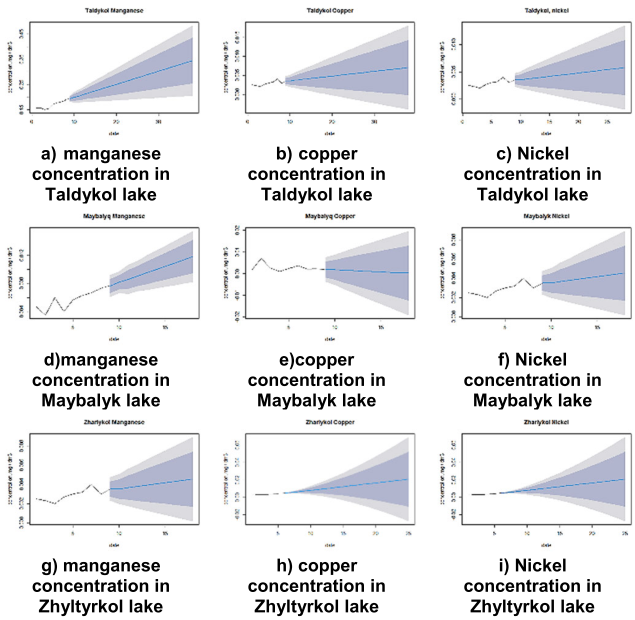

FIGURE 10. Graphical visualization of the forecasting result.

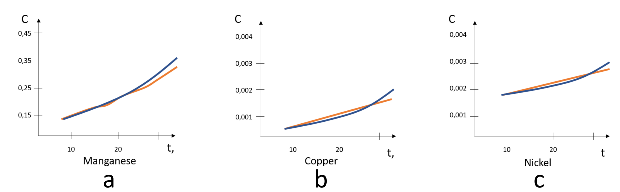

FIGURE 11. Graphical comparison of forecasting data obtained using mathematical modeling and Arima (a – manganese, b-Copper, c- Nickel).

VI. CONCLUSION

Calculations and comparisons have shown that in hydromonitoring, along with traditional methods, we can use mathematical modeling. Mathematical modeling allows not only to determine the qualitative and quantitative composition of natural waters exposed to man-made impacts, but also to predict the course of some chemical and physico-chemical processes that occur in the aquatic ecosystem, taking into account the hydrological and hydrochemical indicators of the aquatic environment. The disadvantage of this method is that the development of such models should be limited to a small number of factors that take into account the spread of pollutants, but mathematical modeling is able to predict the behavior of pollutants not only in a temporary order, but also in the long term.

Special software has been developed that implements the ARIMA family models in the software environment of the R language. More than 700 sets of annual data with a small length were studied in the work. The data were processed using the forecast and forecast.arima libraries from the CRAN repository. To determine the stationarity of the series, the Dickey-Fuller test was used. The stationarity of the time series allows you to make better forecasts. The Akaike information criterion was used to select the best model.

According to the unified classification system, in the 2nd half of 21, the water quality of water sources in Akmola region is estimated as follows: Ishim (Nur – Sultan city), Nura River, Taldykol, Maybalyk and Zhaltyrkol Lakes belong to the 4th class, and the Ishim River (Arnasay reservoir) – (>5th class) corresponds to the norm.

The result of our research is that water samples taken from water bodies are within the norm. It can be seen that in most water samples there is a high concentration of manganese. Sulfates, in turn, are considered as indicators of the level of contamination. Based on the hydrochemical indicators obtained from the research work, we have supplemented the database of aquatic ecosystems in Akmola region.

To reduce the adverse impact on the environment in the industry, it is necessary to adhere to high environmental requirements. To prevent such situations, it is necessary to introduce new performance standards that take into account past negative experience, and promote a culture of safe work. Develop technical and technological means to prevent the risk of such situations. It is also necessary to conduct periodic environmental monitoring: take water samples, control the species composition of aquatic biota. Also, an ecologist should be constantly at the emission sites, who will control all processes and look to see that everything goes within the framework of environmental standards.

Surface and underwater vehicles, sensors are also promising methods of monitoring water. The research group is currently working on the development of a mobile robotic complex for monitoring reservoirs and collecting bathymetric and hydrographic data within the framework of the AP09058557 project. The results of the launch of the mobile robotic complex will be subsequently published in the form of scientific publications.

Industrial complexes are considered as a source of heavy metals entering water sources as pollutants. In order to prevent the impact of anthropogenic factors on the environment, the concentration of each heavy metal was estimated with the participation of the above researchers.

The data obtained from the annual calculation of the results of samples taken from common water sources prove the need to organize measures to protect water sources from external pollutants in the future. According to the forecasting made by researchers, the concentration of heavy metals in water sources is increasing from year to year. According to forecasts, the concentration of heavy metals (manganese, nickel and copper) in the Ishim River may increase to 95% in the next decade. This, in turn, can lead to a decrease in the diversity of aquatic biota and the flowering of water sources. Among the proposed actions, installation of treatment plants in industrial complexes that cause pollution to water sources, timely inspection of treatment plants, and annual quality monitoring of water sources. We supplemented the database of water ecosystems in Akmola region on the basis of hydrochemical indicators obtained from research works.

REFERENCES

[1] F. Thomas, *Rural Water Supplies and Water-Quality/Healthy Housing Reference Manual*. USA: National Center for Environmental Health, 2005, ch. 8. [Online]. Available: https://www.cdc.gov/nceh/publications/books/housing/06_hhm_final_chapter_08.pdf

[2] R. S. Bryan Swistock Clemens and E. William Sharpe, "Drinking water quality in rural Pennsylvania and the effect of management practices/school of forest resources and institutes of energy and the environment," Pennsylvania State Univ., State College, PA, USA, Tech. Rep., Oct. 2007. [Online]. Available: https://www.rural.palegislature.us/drinking_water_quality.pdf

[3] *Water Resources of Kazakhstan in the New Millennium. Review (A Publication in Support of the UN Millennium Development Goals. Goal 7: Ensure Environmental Sustainability)*, United Nations Develop. Program Kazakhstan, Almaty, Kazakhstan, 2004, p. 132. [Online]. Available: http://www.undp.kz

[4] K. L. Valentini, E. E. Orolbaev, and A. K. Abylgazieva, *Water Problems of Central Asia*, Bishkek, Kyrgyzstan, 2004, p. 142.

[5] J. Hornero, M. Manzano, L. Ortega, and E. Custodio, "Integrating soil water and tracer balances, numerical modelling and GIS tools to estimate regional groundwater recharge: Application to the Alcadozo aquifer system (SE Spain)," *Sci. Total Environ.*, vol. 568, pp. 415–432, Oct. 2016, doi: 10.1016/j.scitotenv.2016.06.011.

[6] W. P. Kelley, "Permissible composition and concentration of irrigation water," in *Proc. ASCE*, vol. 66, May 1940, pp. 607–613.

[7] B. K. Handa, "Modified classification procedure fro rating irrigation waters," *Soil Sci.*, vol. 98, no. 4, pp. 264–269, 1964.

[8] *Environmental Protection in the Republic of Kazakhstan 15–19 Statistical Collection*, Dept. Prod. Environ. Statist., Committee Statist. Ministry Nat. Economy Republic Kazakhstan, Astana, Kazakhstan, 2019.

[9] J. I. Ubah, L. C. Orakwe, K. N. Ogbu, J. I. Awu, I. E. Ahaneku, and E. C. Chukwuma, "Forecasting water quality parameters using artificial neural network for irrigation purposes," *Sci. Rep.*, vol. 11, no. 1, p. 24438, Dec. 2021, doi: 10.1038/s41598-021-04062-5.

[10] N. F. Che Nordin, N. S. Mohd, S. Koting, Z. Ismail, M. Sherif, and A. El-Shafie, "Groundwater quality forecasting modelling using artificial intelligence: A review," *Groundwater Sustain. Develop.*, vol. 14, Aug. 2021, Art. no. 100643, doi: 10.1016/j.gsd.2021.100643.

[11] S. V. Volkov, "Technological aspects of water disinfection by UV radiation Text," *Technique*, vol. 1, no. 12, pp. 7–10, May 1996.

[12] J. A. Guzman, A. Shirmohammadi, A. M. Sadeghi, X. Wang, M. L. Chu, M. K. Jha, P. B. Parajuli, R. D. Harmel, Y. P. Khare, and J. E. Hernandez, "Uncertainty considerations in calibration and validation of hydrologic and water quality models," *Trans. ASABE*, 58, no. 6, pp. 1745–1762, 2015, doi: 10.13031/trans.58.10710.
[13] E. D. Spyrou, I. Tsoulos, and C. Stylios, "Applying and comparing LSTM and ARIMA to predict CO levels for a time-series measurements in a port area," *Signals*, vol. 3, no. 2, pp. 235–248, Apr. 2022.
[14] S. Agarwal, S. Sharma, M. H. Rahman, S. Vranckx, B. Maiheu, L. Blyth, S. Janssen, P. Gargava, V. K. Shukla, and S. Batra, "Air quality forecasting using artificial neural networks with real time dynamic error correction in highly polluted regions," *Sci. Total Environ.*, vol. 735, Sep. 2020, Art. no. 139454.
[15] K. Zhang, J. Thé, G. Xie, and H. Yu, "Multi-step ahead forecasting of regional air quality using spatial–temporal deep neural networks: A case study of huaihai economic zone," *J. Cleaner Prod.*, vol. 277, Dec. 2020, Art. no. 123231.
[16] E. Eslami, A. K. Salman, Y. Choi, A. Sayeed, and Y. Lops, "A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks," *Neural Comput. Appl.*, vol. 32, no. 11, pp. 7563–7579, Jun. 2020.
[17] Y. Cheng, H. Zhang, Z. Liu, L. Chen, and P. Wang, "Hybrid algorithm for short-term forecasting of $PM_{2.5}$ in China," *Atmos. Environ.*, vol. 200, pp. 264–279, Mar. 2019.
[18] C. Zhong, Z. Jiang, X. Chu, T. Guo, and Q. Wen, "Water level forecasting using a hybrid algorithm of artificial neural networks and local Kalman filtering," *Proc. Inst. Mech. Eng., M, J. Eng. Maritime Environ.*, vol. 233, no. 1, pp. 174–185, Feb. 2019, doi: 10.1177/1475090217727135.
[19] S. Kageyama, "Water balance analysis considering runoff of ungauged catchments in Iwaki river basin, northern Japan," *Int. J. Geomate*, vol. 9, no. 17, pp. 1434–1440, 2015.
[20] V. Balakrishnan, *Mathematical Physics With Applications, Problems and Solutions Unknown Binding*. New Delhi, India: Ane Books, Jan. 2019.
[21] D. K. Durdiev, "Inverse coefficient problem for the time-fractional diffusion equation," *EURASIAN J. Math. Comput. Appl.*, vol. 9, no. 1, pp. 44–54, 2021, doi: 10.32523/2306-6172-2021-9-1-44-54.
[22] B. R. Kusse and E. A. Westwig, *Mathematical Physics: Applied Mathematics for Scientists and Engineers*, 2nd ed. Berlin, Germany: Wiley, Jan. 2006.
[23] H. Ii, "Characteristic of water chemistry for ARIMA type deep thermal water in the Kinokawa river catchment, KII Peninsula, Japan," *Int. J. Geomate*, vol. 17, no. 62, pp. 158–166, Oct. 2019.
[24] Z. Oralbekova, Z. Khassenova, B. Mynbayeva, M. Zhartybayeva, and K. Iskakov, "Information system for monitoring of urban air pollution by heavy metals," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 22, no. 3, pp. 1590–1600, 2021.
[25] Q. Zhou, N. Yang, Y. Li, B. Ren, X. Ding, H. Bian, and X. Yao, "Total concentrations and sources of heavy metal pollution in global river and lake water bodies from 1972 to 2017," *Global Ecol. Conservation*, vol. 22, Jun. 2020, Art. no. e00925, doi: 10.1016/j.gecco.2020.e00925.

MAKPAL ZHARTYBAYEVA was born in Zhezqazghan, Kazakhstan, in 1988. She received the B.S. and M.S. degrees in applied computer engineering and software from Satbayev Kazakh National Technical University, in 2012, and the Ph.D. degree in applied computer engineering and software from Gumilyov Eurasian National University, in 2020. She is currently an Associate Professor at the Department of Computer and Software Engineering, L. N. Gumilyov ENU, where she is the Head of the Research Project. She has an experience in the direction of the project for nine years. She is a holder of the State Scholarship for Young and Talented Scientists of the Republic of Kazakhstan, in 2018. She is the author of three books and more than 20 articles. Her research interests include green technologies, development of mobile robotic systems, georadar research, and development of an information system for optimizing monitoring of environmental pollution.

NURZHAN MUNTAYEV was born in Karaganda, Kazakhstan, in 1988. He received the B.S. degree in information systems from the Economic University of Karaganda, Kazakhstan, in 2009, and the M.S. degree in IT engineering from Novosibirsk State University, Russia, in 2012.

He is a Research Project Executor. Since 2014, he has been a Commercial Director at IT company, a Half Stack Employee at Karaganda Economic University, the Chairman of the Diploma Defense IT Commission, and a Scientific Supervisor. He was a Founder of the practice base for the Department of Computing Systems.

Mr. Muntayev was a participant and a speaker of the 2021 IEEE International Conference on Smart Information Systems and Technologies at Astana IT University in 2021.

SYMBAT TULEGENOVA received the bachelor's degree in biology from E. A. Buketov Karaganda University, in 2014, and the master's degree in natural sciences and the Doctoral degree in ecology from L. N. Gumilyov Eurasian National University, in 2016 and 2021, respectively. Her research interests include assessment of pollution of treated and untreated wastewater and consideration of ways to clean the environment from eco-pollutants.

ZHANAR ORALBEKOVA was born in Almaty, Kazakhstan, in 1980. She received the B.S. and M.S. degrees in applied mathematics and informatics and the Ph.D. degree in mathematics from Al-Farabi Kazakh National University, in 2003 and 2013, respectively.

From 2005 to 2010, she was a Senior Teacher of information systems at the Education Department, Abay Kazakh National Pedagogical University. Since 2014, she has been an Assistant Professor with the Computer and Software Engineering Department, L. N. Gumilyov Eurasian National University. She is the author of three books and more than 20 articles. Her research interests include computer modeling and applications, robot technics, and ecological monitoring systems.

Dr. Oralbekova's awards and honors include the Bolashak Fellowship (Kazakhstan) and the Ministry of Education and Science of the Republic of Kazakhstan Young Scientist Award for Excellence in 2014.

ZHANAR LAMASHEVA received the Ph.D. degree in information systems from Satbayev University, Almaty, Kazakhstan, in 2015.

From 2015 to 2016, she was an Assistant Lecturer at Satbayev University. She is currently working as a Lecturer at Gumilyov Eurasian National University, Astana, Kazakhstan. Her research interests include simulation and data analysis.

KAZIZAT ISKAKOV received the Doctor of Physical and Mathematical Sciences degree in 2002, and the degree in computational mathematics in 2007.

He was an Associate Professor of informatics and management in 1994 and a Professor of mathematics in 2005. His research interests include software development for the interpretation of radargrams, visualization of geodata, image processing, and numerical methods for solving inverse and applied problems.