ISSN 2518-170X (Online)
ISSN 2224-5278 (Print)

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ
Satbayev University

ХАБАРЛАРЫ

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ
НАУК РЕСПУБЛИКИ
КАЗАХСТАН
Satbayev University

N E W S

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF
KAZAKHSTAN
Satbayev University

SERIES
OF GEOLOGY AND TECHNICAL SCIENCES

3 (459)

MAY – JUNE 2023

THE JOURNAL WAS FOUNDED IN 1940

PUBLISHED 6 TIMES A YEAR

ALMATY, NAS RK

NAS RK is pleased to announce that News of NAS RK. Series of geology and technical sciences scientific journal has been accepted for indexing in the Emerging Sources Citation Index, a new edition of Web of Science. Content in this index is under consideration by Clarivate Analytics to be accepted in the Science Citation Index Expanded, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. The quality and depth of content Web of Science offers to researchers, authors, publishers, and institutions sets it apart from other research databases. The inclusion of News of NAS RK. Series of geology and technical sciences in the Emerging Sources Citation Index demonstrates our dedication to providing the most relevant and influential content of geology and engineering sciences to our community.

Қазақстан Республикасы Ұлттық ғылым академиясы «ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы» ғылыми журналының Web of Science-тің жаңаланған нұсқасы Emerging Sources Citation Index-те индекстелуге қабылданғанын хабарлайды. Бұл индекстелу барысында Clarivate Analytics компаниясы журналды одан әрі the Science Citation Index Expanded, the Social Sciences Citation Index және the Arts & Humanities Citation Index-ке қабылдау мәселесін қарастыруда. Webof Science зерттеушілер, авторлар, баспашылар мен мекемелерге контент тереңдігі мен сапасын ұсынады. ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы Emerging Sources Citation Index-ке енуі біздің қоғамдастық үшін ең өзекті және беделді геология және техникалық ғылымдар бойынша контентке адалдығымызды білдіреді.

НАН РК сообщает, что научный журнал «Известия НАН РК. Серия геологии и технических наук» был принят для индексирования в Emerging Sources Citation Index, обновленной версии Web of Science. Содержание в этом индексировании находится в стадии рассмотрения компанией Clarivate Analytics для дальнейшего принятия журнала в the Science Citation Index Expanded, the Social Sciences Citation Index и the Arts & Humanities Citation Index. Web of Science предлагает качество и глубину контента для исследователей, авторов, издателей и учреждений. Включение Известия НАН РК. Серия геологии и технических наук в Emerging Sources Citation Index демонстрирует нашу приверженность к наиболее актуальному и влиятельному контенту по геологии и техническим наукам для нашего сообщества.

«ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.).

Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 29.07.2020 ж. берілген **№ KZ39VPY00025420** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *геология, мұнай және газды өңдеудің химиялық технологиялары, мұнай химиясы, металдарды алу және олардың қосындыларының технологиясы.*

Мерзімділігі: жылына 6 рет.

Тиражы: 300 дана.

Редакцияның мекен-жайы: 050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19

http://www.geolog-technical.kz/index.php/en/

Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

«Известия НАН РК. Серия геологии и технических наук».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Собственник: Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ KZ39VPY00025420,** выданное 29.07.2020 г.

Тематическая направленность: *геология, химические технологии переработки нефти и газа, нефтехимия, технологии извлечения металлов и их соеденений.*

Периодичность: 6 раз в год.

Тираж: 300 экземпляров.

Адрес редакции: 050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19

http://www.geolog-technical.kz/index.php/en/

Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75.

News of the National Academy of Sciences of the Republic of Kazakhstan. Series of geology and technology sciences.
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)
Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty).
The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. KZ39VPY00025420,** issued 29.07.2020.
Thematic scope: *geology, chemical technologies for oil and gas processing, petrochemistry, technologies for extracting metals and their connections.*
Periodicity: 6 times a year.
Circulation: 300 copies.
Editorial address: 28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19
http://www.geolog-technical.kz/index.php/en/

Address of printing house: ST «Aruna», 75, Muratbayev str, Almaty.

NEWS of the National Academy of Sciences of the Republic of Kazakhstan
SERIES OF GEOLOGY AND TECHNICAL SCIENCES
ISSN 2224-5278
Volume 3, Number 459 (2023), 164–179
https://doi.org/10.32014/2023.2518-170X.307
UDC 556:36

S. Seifullin Kazakh AgroTechnical Research University, Astana, Kazakhstan.
E-mail: raushan85_07@mail.ru

MATHEMATICAL FOUNDATIONS OF ALGORITHMIZATION OF WATER POLLUTION MODELING PROCESSES

Moldasheva R.N. — Doctoral student of the Department of «Information Systems» S. Seifullin Kazakh AgroTechnical Research University, Astana, Kazakhstan
E-mail: raushan85_07@mail.ru, https://orcid.org/0000-0002-4570-0487;
Shazhdekeyeva N.K. — Associate Professor of the Department of Mathematics and Methods of Teaching Mathematic of the Atyrau University named after Kh. Dosmukhamedov. Candidate of Physical and Mathematical Sciences. Republic of Kazakhstan, Atyrau, Kazakhstan
E-mail: n.shazhdekeeva@mail.ru, https://orcid.org/0000-002-6282-8124;
Myrzagereikyzy G. — Senior lecturer of the Department of «Physics and technical disciplines» Atyrau University named after H. Dosmukhamedov, Atyrau, Kazakhstan
E-mail: gulmay_86@mail.ru, http://orcid.org/0000-0003-3245-6662;
Makhatova V.E. — Associate Professor of the Department of software engineering of Kh. Dosmukhamedov Atyrau University. Candidate of Technical Sciences, Atyrau, Kazakhstan
E-mail: mahve@mail.ru, https://orcid.org/0000-0002-4082-9193;
Zadagali A.M. — Doctoral Student of the Department of «General Biology and Genomics» L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: z.a.aizhan1993@gmail.com, https://orcid.org/0000-0002-2537-3538.

Abstract. The paper considers the actual task of developing mathematical foundations for algorithmization of the processes of modeling pollution of reservoirs. In the course of long – term studies of the distribution of phytoplankton of the Kokshetau lakes group, in particular, Lakes Zerendi, Kopa, Shalkar, Imantau, measurements of chemical parameters of water, organoleptic properties, transparency were carried out. These data were used to detail individual results and construct forecast values that depend on fluctuations in indicators that characterize the state of hydrobiota. In modeling, a lake is considered as a complex system, and surface sampling points are considered as sources of information about the state of a water body at certain time intervals. The solution of the task is carried out by constructing a critical area, and the incoming information is ranked by the level of significance. The hypothesis is the statement that a certain forecast value is accepted if it enters a certain critical area limited by the values that are determined as a result of experimental measurements. The advantage of the proposed approach is the possibility of simultaneous comparison of the influence of many factors, as well as the use of both empirical and theoretical frequencies. This approach to algorithmization of reservoir pollution modeling can be used to solve applied problems related to the development of ranking algorithms, assessing the impact of a number of factors on the management object, creating environmental monitoring programs near potentially hazardous and hazardous industrial facilities, creating information technologies for analysis and project activities.

Keywords: critical domain, competing hypothesis, concordance coefficient, time interval, rank, theoretical frequency, empirical frequency, interval, normalization

© Р.Н. Молдашева*, Н.К. Шаждекеева, Г. Мырзагерейкызы, В.Е. Махатова, А.М. Задагали, 2023

С. Сейфуллина атындағы Қазақ агротехникалық зерттеу университеті,
Астана, Қазақстан.
E-mail: raushan85_07@mail.ru

СУ ОБЪЕКТІЛЕРІНІҢ ЛАСТАНУЫН МОДЕЛЬДЕУ ПРОЦЕСТЕРІН АЛГОРИТМДЕУДІҢ МАТЕМАТИКАЛЫҚ НЕГІЗДЕРІ

Аннотация. Жұмыста су объектілерінің ластануын модельдеу процестерін алгоритмдеудің математикалық негіздерін әзірлеудің өзекті міндеті қарастырылады. Көкшетау көлдері тобының, атап айтқанда Зеренді, Қопа, Шалқар, Имантау көлдерінің фитопланктонының таралуын көпжылдық зерттеу барысында судың химиялық көрсеткіштері, органолептикалық қасиеттері, мөлдірлігі өлшенді. Бұл деректер жеке нәтижелерді егжей-тегжейлі көрсету және гидробиотаның күйін сипаттайтын көрсеткіштердің ауытқуына байланысты болжамды мәндерді құру үшін пайдаланылды. Модельдеу кезінде көл күрделі жүйе ретінде қарастырылады, ал жер үсті сынамаларын алу нүктелері белгілі бір уақыт аралығында су объектісінің күйі туралы ақпарат көзі ретінде қарастырылады. Қойылған міндеттерді шешу сыни саланы құру арқылы жүзеге асырылады, ал келіп түскен ақпарат маңыздылық деңгейіне қарай сараланады. Гипотеза, егер ол тәжірибелік өлшеулер нәтижесінде анықталған мәндермен шектелген кейбір маңызды салаға енсе, кейбір болжамды мән қабылданады деген тұжырым болып табылады. Ұсынылған тәсілдің артықшылығы-көптеген факторлардың әсерін бір уақытта салыстыру, сонымен қатар эмпирикалық және теориялық жиіліктерді қолдану мүмкіндігі. Су объектілерінің ластануын модельдеуді алгоритмдеуге бұл тәсіл саралау алгоритмдерін әзірлеуге, басқару объектісіне бірқатар факторлардың әсерін бағалауға, өнеркәсіптің ықтимал қауіпті және қауіпті объектілерінің жанында қоршаған ортаның жай-күйін мониторингтеу бағдарламаларын құруға, талдау мен жобалау қызметіне арналған ақпараттық технологияларды құруға байланысты қолданбалы есептерді шешу үшін пайдаланылуы мүмкін.

Түйін сөздер: сыни сала, бәсекелес гипотеза, конкордация коэффициенті, уақыт сегменті, дәреже, теориялық жиілік, эмпирикалық жиілік, аралық, нормалау

© Р.Н. Молдашева*, Н.К. Шаждекеева, Г. Мырзагерейкызы, В.Е. Махатова, А.М. Задагали, 2023
Казахский агротехнический исследовательский университет
имени С. Сейфуллина, Астана, Қазақстан
E-mail: raushan85_07@mail.ru

МАТЕМАТИЧЕСКИЕ ОСНОВЫ АЛГОРИТМИЗАЦИИ ПРОЦЕССОВ МОДЕЛИРОВАНИЯ ЗАГРЯЗНЕНИЙ ВОДОЕМОВ

Аннотация. В работе рассматривается актуальная задача разработки математических основ алгоритмизации процессов моделирования загрязнений водоемов. В процессе многолетних исследований распространения фитопланктона группы Кокшетауских озёр, в частности – озер Зеренди, Копа, Шалкар, Имантау — проводились замеры химических показателей воды, органолептических свойств, прозрачности. Эти данные использовались для детализации отдельных результатов и построении прогнозных значений, которые зависят от колебаний показателей, что характеризуют состояние гидробиоты. При моделировании озеро рассматривается как сложная система, а точки забора поверхностных проб – как источники информации о состоянии водного объекта на отдельных отрезках времени. Решение поставленной задачи осуществляется посредством построения критической области, а поступающая информация ранжируется по уровню значимости. Гипотезой выступает утверждение, что некоторое прогнозное значение принимается, если оно входит в некоторую критическую область, ограниченную значениями, которые определены в результате опытных замеров. Преимуществом предложенного подхода является возможность одновременного сравнения влияния множества факторов, а также использования как эмпирических, так и теоретических частот. Данный подход к алгоритмизации моделирования загрязнений водоемов может быть использован для решения прикладных задач, связанных с разработкой алгоритмов ранжирования, оценки влияния ряда факторов на объект управления, создания программ мониторинга состояния окружающей среды возле потенциально опасных и опасных объектов промышленности, создания информационных технологий для анализа и проектной деятельности.

Ключевые слова: критическая область, конкурирующая гипотеза, коэффициент конкордации, отрезок времени, ранг, теоретическая частота, эмпирическая частота, интервал, нормирование

Introduction

The task of preserving freshwater sources is relevant for many countries of the world (Bozorg-Haddad et al., 2017). The Republic of Kazakhstan has reserves of this resource (Thevs et al., 2017), however, the globalization of environmental problems denies the negative impact of water sources as a whole, without reference to individual countries or regions (Tang et al., 2019). Researchers are concerned about the presence of inorganic compounds in water, such as agrochemicals, pharmaceuticals, household chemicals. Such pollution is much more disturbing than the presence of organic impurities (Tang et al., 2019), since the mechanism of the latter's propagation, including the possibility of modeling mass transfer processes, is well studied and represented by various models (Bozorg-Haddad et al., 2017; Wen et al., 2017), although it is still one of the global challenges of modern society (Rakhmetov et al., 2022).

Algorithmic description of pollution processes allows reflecting the sequence and interrelation of system features by means of mathematical expressions for subsequent computer implementation. It is thanks to algorithmization that it is possible to present the object of research, the principles of functioning and properties as close as possible to the original (Orazbayev et al., 2021). Statistical data make it possible to obtain a model based on a set of random signals with a given probability density and accuracy, for example, as it is presented for studying the balance of groundwater (Kuanbayeva et al., 2022) or water pollution processes in the cooling reservoir of Ekibastuz GRES-1 (Romanova et al., 2017). There are many software packages for modeling pollution of reservoirs (Ziemińska-Stolarska et al., 2012). However, most of them are designed to process a large number of indicators and are intended for research of oceans, seas, large waterways. Such programs are based on algorithms describing the processes occurring in large masses of water, operations are performed with large data that are difficult to collect and process in the study of small bodies of water. In addition, when studying a small reservoir as a separate biota, it is necessary to consider many disparate factors (Alpyssov et al., 2023), obtained, among other things, as a result of paleontological studies, combining them into a complex system, while subjecting or refuting various hypotheses based on statistical data and the results of field experiments (Kaziyeva et al., 2018; Orazbayev et al., 2018). That is, in this case, it is necessary to algorithmically describe the processes that make it possible to simulate the ecological development or degradation of biota.

That is why the development of mathematical foundations for algorithmization of the processes of modeling pollution of reservoirs is an urgent task of research, especially in relation to the description of the ecological state of small reservoirs: lakes, small rivers, artificial outdoor pools for industrial and recreational purposes.

Methodology of experimental research

Materials and basic methods. In the course of long–term studies of the distribution of phytoplankton of the Kokshetau lakes group, in particular, Lakes Zerendi, Kopa, Shelkar, Imantau, the distribution of unicellular algae was studied depending on various indicators of the state of the reservoir, including pollution. These data were used as constraints or control actions in the development of a software implementation of the phytoplankton propagation method. The essence of the method consisted in determining the target point from which sampling vectors were constructed, based on the measurement results of which the picture of the phytoplankton of the lake was described. In the process of improving the development, the question arose about detailing individual results and constructing forecast values that depend on fluctuations in indicators characterizing pollution in hydrobiota (Abakumov et al., 2014). In particular, the lake is a kind of complex system *S*, from different points of which the researcher receives information that changes over time. Figure 1 shows a visualization of such a system using the example of Lake Zerendi.

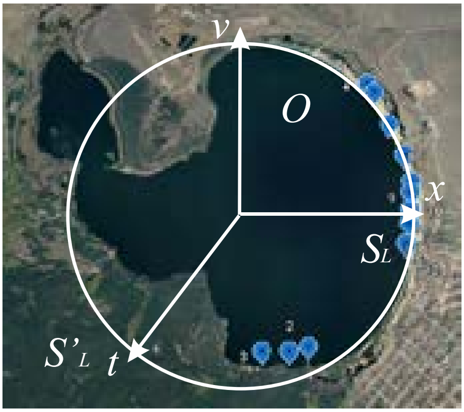

Fig. 1 – A complex system with changing information (on the example of Lake Zerendi)

The viability of the specified system, and, accordingly, the model of this system, can be represented as a function of time. That is, it depends on the information that comes from the system for a certain period of time. The general conclusion about the state of the system will depend on the solution of many tasks, but in general it can be described as the processing of each type of information over a period of time $[t_1, t_2]$. That is, on the basis of the totality of all objects of system *S*, a sample is formed describing the behavior of the system over time periods (Alpyssov et al., 2023).

If we imagine that the information $J'(S,t),\ J'(S,t) = J_k(S,t),\ k = 1,\ldots,r$ has the form

$$J'(S,t) = \|i(S_l,t)\|_{l=1}^{p}, \quad (1)$$

where: $i(S_l,t)$ is the value of some criterion based on the results of surface sampling S_l, $l=1,\ldots,p$. In this case, all the values obtained from the results of sampling, limited to the area *O* in Fig. 1, allow us to conclude about the pollution or purity of the aquatic ecosystem in a certain period of time, taking into account (1). But when describing the surface with vectors, some critical region will arise when the obtained value of the criterion lies outside the region *O* and is not the point of passage of the vector. However, the impact of this criterion on the system can be significant. And if the observed value of the criterion belongs to a critical area, then the hypothesis is put forward that this area is not subject to the influence of pollution processes of a water body. Now the points along the boundary of the description of the region *O* take the values of the critical points of the K_{cp}. They separate the area where the values from the sampling results are obtained from the area described by the model.

When finding a critical area, the significance level α is set and critical points are searched based on the following relations:

a) for the right-hand critical area:

$$P(K > K_{cp}) = \alpha(K_{cp} > 0);$$

b) for the left-hand critical area:

$$P(K < K_{cp}) = \alpha(K_{cp} < 0);$$

c) for a two-sided symmetric domain:

$$P(K > Kcp) = \frac{\alpha}{2}, \quad P(K < -Kcp) = \frac{\alpha}{2},$$

$$(K_{cp} > 0).$$

In this study, the critical area is constructed based on the requirement of the probability of obtaining points with results that coincide with the criterion equal to α, provided that the null hypothesis H_0 is valid. It turns out to be expedient to introduce into consideration the probability of the criterion falling into the critical region, provided that the null hypothesis is incorrect and, therefore, the competing hypothesis is valid. In this case, it is advisable to use the concept of criterion power when there is a probability that points corresponding to a certain criterion will fall into the critical region, provided that the competing hypothesis is valid.

In this case, the critical area model is constructed so that the power of the criterion is maximum. If the probability of an error of the second kind (to accept an incorrect hypothesis) is β, then the power is 1−β. If the power of 1−β increases, then the probability of β making a mistake of the second kind decreases. Thus, the higher the power, the less likely the error of the second kind is. However, in the process of algorithmization, it must be remembered that it is impossible to reduce both and β: if you reduce β, then β will increase. The only way to simultaneously reduce the probabilities of errors of the first and second kind is to increase the volume of samples, that is, taking surface samples. And this will lead to additional financial and labor costs.

Taking into account all the above, it is possible to form an approach to algorithmization of water pollution modeling processes. Let's assume that there are *n* measurements based on the results of sampling on a reservoir. It is necessary to arrange the indicators (*m*) and factors $X_1, X_2, \ldots, X_n$ in descending order of their influence on the result of the process (on the state variable *Y*). To do this, it is necessary to assign ranks to various factors. It is assumed that a_{ij} is the rank assigned to the *j*-th factor $(1 \le a_{ij} \le n,\ i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$ (Table 1).

Table 1 – Table of ratios of results and factors influencing them

Indicators / Factors	1	2	…	*m*
X_1	a_{11}	a_{12}	…	a_{1n}
X_2	a_{21}	a_{22}	…	a_{2n}
…	…	…	…	…
X_n	a_{m1}	a_{m2}	…	a_{mn}

The sum of ranks per row for all rows is the same and equal

$$\frac{(п+1)п}{2} \quad (2).$$

Average value of ranks in a row:

$$\frac{п+1}{2} \quad (3).$$

The average value of the sum of ranks in the column:

$$\bar{a} = \frac{т(п+1)}{2} \quad (4).$$

The coefficient of agreement (coefficient of concordance) is calculated:

$$W = \frac{S(d^2)}{max\, S(d^2)} \quad (5),$$

where: $S(d^2)$ is the sum of the squares of the deviation of the sum of ranks from the average sum. Calculated by the formula:

$$S(d2) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij} - \bar{a}\right)^2 = \sum_{j=1}^{n} d_j^2, \quad (6).$$
$$max\, S\,(d^2) = \frac{1}{12} m^2 (n^3 - n)$$

If the hypothesis is consistent and based on calculated data, then $S(d2) = max\, S\,(d^2)$, i.e. $W = 1$, then an area is formed based on the model data.

If the hypothesis is not consistent, then the value of $S(d^2)$ is close to zero and $W \approx 0$. Accordingly, the region is not constructed.

If fractional ranks were obtained based on several factors, then the concordance coefficient is calculated by the formula

$$\mathrm{W} = \frac{S(d^2)}{max\, S(d^2) - \frac{т}{12}\sum_{i=1}^{m} T_i} = \frac{S(d^2)}{\frac{1}{12} m^2 (n^3 - n) - \frac{т}{12}\sum_{i=1}^{m} T_i} \quad (7),$$

and T_i is represented as

$$T_i = \sum_{k=1}^{r}(t_{ik}^3 - t_{ik}) \quad (8),$$

where i - is the number of the indicator for which the measurement results were obtained;

k - is the repetition number;

t_{ik} - is the number of identical ranks in the k - th repetition of the i - th dimension.

To test the null hypothesis H_0, a random variable x^2- distribution with the number of degrees of freedom $k = n - 1$ is used as a statistical criterion. In this case, a left-sided critical region is constructed:

1. Calculate the observed value of the criterion

$$\chi^2_{obser.} = (n - 1)mW \quad (9).$$

2. According to a given level of significance α and the number of degrees of freedom $k = n - 1$, the critical point $\chi^2_{cr}(\alpha;\ k)$ is determined according to the table of critical points x^2- distributions.

3. If $\chi^2_{obser.} > \chi^2_{cr}$, then the hypothesis H_0 is accepted;

4. If $\chi^2_{obser.} < \chi^2_{cr}$, then the hypothesis H_0 is not accepted.

During the study of Lake Zerendi, six factors affecting the state variable – the pollution of fresh water were identified. The results of measurements of these indicators were carried out from May to August in 2020 – 2022. The indicators for 2022 are shown in Table 2.

Table 2 – Information on the surface water quality of Lake Zerendi in 2022

Physico-chemical parameters	May	June	July	August
Hydrogen index (X_1)	9,00	8,71	8,5	8,86
Oxygen concentration in water (X_2)	8,00 mg/dm³	7,84 mg/dm³	6,91 mg/dm³	13,19 mg/dm³
Biochemical consumption of the reservoir (X_3)	1,70 mg/dm³	2,14 mg/dm³	0,88 mg/dm³	0,88 mg/dm³
Chemical Oxygen consumption (X_4)	47,0 mg/dm³	45,0 mg/dm³	49,7 mg/dm³	64,2 mg/dm³
Suspended solids (X_5)	13,0 mg/dm³	13,0 mg/dm³	5,2 mg/dm³	5,2 mg/dm³
Mineralization (X_6)	631 mg/dm³	761 mg/dm³	1204 mg/dm³	1178 mg/dm³

Using similar data over the past years, average indicators were derived, on the basis of which the algorithmization of the processes of pollution of Lake Zerendi was worked out.

The results of experimental research

Six factors that have a significant impact on the pollution of Lake Zerendi are the hydrogen index (X_1), the concentration of oxygen in the water (X_2), the biochemical consumption of the reservoir (X_3), chemical oxygen consumption (X_4), suspended solids (X_5), mineralization (X_6). The ranking of the influence of factors on water pollution in Lake Zerendi by month is presented in Table 3.

Table 3 – Ranks of the influence of factors on water pollution by month

Indicators / Factors	May	June	July	August
X_1	1,5	2	2	1,5
X_2	5	3	3	3,5
X_3	1,5	1	1	1,5
X_4	4	4,5	5,5	5,5
X_5	3	4,5	5,5	3,5
X_6	6	6	4	6

Based on the presented data, we compile an algorithm for modeling the influence of specified factors on the state of water in the lake for the future. The significance level is taken equal to $\alpha = 0.05$.

Step 1. From Table 3, set n and m.

Step 2. Determine the average value (4) of the sum of ranks in the column.

Step 3. Form the calculation tables (Table. 4 and table. 5) according to the formulas (2), (3), (8) based on the data (Table 3).

Table 4 – Calculated table of ranks by month

Indicators / Factors	t_{i1}	$t^3_{i1} - t_{i1}$	$t\ t_{i2}$	t^3_{i2} - t_{i2}	T_i
May	2	6	0	0	6
June	2	6	0	0	6
July	2	6	0	0	6
August	2	6	2	6	12

Table 5 – Calculated ranking table of factors

Indicators / Factors	X_1	X_2	X_3	X_4	X_5	X_6
$\sum_{i=1}^{4} a_{ij}$	7	14,5	5	19	16,5	22
d_j	−7	0,5	−9	5	2,5	8
d_j^2	49	0,25	81	25	6,25	64

Step 4. Since there are fractional ranks in the table, the concordance coefficient is calculated by formula (7).

Otherwise, by formula (5).

Step 5. Calculate $S(d^2)$ by formula (6). Based on Table 5:

$$S(d^2) = 225,5.$$

Step 6. Using (8) we define $\sum_{i=1}^{T} T_i$.

We get:

$$T_1 = (t_{11}^3 - t_{11}) + (t_{12}^3 - t_{12}) = 6,$$
$$T_2 = (t_{21}^3 - t_{21}) + (t_{22}^3 - t_{22}) = 6,$$
$$T_3 = (t_{31}^3 - t_{31}) + (t_{32}^3 - t_{32}) = 6,$$
$$T_4 = (t_{41}^3 - t_{41}) + (t_{42}^3 - t_{42}) = 12$$

$$\sum_{i=1}^{T} T_i = \sum_{i=1}^{4} T_i = 30$$

Step 7. Based on formula (7), we get the result $W = 0.805$.

Step 8. Proceed to testing the hypothesis H_0.

Step 9. Calculate by (9) the observed value of the criterion:

$$\chi^2_{\text{obser.}} = (6-1)\cdot 4 \cdot 0{,}805 = 16{,}1.$$

Step 10. According to the given significance level $\alpha = 0.05$ and the number of degrees of freedom $k = 6 - 1 = 5$, we determine the critical point $\chi^2_{cr}(\alpha;\ k)$ from the table of critical points x^2-distributions:

$$\chi^2_{cr}(\alpha;\ k) = \chi^2_{cr}(0{,}05;\ 5) = 11{,}07.$$

Step 11. Since $\chi^2_{obser.} > \chi^2_{cr}$, then the hypothesis H_0 is accepted – the area of the model is described based on the forecast data.

Otherwise, the hypothesis is not accepted and the forecast data are not determined (Fig. 2).

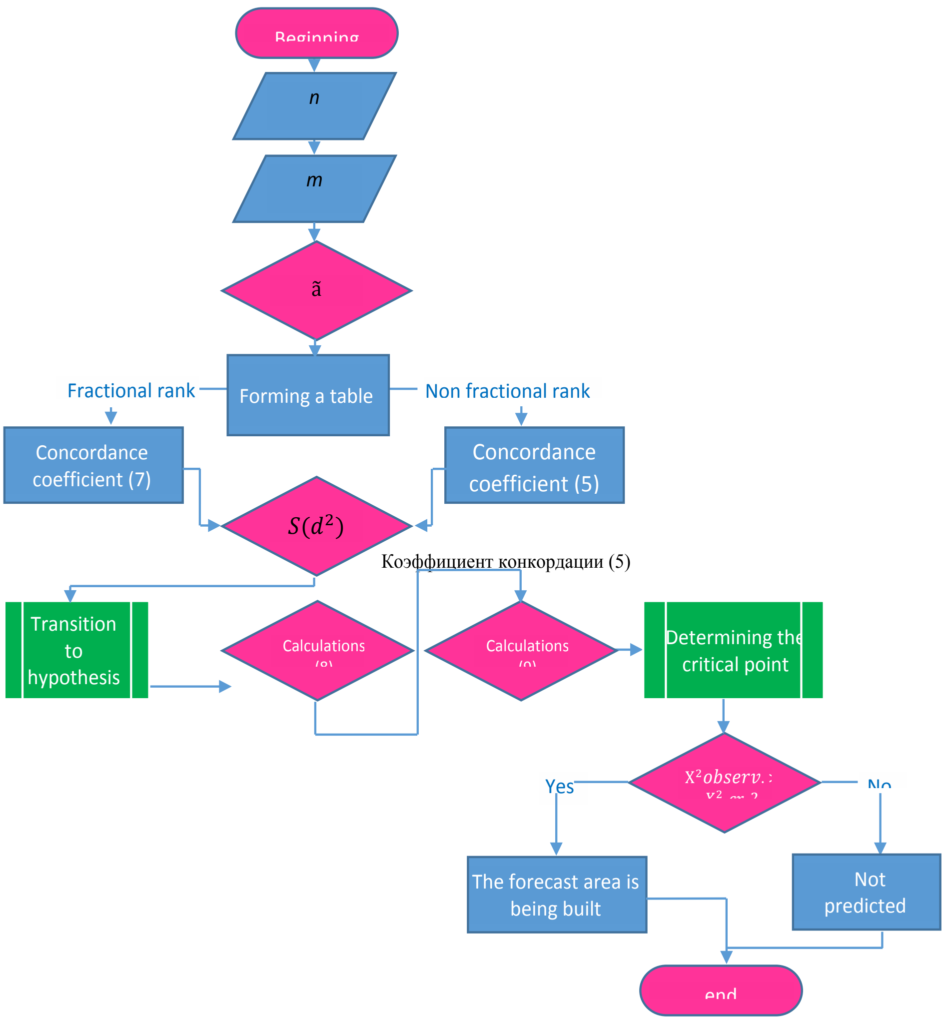

Fig. 2 – Algorithm for modeling the influence of specified factors on the state of water in the lake

After that, you can implement an algorithm for direct visualization of the model or output of calculated results.

As can be seen from the calculations, chemical oxygen consumption is the main indicator of pollution of Lake Zerendi, which is consistent with the results of laboratory studies. Accordingly, according to this indicator, it is possible to simulate lake pollution.

Discussion

To verify the presented mathematical description of the process of spreading pollution of reservoirs, one can use the classical approach in statistics using the Pearson criterion (Orazbayev et al., 2020). In the process of collecting surface samples, it is possible to form some empirical distribution in the form of a sequence of variants, on the basis of which the hypothesis H_0 can be tested: the general population X is distributed according to the normal law. But at the same time, it should not be forgotten that using this criterion will allow you to get more accurate results with an increase in the number of grouped data by criteria. When testing complex hypotheses, this does not give an advantage (Ismailova et al., 2018). Avoiding additional calculations, it is possible to carry out the proof using the following algorithm:

Step 1. Write down the empirical distribution in the form of a sequence of intervals [x_i, x_{i+1}) with the corresponding frequencies n_i (n_i - is the sum of the frequencies that fell into the i - th interval).

$$[x_i, x_{i+1}): [x_1, x_2)\ [x_2, x_3)\ldots[x_s, x_{s+1})$$
$$n_i \quad : \quad n_1 \quad n_2 \quad \ldots \quad n_s$$

Step 2. Calculate the sample average $\overline{x^*}$ and the mean square deviation σ*, and as variants x_i^* we take the arithmetic mean of the ends of the interval:

$$x_i^* = \frac{x_i + x_{i+1}}{2}.$$

Step 3. Normalize X, i.e. we make the transition to a random value

$$Z = \frac{X - \overline{x^*}}{\sigma^*},$$

and calculate the ends of the intervals:

$$Z_i + 1 = \frac{x_{i+1} - \overline{x^*}}{\sigma^*},$$

moreover, the smallest value of Z, i.e. z_1, is assumed to be equal to −∞, and the largest, i.e. z_s, is assumed to be equal to +∞.

Step 4. Calculate the theoretical frequencies

$$\text{п}_i' = n \cdot P_i,$$

where:

n - is the sample size;

$P_i = F(z_{i+1}) - F(z_i)$ is the probability of Z falling into the intervals (z_i, z_{i+1}),

$F(z)$ - is the Laplace function.

Step 5. We compare empirical and theoretical frequencies using the Pearson criterion. To do this:

a) a calculation table is formed, according to which the observed value of the Pearson criterion is found

$$\chi^2_{obser.} = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'};$$

b) according to the table of critical distribution points x^2 [16], according to a given level of significance and the number of degrees of freedom $k = s - 3$ (s - is the number of sampling intervals), the critical point of the right–sided critical region $\chi^2_{cr}(\alpha, k)$ is located.

Step 6. If $\chi^2_{obser.} < \chi^2_{cr}$, then the hypothesis H_0 about the normal distribution of the general population is accepted. In other words, empirical and theoretical frequencies differ insignificantly.

If $\chi^2_{obser.} > \chi^2_{cr}$, then the hypothesis H_0 is rejected. In other words, empirical and theoretical frequencies differ significantly (Fig. 3).

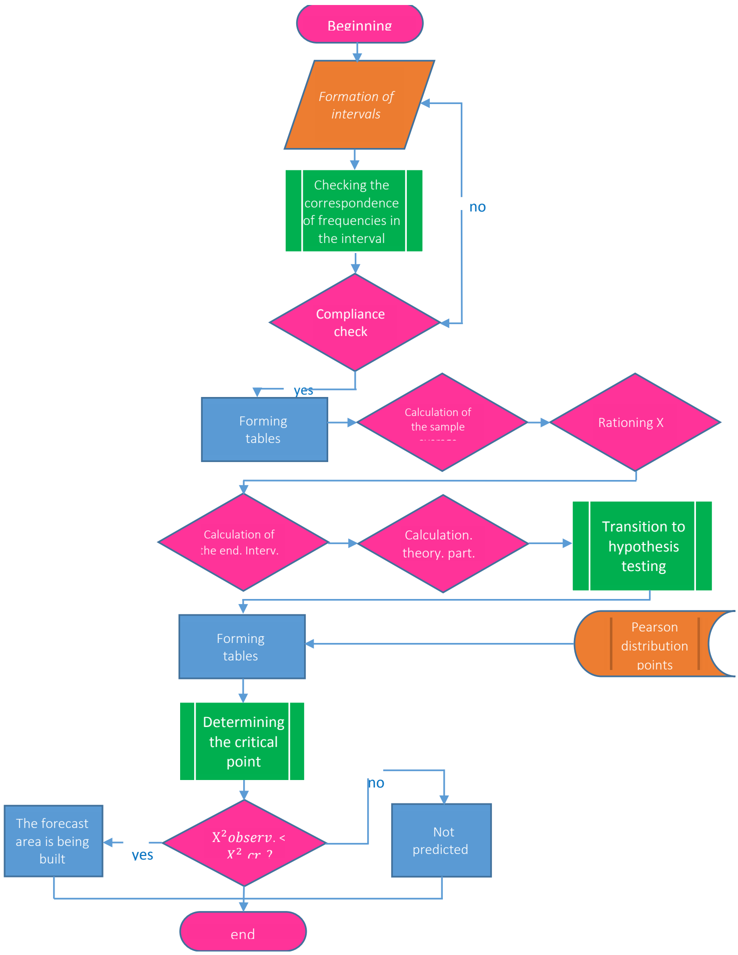

Fig.3 – Algorithm for modeling the influence of specified factors on the state of water in the lake using the Pearson criterion

The presented approach to verification is more voluminous for calculations and computer implementation, involving external resources, as well as tables for intermediate data processing. However, in this case, the presented allows us to prove the rationality of the proposed algorithmization of the processes of modeling pollution of reservoirs. This is especially true of the processes of studying relatively small reservoirs, since, based on (Shopagulov et al., 2016), pollution modeling occurs at various spatial scales with empirical parameters, then such modeling with empirical and theoretical frequencies allows not only to build a model, but also to verify its accuracy. And if we take into account the work of (Nagmetova et al., 2020; Mauina et al., 2021), then this approach allows us to reduce the number of calculations while maintaining the accuracy of the results obtained.

Taking into account what is stated in, the presentation of data by ranks makes it possible to present multi-criteria models, including the involvement of factors that could not be included in the models (Golenko et al., 2022; Bayegizova et al., 2016), but had an impact on the system. This is justified by the use of the significance level and a different number of degrees of freedom.

But if we take into account the results of (Yessenova et al., 2023) and add the possibility of to the one proposed in the work, then it is possible to obtain models of mixing of water layers and pollutants, which will allow us to deepen studies like (Orazbayev et al., 2021).

I would like to pay special attention to Table 3 and steps 1-4 of the algorithm presented in the results. In this case, the ranks of the influence of factors on water pollution in the object of study can be determined both with the help of coefficients, automating this process, and with the help of experts, for example, scientists studying the reservoir. From step 5 to the end of the algorithm, these ranking forms should be calculated according to the proposed formulas (6)–(9), after which they should be compared, analyzing which parameters the estimates differ and why. In this case, using the software implementation of the phytoplankton distribution method, which was mentioned in the materials and methods section of this work, it is possible to obtain the basis for creating a full–fledged information technology - an expert system or a decision support system in the field of water body protection. This indicates the prospects for the development of this development.

Conclusion

The paper presents the development of mathematical foundations for algorithmization of water pollution modeling processes with an emphasis on the study of small water sources, which are not only unique biota with diverse representatives of flora and fauna, but also reserve sources of fresh water. It was in the process of studying the distribution of phytoplankton of the Kokshetau lakes group that a number of data on the chemical and organoleptic properties of water affecting the distribution of unicellular plant organisms were obtained.

The mathematical description of water pollution modeling processes is based directly on the study of information received from a source in a separate period of time. Such information can serve as data obtained from the results of sampling surface water samples. Next, a critical area is constructed based on the requirement of the probability of obtaining points with results that coincide with the criterion of a given level of significance. And then the hypothesis is tested, where the frequencies are compared, which make it possible to rank the influence of individual factors on the pollution of the reservoir.

The advantage of the proposed approach is the possibility of simultaneous comparison of the influence of many factors, as well as the use of both empirical and theoretical frequencies.

This approach to algorithmization of reservoir pollution modeling can be used to solve applied problems related to the development of ranking algorithms, assessing the impact of a number of factors on the management object, creating environmental monitoring programs near potentially hazardous and hazardous industrial facilities, creating information technologies for analysis and project activities.

REFERENCES

Alpyssov A., Uzakkyzy N., Talgatbek A., Moldasheva R., Bekmagambetova G., Yessekeyeva M., Kenzhaliev D., Yerzhan A., Tolstoy A., 2023 — *Alpyssov A., Uzakkyzy N., Talgatbek A., Moldasheva R., Bekmagambetova G., Yessekeyeva M., Kenzhaliev D., Yerzhan A., Tolstoy A.* Assessment of plant disease detection by deep learning. Eastern-European Journal of Enterprise Technologies, 1 (2 (121)). Pp. 41–48. doi: https://doi.org/10.15587/1729-4061.2023.274483.

Abakumov A., Ismailova A., Adamov A., 2014 — *Abakumov A., Ismailova A., Adamov A.* MODELING OF MICROBIAL COMMUNITIES OF PLANT ORGANISMS IN AQUATIC ECOSYSTEM. Information (Japan), 17(1). Pp. 209–218.

Bozorg-Haddad O., Soleimani S., Loáiciga H.A., 2017 — *Bozorg-Haddad O., Soleimani S., Loáiciga H.A.* Modeling Water-Quality Parameters Using Genetic Algorithm–Least Squares Support Vector Regression and Genetic Programming. Journal of Environmental Engineering. 143 (7): 04017021. doi:10.1061/(ASCE)EE.1943-7870.0001217. ISSN 0733-9372.

Bayegizova A., Murzabekova G., Ismailova A., Aitimova U., Mukhanova A., Beldeubayeva Zh., Ainagulova A., Naizagarayeva, A. — *Bayegizova A., Murzabekova G., Ismailova A., Aitimova U., Mukhanova A., Beldeubayeva Zh., Ainagulova A., Naizagarayeva A.* Effectiveness of the use of algorithms and methods of artificial technologies for sign language recognition for people with disabilities. // Eastern-European Journal of Enterprise Technologies, 2022, Volume 4, Issue (2-118). Pp. 25–31. ISSN 1729-3774, DOI: https://doi.org/10.15587/1729-4061.2022.262509/ ttps://www.scopus.com/record/display.uri?eid=2-s2.0-85137728736& origin = resultslist&sort=plf-f.

Golenko Y., Ismailova A., Shaushenova A., Mutalova Zh., Dossalyanov D., Ainagulova A., Naizagarayeva A. — *Golenko Y., Ismailova A., Shaushenova A., Mutalova Zh., Dossalyanov D., Ainagulova A., Naizagarayeva A.* Implementation of machine learning models to determine the appropriate model for protein function prediction. // Eastern-European Journal of Enterprise Technologies Volume 5, Issue 4-119, 2022. Pp. 42–49, https://doi.org/10.15587/1729-4061.2022.263270.

Ismailova A.A., Zhamangara A.K., Sagnayeva S.K., Kaziyeva G.D., Abakumov A.I., Park S.Y., 2018 — *Ismailova A.A., Zhamangara A.K., Sagnayeva S.K., Kaziyeva G.D., Abakumov A.I., Park S.Y.* TECHNOLOGIES OF INFORMATION MONITORING BIOGENS LAKES OF KAZAKHSTAN. // News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences. 3(430). Pp. 69–73. http://nblib.library.kz/elib/library.kz/Jurnal/%D0%93%D0%B5%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%8F_04_2018/Ismailova%20A.A,Zhamangara.pdf

Kaziyeva G.D., Sagnayeva S.K., Sembina G.K., Ismailova A.A. — *Kaziyeva G.D., Sagnayeva S.K., Sembina G.K., Ismailova A.A. Software* tools for environmental monitoring of the Northern part of the Caspian Sea. // 2018 4th International Conference on Computer and Technology Applications, ICCTA. 2018. Pp. 85–90. DOI: 10.1109/CATA.2018.8398661.

Kuanbayeva B., Rakhmetov M., Turkmenbayev A., Abdykerimova, Tumysheva A., — *Kuanbayeva B., Rakhmetov M., Turkmenbayev A., Abdykerimova, Tumysheva A.* The effectiveness of using interactive computer models as methodological tools in science school education. // World Trans. on Engng. and Technol. Educ., 20, 4. Pp.306–311. (2022). URL: http://www.wiete.com.au/journals/WTE&TE/Pages/TOC_V20 N4.html.

Mauina G.M., Chertkova E.A., Nukusheva S.A., Aitimova U.Zh., Ismailova A.A. — *Mauina G.M., Chertkova E.A., Nukusheva S.A., Aitimova U.Zh., Ismailova A.A.* Expert-statistical method of management decision support for agricultural enterprises of northern Kazakhstan. Journal of Theoretical and Applied Information Technology. 30th June 2021. Vol.99. № 12. Pp. 3071–3083. ISSN: 1992-8645. E-ISSN: 1817–3195.

Nagmetova G., Berdimuratova K., Amirgazin A., Shevtsov V., Ismailova A., Ramankulov Y., Shevtsov A., Kurmanbayev A. — *Nagmetova G., Berdimuratova K., Amirgazin A., Shevtsov V., Ismailova A., Ramankulov Y., Shevtsov A., Kurmanbayev A.* Draft genome sequence of a potential commercial biocellulose producer, strain Komagataeibacter europaeus GH1. // Microbiology Resource Announcements, 2020 Oct 15; 9(42): e00963-20. DOI: https://doi.org/10.1128/MRA.00963-20.

Orazbayev B., Zhumadillayeva A, Orazbayeva K, Kurmangaziyeva L, Dyussekeyev K, Iskakova S. — *Orazbayev B., Zhumadillayeva A, Orazbayeva K, Kurmangaziyeva L, Dyussekeyev K, Iskakova S.* Methods for Developing Models in a Fuzzy Environment of Reactor and Hydrotreating Furnace of a Catalytic Reforming Unit. // *Applied Sciences*. 2021; 11(18):8317. –P. 1-22. https://doi.org/10.3390/app11188317

Orazbayev B., Assanova B., Bakiyev M., Krawczyk J., Orazbayeva K. — *Orazbayev B., Assanova B., Bakiyev M., Krawczyk J., Orazbayeva K.* Methods of model synthesis and multicriteria optimization of chemical-engineering systems in the fuzzy environment. // Journal of Theoretical and Applied Information Technology 31st March 2020. Vol.98. № 06. –Pp. 1021–1036. https://www.scopus.com/record/display.uri?eid=2-s2.0-85083790104&origin=resultslist&sort=plf-f

https://www.scopus.com/record/display.uri?eid=2-s2.0-85136288448&origin=resultslist&sort=plf-f.

Orazbayev B., Moldasheva Zh., Orazbayeva K., Makhatova V., Kurmangaziyeva L. and Gabdulova A., — *Orazbayev B., Moldasheva Zh., Orazbayeva K., Makhatova V., Kurmangaziyeva L. and Gabdulova A.* Development of mathematical models and optimization of operation modes of the oil heating station of main oil pipelines under conditions of fuzzy initial information. // Eastern-European J. of Enterprise Technologies, 6, 2 (114). Pp. 147–162 (2021).

Orazbayev B.B., Ospanov E.A., Orazbayeva K.N., Kurmangazieva L.T. — *Orazbayev B.B., Ospanov E.A., Orazbayeva K.N., Kurmangazieva L.T. A* Hybrid Method for the Development of Mathematical Models of a Chemical Engineering System in Ambiguous Condition // Mathematical Models and Computer Simulations, 2018, Vol. 10. № 6, -Pp.748–758. DOI 10.1134/S2070048219010125.

Rakhmetov M., Sadvakassova A., Saltanova G. and Yessekenova A., — *Rakhmetov M., Sadvakassova A., Saltanova G. and Yessekenova A.* Usage and effectiveness of educational platforms in Kazakhstan during the Covid-19 pandemic. // World Trans. on Engng. and Technol. Educ., 20, 3. Pp. 226–231. (2022).URL:

Romanova S.M., Ponomarenko O.I., Niyazbaeva A.I., Amirgaliev N.A., 2017 — *Romanova S.M., Ponomarenko O.I., Niyazbaeva A.I., Amirgaliev N.A.* QUALITY OF WATER COOLER -RESERVOIR OF EKIBASTUZ POWER PLANT-1 NAMED AFTER BULAT NURZHANOV. NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN SERIES OF GEOLOGY AND TECHNICAL SCIENCES. Volume 2. Number 422 (2017). Pp. 90–98. ISSN 2224-5278. http://www.geolog-technical.kz/images/pdf/g20172/9099.pdf

Shopagulov O., Tretyakov I., Ismailova, A. Aitimova U., Beisembayev K., Mukhanbetkaliyeva A. — *Shopagulov O., Tretyakov I., Ismailova, A. Aitimova U., Beisembayev K., Mukhanbetkaliyeva A.* An expert system for diagnosis cow diseases. // Journal of Theoretical and Applied Information Technology. 15th August 2020. Vol. 98. № 15. Pp. 3106–3115. ISSN: 1992–8645. E-ISSN: 1817–3195.

Thevs N., Nurtazin S., Beckmann V., Salmyrzauli R., Khalil A. — *Thevs N., Nurtazin S., Beckmann V., Salmyrzauli R., Khalil A.* Water Consumption of Agriculture and Natural Ecosystems along the Ili River in China and Kazakhstan. Water 2017, 9, 207. https://doi.org/10.3390/w9030207.

Tang T., Strokal M., van Vliet Mchelle T.H., Seuntjens P., Burek P., Kroeze C., Langan S., Wada Y., 2019 — *Tang T., Strokal M., van Vliet Mchelle T.H., Seuntjens P., Burek P., Kroeze C., Langan S.,* *Wada Y.* Bridging global, basin and local-scale water quality modeling towards enhancing water quality management worldwide. Current Opinion in Environmental Sustainability. 36. Pp.39–48. doi:10.1016/j.cosust.2018.10.004.

Yessenova M., Abdikerimova G., Ayazbaev T., Murzabekova G., Ismailova A., Beldeubayeva Zh., Ainagulova A., Mukhanova A. — *Yessenova M., Abdikerimova G., Ayazbaev T., Murzabekova G., Ismailova A., Beldeubayeva Zh., Ainagulova A., Mukhanova A.* The effectiveness of methods and algorithms for detecting and isolating factors that negatively affect the growth of crops. // International Journal of Electrical and Computer Engineering, 2023, 13(2). Pp. 1669–1679. DOI:10.11591/ijece.v13i2.pp1669-1679.

Wen Y., Schoups G. & van de Giesen N. — *Wen Y., Schoups G. & van de Giesen N.* Organic pollution of rivers: Combined threats of urbanization, livestock farming and global climate change. Sci Rep 7, 43289 (2017). https://doi.org/10.1038/srep43289.

Ziemińska-Stolarska A., Skrzypski J., 2012 — *Ziemińska-Stolarska A., Skrzypski J.* Review of Mathematical Models of Water Quality. Ecological Chemistry and Engineering, 19(2). Pp. 197–211. https://doi.org/10.2478/v10216-011-0015-x.

CONTENTS

Publication Ethics and Publication Malpractice in the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
www:nauka-nanrk.kz
http://www.geolog-technical.kz/index.php/en/
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Заместитель директор отдела издания научных журналов НАН РК *Р. Жәлиқызы*
Редакторы: *М.С. Ахметова, Д.С. Аленов*
Верстка на компьютере *Г.Д.Жадыранова*

Подписано в печать 14.06.2023.
Формат 70х90 $^1/_{16}$. Бумага офсетная. Печать – ризограф.
20,0 п.л. Тираж 300. Заказ 3.

Национальная академия наук РК
050010, Алматы, ул. Шевченко, 28, т. 272-13-19