

**REPORTS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN**

ISSN 2224-5227

<https://doi.org/10.32014/2020.2518-1483.89>

Volume 4, Number 332 (2020), 57 – 64

МРНТИ 37.27

УДК 556

G.V. Ayzel¹, A.C. Izhitskiy², A.K. Kurbaniyazov³¹Ph.D, researcher, Institute of Water Problems of the Russian Academy of Sciences, Moscow, Russia;²Ph.D, senior researcher, P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia;³Ph.D., Associate Professor, International Kazakh-Turkish University named after H.A. Yasavi,

Turkestan, Kazakhstan.

E-mail: abigazi@mail.ru

**FORECASTING OF THE STATE OF THE SMALL ARAL SEA
BASED ON OPEN DATA SOURCES**

Abstract. The study of the dynamics of the level and volume of water in the Aral Sea is an urgent scientific task due to the need to understand the mechanisms of natural and anthropogenic processes that have induced a radical change in its water and salt balance over the past 60 years. In particular, the study of the dynamics of water balance components in the basin of the Small Aral Sea is the most important task when planning scenarios for water use in the region. In the proposed work, based on the methods of machine learning (for the implementation of computational functions in the program), two statistical models were developed: a forecast model for the monthly values of river flow in the Syrdarya river and the forecast of variability of the water volume of the Small Aral Sea. Based on the simulation results, forecasts were made for the values of the Syrdarya drainage and the water volume of the Small Aral Sea. In conditions of low availability of field observations data, the operational estimates of the water balance component are the most important source of information on the changes occurring in the basin under investigation. The proposed technique can also be used to obtain initial conditions in experiments on hydrodynamic modeling, as well as to calculate climatic scenarios for the development of the hydrological system of the Aral Sea.

Keywords: Small Aral Sea, sea level, machine learning, river discharge, basins, Syrdarya, hydrodynamic modeling.

Introduction. The research of the dynamics of the level and volume of the Aral Sea is an urgent scientific task due to the need to understand the mechanisms of natural and anthropogenic processes that have induced a radical change in its water and salt balance over the past 50 years.

Since the late 60's, the volume of the sea decreased by 90%, the salinity of its water increased by an order of magnitude [1]. The research of the processes of the hydrological cycle of the Aral Sea is hampered by the almost complete absence of data from modern measurements of the water balance composing it. Since the 60s the Aral Sea is undergoing irreversible changes in the water and salt regimes, cardinal changes have affected not only the ecosystem of the sea itself, but also affected its entire basin. Over the past 15 years, the tendency of the Aral Sea to separate water bodies has deepened: the deep-water western and shallow eastern basins of the Greater Aral, the Small Aral Sea, and the Tushi Baza (figure 1). Thus, the eastern basin of the Greater Aral Sea in recent years has virtually ceased to be a permanent reservoir, becoming an ephemeral lake, whose existence is determined by the magnitude of the seasonal full-water flow of the Amu Darya river.. The northern basin of the Aral Sea (or the Small Aral Sea) has in recent years almost lost contact with other basins - both due to natural causes of drying up of the western and eastern basins, and due to the construction of a dam designed to minimize water exchange between the basins [2].

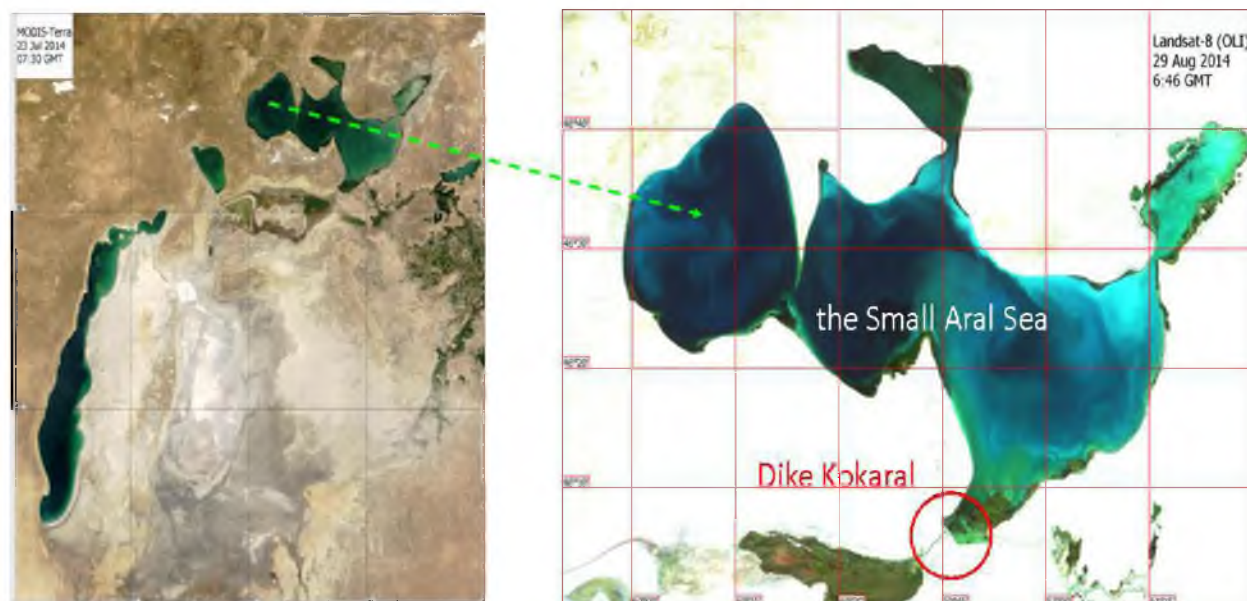


Figure 1 - Satellite image of the Aral Sea in July 2014 (left), satellite image of the Small Aral in August 2014 (right)

In this paper, we have emphasized the research of the dynamics of the water volume of the Small Aral Sea. In recent years, after minimizing its water exchange with other parts of the sea, it began to establish a relatively stable water-salt regime, similar in characteristics to the "conditionally-natural" regime of the Aral Sea, before drying began. Nevertheless, the significant shortage of these direct observations of the water balance component for this region is still felt as sharply. Modern field observations on all parts of the Aral Sea are episodic, most of the regime monitoring hydro-meteorological stations and stations ceased to exist in the first half of the 1990s, so in this work we used the database of indirect water level observations as well as climate reanalysis database. The main goal of this work is to research the possibility of modeling the dynamics of the water volume of a basin located in arid conditions using the data of indirect measurements of open sources (Fig. 2). Key modeling concept is based on implementation of a simple Decision Tree model in case of regression task [3]. Typical Decision Tree model is a "white box" consists of the range of boolean classifiers which split our samples to tiny "leaf" nodes where all samples constantly refers to the one target value [4]. Single tree-based implementation of Decision Tree algorithm faced with the case of overfitting and robustness lack that lead to limited using in real world examples [5]. In our work we used three cutting-edge machine learning techniques based on ensemble approach to predictions: Random Forest, Extra Trees and Gradient Boosting [6]. All of them are based on ensembles of simple Decision Tree models and provide useful tricks such bagging, bootstrapping, pruning etc. which totally reduce overfitting and make our models suitable to provide robust predictions [7-9].

Materials and methods. As the research material, was selected the analysis period from 2002 to 2014, all daily (reanalysis) and decadal (water level) measurements were taken to mean monthly values. Below is given a more specific description of the data used:

1. Sea level variability based on satellite altimetry data from 1992 to 2014 (decadal values), DAHITI project [10].
2. River runoff. Syrdarya (post Kazalinsk), historical monthly values 1979-1986, Global Runoff Data Center (GRDC) [11]
3. Atmospheric forcing: reanalysis Era-Interim, resolution 1x1 degree, daily values [12]

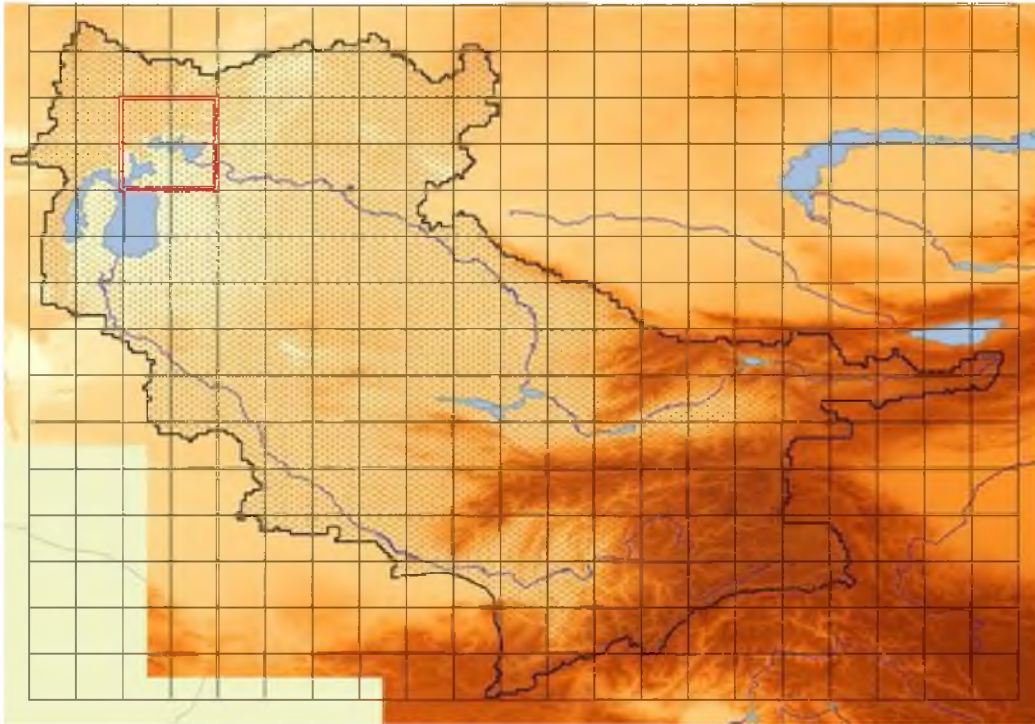


Figure 2 – Study area – the Small Aral Sea and the Aral Sea basin

A hypsometric relationship between the sea surface layer and the water volume was obtained for the entire range of sea level variability according to the DAHITI data for the period under research on the basis of a detailed bathymetric map of the Small Aral [13].

As a model linking the average monthly rates of fluctuations in climatic characteristics with the dynamics of river flow values [14-16], a regression model of solving trees was chosen, which in general is a nonparametric model of machine learning - a simple model of the "white box" solutions of which can be described by a set of simple Boolean functions. Advantages of the regression model of decision trees are: interpretability, quick learning, high tolerance for incomplete data [17]. As shortcomings, usually distinguished: high ability to retrain, instability with the dominance of one decisive class, the difficulty of finding the exact structure of the tree.

Results and discussion. On the basis of the described method, two statistical models were developed: a model for predicting the monthly values of river flow Syrdarya river and the forecast of variability of the water volume of the Small Aral. In the first model, the predictors used air and precipitation values for the current and previous six months averaged over the entire Aral Sea basin. The model was trained on the period from July 1979 to December 1985 according to the monthly values of the river flow at the Kazaly station (Figure 3). As a result, it was used to simulate river flow values for the period from January 1986 to September 2015 - for this period, the actual measurements of the river flow were not available (Fig. 4). These data, together with reanalysis data (air temperature, wind speed, precipitation) were used to train the second statistical model for predicting small-Aral volume fluctuations on satellite altimetry data for the period from October 1992 to November 2014 (Fig. 5). In conclusion, a forecast was made for the monthly values of the water volume of the Small Aral Sea in 2015 (Figure 6). Analysis of the significance of the signs showed that the parameters of mean temperature, precipitation amount and total evaporation are most significant for the results of constructing the model of decision trees, so only they were chosen as predictors of the final model - this made the model less susceptible to noise (robust), and also minimized the risk of retraining.

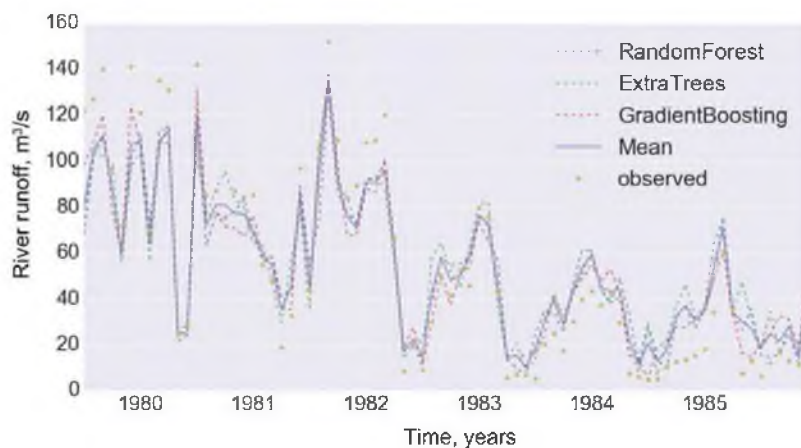


Figure 3 – Training of the model of river flow formation for the period from July 1979 to December 1985

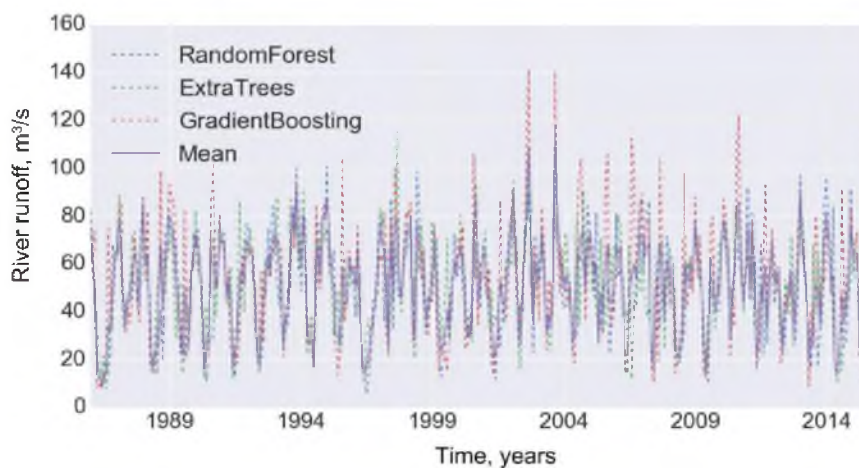


Figure 4 – Simulation of monthly river flow values for the period from January 1986 to September 2015

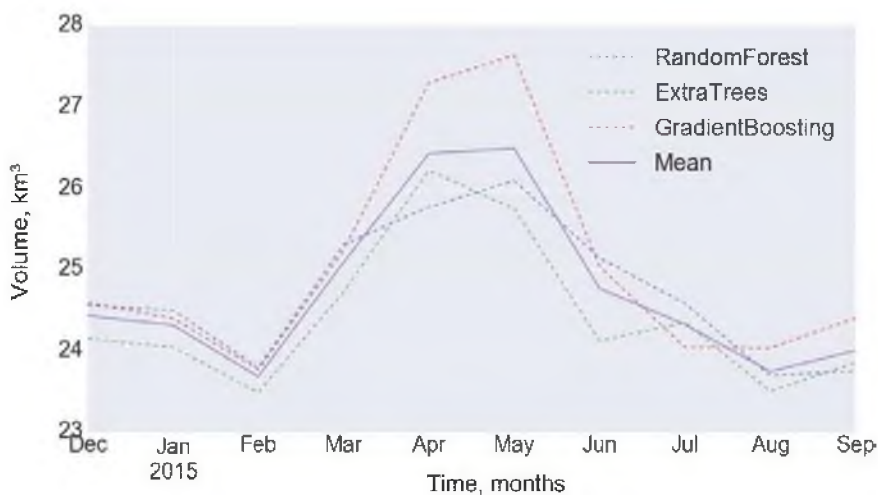


Figure 5 – Training model for forecasting the water volume of the Small Aral Sea, the period from October 1992 to November 2014

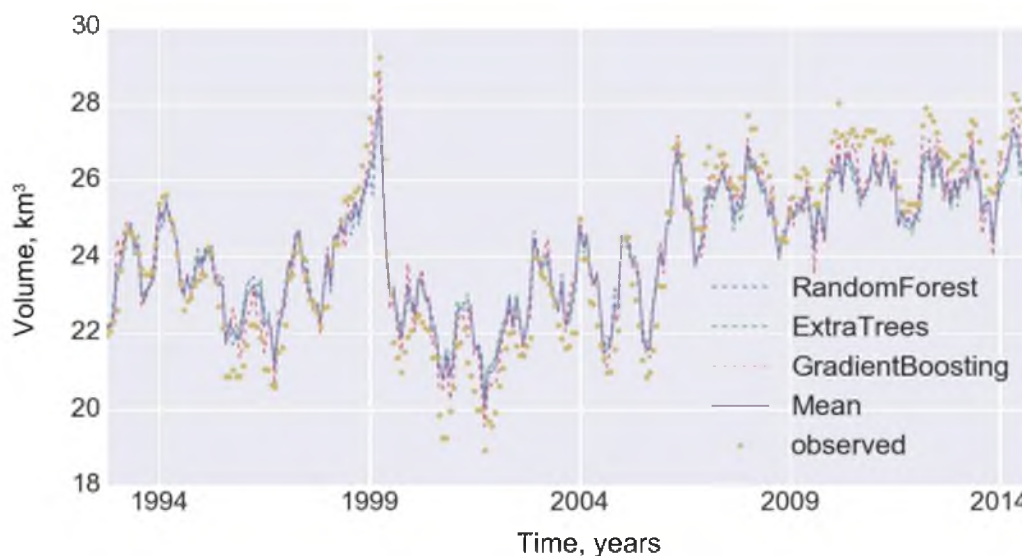


Figure 6 – Modeling of monthly values of the water volume of the Small Aral Sea in 2015

Conclusion. As a result, it should be noted that despite the serious shortage of these direct measurements of the main components of the water balance, such as the sea surface level and river runoff, the proposed method makes it possible to accurately predict the volume of waters of the Small Aral on the basis of open databases. The introduction of the factor of influence of the Kokaral dam into the model will help improve the quality of the forecast. The proposed technique can be used to obtain initial conditions in hydrodynamic simulation experiments, as well as to calculate climatic scenarios for the development of the hydrological system of the Aral Sea. Directions of further studies of the dynamics of the water volume of the Small Aral Sea are seen in the use of episodic data of field measurements for verification and more detailed adjustment of the model used, Syr Darya for the restoration of the values of natural water inflow into the Small Aral Sea.

Acknowledgment

This work was supported by the Russian Foundation for Basic Research, project 17-05-01175 A, and the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan in the framework of the scientific Project № AP05134202 «The current ecological state of endorheic lakes of the arid climate zone and perspectives of their sustainable use: the case of the Aral sea residual basins».

Г.В. Айзель¹, А.С. Ижицкий², А.К. Курбаниязов³

¹К.т.н., научный сотрудник Институт водных проблем Российской академии наук, Москва, Россия;

²К.г.н., старший научный сотрудник, Институт океанологии им. П.П. Ширшова Российской академии наук, Москва, Россия;

³К.г.н., доцент Международный Казахско-турецкий университет им. Х.А. Яссави, Туркестан

ПРОГНОЗИРОВАНИЕ СОСТОЯНИЯ МАЛОГО АРАЛЬСКОГО МОРЯ НА ОСНОВЕ ДАННЫХ ОТКРЫТЫХ ИСТОЧНИКОВ

Аннотация. Исследование динамики уровня и объема вод Аральского моря является актуальной научной задачей в силу необходимости понимания механизмов природных и антропогенных процессов, индуцировавших радикальное изменение его водного и солевого балансов за последние 60 лет. В частности, исследование динамики компонент водного баланса бассейна Малого Аральского моря является важнейшей задачей при планировании сценариев водопользования в регионе. В предлагаемой работе на основе методов машинного обучения (для внедрения вычислительных функции в программу) были разработаны две статистические модели: модель прогноза месячных значений речного стока р. Сырдария и прогноза изменчивости объема вод Малого Арала. По результатам моделирования были получены прогнозы значений

стока Сыр-Дарьи и объема вод Малого Аральского моря. В условиях малой обеспеченности данными натурных наблюдений полученные оперативные оценки компонент водного баланса являются важнейшим источником информации о происходящих изменениях в исследуемом бассейне. Предлагаемая методика также может быть использована для получения начальных условий в экспериментах гидродинамического моделирования, а также для расчета климатических сценариев развития гидрологической системы Арала.

Исследование процессов гидрологического цикла Аральского моря затруднено практически полным отсутствием данных современных измерений составляющих его водных балансов. За последние 15 лет тенденция Аральского моря к разделению водных объектов углубилась: глубоководные западные и мелководные восточные бассейны Большого Аральского моря, Малого Аральского моря и Туш-базы. Таким образом, восточный бассейн Большого Аральского моря в последние годы практически перестал быть постоянным водохранилищем, превратившись в эфемерное озеро, существование которого определяется величиной сезонного полноводного стока реки Амударьи. Северный бассейн Аральского моря (или Малый Арал) в последние годы практически утратил контакт с другими бассейнами как из-за естественных причин пересыхания Западного и Восточного бассейнов, так и из-за строительства плотины, призванной свести к минимуму водообмен между бассейнами.

В данной работе мы акцентировали внимание на исследовании динамики водного объема Малого Аральского моря. В последние годы, сведя к минимуму свой водообмен с другими частями моря, он начал устанавливать относительно стабильный водно-солевой режим, сходный по своим характеристикам с "условно-естественным" режимом Аральского моря, еще до начала высыхания. Тем не менее, столь же остро ощущается и значительный дефицит этих прямых наблюдений за компонентом водного баланса для данного региона. Современные полевые наблюдения на всех участках Аральского моря носят эпизодический характер, большинство режимных гидрометеорологических станций и станций прекратили свое существование в первой половине 1990-х годов, поэтому в данной работе мы использовали базу данных косвенных наблюдений уровня воды, а также базу данных реанализа климата. Основной целью данной работы является исследование возможности моделирования динамики объема воды бассейна, расположенного в засушливых условиях, с использованием данных косвенных измерений открытых источников.

Введение в модель фактора влияния Кокаральской плотины позволит повысить качество прогноза. Предложенная методика может быть использована для получения начальных условий в гидродинамических имитационных экспериментах, а также для расчета климатических сценариев развития гидрологической системы Аральского моря. Модель была подготовлена на период с июля 1979 года по декабрь 1985 года по месячным значениям речного стока на станции Казалы. Направления дальнейших исследований динамики водного объема Малого Аральского моря видятся в использовании эпизодических данных полевых измерений для верификации и более детальной корректировки используемой модели Сырдарьи для восстановления значений естественного притока воды в Малое Аральское море. Анализ значимости признаков показал, что параметры средней температуры, количества осадков и общего испарения являются наиболее значимыми для результатов построения модели деревьев решений, поэтому только они были выбраны в качестве предикторов конечной модели – это сделало модель менее восприимчивой к шуму (робастной), а также минимизировало риск перекалибровки.

Ключевые слова: Малый Арал, уровень моря, машинное обучение, речной сток, бассейны, Сырдарья, гидродинамическое моделирование.

Г.В. Айзель¹, А.С. Ижицкий², А.К. Курбаниязов³

¹Ресей ғылымдар академиясы Су мәселелері институтының ғылыми қызметкері, Мәскеу, Ресей;

²Ресей ғылымдар академиясы П.П. Ширшов атындағы Океанология институтының аға ғылыми қызметкері, Мәскеу, Ресей;

³Қ.А. Ясауи атындағы Халықаралық қазақ-түрік университеті,

Үздіксіз білім беру институтының директоры, г.ғ.к., доцент, Түркістан, Қазақстан

КІШІ АРАЛ ТЕҢІЗІ ЖАҒДАЙЫН АШЫҚ ДЕРЕККӨЗ НЕГІЗІНДЕ БОЛЖАУ

Аннотация. Арал теңізі суларының деңгейі мен көлемінің динамикасын зерттеу соңғы 60 жылда оның су және тұз теңгерімінің түбегейлі өзгеруін индукциялаған табиғи және антропогендік үдеріс тетіктерін түсіну қажеттігіне байланысты өзекті ғылыми міндет болып есептеледі. Атап айтқанда, шағын Арал теңізі бассейнінің су балансы компонентінің динамикасын зерттеу өңірдегі су пайдалану сценарийлерін жоспарлауда маңызды саналады. Ұсынылып отырған жұмыста машиналық оқыту әдістерінің негізінде (бағдарламаға есептеу функциясын енгізу үшін) екі статистикалық модель жасалды: Сырдария өзені

ағысының айлық мәні мен Кіші Арал суының өзгергіштігін болжау моделі. Үлгілеу нәтижелері бойынша Сырдария ағынының мәні мен Кіші Арал теңізі су көлемінің болжамы алынды. Табиғи бақылау деректерімен аз қамтамасыз етілген жағдайда алынған жедел бағалар су балансының компоненті зерттелетін бассейнде болып жатқан өзгерістер туралы ақпараттың маңызды көзі болып саналады. Ұсынылған әдістеме сондай-ақ гидродинамикалық үлгілеу эксперименттеріндегі бастапқы жағдайларды алу үшін, сондай-ақ аралдың гидрологиялық жүйесін дамытудың климаттық сценарийін есептеу үшін пайдаланылуы мүмкін.

Арал теңізінің гидрологиялық никлі үдерістерін зерттеу оның су балансын құрайтын қазіргі заманғы өлшеу деректерінің толық болмағандығынан қиындайды. Соңғы 15 жылда Арал теңізі су нысандарын бөлу беталысы тереңдеді. Үлкен Арал теңізінің, Кіші Арал теңізінің терең батыс және таяз Шығыс бассейндері және Туш-базаларын атап айтуға болады. Осылайша, үлкен Арал теңізінің шығыс бассейні соңғы жылдары эфемерлік көлге айналып, Амудария өзенінің маусымдық толық су ағынының шамасы арқылы анықталады. Арал теңізінің солтүстік бассейні (немесе Кіші Арал) соңғы жылдары Батыс және Шығыс бассейндерінің құрғауының табиғи себептеріне байланысты бассейндер арасындағы су араластыруды барынша азайтуға бағытталған бөгет құрылысы барысында басқа бассейндермен байланысты жоғалтпады.

Жұмысымызда Кіші Арал теңізінің су көлемінің динамикасын зерттеуге назар аудардық. Соңғы жылдары теңіздің басқа бөліктерімен су араластыруды барынша азайтып, Арал теңізінің «шартты-табиғи» режимімен сипаттамалары бойынша ұқсас салыстырмалы түрде тұрақты су-тұз режимін орната бастады. Дегенмен, осы өңір үшін су теңгерімінің компонентін тікелей бақылаудың айтарлықтай тапшылығы сезіледі. Арал теңізінің барлық учаскелерінде қазіргі заманғы далалық бақылаулар эпизодтық сипатта болады, режимдік гидрометеорологиялық станциялар мен станциялардың көпшілігі 1990 жылдардың бірінші жартысында тоқтады, сондықтан осы жұмыста біз су деңгейін жанама бақылау дерекқорын, сондай-ақ климат реанализінің дерекқорын пайдаландық. Жұмыстың негізгі мақсаты ашық дереккөздердің жанама өлшеу деректерін пайдалана отырып, қуаң жағдайда орналасқан бассейн су көлемінің динамикасын модельдеу мүмкіндігін зерттеу болып саналады.

Кокарал бөгетінің әсер ету факторының моделіне енгізу болжамның сапасын арттыруға мүмкіндік береді. Ұсынылған әдістеме гидродинамикалық имитациялық эксперименттерде бастапқы жағдайды алу үшін Арал теңізінің гидрологиялық жүйесін дамытудың климаттық сценарийін есептеуде пайдаланылуы мүмкін. Модель 1979 жылдың шілде айынан 1985 жылдың желтоқсан айына дейін Қазалы стансасындағы өзен ағынының айлық мәні бойынша дайындалды. Кіші Арал теңізі су көлемінің динамикасын одан әрі зерттеу бағыттары Кіші Арал теңізіне судың табиғи ағынының мәнін қалпына келтіру үшін Сырдария моделін верификациялау және неғұрлым егжей-тегжейлі түзету үшін дала өлшемдерінің эпизодтық деректерін пайдалануда көрінеді. Белгілердің маңыздылығын талдау барысы көрсеткендей, орташа температура, жауын-шашын мөлшері және жалпы булану параметрлері ағаштардың шешім моделін құру нәтижелері үшін ең маңызды саналады, сондықтан олар тек соңғы модельдің предикторы ретінде таңдалған, бұл модель шуға (робастты) азырақ бейімдеді, сондай-ақ қайта мамандану қаупін барынша азайтты.

Кілт сөздері: Кіші Арал, теңіз деңгейі, машиналық оқыту, өзен ағысы, бассейндер, Сырдарья, гидродинамикалық моделдеу.

Information about the authors:

Ayzel George Vladimirovich, Ph.D, researcher, Institute of Water Problems of the Russian Academy of Sciences, Moscow, Russia;

Izhitskiy Alexander Sergeevich, Ph.D, senior researcher, P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia;

Kurbaniyazov Abilgazy Koptileuovich, Ph.D., Associate Professor, International Kazakh-Turkish University named after H.A. Yasavi, Turkestan city, the Republic of Kazakhstan, abigazi@mail.ru

REFERENCES

[1] Zavialov, Peter O. Physical oceanography of the dying Aral Sea. Springer Science & Business Media, 2007. Zhang, H., Si, S., and Hsieh, C.-J.: GPU-acceleration for Large-scale Tree Boosting, <http://arxiv.org/abs/1706.08359>, 2017

[2] Izhitskiy A.S., Ayzel G.V., Zavialov P.O., Kurbaniyazov A.K. Estimation of the Aral Sea state predictability based on the open data sources and the unique field observations. GeophysicalResearchAbstracts. Vol.18, EGU2016-6131, 2016a.

[3] Breiman L., Friedman J., Stone C.J., Olshen R.A. Classification and regression trees. CRC press, 1984: 360 p.

[4] Hastie T., Tibshirani R., Friedman J., Franklin J. The elements of statistical learning: data mining, inference and prediction. Springer, 2005: 758 p.

[5] Freund Y., Schapire R.E. A decision-theoretic generalization of on-line learning and an application to boosting. Journ. of computer and system sciences. 1997, 55 (1): 119–139. doi: 10.1007/3-540-59119-2_166

- [6] Ayzel G.V. Use of machine learning techniques for modeling of snow depth. *Ice and Snow*. 2017;57(1):34-44. (In Russ.) DOI:10.15356/2076-6734-2017-1-34-44
- [7] Friedman J.H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*. 2001, 29(5): 1189–1232
- [8] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Intern. Joint Conf. on Artificial Intelligence (Ijcai)*. 1995, 14 (2): 1137–1145
- [9] Schapire R.E., Freund Y. *Boosting: Foundations and algorithms*. MIT press, 2012: 528 p
- [10] Schwatke, C. et al. DAHITI - an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry, *Hydrol. EarthSyst. Sci.*, 2015, 19, 4345-4364
- [11] www.bafg.de
- [12] Dee D.P. et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*(2011), 137: 553–597. doi: 10.1002/qj.828
- [13] Izhitskiy A.S. et al. Present state of the Aral Sea: diverging physical and biological characteristics of the residual basins. *Scientific Reports*, 6, 23906 (2016b) doi:10.1038/srep23906
- [14] Ayzel G.V., Izhitskiy A.S. (2016). Issledovanie dinamiki urovnyavodi Aralskogo moray podannym distancionnogo zondiro vanya I klimaticheskogo reanalaza //Vodnye resursy: izuchenye I upravlenye (limnologogi- cheskaya shkola-praktika), [Dynamics of the Aral Sea Mountaineering in the range of remote sensing and climatic reanalysis]Water resources: learning and governance (limnological school-practice). Tom 2. P.p 189-19
- [15] Mason L., Baxter J., Bartlett P. L., Frean M. R. (2000). Boosting algorithms as gradient descent. In *Advances in neural information processing systems* (pp. 512-518).
- [16] Monthly runoff data. The Global Runoff Data Centre, 56068 Koblenz, Germany. www.bafg.de
- [17] Kurbaniyazov A.K. et.al. Research of transformation of biogenous elements in water ecosystems of the southern Aral Sea area *News Of The National Academy Of Sciences Of The Republic Of Kazakhstan Series Of Agricultural Sciences* ISSN 2224-526X Volume 3, Number 51 (2019), 33 – 36 <https://doi.org/10.32014/2019.2224-526X.32>