INTELLECTUAL PROCESSING OF HYDROGEOLOGICAL DATA

(Using the example of groundwater)

Z. B. Dadajonova [3],
B. D. Abdullaev [1,2],
Kh. N. Zaynidinov [3],
B. B. Akbaraliev [4],
Nasibov B. R. [2]
[1] State Institution "Institute of Hydrogeology and Engineering Geology", City of Tashkent, Republic of Uzbekistan.
[2] National Research University "Tashkent Institute of Irrigation and Agricultural Mechanization Engineers"
[3] Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, city of Tashkent, Republic of Uzbekistan.
[4] Andijan State University of the Republic of Uzbekistan, City of Andijan, Republic of Uzbekistan.

Abstract:
This comprehensive article delves into the development of an automated information reference hydrogeochemical data bank through the systematization and typification of hydrogeological data using intelligent processing methods and algorithms. It scrutinizes various methods of intelligent analysis of hydrogeological data, emphasizing the importance of understanding general patterns in groundwater chemical composition and predicting changes influenced by anthropogenic factors. The study also focuses on enhancing data presentation to reflect the hydrogeochemical landscape of Uzbekistan. By utilizing data mining techniques and mathematical algorithms, the research aims to delineate geological bodies based on groundwater properties and contamination levels. The article highlights the significance of visualizing results and utilizing data mining tools for decision-making processes. The application of Fisher-type criteria and proximity functions in data analysis further enhances the study's depth and accuracy.

Keywords. Data Mining, integration, interpretation, Garbage In Garbage Out, data cleaning, Intelligent data analysis system, Intelligent data analysis, hydrogeological map, groundwater deposits, formation of optimal information systems.

Introduction

At present, as a result of previously performed hydrogeological studies on the territory of the Republic, a huge amount of factual material on the chemical composition of groundwater has been accumulated. This information is supplemented with new data by annual hydrogeological studies [1-3].

The main task of processing chemical data of groundwater is both the study of general patterns of formation of chemical composition of groundwater and forecasting its change under the influence of anthropogenic factors, and the introduction of new

Journal Zone Publishing, Ilford, United Kingdom

methods in their processing [4-6]. An equally important task is the development of forms of presentation of the obtained data that would most fully reflect the hydrogeochemical situation of the underground hydrosphere of the territory of the Republic of Uzbekistan [7-8].

In recent years, a great deal of work has been carried out in the republic to study changes in the hydrogeochemical parameters of groundwater, which included the processing of retrospective information, conducting area hydrogeochemical testing for an expanded composition of components with the use of modern laboratory research [9-12].

Understanding the hydrologic status of the environment and groundwater dynamics requires that the datasets and measurements for modeling be accessible to scientists, planners, and stakeholders to ensure informed decision-making [13-14].

Remote sensing can reveal information in remote areas that cannot be observed on the ground. Optical remote sensing has no or very limited ground penetration capabilities from a few centimeters to several meters above the surface, as it consists of the visible and thermal domains 15-18]. Acoustic signal processing spans from 20 Hz to 20 kHz and exhibits varying responses when encountering different materials in diverse environments. The audible spectrum offers insights into acoustic properties for identifying optimal soundproofing solutions, noise pollution assessment, acoustic modeling, and related factors. The ultrasonic domain focuses on material properties. is linked, offering crucial insights into material defects that are readily identifiable in ultrasonic inspection data [15,19-20].

Research Methodology

Usually, the following typical stages accompanying the solution of data mining tasks are distinguished:

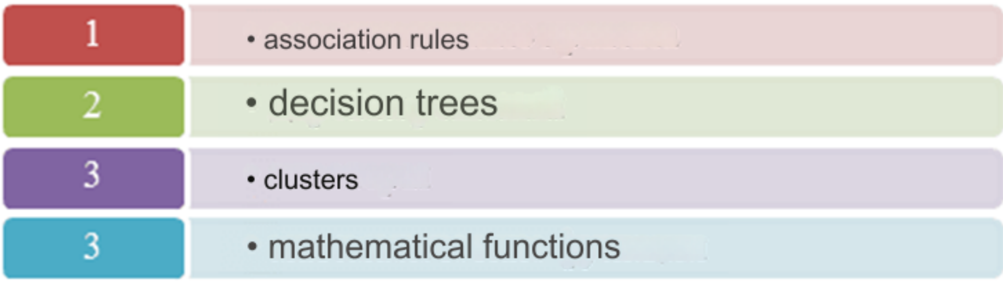
1. Analysis of the subject area, formulation of the goals and objectives of the study.
2. Extraction and storage of data.
3. Preliminary data processing:
4. Substantive analysis of data using Data Mining methods (establishing general patterns or solving more specific, private problems).
5. Interpretation of the obtained results by presenting them in a convenient format (visualization and selection of useful patterns, formation of informative graphs and/or tables).
6. Using new knowledge for decision-making.

The basis of data mining methods is made up of all kinds of classification, modeling and forecasting methods. One of the most important purposes of data mining methods is the visual presentation of calculation results, which allows the use of data mining tools by people without special mathematical training.

Knowledge obtained using data mining methods is usually presented in the form of models.

Journal Zone Publishing, Ilford, United Kingdom

British Journal of Global Ecology and Sustainable Development
Volume-32, September- 2024
ISSN (E): 2754-9291



Most analytical methods used in the DM technology are well-known mathematical algorithms and methods. What is new in their application is the possibility of using them to solve certain specific problems, due to the emerging capabilities of technical and software tools. It should be noted that most DM methods were developed within the framework of artificial intelligence theory [21-13]

Analysis and Results

In the process of constructing hydrogeological maps that require quantitative justification with objective identification of the boundaries of geological bodies, it is advisable to use geological and mathematical methods. Using the latter, it is easy to prove whether the identified geological bodies differ significantly in terms of the properties of groundwater and whether a more detailed division should be made. At the same time, the quantitative nature of the information makes it possible to resolve the problem of a justified and formalized delineation of geological bodies.

The solution to the problem of selecting the most informative indicators (signs) and classification by hydrochemical indicators of groundwater contain sufficient relevant meaning for further differentiation of hydrogeological bodies.

Below are the results of determining the levels of contamination of groundwater deposits in relation to mineralization. Here, groundwater deposits are divided into 6 classes according to the degree of salinity, that is.

$$X = \bigcup_{p=1}^{6} X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1,6},$$

Where

$$X\_1 - \ not\ salted, m\_1 = card(X\_1\ ) = 749;$$
$$X\_2 - \ slightly\ salted, m\_2 = card(X\_2\ ) = 294;$$
$$X\_3 - \ moderately\ salted, m\_3 = card(X\_3\ ) = 178;$$
$$X\_4 - \ salted\ (saline), m\_4 = card(X\_4\ ) = 77;$$
$$X\_5 - \ highly\ salted, m\_5 = card(X\_5\ ) = 151;$$
$$X\_6 - \ very\ highly\ salted, m\_6 = card(X\_6\ ) = 36.$$
$$m_i is\ the\ number\ of\ objects\ in\ class\ i.$$

The initial data are presented in the form of a table containing the values of 9 indicators, i.e. $x_{ij} = (x_{ij}^1, x_{ij}^2, \dots, x_{ij}^9)$, which constitute an a priori dictionary of features that characterize 1485 studied hydrogeological objects.

The indicators of hydrogeological objects have the following meaning, that is. $\forall x = (x^1, x^2, \dots, x^9) \in X$

Journal Zone Publishing, Ilford, United Kingdom

x^1 – dry residue;
x^2 – bicarbonate (HCO_3);
x^3 – carbonate(CO_3);
x^4 – nitrate (NO_3);
x^5 – chloride (Cl);
x^6 – sulfate (SO_4);
x^7 – calcium (Ca);
x^8 – magnesium (Mg);
x^9 – sodium-potassium $(Na + K)$.

When solving the problem, the Fisher-type criterion was used as a measure of the informativeness of indicators or a set of indicators, that is.

$$\text{I.}\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} \to max \\ \lambda \in \Lambda^{\ell} \end{cases}, \qquad \text{II.}\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} \to max \\ \lambda \in \Lambda^{\ell} \\ (c, \lambda) \le c_0 \end{cases}.$$

where I – for the formation of ℓ information systems for describing objects in conditions of unlimited resources; II – for the formation of ℓ information systems for describing objects in conditions of limited resources.

Table 2. After normalization, the training sample has the following form:

	x^1	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9
$\bar{x}_1$	0,17	0,64	0,04	0,09	0,25	1,37	1,38	0,54	0,51
$S^2(X_1)$	0,05	0,05	0,01	0,08	0,12	6,80	5,30	0,59	0,56
$\bar{x}_2$	0,4	0,69	0,12	0,13	0,45	5,02	4,12	0,97	1,32
$S^2(X_2)$	0,23	0,07	1,69	0,17	0,37	51,18	37,3	2,71	0,98
$\bar{x}_3$	0,62	0,57	0,11	0,24	1,25	7,60	5,35	1,64	4,47
$S^2(X_3)$	0,25	0,08	1,00	1,48	3,17	48,66	31,66	10,12	378,53
$\bar{x}_4$	2,39	0,60	0,00	0,22	17,47	13,79	8,06	4,07	19,29
$S^2(X_4)$	7,60	0,02	0,00	3,32	83,26	74,29	59,06	8,94	103,51
$\bar{x}_5$	1,28	0,41	0,03	0,29	2,20	17,76	13,31	2,28	4,68
$S^2(X_5)$	0,45	0,04	0,01	1,71	20,45	88,77	76,78	3,46	23,20
$\bar{x}_6$	1,69	0,43	0,01	1,98	10,86	20,27	14,88	2,46	15,25
$S^2(X_6)$	0,81	0,09	0,00	14,11	284,39	106,88	53,42	4,25	285,52
C	8	4	6	5	5	7	8	6	7

where x ̅_i is the average object of class i, S^2 (X_i) is the average scatter in class X_i (i=(1,9) ̅), c=(c_1,c_2,…,c_9) is the resource vector, c_0=34 is the total resource allocated for measuring the feature.

Then we have

Journal Zone Publishing, Ilford, United Kingdom

№	1	2	3	4	5	6	7	8	9
a	21,78	0,39	0,08	16,11	1523,9	1676,8	847,8	47,38	1819,24
b	9,39	0,35	2,71	20,87	391,76	376,58	263,5	30,07	792,3

To solve the problem of forming optimal information systems (FOIS) when describing objects, we will use the vector $\lambda: X \to X|_\lambda = \{x|\ x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$, that is.
Where $\lambda^j \in \{0,1\}$ $(j = \overline{1,N})$.
If λ^j=1, then the -th feature will participate in the FOIS, otherwise it will not participate.
1. If X⊂R^N, then the proximity function between objects is defined as the distance using the Euclidean metric.
2. If the features of the training sample object X take a value from the set {0,1} or {1,2,...,K}, then the proximity function (ρ) and the distance (d) between ∀x,y∈X objects are defined as follows:

a) $$\rho(x^j, y^j) = \begin{cases} 1, If\ x^j = y^j \\ 0, If\ x^j \neq y^j \end{cases} \quad (2)$$

b) let ε_1,ε_2,... ,ε_N be positive numbers.

$$\rho(x^j, y^j) = \begin{cases} 1, If\ |x^j - y^j| \leq \varepsilon_j \\ 0, If\ |x^j - y^j| > \varepsilon_j \end{cases} \quad (3)$$

$$\rho(x, y) = \sum_{j=1}^{N} \rho(x^j, y^j) \quad (4)$$

$$d(x, y) = N - \rho(x, y) \quad (5)$$

Table 3. Results of the FOIS

ℓ	Without taking into account the resource	$\max_{\lambda\in\Lambda^\ell} I(\lambda)$	Taking into account the resource.	(c, λ)	$\max_{\lambda\in\Lambda^\ell} I(\lambda)$
4	1,2,3,6	4,3674	1,2,3,6	25	4,3674
5	1,2,3,4,6	4,1843	1,2,3,4,6	30	4,1843
6	1,2,3,4,5,6	4,0404	2,3,4,5,6,8	33	3,97
7	1,2,3,4,5,6,8	3,9513	There is no solution	$c > c_0$	

The table below shows the classification results for the selected set of features.

Table 4. Classification results

ℓ	$\Re_{k-yq}(y, Y_p)$	$\Re_{BHA}(y, Y_p)$	$\Re_{XD}(y, Y_p)$	$\Re_{NT}(y, Y_p)$	$\Re_{Im}(y, Y_p)$
4	75,8	80,9	79,5	83,2	84,7
5	80	87,3	87,3	89,5	90
6	85,4	88,9	88	92	94

To solve the classification problem, widely known algorithms were used – decision rule algorithms, where $\Re_{k-yq}(y, Y_p)$, $\Re_{BHA}(y, Y_p)$, $\Re_{XD}(y, Y_p)$, $\Re_{NT}(y, Y_p)$), Impuls ($\Re_{Im}(y, Y_p)$) - respectively k-nearest neighbors , algorithms for calculating estimates, decision trees, neural networks and momentum.

Journal Zone Publishing, Ilford, United Kingdom

Conclusion

One of the most valuable technologies for hydrogeological analysis is GIS technology. In the implementation of the analysis, the main stages, such as data integration, spatial analysis, modeling, visualization and monitoring, ensure that the results are maximally positive. GIS should be used to integrate different types of spatial data such as geological maps, hydrological data, land use data and groundwater monitoring data through data integration. GIS tools for spatial analysis should be used to identify relationships and trends within hydrogeological data through spatial analysis. It helps to understand the spatial distribution of groundwater resources, sources of contamination and potential risks. How does modeling help us? It enables the development of hydrogeological models to simulate groundwater flow, contaminant transport, and aquifer properties using GIS software. These models help predict groundwater behavior under various scenarios and aid in decision-making processes. And by visualization, we use GIS to create visual representations of hydrogeological data through maps, diagrams, and 3D models. Visualization enables us to effectively communicate complex information to stakeholders and decision makers. We achieve monitoring results by implementing GIS to track groundwater levels, water quality parameters, and land use changes over time.
The lack of systematized information on the study of the territory complicates the process of searching for the full range of necessary materials for forecasting the hydrogeochemical situation.
In order to identify patterns and predict changes in the quality of groundwater, there is a need to create an automated information reference hydrogeochemical data bank based on their systematization and typification.

References

1. Abdullaev B.D. Zokirova A. Geoaxborot tizimlari texnologiyalaridan foydalangan holda gidrogeologik xaritalar tuzish// So`ngi ilmiy tadqiqotlar nazariyasi respublika ilmiy-uslubiy jurnali. 9-son. 13.04.2023
2. Абдуллаев Б.Д., Андреев Д,Н., Умарова З.М. Мониторинг подземных вод Республики Узбекистан: достижение, проблемы и пути их решения// Материалы Международной научной конференции, посвященной 15-летию со дня образования ЦАИИЗ Дистанционные и наземные исследования земли в Центральной Азии Бишкек, Кыргызстан, 17-18 сентября 2019 с.145-148
3. Dadajonovich, A. B., Temirkhojaevich, S. A., & Ishkulovich, R. R. (2024). MONITORING OF GROUNDWATER QUALITY CHANGES IN IRRIGATED LANDS OF KASHKADARYA REGION. British Journal of Global Ecology and Sustainable Development, 27, 11-21.
4. Rushton, K. R. (2004). Groundwater hydrology: conceptual and computational models. John Wiley & Sons.
5. Аюпов, В.В. Математическое моделирование технических систем: учебное пособие / М-во с.-х. РФ, федеральное гос. бюджетное образов. учреждение

Journal Zone Publishing, Ilford, United Kingdom

высшего образования «Пермская гос. с.-х. акад. им. акад. Д.Н. Прянишникова». – Пермь : ИПЦ «Прокростъ», 2017. – 242 с.
6. Веселов В.В., Паничкин В.Ю., Мирошниченко О.Л. Решение обратных задач геофильтрации на математических моделях гидрогеологических объектов // Геология Казахстана. Сборник трудов, посвященный XXXII сессии международного геологического конгресса (Флоренция, Италия 2004). МОН РК, Алматы, 2004, С. 417-426
7. Джуманов Ж.Х. Геоинформационные технологии в гидрогеологии. / Ж.Х.Джуманов; - Т.: ГП «Институт ГИДРОИНГЕО», 2016. - 258 с.
8. Емельянова Т.Я. Инженерно-геологические изыскания. 2002г.
9. Мамаев С.А., Ибаев Ж.Г., Гусейнова А.Ш., Курбанисмаилова А.С., Мамаев А.С. Геоинформационные базы данных гидрогеологических параметров артезианских скважин Терскокумского артезианского бассейна// Геология и геофизика Юга России, № 3, 2018 с. 51-63.
10. Мусин Р.Х. Гидрогеологическая карта и объяснительная записка к ней: учебно-методическое пособие /Казань: Изд-во Казан. ун-та, 2020. –38 с.
11. Николаева Н.К. «Создание базы гидрогеологических данных в системе GeODIN». АНО УКЦ «Изыскатель». 17 стр.
12. Нишанов А.Х., Акбаралиев Б.Б., Холиқназаров А.Х., Ахмедов О., Тажибаев Ш. Ер ости сув ҳолатларини аниқлашда муҳим белгиларни танлаш алгоритми//International scientific online conference «Actual problems and prospectes of the development of intelligent information and communication systems», IICS-2020, Tashkent, 2020, 181-185-б.
13. Klug, H., & Kmoch, A. (2014). A SMART groundwater portal: An OGC web services orchestration framework for hydrology to improve data access and visualisation in New Zealand. Computers & Geosciences, 69, 78-86.
14. Kresic, N., & Mikszewski, A. (2012). Hydrogeological conceptual site models: data analysis and visualization. CRC press.
15. Jasurbek, S. (2022). MASOFADAN ZONDLASH, GEOGRAFIK AXBOROT TIZIMLARI VA GLOBAL JOYLASHISHNI ANIQLASH TIZIMI TEXNIKASINI GIDROLOGIK MODELLASHTIRISH BILAN INTEGRATSIYALASH. Yosh Tadqiqotchi Jurnali, 1(4), 77-83.
16. Condon, L. E., Kollet, S., Bierkens, M. F., Fogg, G. E., Maxwell, R. M., Hill, M. C., ... & Abesser, C. (2021). Global groundwater modeling and monitoring: Opportunities and challenges. Water Resources Research, 57(12), e2020WR029500.
17. Asher, M. J., Croke, B. F., Jakeman, A. J., & Peeters, L. J. (2015). A review of surrogate models and their application to groundwater modeling. Water Resources Research, 51(8), 5957-5973.
18. Нишанов А.Х., Акбаралиев Б.Б., Холиқназаров А.Х., Ахмедов О., Тажибаев Ш. Ер ости сувлари давлат мониторинг объектларини интеллектуал таҳлил қилиш// International scientific online conference «Actual problems and

Journal Zone Publishing, Ilford, United Kingdom

prospectes of the development of intelligent information and communication systems», IICS-2020, Tashkent, 2020, 185-190-б.
19. H Zaynidinov, J Nurmurodov, S Qobilov. Application of Machine Learning Methods for Signal Processing in Piecewise-Polynomial Bases//2023 IX International Conference on Information Technology and Nanotechnology (ITNT), IEEE, April 2023, pp. 1-4.
20. Zaynidinov, H. , Mallayev, O. Parallel Algorithm for Calculating the Learning Processes of an Artificial Neural Network// AIP Conference, 2022, 2647, 050006.
21. Zaynidinov, H.N. , Singh, D. , Yusupov, I. , Makhmudjanov, S.U. Algorithms and Service for Digital Processing of Two-Dimensional Geophysical Fields Using Octave Method// IOT, Electronics and Mechatronics Conference (IEMTRONICS), IEEE International, 2022, 13184 LNCS, pp. 460–470.
22. Шерфединов Л.З., Абдуллаев Б.Д., Эколого-экономическая оценка ущерба от загрязнения нефтепродуктами месторождений пресных подземных вод/ /Геология и минеральные ресурсы. - 2017. - № 4. - С. 33-34.
23. Шерфединов Л.З., Абдуллаев Б.Д., Мавлонов А.А., Геохимическая структура потоков грунтовых (трещиных) вод// Доклады Академии Наук Республики Узбекистан. - 2014. - № 2. – С. 33-35.
24. Шерфединов Л.З., Абдуллаев Б.Д., Мавлонов А.А., Систематизация гидрогеологических объектов по уровням организации// Геология и минеральные ресурсы. - 2013. - № 5. - С. 61-63. (04.00.00; № 2).
25. Hakimjon Zaynidinov, Sanjarbek Ibragimov, Botirjon Abdullayev, Shavkat Abdullayev, Zilola Dadajonova, Igor Motorniy, "Creation of A Three-Dimensional Model of The Underground Layer Using Digital Processing of Signals Obtained From Field Geophysical Research", International Conference on Information Science and Communications Technologies (ICISCT2023), Tashkent, Uzbekistan, 29-30 September 2023.
26. B.D. Abdullaev,Q.R. Mingboyev, Z.M. Umarova, Z.B. Dadajonova, B.B. Dadajonov, Factors affecting the hydrodynamic state of the okh groundwater deposit//World Bulletin of Social Sciences (WBSS) Available Online at: https://www.scholarexpress.net Vol. 29, December 2023 ISSN: 2749-361X.
27. Odilovich, K. A. (2024). PROBLEMS IN USING THE TRANSBORDER RIVER. Web of Agriculture: Journal of Agriculture and Biological Sciences, 2(3), 1-8.
28. Nasibov, B. R., & Abdullaev, B. D. (2023). IMPACT OF CLIMATE CHANGE ON GROUNDWATER RESOURCES. Ethiopian International Journal of Multidisciplinary Research, 10(11), 441-449.
29. Abdullaev, B. D., Razzakov, R. I., Okhunov, F. A., & Nasibov, B. R. (2023). Modeling of hydrogeological processes in irrigation areas based on modern programs. In E3S Web of Conferences (Vol. 401, p. 02006). EDP Sciences.
30. Абдуллаев Б.Д. Нефтяное загрязнение подземных вод: проблемы и пути их решения.// Министерство горно-добывающей промышленности и геологии

Journal Zone Publishing, Ilford, United Kingdom

РУз, Университет геологических наук, ГУ «Институт ГИДРОИНГЕО». - Т. ГУ «ИМР», 2023. – 109 с.

31. Усманов Р.Н., Абдуллаев Б.Д., Мавлонов А.А. Интеграция гибридных нейронных сетей для исследования гидрохимического режима водозаборов подземных вод// Геология и минеральные ресурсы. - 2011. - № 5. - С. 39-43.
32. Усманов Р.Н., Абдуллаев Б.Д., Мавлонов А.А. Интеграция гибридных нейронных сетей для исследования гидрохимического режима водозаборов подземных вод// Геология и минеральные ресурсы. - 2011. - № 5. - С. 39-43.
33. Абдуллаев Б.Д., Мавлонов А.А., ШерефединовЛ.З. Актуализация видения месторождений подземных вод// Геология и минеральные ресурсы. - 2012. - № 2. - С. 52-55.
34. Абдуллаев Б.Д., Мавлонов А.А., ШерефединовЛ.З. Систематизация гидрогеологических объектов по уровням организации// Геология и минеральные ресурсы. - 2013. - № 5. - С. 61-63.
35. Абдуллаев Б.Д., Мавлонов А.А., Шерефединов Л.З. Геоэкологические системы Узбекистана// Геология и минеральные ресурсы. - 2014. - № 1. - С. 55-59.
36. Абдуллаев Б.Д., Мавлонов А.А., Шерефединов Л.З. Геохимическая структура потоков грунтовых (трещиных) вод// Доклады Академии Наук Республики Узбекистан. - 2014. - № 2. – С. 33-35.
37. Абдуллаев Б.Д, Саидова С.А. Мониторинг подземных вод – основной фактор сохранения ресурсов// Геология и минеральные ресурсы -2018.- № 1.–С. 46-47.
38. Абдуллаев Б.Д., Отакулова Ю.А. Основные природные факторы, формирующие состав подземных вод Приташкентского региона// Вопросы географии и геоэкологии, Алматы-2019. -№1- С. 63-67.
39. Абдуллаев Б.Д., Мягков С.В. Динамика поверхностных и подземных вод в бассейне Сырдарьи при изменении климата// Геология и минеральные ресурсы. - 2019. - №6. – С. 50-52
40. Абдуллаев Б.Д., Холмирзаев М.Ж, Абдуллаев Бм.Д.Гидродинамические и гидрохимические факторы формирования естественных ресурсов грунтовых вод реки Чирчик// Республиканский научно теоретический журнал «Наука, новые технологии и инновации Кыргызстана», №6, 2019.-С. 6-10.