

УДК 622.276  
МРНТИ 52.47.15

DOI: <https://doi.org/10.54859/kjogi108677>

Получена: 02.10.2023.

Одобрена: 23.02.2024.

Опубликована: 31.03.2024.

## Оригинальное исследование

### Автоматический подбор зон для бурения нагнетательных скважин-кандидатов

**А.А. Бекен, А.Е. Ибраев, Ж.Т. Жетруов, А.С. Елемесов, А.Т. Жолдыбаева**

*КМГ Инжиниринг, г. Астана, Казахстан*

#### АННОТАЦИЯ

**Обоснование.** Возрастающая сложность в поиске точек для бурения нагнетательных скважин на поздних этапах разработки месторождений АО НК «КазМунайГаз», обусловленная уплотнением сетки действующих скважин и неравномерностью выработки запасов, является на сегодняшний день актуальной проблемой. Разработки в области геопространственного анализа и искусственного интеллекта послужили стимулом к поиску новых подходов для решения этой задачи.

**Цель.** Исследование направлено на разработку инновационного подхода к автоматическому определению наиболее перспективных зон для бурения нагнетательных скважин, основанного на комплексном анализе больших объемов данных с применением продвинутого алгоритмов.

**Материалы и методы.** В работе используются методы сбора и анализа производственных и геологических данных, применяются пространственные алгоритмы для мультифакторного анализа и методы нормализации данных, включая скорректированный межквартильный диапазон для определения выбросов.

**Результаты.** Описываются результаты, демонстрирующие ранжирование ячеек по потенциалу для бурения на основе комплексного анализа, а также присвоение уникальных кодов каждой ячейке для улучшения точности принятия решений.

**Заключение.** Заключительная часть подводит итоги, подчеркивая важность интеграции различных методов и подходов. Отмечаются направления для дальнейших исследований, включая анализ неточностей данных, учет дополнительных параметров, выявление эффективных пропластков, применение методов машинного обучения и расширение тестирования подхода на других месторождениях.

**Ключевые слова:** нагнетательные скважины, подбор скважин-кандидатов, сетка разработки, ячейки, первый радиус скважин, гидродинамические исследования скважин.

#### Как цитировать:

Бекен А.А., Ибраев А.Е., Жетруов Ж.Т., Елемесов А.С., Жолдыбаева А.Т. Автоматический подбор зон для бурения нагнетательных скважин-кандидатов // *Вестник нефтегазовой отрасли Казахстана*. 2024. Том 6, №1. С. 74–86. DOI: <https://doi.org/10.54859/kjogi108677>.

UDC 622.276  
CSCSTI 52.47.15

DOI: <https://doi.org/10.54859/kjogi108677>

Received: 02.10.2023.

Accepted: 23.02.2024.

Published: 31.03.2024.

---

## Original article

### Automatic selection of sites for drilling candidate injection wells

**Aidana A. Beken, Aktan Ye. Ibrayev, Zhassulan T. Zhetruov, Azamat S. Yelemessov, Assel T. Zholdybayeva**

*KMG Engineering, Astana, Kazakhstan*

#### ABSTRACT

**Background:** The increasing difficulty in finding sites for drilling injection wells at the later stages of field development by NC “KazMunayGas” JSC, due to infill drilling of the grid of existing wells and uneven reserve production, is a pressing problem today. Developments in geospatial analysis and artificial intelligence have stimulated the search for new approaches to solve this problem.

**Aim:** The research is aimed at developing an innovative approach to automatically identifying the most promising sites for drilling injection wells, based on comprehensive analysis of large volumes of data using advanced algorithms.

**Materials and methods:** The work uses methods for collecting and analyzing production and geological data, uses spatial algorithms for multivariate analysis and data normalization methods, including the adjusted interquartile range to determine outliers.

**Results:** Results are described showing the ranking of cells by drilling potential based on comprehensive analysis, as well as the assignment of unique codes to each cell to improve decision-making accuracy.

**Conclusion:** Directions for further research are noted, including analysis of data inaccuracies, consideration of additional parameters, identification of effective interlayers, application of machine learning methods, and expansion of testing of the approach in other fields.

**Keywords:** *injection wells; selection of candidate wells; well spacing; cells; first radius of wells; hydrodynamic studies of wells.*

#### To cite this article:

Beken AA, Ibrayev AY, Zhetruov ZT, Yelemessov AS, Zholdybayeva A.T. Automatic selection of sites for drilling candidate injection wells. *Kazakhstan journal for oil & gas industry*. 2024;6(1):74–86. DOI: <https://doi.org/10.54859/kjogi108677>.

ӨОЖ 622.276  
FTAХР 52.47.15

DOI: <https://doi.org/10.54859/kjogi108677>

Қабылданды: 02.10.2023.

Мақұлданды: 23.02.2024.

Жарияланды: 31.03.2024.

## Түпнұсқа зерттеу

### Айдау ұңғымалар-кандидаттарын бұрғылауға арналған аймақтарды автоматты түрде таңдау

**А.А. Бекен, А.Е. Ибраев, Ж.Т. Жетруов, А.С. Елемесов, А.Т. Жолдыбаева**  
*ҚМГ Инжиниринг, Астана қаласы, Қазақстан*

#### АННОТАЦИЯ

**Негіздеу.** «ҚазМұнайГаз» ҰК АҚ кен орындарын игерудің соңғы кезеңдерінде айдау ұңғымаларын бұрғылау үшін нүктелерді іздеудегі күрделіліктің артуы, жұмыс істеп тұрған ұңғымалар торының тығыздалуына және қорлардың біркелкі өндірілмеуіне байланысты бүгінгі таңда өзекті мәселе болып табылады. Геокеңістіктік талдау және жасанды интеллект саласындағы әзірлемелер бұл мәселені шешудің жаңа тәсілдерін іздеуге түрткі болды.

**Мақсаты.** Зерттеу озық алгоритмдерді қолдана отырып, үлкен көлемдегі деректерді кешенді талдауға негізделген айдау ұңғымаларын бұрғылауға арналған ең перспективалы аймақтарды автоматты түрде анықтаудың инновациялық тәсілін жасауға бағытталған.

**Материалдар мен әдістер.** Жұмыста өндірістік және геологиялық деректерді жинау және талдау әдістері қолданылады, мультифакторлық талдау үшін кеңістіктік алгоритмдер және шығаруларды анықтау үшін түзетілген квартильаралық диапазонды қоса алғанда, деректерді қалыпқа келтіру әдістері қолданылады.

**Нәтижелері.** Кешенді талдау негізінде бұрғылау потенциалы бойынша ұяшықтардың рейтингін көрсететін нәтижелер, сондай-ақ шешім қабылдау дәлдігін жақсарту үшін әр ұяшыққа бірегей кодтар тағайындау сипатталған.

**Қорытынды.** Қорытынды бөлім әртүрлі әдістер мен тәсілдерді біріктірудің маңыздылығын атап өтіп, қорытындылайды. Деректердің дәлсіздіктерін талдауды, қосымша параметрлерді, тиімді пропласттарды анықтауды, машиналық оқыту әдістерін қолдануды және басқа кен орындарында тәсілді тестілеуді кеңейтуді қоса алғанда, одан әрі зерттеулерге арналған бағыттар белгіленеді.

**Негізгі сөздер:** айдау ұңғымалары, ұңғыма-кандидаттарды іріктеу, игеру торы, ұяшықтар, ұңғымалардың бірінші радиусы, ұңғымаларды гидродинамикалық зерттеу.

#### Дәйексөз келтіру үшін:

Бекен А.А., Ибраев А.Е., Жетруов Ж.Т., Елемесов А.С., Жолдыбаева А.Т. Айдау ұңғымалар-кандидаттарын бұрғылауға арналған аймақтарды автоматты түрде таңдау // *Қазақстанның мұнай-газ саласының хабаршысы*. 2024. 6 том, №1, 74–86 б. DOI: <https://doi.org/10.54859/kjogi108677>.

## Введение

В рамках данной статьи акцентируется внимание на актуальности усовершенствования процесса подбора нагнетательных скважин для месторождений АО НК «КазМунайГаз». Особый интерес вызывает ситуация, когда месторождение достигает более поздних стадий разработки. В этом контексте осуществление точного и эффективного подбора приобретает повышенное значение. Недостаточно оптимизированные стратегии могут негативно повлиять на долгосрочную стабильность и продуктивность добычных операций.

В последние годы наука геопространственного анализа и искусственный интеллект достигли значительных успехов. Это открыло путь к методам автоматизации алгоритмов принятия решений, которые могут кардинально изменить процесс определения перспективных мест для бурения нагнетательных скважин. Новые подходы с использованием больших объемов данных и сложных алгоритмов предоставляют уникальную возможность для оптимизации размещения скважин и раскрытия неиспользованного потенциала коллекторов. Ученые-геологи и инженеры-разработчики активно изучают применение алгоритмов геопространственного анализа и машинного обучения для повышения скорости и эффективности принятия решений о размещении скважин.

Исследования показали успешность применения геопространственных методов в определении зон с высокими значениями фильтрационно-емкостных свойств и близких к существующим скважинам, что помогает в первоначальном отборе потенциальных мест для бурения [1, 2].

Алгоритмы машинного обучения продемонстрировали свою способность прогнозировать поведение коллектора, оптимизировать варианты закачки и адаптироваться к изменяющимся условиям пласта. Исследователи использовали модели обучения с учителем, такие как опорные векторы и нейронные сети, для выявления закономерностей в исторических данных закачки и прогнозирования производительности нагнетательных скважин в различных сценариях пласта [3, 4].

В литературе [5, 6] подчеркивается важность усвоения данных в режиме реального времени для автоматизации процесса размещения скважин. Непрерывный мониторинг и интеграция данных в гидродинамические модели коллектора позволяют своевременно вносить коррективы и оперативно реагировать на изменения в поведении

пласта, что повышает эффективность систем заводнения.

Тем не менее, несмотря на значительный прогресс, остаются вопросы, требующие дополнительного изучения. Обнаружение неточных и зашумленных данных, интеграция множества критериев и целей, а также разработка надежных и интерпретируемых моделей машинного обучения остаются актуальными областями исследований.

Проблемы, описанные выше, а именно неточность, малое количество и зашумленность данных, не позволяют рассчитывать надёжные прогнозные показатели с помощью комплексных моделей, поэтому было принято решение воспользоваться методом ранжирования. Критерием ранжирования кандидатов является произведение безразмерных параметров, характеризующих текущее состояние разработки пластов в предполагаемом районе заложения нагнетательной скважины.

## Этапы исследования

Проведенное исследование можно разделить на три основных этапа:

1. **Сбор данных.** Первоначальный этап включает в себя сбор и фильтрацию необходимых данных. Качество и полнота входных данных являются чрезвычайно важными для обеспечения точности и надёжности последующих анализов.

2. **Разбивка месторождения на ячейки.** Второй этап заключается в разделении всей площади месторождения на участки заданного размера. На данном этапе определяется оптимальное размещение проектных точек с учётом текущего фонда скважин.

3. **Оценка ячейки.** Третий этап предполагает детальную оценку ячеек, выделенных на предыдущем этапе, и сравнительный анализ проектных точек на основе рассчитанного комплексного параметра, характеризующего приоритетность бурения.

## Сбор данных

В рамках данной статьи рассматривалось месторождение X, которое было введено в промышленную разработку в 1965 г.

В контексте юрско-мелового разреза на месторождении выделено 25 горизонтов, характеризующихся преимущественно песчано-алевролитовыми породами. Рельеф местности спокойный, равнинный. Исследования показывают, что давление насыщения на месторождении варьируется в диапазоне от 4 до 11,2 МПа, а газосодержание колеблется между 32,05 и 84,36 м<sup>3</sup>/т. Объёмный коэффициент доли изменяется от 1,078 до 1,264. Пластовая нефть месторождения обладает разнообразными характеристиками:

плотность варьируется от 0,7601 до 0,8352 г/см<sup>3</sup>, а вязкость находится в пределах 2,41–15,47 МПа·с.

Были собраны такие показатели месторождения, как суточная добыча нефти, суточная закачка воды и заключения гидродинамических исследований скважин (далее – ГДИС) за 3 мес., предшествующих дате анализа. Используя собранные данные, были произведены вычисления по среднему дебиту нефти и жидкости, закачки воды, и определены условия работы скважины с низким забойным давлением.

Для подтверждения работы скважины с низким забойным давлением используется бинарный показатель. Если отношение количества заключений ГДИС с низкими динамическими уровнями к общему количеству исследований превышает пороговое значение, то скважина считается работающей с низким давлением, и ей присваивается значение критерия, равное единице, как показано на рис. 1. В противном случае скважина характеризуется удовлетворительным забойным давлением и получает показатель со значением, равным нулю. В данной работе пороговое значение для доли заключений с низким динамическим уровнем принято равным 75% (1–3):

$$q_{avg} = \frac{\sum_{j=1}^t q_j}{t} \tag{1}$$

$$i_{avg} = \frac{\sum_{j=1}^t i_j}{t} \tag{2}$$

$$lowpbh = \begin{cases} 1, & \text{if } \frac{c_{low}}{c_{all}} > 0.75 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

где  $t$  – количество отработанных дней на текущем горизонте за 3 мес., сут;  $q$  – добыча нефти, т;  $i$  – закачка воды, м<sup>3</sup>;  $q_{avg}$  – средний дебит нефти, т/сут;  $i_{avg}$  – среднее значение приемистости, м<sup>3</sup>/сут;  $c_{low}$  – количество заключений ГДИС с низкими значениями забойных давлений, ед.;  $c_{all}$  – общее количество проведенных ГДИС, ед.;  $lowpbh$  – критерий работы скважины с низким забойным давлением.

well	horizon	rvg_gdis	otk_gdis	sum_gdis	low_pbh
0	A	12	0	93	0
1	B	14	0	95	0
2	F	0	0	94	0
3	F	0	0	92	0
4	L	82	11	93	1
...	...	...	...	...	...
3809	C	0	28	93	0
3810	E	0	0	95	0
3811	E	94	0	94	1
3812	E	0	0	94	0

**Рисунок 1. Расчет критерия работы скважин с низким забойным давлением**

**Figure 1. Calculation of the operating criterion for wells with low bottomhole pressure**

*well* – скважина; *horizon* – горизонт; *rvg\_gdis* – количество ГДИС с заключением «РВГ», ед. / number of gdis (well tests) with a “rvg” conclusion, units; *otk\_gdis* – количество ГДИС с заключением «ОТК», ед. / number of gdis (well tests) with an “otk” conclusion, units; *sum\_gdis* – общее количество ГДИС, ед. / sum gdis, units; *low\_pbh* – критерий работы скважины с низким забойным давлением / operating criterion for a well with low bottomhole pressure.

*РВГ / WWGI* – работа с влиянием газа / working with gas influence; *ОТК / PO* – откачка / pumping out.

well	horizon	liquid	oil	injection	work_time_prod	work_time_inj	days_prod	days_inj	liquid_rate	oil_rate	low_pbh
0	C	3694.7	106.8	0.0	2133.1	0.0	88.9	0.0	41.6	1.2	0.0
1	A	854.7	453.2	0.0	1962.8	0.0	81.8	0.0	10.5	5.5	0.0
2	B	3094.7	217.2	0.0	2061.1	0.0	85.9	0.0	36.0	2.5	0.0
3	B	3008.9	150.3	0.0	2312.6	0.0	96.4	0.0	31.2	1.6	0.0
4	B	1515.7	211.6	0.0	2204.8	0.0	91.9	0.0	16.5	2.3	0.0
...	...	...	...	...	...	...	...	...	...	...	...
4942	D	4833.5	207.5	0.0	2155.3	0.0	89.8	0.0	53.8	2.3	0.0
4943	A	0.0	0.0	13097.5	0.0	1655.0	0.0	69.0	NaN	NaN	NaN
4944	B	5116.3	539.4	0.0	2180.1	0.0	90.8	0.0	56.3	5.9	0.0
4945	A	2130.0	987.6	0.0	2273.2	0.0	94.7	0.0	22.5	10.4	0.0
4946	F	884.0	510.0	0.0	2083.8	0.0	86.8	0.0	10.2	5.9	0.0

**Рисунок 2. Расчет производственных показателей за 3 мес. для действующего фонда скважин**  
**Figure 2. Calculation of production indicators for 3 months for the existing well stock**

*liquid* – добыча жидкости, м<sup>3</sup>; *oil* – добыча нефти, т; *injection* – закачка воды, м<sup>3</sup>; *work\_time\_prod* – время работы добычи, ч / production work time, hours; *work\_time\_inj* – время работы нагнетания, ч / injection work time, hours; *days\_prod* – количество дней добычи, сут / number of days of production, days; *days\_inj* – количество дней нагнетания, сут / number of days of injection, days; *liquate\_rate* – дебит жидкости, м<sup>3</sup>/сут; *oil\_rate* – дебит нефти, т

### Разбивка месторождения на ячейки

С целью унификации потенциальных зон для бурения нагнетательных скважин рассматриваемое месторождение было поделено на равные участки. Для данного месторождения был выбран оптимальный размер ячейки, составляющий 200\*200 м при проектной плотности сетки 20 га/скв.

Проектные точки внутри ячеек размещались согласно двум вариантам. В первом варианте учитывалось лишь соответствие текущей плотности сетки скважин. Во втором варианте точки размещались с учётом текущей наземной инфраструктуры.

С целью определения оптимального с точки зрения плотности сетки скважин варианта размещения ячейки равномерно заполнялись точками с шагом в 10 м. Затем для каждой точки определялось расстояние до действующих скважин. После этого значения расстояний нормировались относительно эталонной дистанции между скважинами и выбиралась точка с максимальным значением нормированного расстояния в ячейке.

Размещение с учётом наземной инфраструктуры определялось с помощью библиотеки Georandas и программного языка Python. Данные по инфраструктуре в формате Shapely были импортированы в проект и преобразованы в полигоны. По умолчанию для каждого объекта была добавлена буферная зона, равная 20 м, которая в будущем может изменяться в зависимости от типа объекта. Для каждой точки также добавляется

буферная зона в 25 м, условно обозначающая размещение буровой установки. Проверка соответствия точек наземной инфраструктуре выполняется путем наложения полигонов вокруг точек и объектов. В случае пересечения полигонов точка исключается из дальнейшего рассмотрения. Поскольку в одной ячейке может быть множество точек, подходящих по инфраструктуре, в качестве итогового результата принимается точка, являющаяся наиболее близкой к ранее выявленной оптимальной точке, согласно сетке разработки в рассматриваемой ячейке.

Таким образом, возможны три варианта размещения точек в ячейке:

- найдена оптимальная точка по сетке разработки, но отсутствует точка по наземной инфраструктуре;
- найдена оптимальная точка по сетке разработки и точка по наземной инфраструктуре, при этом эти точки не совпадают;
- найдена оптимальная точка по сетке разработки и точка по наземной инфраструктуре, при этом эти точки совпадают.

На рис. 3 представлена карта, на которой показаны полигоны, созданные вокруг объектов наземной инфраструктуры, а также точки размещения проектных скважин в ячейках. Красным цветом выделены оптимальные точки по сетке разработке, синим цветом – оптимальные точки по наземной инфраструктуре.

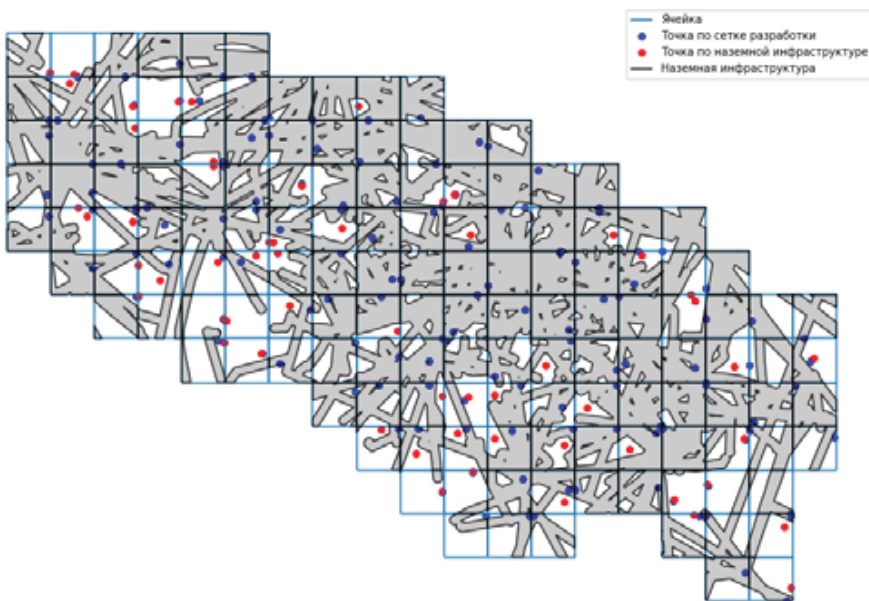
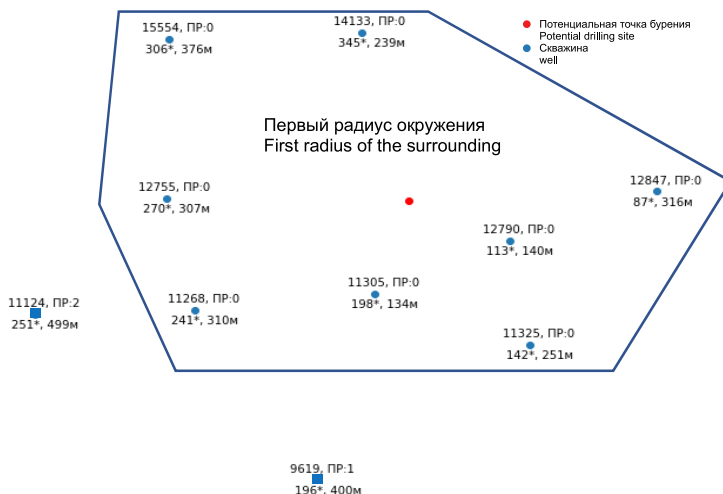


Рисунок 3. Карта наземной инфраструктуры и ячеек с потенциальными точками бурения  
Figure 3. Map of surface infrastructure and cells with potential drilling sites





**Рисунок 4. Определение скважин первого радиуса окружения**  
**Figure 4. Determination of wells of the first radius of the surrounding**

*ПР / FR – первый радиус окружения / first radius of the surrounding.*

**Оценка ячеек**

Для оценки ячеек необходимо рассчитать набор параметров и провести их нормализацию. Параметры ячеек рассчитывались на основе данных скважин первого окружения.

Определение скважин первого радиуса окружения для проектных точек состояло из нескольких шагов. В общий список скважин окружения попали все действующие скважины в радиусе 500 м вокруг проектной точки. Далее проводилось сравнение параметров каждой скважины с остальными скважинами из общего списка. Рассматриваемая скважина относилась к первому радиусу в двух случаях:

- если отрезок от проектной точки до рассматриваемой скважины не пересекает отрезок до сравниваемой скважины, то рассматриваемая скважина относилась к первому радиусу;

- в случае пересечения отрезков, если отрезок от проектной точки до рассматриваемой скважины меньше, чем отрезок до сравниваемой скважины, то рассматриваемая скважина относилась к первому радиусу.

На рис. 4 можно отметить, что показатель «ПР» у скважин с индексами 11124 и 9619 больше 1, следовательно, в первый радиус окружения данные скважины не войдут.

По данным скважин первого радиуса высчитываются средние и интегральные параметры ячеек:

- количество добывающих скважин;
- количество нагнетательных скважин;
- количество проектных скважин;

- доля скважин, работающих с низким забойным давлением;
- сумма добычи скважин;
- сумма закачки скважин;
- сумма дебитов жидкости и нефти скважин;
- сумма дней работы скважин;
- средний дебит жидкости, нефти (4);
- средний темп закачки (5);
- средняя обводненность (6);
- компенсация (7).

$$q_{avg} = \frac{\sum_{w=i}^{n_{well}} Q_w}{\sum_{w=i}^{n_{well}} t_w} \tag{4}$$

$$i_{avg} = \frac{\sum_{w=i}^{n_{well}} I_w}{\sum_{w=i}^{n_{well}} t_w} \tag{5}$$

$$w_{c_{avg}} = \left( 1 - \frac{\sum_{w=i}^{n_{well}} Q_{oilw}}{\sum_{w=i}^{n_{well}} Q_{liqw}} \right) * 100 \tag{6}$$

$$c_{avg} = \frac{\sum_{w=i}^{n_{well}} I_w}{\sum_{w=i}^{n_{well}} Q_{liqw}} \tag{7}$$

где  $n_{well}$  – количество скважин первого окружения, ед.;  $t$  – количество отработанных дней на текущем горизонте скважины за 3 мес., сут;  $w$  – скважина;  $Q$  – добыча нефти скважины за 3 мес., т;  $q$  – средний дебит нефти, т/сут;  $I$  – закачка воды за 3 мес., м<sup>3</sup>;  $i$  – средняя приемистость, м<sup>3</sup>/сут;  $p$  – плотность нефти, т/м<sup>3</sup>;  $wc$  – средняя обводненность, %;  $c$  – компенсация.

sector	horizon	date	liquid	oil	injection	days_prod	days_inj	sum_liquid_rate	sum_oil_rate	count_prod_wells	count_inj_wells
1155		2022-06-01	4126.34	1135.43	0.00	85.82	0.00	48.08	13.23	1.0	0.0
1156		2022-06-01	4126.34	1135.43	0.00	85.82	0.00	48.08	13.23	1.0	0.0
1157		2022-06-01	4126.34	1135.43	16925.15	85.82	85.55	48.08	13.23	1.0	1.0
1158		2022-06-01	7836.97	1503.08	16925.15	177.65	85.55	88.49	17.23	2.0	1.0
1159		2022-06-01	3710.63	367.64	16925.15	91.82	85.55	40.41	4.00	1.0	1.0
...	...	...	...	...	...	...	...	...	...	...	...
23535		2022-06-01	2968.17	1418.90	8045.00	90.54	91.00	32.78	15.67	1.0	1.0

count_inj_wells	count_proj_wells	count_low_pbh_wells	fraction_well_low_pbh	avg_liquid_rate	avg_wc	avg_oil_rate	avg_injection_rate	compensation
0.0	0.0	0.0	0.0	48.08	67.20	13.23	NaN	0.00
0.0	0.0	0.0	0.0	48.08	67.20	13.23	NaN	0.00
1.0	0.0	0.0	0.0	48.08	67.20	13.23	197.84	4.10
1.0	0.0	0.0	0.0	44.11	77.14	8.46	197.84	2.16
1.0	0.0	0.0	0.0	40.41	88.19	4.00	197.84	4.56
...	...	...	...	...	...	...	...	...
1.0	0.0	1.0	1.0	32.78	43.02	15.67	88.41	2.71

Рисунок 5. Параметры ячеек

Figure 5. Cell parameters

*sector* – сектор; *date* – дата расчета; *sum\_liquide\_rate* – сумма дебита жидкости, м<sup>3</sup>/сут (m<sup>3</sup>/day); *sum\_oil\_rate* – сумма дебита нефти, т (tonnes); *count\_prod\_wells* – количество добывающих скважин, ед. / number of production wells, units; *count\_inj\_wells* – количество нагнетательных скважин, ед. / number of injection wells, units; *count\_proj\_wells* – количество проектных скважин, ед. / number of project wells, units; *count\_low\_pbh\_wells* – количество скважин работающих с низким забойным давлением, ед. / number of wells operating with low bottomhole pressure, units; *fraction\_well\_low\_pbh* – соотношение количества скважин, работающих с низким давлением, и количества скважин окружения / ratio of the number of wells operating at low pressure to the number of surrounding wells; *avg\_liquid\_rate* – средний дебит жидкости, м<sup>3</sup>/сут / average liquid flow rate, m<sup>3</sup>/day; *avg\_wc* – средняя обводненность, % / average water cut, %; *avg\_oil\_rate* – средний дебит нефти, т/сут / average oil flow rate, t/day; *avg\_injection\_rate* – средняя приемистость, м<sup>3</sup>/сут / average injectivity, m<sup>3</sup>/day; *compensation* – компенсация

При проведении нормализации данных, включая такие параметры, как средний дебит, общая сумма дебитов и коэффициент обводненности, обычно возникают распространенные проблемы с выбросами, которые представляют собой аномальные значения, выходящие за пределы основных данных. Это может существенно исказить среднее арифметическое и размах, что в свою очередь может повлиять на анализ и результаты. Для решения этой проблемы предлагается заменить размах на интервал, в котором будут находиться основные данные, и затем провести масштабирование данных относительно этого интервала. Такой подход позволит избежать искажений, вызванных выбросами, и более точно представить данные для последующего анализа и интерпретации.

Наиболее распространенным методом автоматического определения выбросов является межквартильный метод. Его принцип заключается в том, что значения, которые на-

ходятся за пределами 1,5 межквартильных диапазонов (далее – IQR) от первого квартиля или третьего квартиля, считаются выбросами.

Однако, когда у распределения данных есть длинные «хвосты», слишком много информации может попасть в категорию «выбросов» – до 7% от общего числа данных. Это может повлиять на точность анализа и привести к неверным выводам. Симметричность межквартильного метода является ещё одной существенной проблемой. Полученный «доверительный интервал» (1,5\*IQR) одинаков для как для малых, так и для больших значений признака. Это означает, что многие аномалии с «короткой» стороны распределения могут быть скрыты этим интервалом.

По причине того, что данные не распределены нормально, использование метода IQR для их нормализации может быть некорректным. В связи с этим был применён другой метод для вычисления доверительного



интервала – скорректированная диаграмма размаха для асимметричных распределений. Данный метод использует медианное отношение (medcouple MC), которое измеряет асимметрию одномерного распределения [7] (8–12):

$$MC = median \{h(x_i, x_j), x_i < Q_2 < x_j\} \tag{8}$$

$$h(x_i, x_j) = \frac{((x_j - Q_2) - (Q_2 - x_i))}{x_j - x_i} \tag{9}$$

$$[min, max] = \begin{cases} [Q_1 - 1.5e^{-4MC}IQR; Q_3 + 1.5e^{3MC}IQR], & \text{if } MC \\ [Q_1 - 1.5e^{-3MC}IQR; Q_3 + 1.5e^{4MC}IQR], & \text{if } MC \end{cases} \tag{10}$$

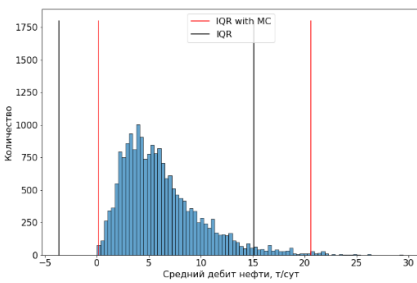
$$IQR = Q_3 - Q_1 \tag{11}$$

$$x_{ni} = \begin{cases} 1, & \text{if } x_i > max \\ \frac{x_i - min}{max - min}, & \text{if } max \leq x_i \leq max \\ 0, & \text{if } x_i < min \end{cases} \tag{12}$$

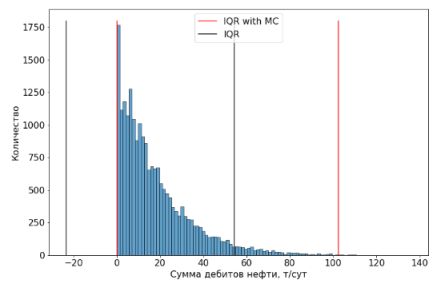
где Q – квартиль; min / max – минимальное / максимальное допустимое значение выборки; x<sub>i</sub> – значение выборки; x<sub>ni</sub> – нормированное значение.

На рис. 6 (г) наглядно видно, что границы доверительных интервалов двух методов совпадают. Однако в случае экспоненциальных распределений или распределений с выраженными хвостами, как показано на рис. 6 (а–в), наблюдается различие между максимальными и минимальными значениями. Метод IQR действительно сокращает длинные «хвосты». Например, по этому методу максимальный дебит составлял бы 15 т/сут, в то время как по скорректированному методу значение составило бы 20 т/сут.

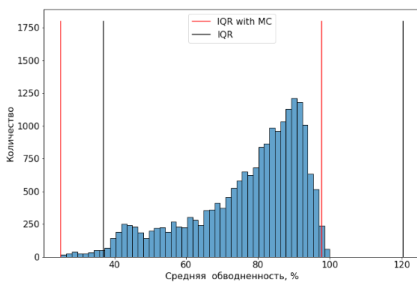
Для компенсации было решено установить максимально допустимое значение, принятое на уровне 120%, минимальное значение – 0%. Это решение обусловлено тем, что, если бы использовались максимальные значения, такие как 4000 или 6,1 (по границе), в этот интервал попало бы много выбросов. Что касается количества нагнетательных скважин, максимально допустимое число было определено как наибольшее значение в выборке, а минимальное значение было установлено равным 0. Это позволяет учесть разнообразие значений и обеспечить корректное определение интервала для данной характеристики.



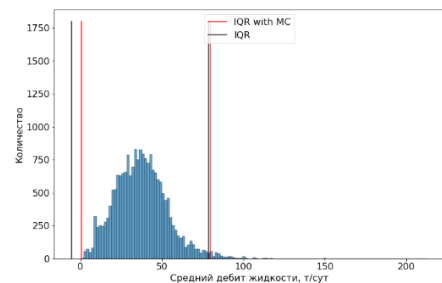
а)



б)



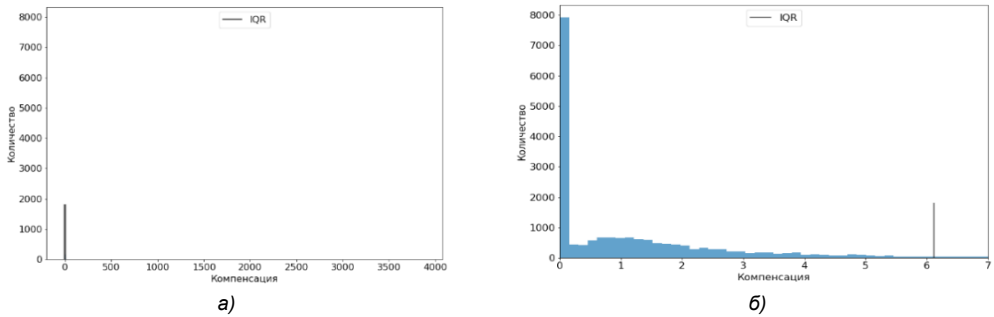
в)



г)

**Рисунок 6. Распределение с границами доверительного интервала двух методов**  
**Figure 6. Distribution with confidence interval limits of two methods**

а) средний дебит нефти / average oil flow rate; б) сумма дебитов нефти / sum of oil flow rates; в) средняя обводненность / average water cut; г) средний дебит жидкости / average liquid flow rate



**Рисунок 7. Распределение компенсации**  
**Figure 7. Compensation distribution**

а) сходное, включающее все значения / initial including all values; б) скорректированное, ограниченное границами доверительного интервала / adjusted, limited by confidence interval limits

Обводненность и количество нагнетательных скважин являются обратными величинами: чем больше значения данных параметров, тем хуже оценка ячейки. Следовательно, была применена следующая формула (13):

$$x_{ni} = 1 - x_{ni} \quad (13)$$

Для того, чтобы построить рейтинг зон, подходящих под нагнетание, была использована результирующая переменная, по которой ранжировались зоны. Данная переменная – это произведение нормированных признаков: суммы дебитов, компенсации, количества нагнетательных скважин и доли скважин, работающих с низким забойным давлением.

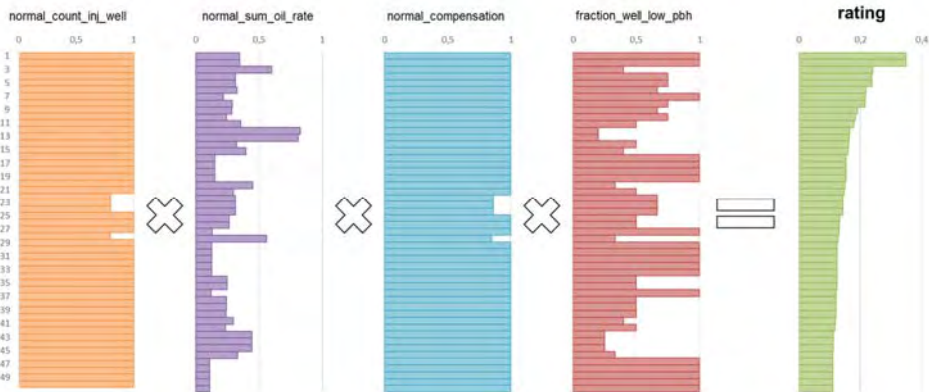
Каждой ячейке присваивается специальный код, который указывает на её пригодность с точки зрения разработки или инфраструктуры. Всего существует 7 различных кодов, из которых 4 связаны с уровнем средней компенсации

в буферной зоне. При этом приоритетность кодов увеличивается по мере убывания значения кода:

- 1 – компенсация менее 80%;
- 2 – компенсация менее 120%;
- 3 – компенсация более 120%;
- 4 – отсутствует добыча;
- 5 – в ячейке возможны возвраты с других горизонтов;
- 6 – в ячейке присутствуют действующие скважины;
- 7 – в ячейке присутствуют проектные скважины.

Эти коды будут дополнительно служить для ранжирования окончательной переменной. В начале списка будут располагаться зоны с кодами от 1 до 4, отсортированные по заключительному параметру. Затем идут зоны с кодами 5 и 6, а в конце списка будут зоны с кодом 7.

Рис. 10 иллюстрирует карту секторов, которая была упорядочена по рейтингу,



**Рисунок 8. 50 наиболее подходящих секторов по точке сетки разработки**  
**Figure 8. 50 most suitable sectors by well spacing point**

*normal\_count\_inj\_well* – нормированная и обратная величина количества нагнетательных скважин / *normalized and reciprocal values of the number of injection wells*; *normal\_compensation* – нормированная компенсация / *normalized compensation*; *rating* – итоговый показатель ранжирования / *summary ranking indicator*

sector	horizon	normal_count_inj_wells	code_inj	normal_sum_oil_rate	normal_avg_wc	normal_compensation	fraction_well_low_pbh	rating
4762	E	1.0	1.0	0.348517	0.210898	1.0	1.000000	0.349
5001	E	1.0	1.0	0.348517	0.210898	1.0	1.000000	0.349
7374	F	1.0	1.0	0.600674	0.720493	1.0	0.400000	0.240
19252	L	1.0	1.0	0.315385	0.724863	1.0	0.750000	0.237
19492	L	1.0	1.0	0.315385	0.724863	1.0	0.750000	0.237
14005	H	1.0	1.0	0.326159	0.744240	1.0	0.666667	0.217
13765	H	1.0	1.0	0.216249	0.776156	1.0	1.000000	0.216
12403	C	1.0	1.0	0.287168	0.646142	1.0	0.750000	0.215
16642	B	1.0	1.0	0.286520	0.319728	1.0	0.666667	0.191
15446	H	1.0	1.0	0.244514	0.696063	1.0	0.750000	0.183
18767	L	1.0	1.0	0.355890	0.430330	1.0	0.500000	0.178
5723	C	1.0	1.0	0.827396	0.672050	1.0	0.200000	0.165
12371	K	1.0	1.0	0.811503	0.555104	1.0	0.200000	0.162
6442	C	1.0	1.0	0.322870	0.527399	1.0	0.500000	0.161
18768	L	1.0	1.0	0.397337	0.449755	1.0	0.400000	0.159
23536	I	1.0	1.0	0.152847	0.752765	1.0	1.000000	0.153
23773	I	1.0	1.0	0.152847	0.752765	1.0	1.000000	0.153
23774	I	1.0	1.0	0.152847	0.752765	1.0	1.000000	0.153
23775	I	1.0	1.0	0.152847	0.752765	1.0	1.000000	0.153
5965	C	1.0	1.0	0.450227	0.864142	1.0	0.333333	0.150

Рисунок 9. 20 наиболее подходящих секторов по точке сетке разработки  
 Figure 9. 20 most suitable sectors by well spacing point

code\_inj – код ячейки / cell code; normal\_sum\_oil\_rate – нормированная сумма дебитов / normalized sum of flow rates; normal\_avg\_wc – нормированная средняя обводненность / normalized average water cut

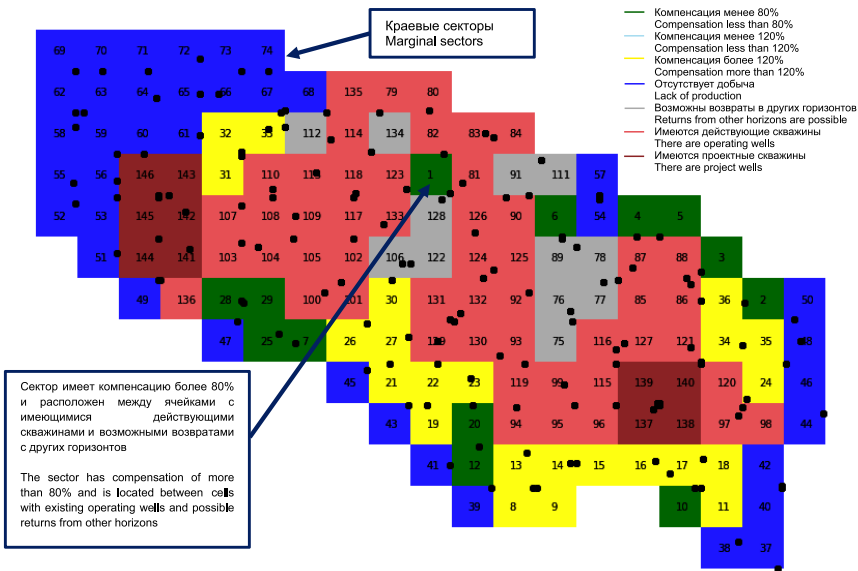


Рисунок 10. Секторы на X горизонте  
 Figure 10. Sectors on the X horizon

где числа внутри ячейки указывают на ранг сектора на данном горизонте. Важно отметить, что синие ячейки, представляющие области без добычи, совпадают с краевыми секторами.

**Выводы и дальнейшие перспективы исследования**

В данном исследовании был рассмотрен подход оценки зон бурения нагнетательных скважин на основе показателей работы действующего фонда. Внимание было уде-

лено изучению алгоритмов для сбора и обработки данных, проведению анализа, а также разработке методов для определения оптимальных зон для бурения.

Техники нормализации данных и определения выбросов играют важную роль в обеспечении точности анализа, а метод скорректированного межквартильного диапазона представляется наиболее эффективным для определения выбросов в данных с различными характеристиками распределения.

Присвоение кодов каждой ячейке позволило дополнительно ранжировать результаты. Это сделало финальную переменную более информативной и способствовало более точному процессу принятия решений.

Несмотря на потенциал автоматического подбора зон для нагнетания, существуют определенные проблемы. Предполагается выполнить следующие дополнительные этапы:

- ранжирование ячеек в разрезе всего месторождения, а не только одного горизонта;
- включение дополнительных параметров ячеек и учёт их важности при ранжировании: оценка неоднородности коллектора и других геолого-геофизических параметров;

### ДОПОЛНИТЕЛЬНО

**Источник финансирования.** Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Вклад авторов.** Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: Бекен А.А. – сбор, обработка и анализ экспериментальных данных, написание рукописи, Ибраев А.Е. – детальный анализ и интерпретация результатов, проведение исследования, редактирование рукописи, Жетруов Ж.Т. – анализ и проверка результатов, предоставление консультаций, разработка методологических подходов, Елемесов А.Т. – анализ и проверка результатов, организация

- выявление эффективных пропластков в соседних добывающих скважинах с целью определения, какие именно пакки следует выделить при проведении взрывных работ;

- тестирование данного подхода подбора зон для бурения нагнетательных скважин на других месторождениях АО НК «КазМунайГаз»;

- использование методов машинного обучения для выявления новых закономерностей и тенденций, влияющих на выбор зон для бурения.

Таким образом, данная исследовательская работа предоставила новые инструменты и методы для автоматического выбора точек бурения нагнетательных скважин.

исследования, Жолдыбаева А.Т. – анализ результатов, предоставление консультаций.

### ADDITIONAL INFORMATION

**Funding source.** This study was not supported by any external sources of funding.

**Competing interests.** The authors declare that they have no competing interests.

**Authors' contribution.** All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. The greatest contribution is distributed as follows: Aidana A. Beken – collection, processing, and analysis of experimental data, manuscript writing; Aktan Ye. Ibrayev – detailed analysis and interpretation of results, conducting research and manuscript editing; Zhassulan T. Zhettuov – analysis and verification of results, providing consultations, development of methodological approaches; Azamat T. Yelemesseov – analysis and verification of results, research organization, Assel T. Zholdybayeva – verification of results.

### СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. *Wei B.* Well Production Prediction and Visualization Using Data Mining and Web GIS : master's thesis. Calgary : University of Calgary, 2016. doi:10.11575/PRISM/28686.
2. *Xu X., Shao Y., Fu J., et al.* The Application of GIS in The Digital Oilfield Construction // 2nd International Conference on Computer Science and Electronics Engineering; Март 2013. Режим доступа: <https://www.atlantis-press.com/proceedings/iccsee-13/4443>. Дата обращения 12.08.2023.
3. *Khan H., Srivastav A., Kumar Mishra A., Anh Tran T.* Machine learning methods for estimating permeability of a reservoir // *Int J Syst Assur Eng Manag.* 2022. Vol. 13. P. 2118–2131. doi:10.1007/s13198-022-01655-9.
4. *Ruizhi Z., Cyrus S., Ray J.* Machine learning for drilling applications: A review // *Journal of Natural Gas Science and Engineering.* 2022. Vol. 108. doi:10.1016/j.jngse.2022.104807.
5. *Ramzey H., Badawy M., Elhosseini M., A. Elbaset A.* I<sup>2</sup>OT-EC: A Framework for Smart Real-Time Monitoring and Controlling Crude Oil Production Exploiting IIOT and Edge Computing // *Energies.* 2023. Vol. 16, N 4. doi:10.3390/en16042023.

6. Schiozer D.J., Souza dos Santos A.A., Graça Santos S.M., Von Hohendorff Filho J.C. Model-based decision analysis applied to petroleum field development and management // *Oil & Gas Science and Technology – Revue d'IFP Energies Nouvelles*. 2019. Vol. 74. doi:10.2516/ogst/2019019.

7. Hubert M., Vandervieren E. An adjusted boxplot for skewed distributions // *Computational Statistics & Data Analysis*. 2008. Vol. 52, N 12. P. 5186–5201. doi:10.1016/j.csda.2007.11.008.

## REFERENCES

1. Wei B. *Well Production Prediction and Visualization Using Data Mining and Web GIS* [master's thesis]. Calgary: University of Calgary; 2016. doi:10.11575/PRISM/28686.

2. Xu X, Shao Y, Fu J, et al. The Application of GIS in The Digital Oilfield Construction. 2nd International Conference on Computer Science and Electronics Engineering; March 2013. Available from: <https://www.atlantispress.com/proceedings/icsee-13/4443>.

3. Khan H, Srivastava A, Kumar Mishra A, Anh Tran T. Machine learning methods for estimating permeability of a reservoir. *Int J Syst Assur Eng Manag*. 2022;13:2118–2131. doi:10.1007/s13198-022-01655-9.

4. Ruizhi Z, Cyrus S, Ray J. Machine learning for drilling applications: A review. *Journal of Natural Gas Science and Engineering*. 2022;108. doi:10.1016/j.jngse.2022.104807.

5. Ramzey H, Badawy M, Elhosseini M, A. Elbaset A. I2OT-EC: A Framework for Smart Real-Time Monitoring and Controlling Crude Oil Production Exploiting IIOT and Edge Computing. *Energies*. 2023;16(4). doi:10.3390/en16042023.

6. Schiozer DJ, Souza dos Santos AA, Graça Santos SM, Von Hohendorff Filho JC. Model-based decision analysis applied to petroleum field development and management. *Oil & Gas Science and Technology – Revue d'IFP Energies Nouvelles*. 2019;74. doi:10.2516/ogst/2019019.

7. Hubert M., Vandervieren E. An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis*. 2008;52(12):5186–5201. doi:10.1016/j.csda.2007.11.008.

## ИНФОРМАЦИЯ ОБ АВТОРАХ

**\*Бекен Айдана Асқарқызы**

e-mail: [a.beken@kmge.kz](mailto:a.beken@kmge.kz).

**Ибраев Актан Ермакович**

e-mail: [ak.ibrayev@kmge.kz](mailto:ak.ibrayev@kmge.kz).

**Жетруов Жасулан Талгатбекович**

ORCID 0000-0003-3639-4390

e-mail: [zh.zhetruov@kmge.kz](mailto:zh.zhetruov@kmge.kz).

**Елемесов Азамат Серикович**

e-mail: [ayelemessov@kmge.kz](mailto:ayelemessov@kmge.kz).

**Жолдыбаева Асель Талгатовна**

ORCID 0000-0002-1015-0593

e-mail: [a.zholdybayeva@kmge.kz](mailto:a.zholdybayeva@kmge.kz).

## AUTHORS' INFO

**\*Aidana A. Beken**

e-mail: [a.beken@kmge.kz](mailto:a.beken@kmge.kz).

**Aktan Ye. Ibrayev**

e-mail: [ak.ibrayev@kmge.kz](mailto:ak.ibrayev@kmge.kz).

**Zhassulan T. Zhetruov**

ORCID 0000-0003-3639-4390

e-mail: [zh.zhetruov@kmge.kz](mailto:zh.zhetruov@kmge.kz).

**Azamat S. Yelemessov**

e-mail: [ayelemessov@kmge.kz](mailto:ayelemessov@kmge.kz).

**Assel T. Zholdybayeva**

ORCID 0000-0002-1015-0593

e-mail: [a.zholdybayeva@kmge.kz](mailto:a.zholdybayeva@kmge.kz).

\*Автор, ответственный за переписку/Corresponding Author