

УДК [556.555:574.5]:51-7  
AGRIS P10

https://doi.org/10.33619/2414-2948/85/23

## ИДЕНТИФИКАЦИЯ УНИКАЛЬНЫХ ОЗЕР РАЗЛИЧНОГО ПРОИСХОЖДЕНИЯ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

©*Расулова А. М.*, ORCID: 0000-0003-4400-2000, SPIN-код: 3888-6462, канд. ф.-м. наук,  
Институт озероведения Российской академии наук - Санкт-Петербургский Федеральный  
исследовательский центр РАН, г. Санкт-Петербург, Россия, ARasulova@limno.ru

©*Измайлова А. В.*, ORCID: 0000-0003-1705-6863, SPIN-код: 9660-9805, д-р геогр. наук,  
Институт озероведения Российской академии наук - Санкт-Петербургский Федеральный  
исследовательский центр РАН; Государственный гидрологический институт, г. Санкт-  
Петербург, Россия, ianna64@mail.ru

## IDENTIFICATION OF UNIQUE LAKES OF DIFFERENT ORIGIN BY MACHINE LEARNING METHODS

©*Rasulova A.*, ORCID: 0000-0003-4400-2000, SPIN-code: 3888-6462, Ph.D., St. Petersburg  
Federal Research Center of the Russian Academy of Sciences,  
St. Petersburg, Russia, ARasulova@limno.ru

©*Izmailova A.*, ORCID: 0000-0003-1705-6863, SPIN-code: 9660-9805, Dr. habil.,  
St. Petersburg Federal Research Center of the Russian Academy of Sciences; State Hydrological  
Institute, St. Petersburg, Russia, ianna64@mail.ru

*Аннотация.* В настоящее время как никогда стал актуален вопрос разработки критериев отбора озер для внесения их в списки особо охраняемых природных территорий (ООПТ), а также оценки природных экосистем, подвергшихся существенному антропогенному влиянию и требующих особого внимания со стороны экологов. Однако, экспертная оценка индивидуально каждой экосистемы требует значительных исследовательских и экономических ресурсов. С учетом площади России и труднодоступности некоторых регионов она становится практически невозможной. Для предварительных оценок и сужения круга поиска кандидатов на ООПТ могут выступать камеральные методы. К одним из них относятся различные методы идентификации аномалий по базам данных морфометрических, гидрохимических, гидрологических и гидробиологических характеристик озер. В данной работе рассмотрены некоторые методы машинного обучения, направленные на выявления аномальных значений для озер карстового, вулканического и ледникового происхождения. Основной целью данного исследования является поиск оптимальных математических методов установления уникальности той или иной озерной экосистемы. В работе рассмотрены тестовые выборки озер, полученные на основе базы данных WORLDLAKE. При анализе использовались методы: 1) фактор локального выброса, 2) изолированный лес и 3) одноклассовый метод опорных векторов. Выявлены особенности применения различных методов в зависимости от морфометрического происхождения озерных котловин. Полученные аномальные объекты подвергались сравнению и последующей экспертной оценке на наличие у них уникальных свойств по различным параметрам. Экспертная оценка подтвердила, что большинство выделенных озер могут рассматриваться как уникальные с учетом и других признаков, характеризующих озерные экосистемы.

*Abstract.* At present, more than ever, the issue of developing criteria for selecting lakes for inclusion in the lists of protected areas, as well as assessing natural ecosystems that have undergone significant anthropogenic impact and require special attention from environmentalists, has become relevant. However, peer review of each individual ecosystem requires significant research and economic resources. Taking into account the area of Russia and the inaccessibility of some regions, it becomes almost impossible. For preliminary assessments and narrowing the search for candidates for protected areas, cameral methods can be used. One of them includes various methods for identifying anomalies in databases of morphometric, hydrochemical, hydrological, and hydrobiological characteristics of lakes. This paper discusses some machine learning methods aimed at identifying anomalous values for lakes of karst, volcanic and glacial origin. The main goal of this study is to find optimal mathematical methods for establishing the uniqueness of a particular lake ecosystem. The paper considers test samples of lakes obtained on the basis of the WORLDBASE database. The following methods were used in the analysis: 1) local outlier factor, 2) isolated forest, and 3) one-class support vector machine. The features of the application of various methods depending on the morphometric origin of lake basins are revealed. The resulting anomalous objects were compared and then subjected to expert evaluation for their unique properties in various parameters. The expert assessment confirmed that most of the identified lakes can be considered unique, taking into account other features that characterize lake ecosystems.

*Ключевые слова:* охраняемые территории, охрана экосистем, озера, идентификация аномалий, локальный фактор выброса, изолированный лес, одноклассовый метод опорных векторов.

*Keywords:* protected areas, ecosystem conservation, lakes, identification of anomalies, Local Outlier Factor, Isolated Forest, One-Class Support Vector Machine.

### *Введение*

Уникальными принято считать природные объекты, характеризующиеся специфическим набором свойств, особым образом выделяющихся среди природного многообразия. Логично, что любой такой объект требует особой охраны. Вместе с тем часто остается открытым вопрос об отсутствии объективных критериев, позволяющих отнести ту или иную территорию или акваторию к категории «уникальных». Чаще всего для этого используются экспертные оценки, справедливость и независимость, которых может быть подтверждена за счет алгоритмизации процесса выявления уникальных свойств и характеристик. С этой целью ранее была предложена возможность выявления уникальных свойств природных объектов, в нашем случае — озерных экосистем, с использованием статистических критериев идентификации выбросов [1]. В основу анализа было положено рассуждение, что если озера представить, как статистическую совокупность, обладающую набором признаков (параметров), то большинство выборок, построенных по этим параметрам, может быть описано с помощью теоретического распределения. При этом по некоторым параметрам будут наблюдаться аномальные значения, т.е. объекты, резко выделяющиеся из общей совокупности. В большинстве случаев именно такие аномалии свидетельствуют об уникальности озерной экосистемы, тем более что в силу взаимообусловленности происходящих в озерах процессов, если водная экосистема резко отлична от остальных на определенной территории по одному признаку, то она может выделяться и по ряду других признаков.

Справедливость данной гипотезы была обоснована нами при поиске уникальных озер с применением различных математических методов. Исследовались озера вулканического [2] и карстового [3] происхождения, также было проведено исследование, где озера объединялись в группы не по происхождению озерной котловины, а по их принадлежности к одной физико-географической стране [4]. В простейшем случае применялся статистический поиск аномалий в распределении, который сводится к задаче о выявлении выбросов. Под выбросом понимается наблюдение, резко выделяющееся из набора данных. Для идентификации аномальных значений использовались классические статистические критерии поиска выбросов в нормальном распределении (такие, как Смирнова-Граббса, Хоглина-Иглевича, Титьена-Мура, Шовене, Ирвина и метод Томпсона (критерий Рошера)). Однако, как показала работа над выборками озер различного происхождения, применение статистических методов поиска аномалий возможно далеко не всегда. При исследовании выборок, не поддающимся статистической идентификации выбросов, нами был применен древовидный алгоритм изоляции аномалий – «Изолированный лес» (Isolation Forest).

Полученные в каждом случае итоговые наборы озер с аномальными параметрами проходили экспертную оценку на предмет исключительности их экосистем. Результаты свидетельствовали, что полученные разными методами наборы содержат большое количество озер, отличающихся не только аномалиями единичных параметров, но и обладающих своеобразием других абиотических характеристик, а подчас — характеризующихся уникальной биотой. В то же время при анализе различных типов озер мы не смогли применить единый математический аппарат поиска выбросов. Было отмечено, что эффективность тех или иных подходов к поиску аномальных значений и выделению водоемов с уникальными свойствами, прежде всего, определяется характером анализируемой выборки — единством генетического происхождения озерных котловин в выборке, принадлежности к физико-географической стороне, биому и т.п.

В данной работе представлены результаты исследования применимости методов поиска аномальных выборок морфометрических значений озер ледникового, карстового и вулканического происхождения с помощью методов машинного обучения на основе библиотеки python Scikit-learn (ранее scikits.learn, также известная как sklearn). Произведен анализ итоговых наборов озер с аномальными характеристиками, полученных несколькими методами. Среди алгоритмов поиска аномалий, доступных в библиотеке Scikit-learn наибольшую эффективность показали:

1. Локальный фактор выброса (Local Outlier Factor).
2. Изолированный лес (Isolation Forest).
3. Одноклассовый метод опорных векторов (One-Class Support Vector Machine).

#### *Материалы и методы исследования*

В настоящем исследовании была использована дополненная авторами база данных (БД) WORDLAKE [5]. Поскольку БД представляет собой уже обработанный массив натуральных наблюдений, при работе на выявление выбросов априори считалось, что данные не содержат ошибок измерений. Таким образом, идентифицируются исключительно аномальные значения.

Репрезентативные выборки по озерам различного происхождения, отобранные для дальнейшего анализа, строились на основе наибольшего заполнения всех морфометрических параметров в группе водоемов единого происхождения. Морфометрические параметры БД WORDLAKE представлены в Таблице 1.

Таблица 1

МОРФОМЕТРИЧЕСКИЕ ПАРАМЕТРЫ, СОДЕРЖАЩИЕСЯ В БД WORDLAKE

| Обозначение         | Характеристика и единицы измерения   |
|---------------------|--|
| $L$                 | Длина озера (км)   |
| $W$                 | Ширина озера (км)  |
| $S$                 | Площадь зеркала озера (км <sup>2</sup> )   |
| $\ell$              | Длина береговой линии озера (км)   |
| $H_{avg}$           | Средняя глубина озера (м)  |
| $H_{max}$           | Максимальна глубина озера (м)  |
| $V$                 | Объем озера (км <sup>3</sup> )   |
| $C$                 | Площадь водосбора (км <sup>2</sup> )   |
| $k_s$               | Коэффициент сжатия зеркала озера: $k_s = \frac{W}{L}$  |
| $1/k_s$             | Показатель удлиненности, обратно пропорциональный к коэффициенту сжатия $k$  |
| $c$                 | Коэффициент формы озерной котловины (коэффициент емкости) $c = \frac{H_{avg}}{H_{max}}$  |
| $H_{avg}^*$         | Безразмерная средняя глубина (по отношению к средней глубине озера Байкал $H_{avg} = 740$ м) $H_{avg}^* = \frac{H_{avg}}{H_{avgB}}$            |
| $H_{max}^*$         | Безразмерная максимальная глубина (по отношению к максимальной глубине озера Байкал $H_{max} = 1650$ м) $H_{max}^* = \frac{H_{max}}{H_{maxB}}$ |
| $S^*$               | Безразмерная площадь озера (относительно площади Каспийского моря $S_{KS} = 3,86 \cdot 10^5$ км <sup>2</sup> ) $S^* = \frac{S}{S_{KS}}$        |
| $V^*$               | Безразмерный объем озера (относительно объема Каспийского моря $V_{KS} = 7,87 \cdot 10^5$ км <sup>3</sup> ) $V^* = \frac{V}{V_{KS}}$           |
| $H^{**}$            | Относительная глубина, равная отношению между максимальным вертикальным масштабом и площадью зеркала   |
| $\frac{S}{H_{avg}}$ | Показатель открытости, равный отношению площади озера к средней глубине $H^{**} = 10^{-3} \frac{H_{max}}{\sqrt{S}}$                            |
| $\alpha$            | Коэффициент «глубинности» озера $\alpha = \frac{H_{avg}}{S^{1/3}}$   |
| $\varphi$           | Широта (в градусах)  |
| $\lambda$           | Долгота (в градусах)   |
| $Z_{asl}$           | Высота озера над уровнем моря (м)  |
| $N$                 | Количество валидных объектов в выборке   |

В итоговых тестовых выборках получилось следующее количество объектов: ледниковые озера — 1777, карстовые озера — 251, вулканические озера — 129.

Поскольку все полученные выборки не относятся к нормальному распределению и не сводятся к нему универсальным преобразованием Бокса-Кокса [7], к ним невозможно применять классические статистические критерии поиска выбросов типа 3-сигма, Смирнова-

Граббса, Хоглина-Иглевича, Титьена-Мура, Шовене, Ирвина и метод Томпсона (критерий Рошера). В связи с этим для поиска аномалий были исследованы альтернативные методы, основанные на синтезе различных подходов, которые включают в себя теорию графов, метрические подходы, математическую оптимизацию и статистические методы. Среди различных подходов неконтролируемого обучения были выбраны группы методов на основе: ближайшего соседа (алгоритм LOF и Isolation Forest) и классификации (One-class SVM).

### Локальный фактор выброса (Local Outlier Factor, LOF)

LOF относится к неконтролируемым методам машинного обучения, что является важным в рамках нашего исследования, т.к. изначально не обязательно знать являются ли анализируемые выборки данных нормальными, что упрощает расчеты. Алгоритм LOF базируется на вычислении отклонения локальной плотности точки по отношению к ее  $k$ -ближайшим соседям [8, 9]. Основным параметром алгоритма является количество ближайших соседей  $N_k(p)$ , где  $p$  — объект, от которого измеряется расстояние. Здесь под «объектом» понимается совокупность морфометрических параметров одного озера.

Пусть  $d(p, q)$  — расстояние между объектами  $p$  и  $q$ ,  $C$  — набор объектов,  $d(p, C)$  — минимальное расстояние между объектами  $p$  и объектом  $q$ , принадлежащим кластеру  $C$ . Выбросом будем называть такой объект  $DB(pct, d_{min})$  набора данных  $D$ , у которого хотя бы процентная доля  $pct$  объектов в  $D$  лежит на расстоянии больше, чем  $d_{min}$  от  $p$ . Однако, под это определение могут не попадать объекты, удаленные от своих локальных кластеров на расстояния меньше, чем  $d_{min}$ , но при этом являющиеся аномалиями для определенного набора данных. Таким образом, определение  $DB(pct, d_{min})$  является необходимым, но недостаточным. Для дополнения этого определения используется количественная оценка каждого выброса на основе присвоения объектам степени отклонения (т.н. локальный фактора выброса). Степень отклонения объекта строится на основе понятий  $k$ -расстояния объекта  $p$  и  $k$ -ближайших соседей  $p$ :

1. По определению  $k$ -расстоянием объекта  $p$  ( $k$ -distance( $p$ ) или просто  $k(p)$ ) называется расстояние  $d(p, o)$  между объектом  $p$  и самой дальней точкой объекта  $s \in D$ , такое что: 1) не меньше, чем для  $k$  объектов,  $s \in D$  верно, что  $d(p, s') \leq d(p, s)$  и 2) для больше чем  $k-1$  объекту  $s' \in D$  верно что  $d(p, s') < d(p, s)$ .

2.  $k$ -ближайшими соседями  $p$  ( $N_k(p)$ ) иногда обозначают как  $kNN$  (от англ.  $k$ -Nearest Neighbors) называются объекты  $q \in D$ , расстояние которых от  $p$  не превышает  $k(p)$ , т.е.  $N_k(p)$ :  $d(p, q) \leq k(p)$ .

Расстояние  $d(p, o)$  — это расстояние в  $n$ -мерном пространстве, которое может вычисляться различными способами. Мерой расстояния может быть Евклидова, Минковского метрики, манхэттенское расстояние или любая другая метрика. В настоящей работе наилучший результат показала метрика Минковского.

На основе определения  $k(p)$  строится понятие достижимого расстояния объекта  $p$  относительно объекта  $s$  ( $dr(p, s)$ ), как максимального расстояния объекта  $s$  от объекта  $p$ , т.е.  $d(p, s) = \max\{k(s), d(p, s)\}$ .

Для поиска аномальных значений необходимо сравнить плотность различных групп объектов, для этого вводится локальная плотность достижимости объекта  $p$ , определяемой следующим образом:

$$LRD_k(p) = \left( \frac{\sum_{q \in N_k(p)} dr_k(p, q)}{|N_k(p)|} \right)^{-1} \quad (1)$$

Сравнивая локальные плотности достижимости объекта  $p$  с его соседним объектом  $q$ , найдем локальный фактор выброса как:

$$LOF_k(p) = \frac{\sum_{q \in N_k(p)} \frac{LRD_k(q)}{LRD_k(p)}}{|N_k(p)|} \quad (2)$$

$LOF_k(p)$  показывает среднее отношение плотности локальной достижимости  $p$  к плотности  $k$ -ближайших соседей  $p$ . Из (2) следует, что с уменьшением локальной плотности достижимости  $p$  и увеличением локальной плотности  $k$ -ближайших соседей  $p$ , увеличивается  $LOF_k(p)$ , т.е. с большей вероятностью объект является аномалией. Для данного метода значение параметра  $LOF_k(p) \approx 1$  означает, что локальные плотности достижимости примерно равны и объект  $p$  не является выбросом. Значение  $LOF_k(p) < 1$  указывают на более плотную область, т.е. объект принадлежит к группе, а значения  $LOF_k(p) \gg 1$  говорит о том, что объект является аномальным.

#### Алгоритм *iForest*

Алгоритм Isolation Forest (*iForest*) основан на изоляции экземпляров выборки [10, 11]. В данном случае понятие "экземпляр" служит аналогом понятия «объект» для алгоритма LOF. Разница в понятиях единичной строки данных связана с разным математическим аппаратом, на которых базируются алгоритмы. Isolation Forest не требует при вычислении меры расстояния или локальной плотности и базируется на основополагающих свойствах аномалий: 1) количество аномалий намного меньше размеров выборки; 2) значения атрибутов аномальных экземпляров всегда разнообразно и резко отличаются от нормальных. Реализация алгоритма *iForest* основана на построении ансамбля деревьев решений – совокупности непересекающихся неориентированных бинарных деревьев решений (или деревьев изоляции, Isolation Tree, далее *iTree*). Каждое дерево решений представляет собой связный ациклический граф.

Изоляционное дерево строится следующим образом: пусть  $X = \{x_1, \dots, x_n\}$  – набор данных из  $n$  точек с количеством вариационных атрибутов  $d$ . Для построения *iTree* используется подвыборка из  $\psi$ -экземпляров, таких что  $X' \subset X$ .  $X'$  делится рекурсивно, случайным образом выбирается атрибут  $q$  и значение разделения  $p$  пока: 1) узел имеет только один экземпляр или 2) все данные на узле не принимают одинаковые значения. Полностью выросшее дерево решений имеет  $\psi$  – внешних,  $\psi-1$  – внутренних и  $2\psi-1$  – общего количества узлов. Аномалией признается экземпляр, который наиболее просто изолировать.

Для количественной оценки каждого экземпляра на аномальность рассчитывается длина пути (количество ребер) от концевой до внешнего узла дерева решений  $h(x)$ . Чем меньше длина пути  $h(x)$ , тем выше восприимчивость к изоляции, т.е. экземпляр с минимальным  $h(x)$  является аномальным.

Однако, чисто на основе  $h(x)$  расчет длины пути сделать затруднительно из-за разницы в скорости роста средней высоты *iTree*, пропорциональной  $\log \psi$ , и максимально возможной высоты, пропорциональной  $\psi$ . Поэтому используется оценка среднего  $h(x)$  взятая аналогично неудачному поиску внешних узлов в двоичном дереве поиска [11], имеющая вид:

$$c(\psi) = \begin{cases} 2H(\psi - 1) - 2(\psi - 1)/n, & \text{при } \psi > 2 \\ 1, & \text{при } \psi = 2 \\ 0, & \text{при } \psi < 2 \end{cases} \quad (3)$$

где  $H(i)$  — гармоническое число, равное  $\ln(i) + \gamma$ , где  $\gamma$  — постоянная Эйлера. Тогда оценка аномальности экземпляра  $x$  имеет вид:

$$s(x, \psi) = 2 \frac{E(h(x))}{c(\psi)} \quad (4)$$

где  $E(h(x))$  — среднее значение  $h(x)$  из ансамбля  $i$ Tree.

Оценки аномалий по формуле (4) показывают, что экземпляр признается аномальным если  $s \rightarrow 1$ . Для  $s \ll 0,5$  экземпляра можно с полной уверенностью назвать нормальным.

#### *Одноклассовый метод опорных векторов (One Class Support Vector Machine, OCSVM)*

Метод опорных векторов (Support Vector Machine, SVM) базируется на классификации данных, основанной на разделении  $n$ -мерного пространства гиперплоскостью, такой, что расстояние между классами будет максимальным. Каждое наблюдение соответствует многомерному вектору. Пусть  $\{(x_1, y_1), \dots, (x_i, y_i)\} \in X$ , — входные данные, представляющие собой векторы размерности  $n$ , а  $X \in R$ , где  $R$  — пространство признаков. Каждый  $x_i$  — точка данных, а  $y_i$  — отвечает за класс, к которому принадлежит точка. Основной задачей алгоритма SVM является нахождение функции  $f(x)$ , такой, что отклонение от  $y_i$  меньше  $\varepsilon$  и  $f(x)$  — максимально плоская. В простейшем случае  $f(x)$  — линейная [12], т.е.:

$$f(x) = \langle w, x \rangle + b \quad (5)$$

где  $\langle \cdot, \cdot \rangle$  — скалярное произведение в  $X$ ,  $w$  — вектор нормали к гиперплоскости, параметр  $b$  отвечает за смещение гиперплоскости относительно начала координат. Поскольку мы приняли допущение, что гиперплоскость описывается линейным уравнением (5), то для этого норма  $\|w\|^2 = \langle w, w \rangle$  должна быть минимальна. Тогда найдем такую  $f(x)$ , которая аппроксимирует все  $\{(x_1, y_1), \dots, (x_i, y_i)\}$  с точностью до  $\varepsilon$  можно сформулировать как:

$$\min \left\{ \frac{1}{2} \|w\|^2 \right\}, \text{ при } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \quad (6)$$

Т.е. решается задача выпуклой оптимизации. Иногда для разрешения задачи вводятся дополнительные переменные  $\xi_i$  [13].

Алгоритм One Class Support Vector Machine (OC-SVM) используется в конкретном случае для отделения одного конкретного класса от всех остальных данных [14]. Результатом такой аппроксимации является бинарная функция. В этом случае задачу выпуклой оптимизации будет несколько отличаться от (6):

$$\min \left\{ \frac{1}{2} \|w\|^2 \right\} + \frac{1}{\eta n} \sum_{i=1}^n \xi_i - \rho \quad (7)$$

где  $\eta$  — параметр регуляризации, который отвечает за верхнюю границу доли выбросов,  $\rho$  — параметр, характеризующий максимальное расстояние гиперплоскости от

начала координат. В этом случае функция решения (5) приобретает более сложный вид, через кусочно-постоянную функцию  $sgn(x)$  и т. н. функцию ядра, отвечающую за форму гиперплоскости.

В этом алгоритме OCSVM аномальным объектом считается точка в  $n$ -мерном пространстве, которая не проходит за гиперплоскость.

### Результаты и обсуждение

Для всех алгоритмов мы ставили загрязнение выборки не более 0,1, таким образом ограничивая число аномалий в каждой группе озер 10%. Предыдущие исследования показали, что на выборку озер по морфометрическим параметрам приходится примерно 5–10% аномальных значений.

Анализ озер карстового происхождения показал, что из 251 озера во все алгоритмы попало 13 водоема (Таблица 2). При этом, 4 озера с аномальными значениями обнаружены только алгоритмом *iForest*, 6 — алгоритмом LOF и 9 — OCSVM. Также есть ряд озер, которые были идентифицированы, как аномалии только двумя из трех алгоритмов: 2 озера алгоритмами *iForest* и OCSVM и 6 озер алгоритмами *iForest* и LOF. Между алгоритмами OCSVM и LOF совпадающих озер не найдено.

Таблица 2

#### КАРСТОВЫЕ ОЗЕРА, ПРИЗНАННЫЕ АНОМАЛЬНЫМ ПО РАЗЛИЧНЫМ АЛГОРИТМАМ

|                         | Код озера | Название на латинице | Русское название | Страна               |
|-------------------------|-----------|----------------------|------------------|----------------------|
| 1                       | 14        | Band-i-Amir          | Банде-Амир       | Афганистан           |
| 2                       | 1311      | Luner See            | Люнерзе          | Австрия              |
| 3                       | 5174      | Livanjsko polje      | Ливаньско-поле   | Босния и Герцеговина |
| 4                       | 13731     | Vransko              | Вранско          | Хорватия             |
| 5                       | 13734     | Vransko              | Врана            | Хорватия             |
| 6                       | 13772     | Tsrveno              | Чрвено (Красное) | Хорватия             |
| 7                       | 15463     | Girotte              | Жирот            | Франция              |
| 8                       | 17352     | Spiekeroog           | Спикеруг         | Германия             |
| 9                       | 21369     | Kush-Murun, ozera    | Кушмурун, озера  | Казахстан            |
| 10                      | 39744     | Yarkul'              | Яркуль           | Россия               |
| 11                      | 39922     | Chusovskoye          | Чусовское        | Россия               |
| 12                      | 44007     | Karaviran            | Каравиран        | Турция               |
| 13                      | 53125     | Proval'noye          | Провальное       | Россия               |
| <i>iForest u OS-SVM</i> |           |                      |                  |                      |
| 1                       | 1847      | Dolgoye              | Долгое           | Беларусь             |
| 2                       | 17402     | Arendsee             | Арендзе          | Германия             |
| <i>iForest u LOF</i>    |           |                      |                  |                      |
| 1                       | 17343     | Dummer See           | Дюммер Зее       | Германия             |
| 2                       | 24891     | Shavart nuur         | Шаварт нуур      | Монголия             |
| 3                       | 24905     | Sumiyn nuur          | Сумийн нуур      | Монголия             |
| 4                       | 31493     | Dukhovoye            | Духовое          | Россия               |
| 5                       | 41364     | Druzhinnoye          | Дружинное        | Россия               |
| 6                       | 45170     | Svityazskoye         | Свитязское       | Украина              |
| <i>OCSVM</i>            |           |                      |                  |                      |
| 1                       | 18        | Band-i-Jedacel       | Джидачиль        | Афганистан           |
| 2                       | 13754     | Desne                | Десне            | Хорватия             |
| 3                       | 18958     | Tovel                | Лаго-ди-Товель   | Италия               |

|                | Код озера | Название на латинице | Русское название | Страна     |
|----------------|-----------|----------------------|------------------|------------|
| 4              | 25210     | Nogoon nuur          | Ногоон нуур      | Монголия   |
| 5              | 25616     | Tsagaan nuur         | Цаган нуур       | Монголия   |
| 6              | 27835     | Obradowskie          | Обрадовски       | Польша     |
| 7              | 42258     | Karash               | Караш            | Россия     |
| 8              | 45169     | Somino               | Сомино           | Украина    |
| 9              | 52235     | Kanbeshbulak         | Канбешбулак      | Узбекистан |
| <i>LOF</i>     |           |                      |                  |            |
| 1              | 2228      | Cherstvyaty          | Черствяти        | Беларусь   |
| 2              | 31697     | Kaldy                | Калды            | Россия     |
| 3              | 31707     | Kasargy              | Касарги          | Россия     |
| 4              | 41107     | Sabro                | Сабро            | Россия     |
| 5              | 41447     | Kushtozero           | Куштозеро        | Россия     |
| 6              | 41566     | Shimozero            | Шимозеро         | Россия     |
| <i>iForest</i> |           |                      |                  |            |
| 1              | 1396      | Wildgerlossee        | Вильдерлозе      | Австрия    |
| 2              | 2244      | Krivoye              | Кривое           | Беларусь   |
| 3              | 10228     | Deadmoose            | Дед Муз          | Канада     |
| 4              | 13742     | Kozjak               | Козьяк           | Хорватия   |

Выборка озер ледникового происхождения с аномальными значениями морфометрических параметров показала, что все алгоритмы выделили 37 водоемов (табл. 3) из 1776 (при 10% допущении обнаружения аномалий в конечных выборках по разным алгоритмам попали: 178 озер (для *iForest*), 146 (LOF), 176 (OCSVM). Так же, как и в случае с ледниковыми озерами был выделен ряд водоемов с аномальными значениями, обнаруженных только двумя из трех алгоритмов. В частности: 47 объектов идентифицированы только алгоритмами *iForest* и OCSVM (не приведены в Таблице 3 в силу большого объема всей выборки), 12 — алгоритмами LOF и OCSVM, 28 – алгоритмами *iForest* и LOF. В отличие от случая озер карстового происхождения здесь есть объекты, идентифицированные как алгоритмом LOF, так и OCSVM. Только алгоритмом *iForest* было выделено 66 объектов, OCSVM — 76 и LOF — 69.

Таблица 3

ЛЕДНИКОВЫЕ ОЗЕРА, ПРИЗНАННЫЕ АНОМАЛЬНЫМ ПО ВСЕМ АЛГОРИТМАМ

|    | Код озера | Название на латинице | Русское название                | Страна         |
|----|-----------|----------------------|---------------------------------|----------------|
| 1  | 401       | Buenos Aires         | Буэнос Айрес (Хенераль Каррера) | Аргентина/Чили |
| 2  | 405       | San Martin           | Сан-Мартин                      | Аргентина/Чили |
| 3  | 5786      | Claire               | Клэр                            | Канада         |
| 4  | 6140      | Athabaska            | Атабаска                        | Канада         |
| 5  | 6156      | Adams                | Адамс-Лейк                      | Канада         |
| 6  | 6445      | Cross                | Кросс                           | Канада         |
| 7  | 6552      | Manitoba             | Мапнитоба                       | Канада         |
| 8  | 6669      | Winnipeg             | Виннипег                        | Канада         |
| 9  | 6670      | Winnipegosis         | Виннипегосис                    | Канада         |
| 10 | 6693      | Reindeer             | Олень                           | Канада         |
| 11 | 6767      | Melville             | Мелвилл                         | Канада         |
| 12 | 6848      | Brass d'Or           | Бра-д-Ор                        | Канада         |
| 13 | 6969      | Baker                | Бейкер                          | Канада         |

|    | <i>Код озера</i> | <i>Название на латинице</i> | <i>Русское название</i> | <i>Страна</i>  |
|----|------------------|-----------------------------|-------------------------|----------------|
| 14 | 7064             | Great Bear Lake             | Большое Медвежье        | Канада         |
| 15 | 7066             | Great Slave                 | Большое Невольничье     | Канада         |
| 16 | 7284             | Martre                      | Мартр                   | Канада         |
| 17 | 7304             | Nettilling                  | Неттилинг               | Канада         |
| 18 | 9063             | Nipigon Lake                | Нипигон                 | Канада         |
| 19 | 10466            | Wollaston                   | Вулластон               | Канада         |
| 20 | 10515            | Lake-Of-the-Woods           | Лесное                  | Канада/США     |
| 21 | 10544            | Caburgua                    | Кабургуа                | Чили           |
| 22 | 14365            | Pskovsko-Chudskoye ozero    | Псковско-Чудское        | Эстония/Россия |
| 23 | 14547            | Pielinen-jarvi              | Пиелинен                | Финляндия      |
| 24 | 14548            | Päijänne                    | Паийянне                | Финляндия      |
| 25 | 14583            | Saima                       | Сайма                   | Финляндия      |
| 26 | 26175            | Hawea                       | Хавеа                   | Новая Зеландия |
| 27 | 26210            | Manapouri                   | Манапоури               | Новая Зеландия |
| 28 | 26511            | Mjosa                       | Мьеса                   | Норвегия       |
| 29 | 27017            | Tinnsjoen                   | Тинше                   | Норвегия       |
| 30 | 39554            | Il'men'                     | Ильмень                 | Россия         |
| 31 | 39558            | Seliger                     | Селигер                 | Россия         |
| 32 | 43000            | Stora Lulevatten            | Стора Лулеваттен        | Швеция         |
| 33 | 43120            | Malaren                     | Меларен                 | Швеция         |
| 34 | 43183            | Vanern                      | Венерн                  | Швеция         |
| 35 | 43443            | Brienzer See                | Бриенцское озеро        | Швейцария      |
| 36 | 44867            | Morar                       | Лох-Морар               | Великобритания |
| 37 | 50652            | Chelan                      | Шелан                   | США            |

Выборка озер вулканического происхождения, состоящая из 129 объектов, показывает 13 водоемов с аномальными значениями в каждом из алгоритмов. При этом совпадающих по всем трем алгоритмам озер 8 (Таблица 4). Так же, как и в предыдущих случаях были обнаружены озера, идентифицируемые только двумя из трех алгоритмов: 2 — алгоритмами *iForest* и *OCSVM*, 1 — алгоритмами *iForest* и *LOF*, и 1 алгоритмами *OCSVM* и *LOF*. Остальные озера были идентифицированы только одним из трех алгоритмов.

Таблица 4

ВУЛКАНИЧЕСКИЕ ОЗЕРА, ПРИЗНАННЫЕ АНОМАЛЬНЫМ  
 ПО РАЗЛИЧНЫМ АЛГОРИТМАМ

| <i>№</i>                | <i>Код озера</i> | <i>Название на латинице</i> | <i>Русское название</i> | <i>Страна</i>  |
|-------------------------|------------------|-----------------------------|-------------------------|----------------|
| 1                       | 928              | Corangamite                 | Корангамайт             | Австралия      |
| 2                       | 10564            | Todos Los Santos            | Тодос-лос-Сантос        | Чили           |
| 3                       | 10988            | Jingbo hu                   | Цзинбо                  | Китай          |
| 4                       | 14368            | Ziway                       | Зваи (Цваи)             | Эфиопия        |
| 5                       | 18422            | Toba                        | Тоба                    | Индонезия      |
| 6                       | 19082            | Tazawa-ko                   | Тадзава                 | Япония         |
| 7                       | 19307            | Shikotsu-ko                 | Сикоцу                  | Япония         |
| 8                       | 26085            | Taupo                       | Таупо                   | Новая Зеландия |
| <i>iForest и OC-SVM</i> |                  |                             |                         |                |
| 1                       | 19056            | Towada-ko                   | Товада                  | Япония         |

| №                    | Код озера | Название на латинице | Русское название   | Страна             |
|----------------------|-----------|----------------------|--------------------|--------------------|
| 2                    | 27444     | Taal                 | Тааль              | Филипины           |
| <i>iForest u LOF</i> |           |                      |                    |                    |
| 1                    | 27344     | Wisdom               | Уисдом             | Новая Гвинея\Папуа |
| <i>OCSVM u LOF</i>   |           |                      |                    |                    |
| 1                    | 44142     | Kanyangeye           | Каньянгейе         | Уганда             |
| <i>iForest</i>       |           |                      |                    |                    |
| 1                    | 18416     | Singkarak            | Сингкарарак        | Индонезия          |
| 2                    | 18421     | Ranau                | Ранану             | Индонезия          |
| <i>OCSVM</i>         |           |                      |                    |                    |
| 1                    | 15565     | Bugdasheni           | Бугдашени          | Грузия             |
| 2                    | 18339     | Pakis                | Пакис              | Индонезия          |
| <i>LOF</i>           |           |                      |                    |                    |
| 1                    | 17650     | Yojoa (Taulebe)      | Ея (Таулебе)       | Гондурас           |
| 2                    | 24590     | Terkhin Tzagan-Nuur  | Терхин-Тцаган-Нуур | Монголия           |

Экспертная оценка показывает, что большинство выделенных озер могут рассматриваться как уникальные с учетом и других признаков, характеризующих озерные экосистемы.

Известно, что самым глубоким карстовым озером считается Чрвено (Красное), одно из Имотских озер Хорватии, глубина которого оценивается в 287 метров [15]. Озеро лежит в глубокой карстовой воронке, сформированной в результате разрушения подземной пещеры. В озере обитает эндемичная рыба *Delminichthys adspersus*, в сухое время года встречающаяся также и в окрестных реках и озерах, что предполагает наличие между ними подземных протоков. Среди хорватских озер интересно Вранское озеро глубиной 74 м, расположенное на о-ве Црес и представляющее собой затопленную водой криптодепресссионную впадину [16]. Глубочайшая точка озера лежит на 58 метров ниже уровня моря. В итоговый список попало и искусственное озеро Бушко Блато, расположенное на территории Боснии и Герцеговины в карстовом Ливанском поле, характеризующимся множеством уникальных природных явлений и карстовых особенностей.

Среди карстовых озер России — озеро Провальное, природный колодец карстово-тектонического происхождения на южном склоне горы Машук в Пятигорске, лежащий на дне конусообразной воронки высотой 41 м [17]. Глубина озера составляет 11 м, красивый бирюзовый цвет, наполняющей водоем минеральной воды, обусловлен высоким содержанием серы и присутствием в ней определенных бактерий. Расположенное на дне заболоченной эрозионно-карстовой котловин оз. Чусовское (площадь зеркала 27,8 км<sup>2</sup>) находится в Пермском крае и включено в ООПТ «Озеро Чусовское». Озеро является местом скопления водоплавающих и околоводных птиц во время весенних и осенних миграций. Гнездятся виды, занесенные в Красную книгу, такие как скопа и чернозобая гагара.

В список аномальных озер ледникового происхождения попало большое количество хорошо известных, крупнейших озер Канадского кристаллического щита, среди которых: Б. Медвежье, Б. Невольничье, Виннипег, Атабаска, Оленье, а также Балтийского кристаллического щита: Венерн, Меларен, Мьеса и Сайма. Среди российских озер необычными оказались морфометрические характеристики озер Ильмень, Селигер и Псковско-Чудского, все они были расположены на периферии последнего четвертичного оледенения и характеризуются небольшими глубинами. Озеро Ильмень находится на месте огромного древнего Приильменского водоема и в современный период представляет собой

мелководный водоем с плоским дном, размеры и форма которого сильно изменяются вследствие значительных внутригодовых и межгодовых колебаний уровня воды в условиях плоской низменной поймы. При высоких уровнях берега озера затопляются на протяжении 2–15 км, а его площадь может увеличиваться вдвое, что крайне редко для озер ледникового происхождения. Озеро Селигер примечательно своей формой, это не совсем озеро в привычном понятии, а скорее цепочка озер, протянувшихся с севера на юг на 100 км и связанных между собой короткими узкими протоками. Среди южноамериканских в список вошли, напротив, наиболее глубокие озера. Интересны оз. Сан-Мартин — глубочайшее (максимальная глубина 836 м) в Южной Америке и оз. Буэнос Айрес с максимальной глубиной — 586 м [18]. Основной отток из озера Буэнос-Айрес происходит по р. Бейкер в Тихий океан, однако существует и неустойчивый отток в направлении р. Рио-Десеадо, несущей свои воды в Атлантический океан.

Аномалии выявлены приблизительно в равной пропорции среди всех рассмотренных подвидов озерных котловин, имеющих вулканическое происхождение (кратерные, лавово-подпрудные, смешанного происхождения). Среди кратерных озер выделены водоемы, занимающие наиболее интересные с геологической точки зрения кальдеры — Тобо и Таупо, расположенные в кальдерах вулканов, извержения которых считаются наиболее разрушительными в истории Земли [19]. Весьма необычно по целому ряду параметров попавшие в итоговый список озеро смешанного происхождения Тодос Лос Сантос. Это одно из немногих в мире озер ледниково-вулканического происхождения, в то время как большинство озер смешанного происхождения обязаны появлением своих котловин наряду с вулканическими также тектоническим процессам. Интересно озеро Сикоцу, которое практически никогда не замерзает и является самым северным незамерзающим водоемом Японии. В случае достаточно древних котловин, расположенных в вулканических регионах как у оз. Корангамайт, попадание озера в итоговый список могло быть обусловлено значительным преобразованием его котловины с течением времени, которое отразилось на изменении характерной для вулканического озера формы, улавливаемой с применением проведенного нами анализа.

Экспертная оценка показала, что наиболее необычные озера были выделены всеми использованными методами, однако значительный интерес представляют и ряд водоемов, которые были идентифицированы лишь одним или двумя методами. Например, с помощью методов *iForest* и *OCSVM* были идентифицированы такие интереснейшие вулканические озера как Тааль и Товада-ко. Озеро Тааль расположено в крупной вулканической кальдере на острове Луссон на Филиппинах. Его котловина была образована очень большим извержением примерно 500 000-100 000 лет назад. В 1967 году бассейн озера был объявлен национальным парком. В озере найдено 4 эндемичных вида рыб. Озеро Товада-ко является самым крупным кратерным озером на о-ве Хонсю и третьим по глубине озером Японии. Благодаря своей необычности оно также расположено в национальном парке. Методом *iForest* были идентифицированы озера Сингкарарк и Ранау, а двумя методами, *iForest* и *LOF*, — озеро Уисдом. Все эти озера наряду с морфометрическими характеристиками интересны и по своим гидрохимическим свойствам. Для озер Сингкарарк и Ранау характерно наличие значительного по объему аноксичного гипоплимниона, на глубинах обнаруживается гидроген сульфид [20]. Для озера Уисдом (глубина 360 м) напротив, характерна относительно высокая концентрация кислорода в его глубинной части. На данной выборке по экспертной оценке, метод *iForest* показал свою наилучшую применимость при идентификации аномальных озер вулканического происхождения. Как показала экспертиза, все водоемы, выделенные с его помощью, оказались уникальными не только по своим морфометрическим характеристикам.

### Заключение

Проведенное исследование свидетельствует, что использование математического аппарата поиска выбросов (аномалий) приводит к хорошим результатам при попытке идентификации озер, обладающих уникальными свойствами, и может служить дополнением и объективным подтверждением к экспертным оценкам, используемым в настоящее время при обосновании необходимости придания объекту охранного статуса. Наш анализ базировался на широких выборках, охватывающих различные регионы мира и позволил идентифицировать ряд наиболее известных и необычных водных объектов. При ограничении выборки конкретным небольшим регионом с помощью предложенного аппарата можно получить список озер необычных для данной конкретной территории. В дальнейшем такие водоемы могут рассматриваться, как претенденты на охранный статус.

*Финансирование:* Работа выполнена в рамках проекта РФФИ 20-05-00303\22.

*Financing:* The work was carried out within the framework of RFBR project 20-05-00303\22.

### Список литературы:

1. Поздняков Ш. Р., Измайлова А. В., Расулова А. М. Уникальные озера как объект научного интереса // Известия Русского географического общества. 2020. Т. 152. №3. С. 17-31. <https://doi.org/10.31857/S0869607120030088>
2. Измайлова А. В. Расулова А. М., Шмакова В. Ю. Выделение озер, обладающих уникальными свойствами, статистическими методами // Гидрометеорология и экология. 2021. №62. С. 27-51. <https://doi.org/10.33933/2074-2762-2021-62-27-51>
3. Расулова А. М., Измайлова А. В. Применение алгоритма Isolation Forest для обоснования уникальности водоемов в группе карстовых озер // Бюллетень науки и практики. 2021. Т. 7. №11. С. 63–79. <https://doi.org/10.33619/2414-2948/72/08>
4. Расулова А. М., Измайлова А. В. Методы поиска аномальных характеристик озерных экосистем на примере трансграничных водоемов // Трансграничные водные объекты: использование, управление, охрана: Материалы Всероссийской научно-практической конференции. Новочеркасск: Лик, 2021. С. 309–314.
5. Рянжин С. В., Ульянова Т. Ю. Геоинформационная система “Озера мира”-GIS WORLDLAKE // Доклады Академии наук. 2000. Т. 370. №4. С. 542-545.
6. Кочков Н. В., Рянжин С. В. Озера мира WORLDLAKE. Свидетельство о государственной регистрации базы данных № 2015621549.
7. Vox G. E. P., Cox D. R. An analysis of transformations // Journal of the Royal Statistical Society: Series B (Methodological). 1964. V. 26. №2. P. 211-243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
8. Breunig M. M., Kriegel H. P., Ng R. T., Sander J. LOF: identifying density-based local outliers // Proceedings of the 2000 ACM SIGMOD international conference on Management of data. 2000. P. 93-104. <https://doi.org/10.1145/342009.335388>
9. Alghushairy O., Alsinir R., Soule T., Ma X. A review of local outlier factor algorithms for outlier detection in big data streams // Big Data and Cognitive Computing. 2020. V. 5. №1. P. 1. <https://doi.org/10.3390/bdcc5010001>
10. Liu F. T., Ting K. M., Zhou Z. H. Isolation forest // 2008 eighth IEEE international conference on data mining. IEEE, 2008. P. 413-422. <https://doi.org/10.1109/ICDM.2008.17>

11. Liu F. T., Ting K. M., Zhou Z. H. Isolation-based anomaly detection // ACM Transactions on Knowledge Discovery from Data (TKDD). 2012. V. 6. №1. P. 1-39. <https://doi.org/10.1145/2133360.2133363>
12. Smola A. J., Schölkopf B. A tutorial on support vector regression // Statistics and computing. 2004. V. 14. №3. P. 199-222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
13. Cortes V. Cortes C., Vapnik V // Support-vector networks, Machine learning. 1995. V. 20. №3. P. 273-297. <https://doi.org/10.1007/BF00994018>
14. Tax D. M. J., Duin R. P. W. Support vector data description // Machine learning. 2004. V. 54. №1. P. 45-66. <https://doi.org/10.1023/B:MACH.0000008084.60811.49>
15. Ozimec R., Jalžić B., Jelić D. Preliminarni izvještaj prirodoslovnih istraživanja u okviru Speleoronilačke ekspedicije Crveno jezero 2017 // Subterranea Croatica. 2017. V. 15. №2. P. 34-41.
16. Katalinic A., Rubinic J., Buselic G. Hydrology of two coastal karst cryptodepressions in Croatia: Vrana lake vs Vrana lake // Proceedings of Taal2007: The 12th World Lake Conference. 2008. V. 732. P. 743.
17. Ефремов Ю. В. Горные озера Западного Кавказа. Л.: Гидрометеиздат, 1984. 200 с.
18. Данилов-Данильян В. И. Реки и озера мира: энциклопедия. М.: Энциклопедия, 2012. 927 с.
19. Wilson C. J. N., Walker G. P. L. The Taupo eruption, New Zealand I. General aspects // Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences. 1985. V. 314. №1529. P. 199-228. <https://doi.org/10.1098/rsta.1985.0019>
20. Gopal B., Wetzel R. G. Limnology in developing countries. 1995.

#### References:

1. Pozdnyakov, Sh. R., Izmailova, A. V., & Rasuloва, A. M. (2020). Unique lakes as an object of scientific interest. *Proceedings of the Russian Geographical Society*, 152(3), 17-31. (in Russian). <https://doi.org/10.31857/S0869607120030088>
2. Izmailova, A. V. Rasuloва, A. M., & Shmakova, V. Yu. (2021). Identification of lakes with unique properties using statistical methods. *Gidrometeorologiya i Ekologiya. Journal of Hydrometeorology and Ecology [Proceedings of the Russian State Hydrometeorological University]*, (62), 27-51. (in Russian). <https://doi.org/10.33933/2074-2762-2021-62-27-51>
3. Rasuloва, A., & Izmailova, A. (2021). Application of the Isolation Forest Algorithm to Substantiate the Uniqueness of Water Bodies in the Group of Karst Lakes. *Bulletin of Science and Practice*, 7(11), 63-79. (in Russian). <https://doi.org/10.33619/2414-2948/72/08>
4. Rasuloва, A. M., & Izmailova, A. V. (2021). Metody poiska anomal'nykh kharakteristik ozernykh ekosistem na primere transgranichnykh vodoemov. In *Transgranichnye vodnye ob'ekty: ispol'zovanie, upravlenie, okhrana: Materialy Vserossiiskoi nauchno-prakticheskoi konferentsii. Novoчеркассk: Lik*, 309–314. (in Russian).
5. Ryzanzhin, S. V., & Ul'yanova, T. Yu. (2000). Geoinformatsionnaya sistema "Ozera mira"-GIS WORLDLAKE. *Doklady Akademii nauk*, 370(4), 542-545.
6. Kochkov, N. V., & Ryzanzhin, S. V. Ozera mira WORLDLAKE. Svidetel'stvo o gosudarstvennoi registratsii bazy dannykh № 2015621549.
7. Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211-243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
8. Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000, May). LOF: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data* (pp. 93-104). <https://doi.org/10.1145/342009.335388>

9. Alghushairy, O., Alsini, R., Soule, T., & Ma, X. (2020). A review of local outlier factor algorithms for outlier detection in big data streams. *Big Data and Cognitive Computing*, 5(1), 1. <https://doi.org/10.3390/bdcc5010001>
10. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth IEEE international conference on data mining* (pp. 413-422). IEEE. <https://doi.org/10.1109/ICDM.2008.17>
11. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 6(1), 1-39. <https://doi.org/10.1145/2133360.2133363>
12. Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3), 199-222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
13. Cortes, C., Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273-297. <https://doi.org/10.1007/BF00994018>
14. Tax, D. M., & Duin, R. P. (2004). Support vector data description. *Machine learning*, 54(1), 45-66. <https://doi.org/10.1023/B:MACH.0000008084.60811.49>
15. Ozimec, R., Jalžić, B., & Jelić, D. (2017). Preliminarni izvještaj prirodoslovnih istraživanja u okviru Speleoronilačke ekspedicije Crveno jezero 2017. *Subterranea Croatica*, 15(2), 34-41.
16. Katalinic, A., Rubinic, J., & Buselic, G. (2008). Hydrology of two coastal karst cryptodepressions in Croatia: Vrana lake vs Vrana lake. In *Proceedings of Taal2007: The 12th World Lake Conference* (Vol. 732, p. 743).
17. Efremov, Yu. V. (1984). Gornye ozera Zapadnogo Kavkaza. Leningrad. (in Russian).
18. Danilov-Danil'yan, V. I. (2012). Reki i ozera mira: entsiklopediya. Moscow. (in Russian).
19. Wilson, C. J. N., & Walker, G. P. L. (1985). The Taupo eruption, New Zealand I. General aspects. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 314(1529), 199-228. <https://doi.org/10.1098/rsta.1985.0019>
20. Gopal, B., & Wetzel, R. G. (1995). Limnology in developing countries.

Работа поступила  
в редакцию 13.11.2022 г.

Принята к публикации  
19.11.2022 г.

Ссылка для цитирования:

Расулова А. М., Измайлова А. В. Идентификация уникальных озер различного происхождения методами машинного обучения // Бюллетень науки и практики. 2022. Т. 8. №12. С. 180-194. <https://doi.org/10.33619/2414-2948/85/23>

Cite as (APA):

Rasulova, A., & Izmailova, A. (2022). Identification of Unique Lakes of Different Origin by Machine Learning Methods. *Bulletin of Science and Practice*, 8(12), 180-194. (in Russian). <https://doi.org/10.33619/2414-2948/85/23>