

МЕТОДОЛОГИЧЕСКИЕ ВОПРОСЫ ИСПОЛЬЗОВАНИЯ РАНДОМИЗИРОВАННОГО МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ДИНАМИКИ ТЕРМОКАРСТОВЫХ ОЗЕР АРКТИКИ

Ю.С. Попков¹, З. Волкович², А.В. Мельников³, Ю.М. Полищук³

¹ Институт системного анализа ФИЦ ИУ РАН, г. Москва, Россия,

² Колледж ОРТ им. Брауде, г. Кармиэль, Израиль,

³ Югорский НИИ информационных технологий, г. Ханты-Мансийск, Россия

Проведен анализ состояния проблем моделирования пространственно-временной динамики озерных полей в условиях современных климатических изменений. Показано, что аналитические методы, используемые при изучении динамики термокарстовых процессов в отдельных озерах, не подходят для изучения пространственно-временных изменений полей термокарстовых озер. Предложенный метод гео-имитационного моделирования для изучения динамики полей термокарстовых озер не обеспечивает достаточной точности прогнозирования. Рассмотрены проблемы применения нового подхода к прогнозированию пространственно-временной динамики полей в условиях современных климатических изменений на основе методов и алгоритмов энтропийно-рандомизированного машинного обучения. Экспериментальные результаты удаленных исследований динамики полей термокарстовых озер в арктической зоне вечной мерзлоты Западной Сибири были получены с использованием спутниковых снимков за период в несколько десятилетий, начиная с 1973 года. Климатические данные за тот же период были получены путем реанализа на основе хорошо известных систем ERA-40, ERA-Interim и APHRODITE JMA. Был составлен массив экспериментальных данных об изменениях площади озер, среднегодовой температуры и годовых осадков в зоне вечной мерзлоты Западной Сибири за период исследований. Регрессионный анализ геокриологических и климатических данных показал, что сокращение площади озер можно объяснить, главным образом, увеличением температуры поверхности и изменением осадков. Структура модели рандомизированного прогноза динамики полей термокарстовых озер определяется с учетом параметров, отражающих изменения площади озер, среднегодовой температуры и уровня осадков. Рассмотрены особенности использования экспериментальных данных в рамках энтропийно-рандомизированного подхода к прогнозированию пространственно-временной динамики полей термокарстовых озер в условиях современных климатических изменений.

Ключевые слова: машинное обучение, энтропийно-рандомизированный подход, рандомизированная модель, прогнозирование, пространственно-временная динамика, вечная мерзлота, термокарстовые озера, спутниковые снимки, реанализ метеоданных, климатические изменения, глобальное потепление.

Введение

Особую важность рандомизированный подход представляет для решения задач прогнозирования динамики накопления парниковых газов в термокарстовых озерах арктической зоны в связи с их влиянием на глобальные климатические изменения, что может явиться основой разработки и функционирования систем адаптации к меняющимся условиям среды обитания на различных управленческих уровнях. Потепление климата в ближайшие десятилетия будет приводить к таянию мерзлых пород и дополнительному высвобождению углекислого газа, а также метана как продукта жизнедеятельности микроорганизмов, перерабатывающих оттаявшее органическое вещество. Это способно внести дополнительный ощутимый вклад в потепление климата, что вызывает озабоченность мировой общественности. Осознание этого послужило поводом для принятия

Парижского соглашения по климату (Париж, 12 декабря 2015), направленного на разработку мер по ограничению роста среднегодовой температуры Земли до $1,5\text{ }^{\circ}\text{C}$ до 2050 года. Разработка таких мер на региональном уровне для арктических регионов невозможна без формирования обоснованных прогнозов объемов эмиссии метана и углекислого газа, которые могут быть основаны на знаниях о пространственно-временной динамике полей термокарстовых озер как важных источников эмиссии природного метана и углекислого газа на территории регионов. Известно, что наибольший вклад в глобальную эмиссию парниковых газов вносят территории вечной мерзлоты арктических регионов, что подтверждается многочисленными экспериментальными исследованиями в Сибири и в других северных регионах [1–8].

Рассмотрим вопросы моделирования пространственно-временной динамики полей термокарстовых озер. Обзор результатов исследования термокарстовых процессов с использованием ряда аналитических моделей, приведенный в [9], показал, что эти модели позволяют достаточно эффективно исследовать процессы в одиночных термокарстовых озерах. Использование аналитических моделей для моделирования пространственно-временной структуры полей термокарстовых озер, характерных для арктических территорий, в литературе не рассматриваются. В связи с этим определенный интерес представляют разработанные в [10] методы математической морфологии ландшафта, позволяющие проводить точный теоретико-математический анализ динамики термокарстово-озерных равнин. Методы и модели математической морфологии ландшафта могут использоваться для оценки геоэкологических рисков при развитии термокарстовых процессов в условиях продолжающегося потепления климата. Однако эти методы также не позволяют исследовать пространственно-временные изменения полей термокарстовых озер.

В [11, 12] рассмотрены методы и средства моделирования динамики полей термокарстовых озер на основе геоимитационного подхода, позволяющие прогнозировать динамику полей термокарстовых озер и объемов накопления метана и углекислого газа в озерах Арктики. Однако разработанные средства геоимитационного моделирования использовали традиционные математические методы (в частности, регрессионный анализ данных о взаимосвязи климатических и геокриологических изменений), что не позволило обеспечить достаточно высокую точность результатов моделирования и прогнозирования динамики термокарстовых озер. Так, исследования адекватности геоимитационной модели на основе сравнения с экспериментальными данными для арктической зоны Западной Сибири показали [9], что погрешность моделирования в среднем составила около 18 %. Поэтому возникает необходимость в поиске нового подхода к моделированию пространственно-временной динамики полей термокарстовых озер, который позволил бы получать более точные прогнозные оценки.

По нашему мнению, наиболее перспективным для разработки эффективной модели представляется подход, основанный на методах рандомизированного машинного обучения, которые показали высокую эффективность в решении проблем прогнозирования мировой экономики, демографии и др. [13]. Предлагаемые методы позволяют получить энтропийно-оптимальную апостериорную плотность распределения вероятностей оцениваемых параметров модели при «наихудших», в терминах энтропии, шумах измерений. Такой подход позволяет получать состоятельные и эффективные оценки параметров модели в условиях ограниченных объемов эмпирических данных. Однако вопросы применения рандомизированного подхода к моделированию пространственно-временной динамики полей термокарстовых озер в настоящее время не разработаны, что определяет важность и актуальность поставленной задачи. В связи с этим целью настоящей работы является рассмотрение методологических вопросов использования этого подхода для задач прогнозирования пространственно-временной динамики полей термокарстовых озер на основе анализа экспериментальных данных о свойствах озерных полей, полученных дистанционными методами по спутниковым снимкам.

1. Вопросы энтропийно-рандомизированного моделирования пространственно-временной динамики полей термокарстовых озер

Проблема извлечения полезной информации, под которой понимается оценивание параметров модели и прогнозирование с использованием реальных данных, является одной из основных в современной науке. В разных научных дисциплинах разработаны многочисленные методы решения этой проблемы, основанные на определенных для каждого метода гипотезах о свойствах

данных, собранных в процессе нормального функционирования исследуемого объекта. Наиболее продвинутыми в этом направлении являются математическая статистика, эконометрика, финансовая математика, теория управления, распознавание образов и другие научные дисциплины. Методы, разработанные в указанных научных дисциплинах, базируются на двух группах фундаментальных гипотез, одна из которых относится к моделям, а другая – к данным. Относительно моделей предполагается, что они имеют вполне определенные параметры (будем называть их детерминированными), значения которых неизвестны и не доступны прямому измерению. Вторая группа гипотез, касающихся данных, более существенна, так как эти гипотезы формулируются в терминах статистических свойств массивов данных, например, достаточное их количество, свойство выборки из генеральной совокупности, нормальность плотностей распределения вероятностей. Наличие этих свойств и предположений, касающихся моделей, в конкретных задачах проверить практически невозможно. Такая ситуация характерна, прежде всего, для задач, в которых объемы массивов реальных данных скудны, а сами данные содержат ошибки. Это полностью относится к задачам моделирования и прогнозирования динамики полей термокарстовых озер. В этих условиях оценивание значений характеристик (параметров) модели осуществляется по малому количеству не вполне достоверных данных. Если относиться к ним как к случайным объектам, то и оценки характеристик модели приобретают свойства случайных переменных.

Поэтому естественным образом возникает предложение рассматривать параметры модели как случайные величины. Это предложение приводит к трансформации модели с детерминированными параметрами в модель со случайными параметрами, которую будем, согласно [13], называть рандомизированной моделью (РМ). Характеристиками РМ являются функции плотности распределения вероятностей (ПРВ) параметров. Поэтому для РМ нужно определять на основе экспериментальных данных оценки ПРВ параметров, а не оценки их значений. Имея эти оценки, можно использовать РМ для прогнозирования, если уметь генерировать последовательности случайных векторов с заданными ПРВ.

В данной работе предлагается использовать принцип максимизации информационной энтропии как меры неопределенности. Ее максимизация гарантирует получение наилучших решений при максимальной неопределенности. Поскольку информационная энтропия характеризует неопределенность, связанную не только со случайными параметрами РМ, но с шумами наблюдений, то последнее ее качество гарантирует получение наилучших оценок для максимально неопределенных (в единицах энтропии) шумов. Таким образом, получаемые в результате максимизации информационной энтропии оценки ПРВ можно трактовать как робастные.

Прогнозирование в виде рандомизированных ансамблей – новая парадигма в прогнозировании, альтернативная классическому понятию детерминированного прогноза. Рандомизированный ансамбль содержит прогнозные траектории с оптимальными оценками параметров. Здесь оптимальность понимается в энтропийных терминах. Важная проблема в рандомизированном прогнозировании – генерация последовательностей случайных векторов с заданной ПРВ (в частности, энтропийно-оптимальной ПРВ).

2. Анализ экспериментальных данных о пространственно-временной динамике полей термокарстовых озер по спутниковым снимкам

Информационной основой для построения рандомизированной модели динамики полей термокарстовых озер являются экспериментальные данные об изменениях численности и площадей термокарстовых озер в арктической зоне. Получение этих данных ввиду труднодоступности и высокой степени заболоченности территорий в зоне мерзлоты невозможно без применения дистанционных методов. Исследуемая территория в арктической зоне Западной Сибири имеет площадь около 1,05 млн км². Озера в зоне мерзлоты характеризуются чрезвычайно широким интервалом изменений их размеров – от единиц м² до сотен тысяч гектаров. В наших дистанционных исследованиях [1–3, 14–16] использовались оптические и радарные снимки среднего разрешения Landsat, ERS-2, ENVISAT, снимки высокого разрешения Канопус-В, БКА, Alos и снимки сверхвысокого разрешения Ресурс-П, QuickBird, GeoEye-1. Для минимизации влияния сезонных колебаний уровня воды в озерах снимки выбирались в достаточно короткий период летнего сезона (июль – август). В этот период полностью исчезает ледовый покров на озерах, мешающий их выделению при дешифрировании снимков. Обработка космических снимков, проведенная с исполь-

зованием стандартных средств геоинформационных систем ArcGIS и ENVI, была направлена на получение данных о динамике численности и площадей озер и их пространственном распределении.

Рассмотрим результаты исследования пространственной структуры полей озер. Исследования статистики параметров местоположения озер проведены на исследуемой территории с использованием снимков Landsat. Результаты исследования законов распределения координат центров термокарстовых озер приведены в [9]. Анализ показал, что гистограммы распределения координат центров озер на плоскости, согласно критерию χ^2 , соответствуют закону равномерной плотности с вероятностью 95 %.

Важным показателем пространственной структуры полей озер является распределение озер по размерам. Дистанционным исследованиям пространственной структуры полей озер в зонах мерзлоты различных регионов мира посвящены большое число работ, в которых отмечается, что распределение озер по размерам соответствует степенному закону. В большинстве проведенных исследований использовались космические снимки среднего разрешения Landsat (разрешение 30 м), на которых однако не обнаруживаются озера малых размеров. Такие озера обладают повышенной концентрацией метана, значительно превышающей его концентрацию в озерах больших размеров [4]. Поэтому для изучения закономерностей распределения термокарстовых озер в широком диапазоне размеров с учетом вклада малых озер были применены спутниковые снимки сверхвысокого разрешения в сочетании со снимками среднего разрешения. С использованием полученных данных о числе и площадях озер была построена гистограмма распределения площадей озер в чрезвычайно широком диапазоне изменения их размеров (от 5 до 2×10^8 м²). Статистический анализ гистограммы распределения озер по размерам с использованием критерия Пирсона показал, что гистограмма распределения соответствует логнормальному закону с высокой вероятностью 0,99. Заметим, что оценки запасов метана и углекислого газа в малых озерах (и особенно в сверхмалых) при обычно используемом для этих целей степенном законе распределения [17–19] оказываются значительно завышенными по сравнению с логнормальным законом [2].

Анализ временных рядов суммарных площадей термокарстовых озер на каждом тестовом участке позволил определить линейные тренды изменений средней площади озер. Коэффициент линейного тренда, согласно [9], на разных ТУ за период более 40 лет, начиная с 1973 г., изменяется в пределах от $-0,28$ га/год до $+0,003$ га/год со средним значением $-0,09$ га/год, т. е. в среднем в зоне мерзлоты Западной Сибири наблюдается сокращение площади озер. Так, за период исследований, начиная с 1973 г., общая площадь озер в зоне мерзлоты Западной Сибири в среднем сократилась [3, 9] приблизительно на 8 %.

Климатические данные за тот же период были получены путем реанализа на основе хорошо известных систем ERA-40, ERA-Interim и APHRODITE JMA. Временной ряд среднегодовых значений температуры воздуха по данным реанализа показал рост среднегодовой температуры со временем как результат потепления. Установлен линейный тренд временного ряда среднегодовой температуры с коэффициентом тренда, равным $0,051$ °C/год.

В [9] был составлен массив экспериментальных данных об изменениях площадей озер, среднегодовой температуры и годовых осадков в зоне вечной мерзлоты Западной Сибири за период исследований. Исследование корреляционных связей изменений площадей озер, среднегодовой температуры и суммы осадков [9] позволило установить наличие взаимосвязи между размерами озер и климатическими изменениями. Регрессионный анализ массива этих данных показал [9, 12] преобладающее влияние роста температуры в изменениях площадей озер.

На основе экспериментально установленных основных свойств полей термокарстовых озер применительно к прогнозированию пространственно-временной динамики термокарстовых озер в условиях климатических изменений используется разработанная в [13] методика рандомизированного моделирования динамики площадей озер, основанная на генерации оптимизированного ансамбля предсказательных траекторий. В нашей работе предлагается динамическая модель марковского типа со случайными параметрами, характеризующая изменение размеров озер под влиянием изменений среднегодовых температур воздуха и годовой суммы осадков. Для идентификации модели будет использована процедура энтропийно-рандомизированного машинного обучения, позволяющая построить оценки плотностей распределения вероятностей параметров и

шумов измерений с использованием рассмотренных выше данных о динамике площадей озер со спутниковых снимков и об изменениях климатических параметров за период последних четырех десятилетий.

Выводы

Опыт использования геоимитационного подхода к моделированию и прогнозированию пространственно-временной динамики полей термокарстовых озер в зоне мерзлоты показал недостаточную точность прогноза из-за ряда особенностей экспериментальных данных, в частности, из-за значительного числа пропущенных значений во временных рядах площадей озер вследствие большого числа пасмурных дней на территориях мерзлоты, не позволяющих проводить регулярную космическую съемку поверхности. Поэтому единственной альтернативой выступает энтропийно-рандомизированный подход к прогнозированию динамики озерных полей на основе данных спутниковых измерений численности и площадей озер. Проведен комплексный анализ экспериментальных данных о климатических показателях (среднегодовая температура поверхности и годовая сумма осадков) и изменениях площадей озер по спутниковым данным на территории мерзлоты Западной Сибири. Установлены закономерности случайного распределения местоположения центров озер и их размеров, с учетом которых предлагается рандомизированная модель динамики озерных полей в виде уравнения марковского типа со случайными параметрами, характеризующая изменения размеров озер под влиянием изменений температуры воздуха и уровня осадков. Идентификация модели будет проводиться с использованием разработанной ранее процедуры энтропийно-рандомизированного машинного обучения, позволяющей построить оценки плотностей распределения вероятностей параметров и шумов измерений на основе рассмотренных в работе экспериментальных данных, полученных за период последних десятилетий, начиная с 1973 г.

Исследование проводилось при финансовой поддержке частично грантов Российского фонда фундаментальных исследований по проектам № 19-07-00282, № 18-45-860002, № 18-45-703001, № 18-47-860013 и № 18-47-700001.

Литература

1. Опыт и результаты дистанционного исследования озер криолитозоны Западной Сибири по космическим снимкам различного разрешения за 50-летний период / Ю.М. Полищук, А.Н. Богданов, Н.А. Брыксина и др. // *Современные проблемы дистанционного зондирования Земли из космоса*. – 2017. – Т. 14. – № 6. – С. 42–55. DOI: 10.21046/2070-7401-2017-14-6-42-55
2. *Minor contribution of small thaw ponds to the pools of carbon and methane in the inland waters of the permafrost - affected part of the Western Siberian lowland* / Y.M. Polishchuk, A.N. Bogdanov, I.N. Muratov et al. // *Environmental Research Letters*. – 2018. – Vol. 13. – P. 1–16. DOI: 10.1088/1748-9326/aab046
3. *Hydrological Changes: Historical Analysis, Contemporary Status, and Future Projections* / A.I. Shiklomanov, R.B. Lammers, D.P. Lettermaier et al. // *Regional Environmental Changes in Siberia and Their Global Consequences* / Eds. P.Ya. Groisman and G. Gutman. – Dordrecht; Heidelberg; New-York; London: Springer, 2013. – P. 111–154. DOI: 10.1007/978-94-007-4569-8_4
4. *Effect of permafrost thawing on the organic carbon and metal speciation in thermokarst lakes of Western Siberia* / O.S. Pokrovsky, L.S. Shirokova, S.N. Kirpotin et al. // *Biogeosciences*. – 2011. – Vol. 8. – P. 565–583. DOI: 10.5194/bg-8-565-2011
5. Кравцова, В.И. Изменение размеров термокарстовых озер в различных районах России за последние 30 лет / В.И. Кравцова, А.Г. Быстрова // *Криосфера Земли*. – 2009. – Т. 13, № 2. – С. 16–26.
6. Kirpotin, S. *Abrupt changes of thermokarst lakes in Western Siberia: impacts of climatic warming on permafrost melting* / S. Kirpotin, Y. Polishchuk, N. Bryksina // *International Journal of Environmental Studies*. – 2009. – Vol. 66. – No. 4. – P. 423–431. DOI: 10.1080/00207230902758287
7. Riordan, B. *Shrinking ponds in subarctic Alaska based on 1950–2002 remotely sensed images* / B. Riordan, D. Verbyla, A.D. McGuire // *J. Geophys. Res.* – 2006. – Vol. 111. – G04002. DOI: 10.1029/2005JG000150

8. Zuidhoff, F.S. Changes in palsa distribution in relation to climate change in Laivadalen, Northern Sweden, especially 1960-1997 / F.S. Zuidhoff, E. Kolstrup // *Permafrost and Periglacial Processes*. – 2000. – Vol. 11. – P. 55–69. DOI: 10.1002/(SICI)1099-1530(200001/03)11:1<55::AID-PPP338>3.0.CO;2-T
9. Полищук, В.Ю. Геоимитационное моделирование полей термокарстовых озер в зонах мерзлоты / В.Ю. Полищук, Ю.М. Полищук. – Ханты-Мансийск: УИП ЮГУ, 2013. – 129 с.
10. Викторов, А.С. Основные проблемы математической морфологии ландшафтов / А.С. Викторов. – М.: Наука, 2006. – 252 с.
11. Polishchuk, Y. Forecast of thermokarst lakes dynamics in permafrost based on geo-simulation modeling and remote sensing data / Y. Polishchuk, V. Polishchuk // *Proc. of Conf. “Mathematical and Information Technologies MIT-2016” (Vrnjacka Banja, Serbia – Budva, Montenegro, Aug. 28 – Sept. 5, 2016)* / Eds. Yu. Shokin, H. Miloshevich and D. Esipov. Published on CEUR-Workshop Proceedings. – 2017. – Vol. 1839. – P. 393–405. – <http://ceur-ws.org/Vol-1839>.
12. Polishchuk, V.Y. Modeling of thermokarst lake dynamics in West-Siberian permafrost. Chapter 6 / V.Y. Polishchuk, Y.M. Polishchuk // *Permafrost: Distribution, Composition and Impacts on Infrastructure and Ecosystems* / Ed. by O. Pokrovsky. – New York: Nova Science Publishers, 2014. – P. 205–234.
13. Попков, Ю.С. Рандомизированное машинное обучение при ограниченных объемах данных / Ю.С. Попков, А.Ю. Попков, Ю.А. Дубнов. – М.: ЛЕНАНД. – 320 с.
14. Полищук, Ю.М. Методические вопросы построения обобщенных гистограмм распределения площадей озер в зоне мерзлоты на основе космических снимков среднего и высокого разрешения / Ю.М. Полищук, А.Н. Богданов, И.Н. Муратов // *Современные проблемы дистанционного зондирования Земли из космоса*. – 2016. – Т. 13, № 6. – С. 224–232. DOI: 10.21046/2070-7401-2016-13-6-224-232
15. Size distribution, surface coverage, water, carbon, and metal storage of thermokarst lakes in the permafrost zone of the Western Siberia lowland / Y.M. Polishchuk, A.N. Bogdanov, V.Y. Polishchuk et al. // *Water*. – 2017. – Vol. 9, iss. 3. – 18 p. DOI: 10.3390/w9030228
16. A Canopus-V imagery-based study of the size-distribution of small lakes in the discontinuous permafrost zone of the Western Siberia / Y.M. Polishchuk, A.N. Bogdanov, I.N. Muratov, V.Y. Polishchuk // *Kriosfera Zemli*. – 2017. – Vol. 21. – No. 2. – P. 80–87.
17. Lehner, B. Development and validation of a global database of lakes, reservoirs and wetlands / B. Lehner, P. Doll // *J. Hydrol.* – 2004. – Vol. 296. – P. 1–22. DOI: 10.1016/j.jhydrol.2004.03.028
18. Downing, J.A. The global abundance and size distribution of lakes, ponds, and impoundments / J.A. Downing, Y.T. Prairie // *Limnol. Oceanogr.* – 2006. – Vol. 51. – P. 2388–2397. DOI: 10.4319/lo.2006.51.5.2388
19. A global inventory of lakes based on high resolution satellite imagery / C. Verpoorter, T. Kutser, D.A. Seekel, L.J. Tranvik // *Geophys. Res. Lett.* – 2014. – Vol. 41. – P. 1–7. DOI: 10.1002/2014GL060641

Попков Юрий Соломонович, д-р техн. наук, главный научный сотрудник, Институт системного анализа ФИЦ ИУ РАН, г. Москва; popkov.yuri@gmail.com.

Волкович Зеев, д-р физ.-мат. наук, профессор, Колледж ОРТ им. Брауде, г. Кармиэль, Израиль; vvolkov@braude.ac.il.

Мельников Андрей Витальевич, д-р техн. наук, директор, Югорский НИИ информационных технологий, г. Ханты-Мансийск; MelnikovAV@uriiit.ru.

Полищук Юрий Михайлович, д-р физ.-мат. наук, главный научный сотрудник, Югорский НИИ информационных технологий, г. Ханты-Мансийск; yurpolishchuk@gmail.com.

Поступила в редакцию 5 сентября 2019 г.

METHODOLOGICAL ISSUES OF USING THE RANDOMIZED MACHINE LEARNING FOR FORECASTING THE DYNAMICS OF THERMOKARST ARCTIC LAKES

Yu.S. Popkov¹, popkov.yuri@gmail.com,
Z. Volkovich², vlvolkov@braude.ac.il,
A.V. Melnikov³, MelnikovAV@uriit.ru,
Yu.M. Polishchuk³, yupolishchuk@gmail.com

¹ Institute of System Analysis, Federal Research Center "Information and Control", Moscow, Russian Federation,

² ORT Braude College, Karmiel, Israel,

³ Ugra Research Institute of Information Technologies, Khanty-Mansiysk, Russian Federation

The analysis of the state of the problems of modeling the spatio-temporal dynamics of the lake fields under the conditions of modern climate changes is carried out. It is shown that the analytical methods used in studying the dynamics of thermokarst processes in individual lakes are not suitable for studying the spatiotemporal changes in the fields of thermokarst lakes. The geo-simulation modeling method proposed for studying the dynamics of fields of thermokarst lakes does not provide sufficient forecasting accuracy. The problems of applying a new approach to the prediction of the spatio-temporal dynamics of fields under the conditions of modern climatic changes based on methods and algorithms of entropy-randomized machine learning are considered. The experimental results of remote studies of the dynamics of fields of thermokarst lakes in the Arctic permafrost zone of Western Siberia were obtained using satellite images for the period of several decades starting in 1973. Climatic data for the same period were obtained by reanalysis based on the well-known ERA-40, ERA-Interim systems and APHRODITE JMA. An array of experimental data has been compiled on changes in lake areas, average annual temperature and annual precipitation in the permafrost zone of Western Siberia over the period of research. Regression analysis of geocryological and climatic data showed that the reduction in the area of lakes can be explained mainly by an increase in surface temperature and a change in precipitation. The structure of a randomized forecast model for the dynamics of fields of thermokarst lakes is determined taking into account parameters reflecting changes in lake areas, average annual temperature and precipitation level. The features of using experimental data in the framework of an entropy-randomized approach to forecasting the spatio-temporal dynamics of fields of thermokarst lakes under the conditions of modern climate changes are considered.

Keywords: machine learning, entropy-randomized approach, randomized model, forecasting, spatio-temporal dynamics, permafrost, thermokarst lakes, satellite images, meteorological data reanalysis, climate change, global warming.

References

1. Polishchuk Y.M., Bogdanov A.N., Bryksina N.A., Polishchuk V.Y., Muratov I.N., Kupriyanov M.A., Baisalyamova O.A., Dneprovskaya V.P. [Experience and Results of a Remote Study of Lakes in the Permafrost Zone of Western Siberia from Satellite Images of Various Resolutions over a 50-year Period]. *Current Problems of Remote Sensing of Earth from Space*, 2017, vol. 14, no. 6, pp. 42–55. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85041059818&doi=10.21046%2f2070-7401-2017-14-6-42-55&partnerID=40&md5=ffb8b939f0be5f60e78586b955b48a84>. DOI: 10.21046/2070-7401-2017-14-6-42-55
2. Polishchuk Y.M., Bogdanov A.N., Muratov I.N., Polishchuk V.Y., Lim A., Manasyrov R.M., Shirokova L.S., Pokrovsky O.S. Minor Contribution of Small Thaw Ponds to the Pools of Carbon and Methane in the Inland Waters of the Permafrost-Affected Part of the Western Siberian Lowland. *Environmental Research Letters*, 2018, vol. 13, pp. 1–16. DOI: 10.1088/1748-9326/aab046
3. Shiklomanov A.I., Lammers R.B., Lettermaier D.P., Polishchuk Y.M., Savichev O.G., and Smith L.C. Hydrological Changes: Historical Analysis, Contemporary Status, and Future Projections. *In Book: Regional Environmental Changes in Siberia and Their Global Consequences*. Eds. P.Ya. Groisman and G.Gutman, Dordrecht – Heidelberg – New-York – London: Springer, 2013, pp. 111–154. DOI: 10.1007/978-94-007-4569-8_4
4. Pokrovsky O.S., Shirokova L.S., Kirpotin S.N., Audry S., Viers J., Dupre B. Effect of Perma-

frost Thawing on the Organic Carbon and Metal Speciation in Thermokarst Lakes of Western Siberia. *Biogeosciences*, 2011, vol. 8, pp. 565–583. DOI: 10.5194/bg-8-565-2011

5. Kravtsova V.I., Bystrova A.G., [Size Variation of Thermokarst Lakes in Various Regions of Russia over the Past 30 Years]. *Cryosphere of Earth*, 2009, vol. 13, no. 2, pp. 16–26. (in Russ.)

6. Kirpotin S., Polishchuk Y., Bryksina N. Abrupt Changes of Thermokarst lakes in Western Siberia: Impacts of Climatic Warming on Permafrost Melting. *International Journal of Environmental Studies*, 2009, vol. 66, no. 4, pp. 423–431. DOI: 10.1080/00207230902758287

7. Riordan B., Verbyla D., McGuire A.D. Shrinking Ponds in Subarctic Alaska Based on 1950–2002 Remotely Sensed Images. *J. Geophys. Res.*, 2006, vol. 111, G04002. DOI: 10.1029/2005JG000150

8. Zuidhoff F.S., Kolstrup E. Changes in Palsa Distribution in Relation to Climate Change in Laivadalén, Northern Sweden, Especially 1960–1997. *Permafrost and Periglacial Processes*, 2000, vol. 11, pp. 55–69. DOI: 10.1002/(SICI)1099-1530(200001/03)11:1<55::AID-PPP338>3.0.CO;2-T

9. Polishchuk V.Y., Polishchuk Y.M. *Geoimitatsionnoe modelirovanie polei termokarstovykh ozer v zonakh merzloty* [Geo-Simulation Modeling of Thermokarst Lakes Fields in Permafrost Zones]. Khanty-Mansiysk, UIP YuGU, 2013. 129 p.

10. Viktorov A.S. *Osnovnye problemy matematicheskoi morfologii landshafta* [The Main Problems of Mathematical Morphology of the Landscape]. Moscow, Nauka, 2006. 252 p.

11. Polishchuk Y., Polishchuk V. Forecast of Thermokarst Lakes Dynamics in Permafrost Based on Geo-Simulation Modeling and Remote Sensing Data. *Proc. of Conf. "Mathematical and Information Technologies MIT-2016" (Vrnjacka Banja, Serbia – Budva, Montenegro, Aug. 28 – Sept. 5, 2016)*. Eds. Yu. Shokin, H. Miloshevich and D. Esipov. Published on CEUR-Workshop Proceedings, 2017, vol. 1839, pp. 393–405. Available at: <http://ceur-ws.org/Vol-1839>.

12. Polishchuk V.Y., Polishchuk Y.M. Modeling of Thermokarst Lake Dynamics in West-Siberian Permafrost. Chapter 6. In: *Permafrost: Distribution, Composition and Impacts on Infrastructure and Ecosystems*. Ed. by O. Pokrovsky. New-York, Nova Science Publishers, 2014, pp. 205–234.

13. Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A. *Randomizirovannoe mashinnoye obucheniye s ograniченными данными* [Randomized Machine Learning with Limited Data]. Moscow, LENAND, 2018. 320 p.

14. Polishchuk Yu.M., Bogdanov A.N., Muratov I.N., [Methodological Issues of Constructing Generalized Histograms of the Distribution of lake Areas in the Permafrost Zone Based on Medium and High Resolution Satellite Images]. *Current Problems of Remote Sensing of Earth from Space*, 2016, vol. 13, no. 6, pp. 224–232. (in Russ.) DOI: 10.21046/2070-7401-2016-13-6-224-232

15. Polishchuk Y.M., Bogdanov A.N., Polishchuk V.Y., Manasypov R.M., Shirokova L.S., Kirpotin S.N., Pokrovsky O.S. Size Distribution, Surface Coverage, Water, Carbon, and Metal Storage of Thermokarst Lakes in the Permafrost Zone of the Western Siberia Lowland. *Water*, 2017. vol. 9, iss. 3, 18 p. DOI: 10.3390/w9030228

16. Polishchuk Y.M., Bogdanov A.N., Muratov I.N., Polishchuk V.Y. A Canopus-V Imagery-based Study of the Size-Distribution of Small Lakes in the Discontinuous Permafrost Zone of the Western Siberia. *Cryosphere of Earth*, 2017, vol. 21, no. 2, pp. 80–87.

17. Lehner B., Doll P. Development and Validation of a Global Database of Lakes, Reservoirs and Wetlands. *J. Hydrol.*, 2004, no. 296, pp. 1–22. DOI: 10.1016/j.jhydrol.2004.03.028

18. Downing J.A., Prairie Y.T. The Global Abundance and Size-Distribution of Lakes, Ponds, and Impoundments. *Limnol. Oceanogr.*, 2006, vol. 51, pp. 2388–2397. DOI: 10.4319/lo.2006.51.5.2388

19. Verpoorter C., Kutser T., Seekel D.A., Tranvik L.J. A Global Inventory of Lakes Based on High Resolution Satellite Imagery. *Geophys. Res. Lett.*, 2014, no. 41, pp. 1–7. DOI: 10.1002/2014GL060641

Received 5 September 2019

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Методологические вопросы использования рандомизированного машинного обучения для прогнозирования динамики термокарстовых озер Арктики / Ю.С. Попков, З. Волкович, А.В. Мельников, Ю.М. Полищук // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2019. – Т. 19, № 4. – С. 5–12. DOI: 10.14529/ctcr190401

FOR CITATION

Popkov Yu.S., Volkovich Z., Melnikov A.V., Polishchuk Yu.M. Methodological Issues of Using the Randomized Machine Learning for Forecasting the Dynamics of Thermokarst Arctic Lakes. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2019, vol. 19, no. 4, pp. 5–12. (in Russ.) DOI: 10.14529/ctcr190401