

УДК 551.582

МОДЕЛИРОВАНИЕ НЕТТО-ЭКОСИСТЕМНОГО ОБМЕНА CO<sub>2</sub>  
С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ  
(на примере острова Сахалин)

А. В. ГОРБАРЕНКО <sup>1\*</sup>, Е. М. ГОРБАРЕНКО <sup>2,3\*\*</sup>

<sup>1</sup> Сколковский институт науки и технологий, Москва, Россия

<sup>2</sup> МГУ им. М.В. Ломоносова, Москва, Россия

<sup>3</sup> Институт географии РАН, Москва, Россия

В условиях глобальных климатических изменений оценка углеродного баланса на региональном уровне приобретает все большую значимость для разработки эффективных технологий для мониторинга парниковых газов и смягчения последствий современных климатических изменений. Наземные экосистемы играют ключевую роль в глобальном углеродном цикле, внося существенный вклад в обмен диоксидом углерода между подстилающей поверхностью и атмосферой. В данном исследовании была разработана модель на основе методов машинного обучения, в частности алгоритма градиентного бустинга CatBoost, для комплексной оценки пространственно-временной изменчивости нетто экосистемного обмена CO<sub>2</sub> (NEE) для наземных экосистем. Результаты модельных экспериментов для 2023 г. для острова Сахалин показали, что данный подход позволяет учесть множество факторов, влияющих на углеродный обмен, и получить пространственное распределение потоков CO<sub>2</sub> в региональном масштабе с временным разрешением в один месяц. Разработанная модель продемонстрировала высокую точность прогнозирования потоков с коэффициентом детерминации ( $R^2$ ) по всем экосистемам около 0,76. Полученные результаты могут быть использованы для оценки потоков CO<sub>2</sub> в различных географических регионах в региональном масштабе для оценки реального вклада в углеродный баланс.

<https://doi.org/10.46991/PYSUC.2025.59.2.240>

**Keywords:** net ecosystem exchange of CO<sub>2</sub>, natural ecosystems, carbon dioxide, gradient boosting, machine learning.

**Введение.** Точная оценка и прогнозирование нетто экосистемного обмена (NEE) CO<sub>2</sub> приобретает все большее значение для понимания глобальной динамики углеродного баланса и воздействия на климат [1, 2]. NEE, представляющий собой баланс между поглощением диоксида углерода экосистемой через фотосинтез и выделением CO<sub>2</sub> через дыхание, также служит важным показателем функционирования экосистем и их реакции на изменения

\* E-mail: [aar44i@yandex.ru](mailto:aar44i@yandex.ru)

\*\* E-mail: [lisan.sat@gmail.com](mailto:lisan.sat@gmail.com)

окружающей среды [3]. Недавние исследования подчеркивают важность измерений потоков  $\text{CO}_2$  на региональном масштабе для улучшения нашего понимания круговорота углерода в наземных экосистемах и механизмов его обратной связи с изменением климата [4, 5].

Измерения NEE в природных экосистемах обычно проводятся с помощью метода турбулентных пульсаций [1, 6]. По данной методике рассчитывается обмен  $\text{CO}_2$  между экосистемами и атмосферой путем анализа высокочастотных колебаний вертикальной скорости ветра и концентрации  $\text{CO}_2$  в приземном слое [7]. Метод обеспечивает прямые измерения потоков  $\text{CO}_2$  над пологом однородных участков растительности. Станции пульсационных наблюдений располагаются по всему миру и объединены в сети мониторинга, такие как FLUXNET, собирающие и обменивающиеся данными с сотен станций мониторинга в различных географических регионах с разнообразным климатом и структурой растительности [6, 8]. Широкое внедрение пульсационных систем измерений потоков создало обширную сеть станций с продолжительными рядами наблюдений, которые можно использовать для оценки NEE как на локальном, так и на региональных уровнях. Этот процесс масштабирования обычно основан на физико-математическом моделировании [9, 10]. Хотя процесс-ориентированные модели помогают понять динамику углеродного цикла в наземных экосистемах и отслеживать его изменения во времени и пространстве [11, 12], они сталкиваются с несколькими проблемами: неопределенность в представлении конкретных экосистемных процессов и чувствительность к выбору параметров для модели. Кроме того, данный метод требует значительных вычислительных ресурсов и различных обширных наборов входных данных.

Для решения сложных задач методы машинного обучения стали мощными инструментами в моделировании сложных экологических процессов, открывая новые возможности для лучшего понимания и прогнозирования потоков  $\text{CO}_2$  между наземными экосистемами и атмосферой. Способность моделей машинного обучения обрабатывать большие объемы данных и улавливать сложные закономерности сделала их особенно ценными для моделирования экосистемных потоков [13, 14]. Недавние исследования продемонстрировали превосходную производительность алгоритмов машинного обучения по сравнению с традиционными эмпирическими моделями в анализе и воспроизведении временной и пространственной изменчивости NEE [15, 16]. Различные методы машинного обучения, включая случайные леса (RF), нейронные сети и SVM (Support Vector Machines), были успешно применены для прогнозирования потоков  $\text{CO}_2$  в различных экосистемах [17, 18]. Эти методы могут эффективно обрабатывать нелинейные зависимости и объединять для анализа множество экологических и метеорологических переменных, влияющих на NEE, таких как температура, осадки, солнечная радиация, параметры растительности и др.

В контексте развития применения машинного обучения в моделировании потоков  $\text{CO}_2$ , остров Сахалин представляет собой уникальную область для изучения. Как один из крупнейших островов на Дальнем Востоке России, он состоит из разнообразных экосистем, включая бореальные леса, водно-

болотные угодья и прибрежные районы, которые играют важную роль в региональном круговороте парниковых газов [19, 20]. Расположение острова в зоне умеренного климата делает его особенно чувствительным к изменению климата, предоставляя ценную информацию о реакции экосистем в различных биомах [21].

Целью данного исследования является разработка статистической модели на основе методов машинного обучения, в частности алгоритма градиентного бустинга CatBoost, для комплексной оценки пространственно-временной изменчивости NEE.

**Материалы и методы исследования.** В качестве входных данных в модели использовались наборы данных пульсационных наблюдений из глобальной базы данных FLUXNET (<https://fluxnet.org/>). Для обучения модели были взяты данные 131 станции по всему миру с наиболее продолжительными, полными и репрезентативными рядами данных (рис. 1).

Данные плохого качества исключались из датасета, далее полученные суточные значения NEE по каждой станции суммировались до месячных значений. Обучающая выборка включала в себя станции по нескольким экосистемам согласно классификации, принятой в международной геосферно-биосферной программе (International Geosphere-Biosphere Programme, IGBP) [22]: хвойные вечнозеленые леса (ENF), широколиственный лес (DBF), смешанный лес (MF), лиственные и хвойные вечнозеленые кустарники (OSH), водно-болотные угодья (WET), луга (GRA) и сельскохозяйственные угодья (CRO).

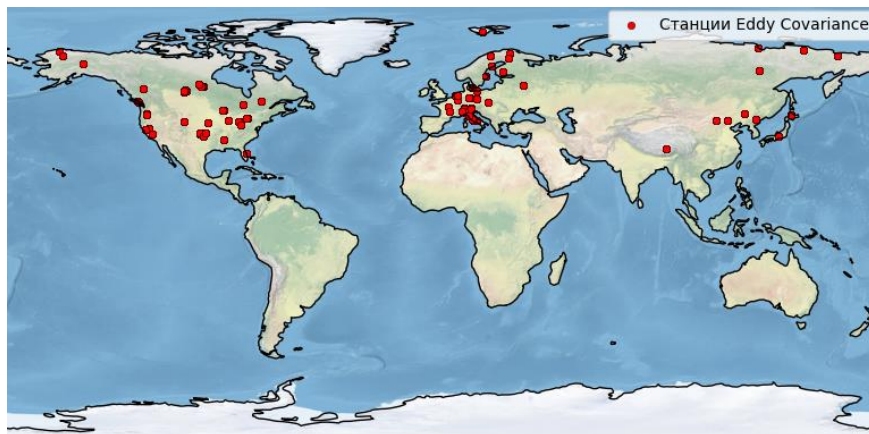


Рис. 1. Распределение станций пульсационных измерений, данные которых были использованы при обучении модели.

Метеорологические параметры были взяты из реанализа ERA5 Land от Европейского центра среднесрочных прогнозов погоды (ECMWF) [23] с пространственным разрешением  $0,1^\circ$  и разрешением по времени 1 месяц. Также были использованы данные о характере земной поверхности согласно классификации IGBP со спутника MODIS (Moderate Resolution Imaging Spectroradiometer) с разрешением 500 м [24]. Для обучения модели также

использовались биофизические параметры с MODIS такие как: фотосинтетически активная радиация (fPAR), индекс листовой поверхности (LAI) и валовая первичная продукция (GPP).

После сбора всех необходимых данных проводилась их предобработка: все пространственные данные (реанализ и спутниковые данные) интерполировались на сетку  $500 \times 500$  м, и усреднялись до 1 месяца. В результате все данные имели единый масштаб по времени и пространству.

Далее для расчета NEE разрабатывались модели на основе градиентного бустинга (CatBoost). Для каждого типа экосистемы согласно классификации IGBP была разработана отдельная модель. Все используемые станции были сгруппированы согласно типам экосистем. Временные ряды данных по каждой станции были разделены на 3 части с учетом хронологии данных: тренировочную, валидационную и тестовую выборки. Таким образом, для каждого типа экосистемы было создано 3 набора данных, состоящих из данных станций наблюдений, принадлежащих данному типу экосистемы, которые в дальнейшем использовались для обучения модели.

Для модели каждой экосистемы подбирались наиболее релевантные параметры условий внешней среды (среди данных реанализа ERA5 Land и MODIS). Для каждой экосистемы было отобрано от 7 до 12 параметров для моделирования NEE, оказывающих наибольшее воздействие на каждый конкретный тип экосистемы. Также для каждой модели были подобраны свои гиперпараметры с помощью фреймворка Optuna.

Далее разработанные модели применялись в эксперименте для территории о. Сахалин для каждого месяца 2023 г. Для инференса необходимо было собрать аналогичный набор данных, который использовался при обучении моделей. Для всей территории о. Сахалин были получены данные реанализа ERA5 Land и MODIS за 2023 г., которые затем прошли тот же этап предобработки, что и данные для обучения. Таким образом, был получен тензор значений параметров для каждого месяца 2023 г. с разрешением  $500 \times 500$  м. Для каждого пикселя имеются вектора предикторов, к которым применялись разработанные модели. Выбор модели, которая будет применена к тому или иному пикселю осуществлялся на основе классификации IGBP (например, для векторов значений, расположенных в типе смешанных лесов применялась модель, обученная на данных станций этой же экосистемы).

После применения модели для каждого вектора в результате получалось одно значение NEE. Агрегируя полученные результаты для каждого пикселя вместе, мы получаем карту пространственного распределения NEE для о. Сахалин за определенный месяц. Результатом работы модели являются карты пространственного распределения NEE для каждого месяца и интегральные значения NEE для определенной территории.

**Анализ результатов и их обсуждение.** Проведенный анализ выявил значительную пространственно-временную изменчивость нетто экосистемного обмена  $\text{CO}_2$ , которая характеризуется выраженной сезонной динамикой и пространственной неоднородностью, обусловленной типом экосистемы, характеристиками экосистемы и метеорологическими параметрами.

Визуализация годовой динамики NEE (рис. 2) демонстрирует четкую сезонность процессов поглощения и эмиссии  $\text{CO}_2$ . В летний период (июнь–август) наблюдается интенсивное поглощение  $\text{CO}_2$  (отрицательные значения NEE, отображаемые зеленым цветом), которое достигает максимума в июле с суммарными показателями до  $-150 \text{ гС} \cdot \text{м}^{-2} \cdot \text{месяц}^{-1}$ . Напротив, зимний период (ноябрь–февраль) характеризуется преобладанием процессов эмиссии  $\text{CO}_2$  (положительные значения NEE, отображаемые желтым и красноватым цветами) до  $50 \text{ гС} \cdot \text{м}^{-2} \cdot \text{месяц}^{-1}$ .

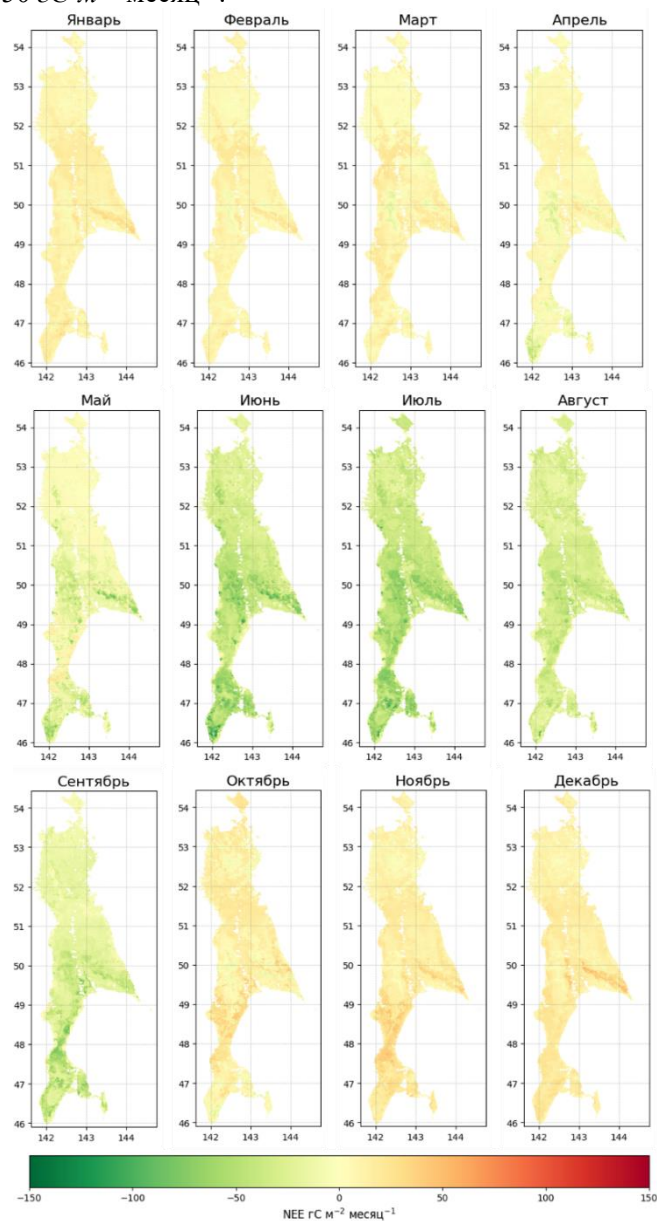


Рис. 2. Пространственное распределение предсказанных среднемесячных значений NEE для о. Сахалин за каждый месяц 2023 г.

Пространственный анализ выявил выраженную широтную закономерность в распределении потоков CO<sub>2</sub>. Южная часть исследуемой территории, где преобладают широколиственные и хвойные лесные массивы, демонстрирует наиболее высокие показатели поглощения CO<sub>2</sub> в течение вегетационного периода. Северные районы, отличающиеся меньшей залесенностью, характеризуются более низкой интенсивностью поглощения CO<sub>2</sub> даже в летние месяцы, что подтверждает ключевую роль лесных экосистем в региональном углеродном балансе.

Весенний и осенний переходные периоды характеризуются постепенной трансформацией от эмиссии к поглощению CO<sub>2</sub> и обратно, с заметным усилением скорости поглощения CO<sub>2</sub> уже в апреле–мае и сохранением значительного поглощения до сентября–октября, особенно в южных лесных экосистемах.

Подобную динамику сезонного хода NEE, описанную выше, можно также отметить в распределении интегральных значений NEE для территории всего о. Сахалин (см. таблицу). По результатам модельных расчетов поглощение CO<sub>2</sub> на о. Сахалин в 2023 г. составило –11162,9 тыс. *т* С.

*Интегральные значения NEE для территории всего острова Сахалин за каждый месяц 2023 г.*

Месяц	Интегральное значение NEE, тыс. <i>т</i> С в месяц
Январь	825,8
Февраль	544,0
Март	570,0
Апрель	–55,3
Май	–1354,4
Июнь	–4522,6
Июль	–4897,9
Август	–3239,4
Сентябрь	–2742,9
Октябрь	1014,6
Ноябрь	1422,2
Декабрь	1273,0
Всего за год	–11162,9

Точность моделирования NEE существенно варьировала в зависимости от типа экосистемы. Тестирование модели каждой из экосистем проводилось на независимой выборке данных измерений потоков CO<sub>2</sub>, полученных со станций пульсационных наблюдений. Тестовый набор формировался путем выделения последних 20% наблюдений из временных рядов каждой станции при сохранении хронологической последовательности данных. Наиболее высокая точность прогнозирования была достигнута для смешанных лесов (MF,  $R^2 = 0,88$ ) и лиственных и хвойных вечнозеленых кустарников (OSH,  $R^2 = 0,89$ ), а также для лиственных лесов (DBF,  $R^2 = 0,87$ ). Несколько ниже оказалась точность моделей для водно-болотных угодий (WET,  $R^2 = 0,76$ ) и сельскохозяйственных земель (CRO,  $R^2 = 0,74$ ). Наименьшая предсказательная способность моделей отмечена для вечнозеленых хвойных лесов (ENF,  $R^2 = 0,64$ ) и травянистых экосистем (GRA,  $R^2 = 0,51$ ). Следует отметить, что достигнутая точность моделирования во многом зависела от качества и

количества данных, используемых для обучения модели. Экосистемы с более обширными и репрезентативными наборами данных демонстрировали лучшие результаты прогнозирования, в то время как недостаточная представленность или неоднородность исходных данных для некоторых типов экосистем могла существенно ограничивать предиктивную способность моделей. Тем не менее, целенаправленное обучение отдельных моделей для каждого типа экосистем позволило оптимизировать прогностические возможности, учитывая специфические особенности функционирования различных биомов.

Полученные результаты демонстрируют не только физическую обоснованность разработанных моделей, но и их способность отражать тонкие механизмы взаимодействия растительности и атмосферы в контексте глобального углеродного цикла.

Выявленная сезонная динамика NEE согласуется с фундаментальными принципами функционирования природных экосистем умеренного пояса и подтверждает результаты предыдущих исследований [1, 8, 19]. Максимальное поглощение  $\text{CO}_2$  соответствует пику фотосинтетической активности растительности, когда благоприятные температурные условия и максимальная продолжительность светового дня создают оптимальные условия для поглощения  $\text{CO}_2$ . Зимняя эмиссия  $\text{CO}_2$  обусловлена преобладанием процессов дыхания экосистемы над фотосинтезом в условиях низких температур и ограниченной приходящей солнечной радиации.

Пространственная неоднородность потоков  $\text{CO}_2$  отражает фундаментальные различия в структуре и функционировании различных типов экосистем [13, 15]. Высокая поглощающая способность южных лесных массивов объясняется большей биомассой древостоя, развитой корневой системой и более продолжительным вегетационным периодом по сравнению с северными районами. Это подчеркивает важную роль лесных экосистем как основных поглотителей атмосферного  $\text{CO}_2$  в региональном масштабе.

Различия в точности моделирования между типами экосистем ( $R^2$  от 0,51 до 0,89) отражают как внутреннюю сложность функционирования различных биомов, так и качество доступных данных для обучения моделей. Высокая точность для лесных экосистем ( $R^2 = 0,64\text{--}0,89$ ) может быть связана с относительной стабильностью и предсказуемостью их функционирования, а также с большим объемом накопленных данных наблюдений. Более низкая точность для травянистых экосистем ( $R^2 = 0,51$ ) и сельскохозяйственных земель ( $R^2 = 0,74$ ) обусловлена их высокой динамичностью, подверженностью антропогенному воздействию и большей временной изменчивостью характеристик растительного покрова [25, 26].

Важно отметить ограничения разработанных моделей, которые могут влиять на точность прогнозирования. Одним из ключевых источников погрешности является недостаточно детальный учет видового состава растительности в пределах каждого типа экосистемы. Современные классификации землепользования, используемые в качестве входных данных, объединяют экосистемы в широкие категории, не учитывая существенные различия в физиологических характеристиках, фенологии и поглощающей способности отдельных видов растений. Дополнительным источником

неопределенности является пространственное разрешение используемых данных, которое может не отражать мозаичность растительного покрова и микроклиматические различия в пределах отдельных пикселей. Кроме того, модели не учитывают влияние экстремальных погодных явлений, нарушений экосистем (пожары, вспышки вредителей, рубки) и долгосрочные сукцессионные изменения, что может приводить к систематическим погрешностям в отдельные периоды.

Несмотря на указанные ограничения, достигнутая точность моделирования сопоставима с результатами аналогичных исследований [15, 16] и демонстрирует эффективность использования алгоритмов градиентного бустинга в моделировании NEE для различных регионов с учетом его особых географических и экологических характеристик. Объединение данных дистанционного зондирования, наземных измерений и метеорологических параметров с использованием методов машинного обучения обеспечивает комплексную основу для понимания региональных закономерностей углеродного баланса [27].

**Заключение.** В ходе исследования была успешно разработана и применена модель на основе алгоритма градиентного бустинга CatBoost для оценки NEE на территории о. Сахалин. Анализ полученных результатов позволяет сформулировать следующие ключевые выводы:

1. Разработанный подход на основе машинного обучения продемонстрировал высокую эффективность в моделировании пространственно-временной динамики NEE с коэффициентом детерминации ( $R^2$ ) в среднем 0,76 по всем типам экосистем, что подтверждает возможность использования методов машинного обучения в качестве альтернативы традиционным физико-математическим моделям.

2. Выявлена выраженная сезонная и пространственная неоднородность потоков  $\text{CO}_2$  на территории Сахалина с максимальным поглощением в летний период (до  $-150 \text{ гС} \cdot \text{м}^{-2} \cdot \text{месяц}^{-1}$  в июле) и доминированием эмиссии в зимние месяцы, что согласуется с общими закономерностями функционирования экосистем.

3. Точность моделирования существенно варьировала в зависимости от типа экосистемы: наилучшие результаты были получены для смешанных лесов ( $R^2 = 0,88$ ) и вечнозеленых кустарников ( $R^2 = 0,89$ ), тогда как для травянистых экосистем ( $R^2 = 0,51$ ) и хвойных лесов ( $R^2 = 0,64$ ) точность была относительно ниже.

4. Дифференцированный подход к моделированию различных типов экосистем позволил учесть специфику их функционирования и повысить общую точность прогнозирования NEE, что подтверждает эффективность стратегии разработки отдельных моделей для каждого биома.

Результаты исследования имеют важное прикладное значение для оценки углеродного бюджета региона и могут быть использованы при разработке стратегий по снижению углеродного следа и адаптации к климатическим изменениям на региональном уровне. Методика, апробированная на примере



о. Сахалин, демонстрирует потенциал для масштабирования и применения к другим регионам с различными биоклиматическими условиями.

Поступила 26.01.2025

Получена с рецензии 14.03.2025

Утверждена 15.08.2025

#### ЛИТЕРАТУРА

1. Baldocchi D.D. Assessing the Eddy Covariance Technique for Evaluating Carbon Dioxide Exchange Rates of Ecosystems: Past, Present and Future. *Global Change Biology* **9** (2003), 479–492.  
<https://doi.org/10.1046/j.1365-2486.2003.00629.x>
2. Valentini R., Matteucci G., et al. Respiration as the Main Determinant of Carbon Balance in European Forests. *Nature* **404** (2000), 861–865.  
<https://doi.org/10.1038/35009084>
3. Chapin F.S., Woodwell G.M., et al. Reconciling Carbon-Cycle Concepts, Terminology, and Methods. *Ecosystems* **9** (2006), 1041–1050.  
<https://doi.org/10.1007/s10021-005-0105-7>
4. Beer C., Reichstein M., et al. Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science* **329** (2010), 834–838.  
<https://doi.org/10.1126/science.1184984>
5. Hilton T. W., Davis K. J., Keller K. Evaluating Terrestrial CO<sub>2</sub> Flux Diagnoses and Uncertainties from a Simple Land Surface Model and its Residuals. *Biogeosciences* **11** (2014), 217–235.  
<https://doi.org/10.5194/bg-11-217-2014>
6. Pastorello G., Trotta C., et al. The FLUXNET2015 Dataset and the ONEFlux Processing Pipeline for Eddy Covariance Data. *Scientific data* **7** (2020), 225.  
<https://doi.org/10.1038/s41597-020-0534-3>
7. Aubinet M., Vesala T., Papale D. (ed.). *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*. Springer Science & Business Media (2012), 438.
8. Baldocchi D., Falge E., et al. FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. *Bulletin of the American Meteorological society* **82** (2001), 2415–2434.  
[https://doi.org/10.1175/1520-0477\(2001\)082<2415:FANTTS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2)
9. Krinner G., Viovy N., et al. A dynamic Global Vegetation Model for Studies of the Coupled Atmosphere-Biosphere System. *Global Biogeochemical Cycles* **19** (2005).  
<https://doi.org/10.1029/2003GB002199>
10. Lawrence D.M., Oleson K.W., et al. Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems* **3** (2011).  
<https://doi.org/10.1029/2011MS00045>
11. Tian H., Chen G., et al. Century-Scale Responses of Ecosystem Carbon Storage and Flux to Multiple Environmental Changes in the Southern United States. *Ecosystems* **15** (2012), 674–694.  
<https://doi.org/10.1007/s10021-012-9539-x>
12. Li X., Hu X.-M., et al. Terrestrial CO<sub>2</sub> Fluxes, Concentrations, Sources and Budget in Northeast China: Observational and Modeling Studies. *Journal of Geophysical Research: Atmospheres* **125** (2020), e2019JD031686.  
<https://doi.org/10.1029/2019JD031686>
13. Papale D., Black T.A., et al. Effect of Spatial Sampling from European Flux Towers for Estimating Carbon and Water Fluxes with Artificial Neural Networks. *Journal of Geophysical Research: Biogeosciences* **120** (2015), 1941–1957.  
<https://doi.org/10.1002/2015JG002997>
14. Knox S. FLUXNET-CH 4 Synthesis Activity. 2020 ESA Annual Meeting (August 3–6). ESA (2020).  
<https://doi.org/10.1175/BAMS-D-18-0268.1>

15. Jung M., Schwalm Ch., al. Scaling Carbon Fluxes from Eddy Covariance Sites to Globe: Synthesis and Evaluation of the FLUXCOM Approach. *Biogeosciences* **17** (2020), 1343–1365. <https://doi.org/10.5194/bg-17-1343-2020>
16. Tramontana G., Jung M., et al. Predicting Carbon Dioxide and Energy Fluxes Across Global FLUXNET Sites with Regression Algorithms. *Biogeosciences* **13** (2016), 4291–4313. <https://doi.org/10.5194/bg-13-4291-2016>
17. Fu Z., Stoy P.C., et al. Climate Controls over the Net Carbon Uptake Period and Amplitude of Net Ecosystem Production in Temperate and Boreal Ecosystems. *Agricultural and Forest Meteorology* **243** (2017), 9–18. <https://doi.org/10.1016/j.agrformet.2017.05.009>
18. Reichstein M., Camps-Valls G., et al. Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature* **566** (2019), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
19. Dolman A.J., Shvidenko A., et al. An Estimate of the Terrestrial Carbon Budget of Russia Using Inventory-Based, Eddy Covariance and Inversion Methods. *Biogeosciences* **9** (2012), 5323–5340. <https://doi.org/10.5194/bg-9-5323-2012>
20. Schulze E.D., Lloyd J., et al. Productivity of Forests in the Euro Siberian Boreal Region and Their Potential to Act as a Carbon Sink – a Synthesis. *Global Change Biology* **5** (1999), 703–722. <https://doi.org/10.1046/j.1365-2486.1999.00266.x>
21. Tchepakova N.M., Parfenova E., Soja A.J. The Effects of Climate, Permafrost and Fire on Vegetation Change in Siberia in a Changing climate. *Environmental Research Letters* **4** (2009), 045013. <https://doi.org/10.1088/1748-9326/4/4/045013>
22. Belward A. S., Estes J. E., Kline K. D. *The IGBP-DIS global 1-km land-cover data set DISCover*. A Project Overview (1999).
23. Muñoz-Sabater J., Dutra E., et al. ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications. *Earth System Science Data* **13** (2021), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
24. Friedl M., Sulla-Menashe D. MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V061 [Data set]. *NASA EOSDIS Land Processes Distributed Active Archive Center* (2022). <https://doi.org/10.5067/MODIS/MCD12Q1.061>
25. Kutsch W.L., Aubinet M., et al. The Net Biome Production of Full Crop Rotations in Europe. *Agriculture, Ecosystems & Environment* **139** (2010), 336–345. <https://doi.org/10.1016/j.agee.2010.07.016>
26. Buysse P., Bodson B., et al. Carbon Budget Measurement over 12 Years at a Crop Production Site in the Silty-Loam Region in Belgium. *Agricultural and Forest Meteorology* **246** (2017), 241–255. <https://doi.org/10.1016/j.agrformet.2017.07.004>
27. Zhuravlev R., Dara A., et al. Globally Scalable Approach to Estimate Net Ecosystem Exchange Based on Remote Sensing, Meteorological Data, and Direct Measurements of Eddy Covariance Sites. *Remote Sensing* **14** (2022), 5529. <https://doi.org/10.3390/rs14215529>

Ա. Վ. ԳՈՐԲԱՐԵՆԿՈ, Ե. Մ. ԳՈՐԲԱՐԵՆԿՈ

CO<sub>2</sub>-Ի ՓՈԽԱՆԱԿՄԱՆ ՆԵՏՏՈ ԷԿՈՀԱՄԱԿԱՐԳՈՎ ՄՈԴԵԼԱՎՈՐՈՒՄ  
ՄԵՔԵՆԱՅԱԿԱՆ ՈՒՍՈՒՑՄԱՆ ՄԵԹՈԴՈՎ  
(Սախային կղզու օրինակով)

Ամփոփում

Կլիմայի գլոբալ փոփոխության համատեքստում ածխածնի հաշվեկշռի գնահատումը տարածաշրջանային մակարդակում զնայով ավելի կարևոր է

դառնում ջերմոցային գազերի արտանետումների մոնիթորինգի և կլիմայական ժամանակակից փոփոխությունների հետևանքների մեղմման համար արդյունավետ տեխնոլոգիաների մշակման տեսանկյունից: Երկրային էկոհամակարգերը առանցքային դեր են խաղում ածխածնի գլոբալ ցիկլում՝ զգալի ներդրում ունենալով գործուն մակերևույթի և մթնոլորտի ածխաթթու գազի միջև փոխանակման գործում: Այս ուսումնասիրության մեջ մշակվել է մեքենայացման ուսուցման մեթոդների վրա հիմնված մոդել, մասնավորապես՝ CatBoost գրադիենտային խթանման ալգորիթմը, CO<sub>2</sub> նետտո էկոհամակարգային փոխանակման տարածաժամանակային փոփոխականության համալիր գնահատելու համար: Սախալին կղզու համար 2023 թ. մոդելի փորձերի արդյունքները ցույց են տվել, որ այս մոտեցումը թույլ է տալիս հաշվի առնել ածխածնի փոխանակման վրա ազդող բազմաթիվ գործոններ և ստանալ CO<sub>2</sub>-ի հոսքերի տարածական բաշխումը տարածաշրջանային մասշտաբով՝ մեկ ամսվա ժամանակային լուծաչափով: Մշակված մոդելը ցույց է տվել հոսքերի կանխատեսման բարձր ճշգրտություն՝ դետերմինացիայի գործակիցը ( $R^2$ ) բոլոր էկոհամակարգերի համար միջինը մոտ 0,76 է: Ստացված արդյունքները կարող են օգտագործվել տարբեր աշխարհագրական շրջաններում CO<sub>2</sub> հոսքերը տարածաշրջանային մասշտաբով գնահատելու, ածխածնային հաշվեկշռում իրական ներդրումը գնահատելու համար:

A. V. GORBARENKO, E. M. GORBARENKO

MODELING NET ECOSYSTEM EXCHANGE OF CO<sub>2</sub>  
USING MACHINE LEARNING TECHNIQUES  
(a case study of Sakhalin Island)

Summary

In the context of global climate change, the assessment of carbon balance at the regional level is becoming increasingly important for developing effective strategies for greenhouse gas emissions management and adaptation to current climate trends. Terrestrial ecosystems play a crucial role in the global carbon cycle, making varying contributions to carbon dioxide exchange between the underlying surface and the atmosphere. In this study, a machine learning-based model was developed, specifically utilizing the CatBoost gradient boosting algorithm, for comprehensive assessment of spatiotemporal variability in net ecosystem exchange of CO<sub>2</sub> (NEE). The results of model experiments for Sakhalin Island in 2023 demonstrated that this approach effectively accounts for multiple factors affecting carbon exchange and provides spatial distribution of CO<sub>2</sub> fluxes at a regional scale with monthly temporal resolution. The developed model showed high prediction accuracy with a coefficient of determination ( $R^2$ ) averaging 0.76 across all ecosystems. The obtained results can be applied for carbon balance assessment in other regions and development of measures to mitigate anthropogenic impact on the climate system.