

**Искусственный интеллект в дистанционном зондировании из космоса потоков углерода из атмосферы в экосистемы**

**Розанов Алексей Павлович**

Институт естественных наук, Екатеринбург  
E-mail: alexey.rozanov@urfu.ru

Хорошо известно, что наземные экосистемы являются одним из основных мест стока углерода на планете, однако существующие численные оценки углеродного баланса различных территорий все еще содержат значительную неопределенность и требуют улучшения. Для решения этой проблемы многие исследователи пытались масштабировать измерения потоков CO<sub>2</sub> методом турбулентных пульсаций, используя методы машинного обучения и наблюдения из космоса. Множество публикаций было посвящено применению классических алгоритмов, таких как "Случайный лес" (RF) [1], Искусственные нейронные сети (ANNs) [2], Метод опорных векторов (SVM) [3] и другие [4], где модели получали на вход данные, полученные с различных орбитальных приборов. Однако, несмотря на то, что эти исследования продемонстрировали относительно хорошую точность оценок, хорошую масштабируемость и значительно превзошли физически модели по времени, все еще существуют проблемы, связанные с ошибками, неопределенностью и объяснимостью предсказаний моделей.

Цель текущей работы - предложить новый набор данных, NorthFlux, содержащий оценки потоков углерода в северном полушарии, путем объединения уникального набора предварительно обработанных предикторов, полученных из продуктов MODIS Terra и Aqua MOD09 CMG и MCD12C1, а также реанализа ERA5 on single levels, и ансамбля моделей машинного обучения: ANN, RF, Extreme Gradient Boosting (XGBoost) и Adaptive Boosting (AdaBoost). В качестве целевых параметров использовались временные ряды Валовой первичной продукции (GPP), Респирации экосистемы (RECO) и Чистого экосистемного обмена (NEE), предоставленные сетями FLUXNET и AmeriFlux для обучения и валидации результатов. При проверке результатов обучения было показано, что в целом все четыре типа моделей достаточно точно оценивали GPP, RECO и NEE, однако медианный ансамблевый прогноз всегда имел наименьшую среднеквадратичную ошибку (MSE) и наивысший коэффициент детерминации (R<sup>2</sup>): для GPP MSE = 2.012 гC/m<sup>2</sup>/день и R<sup>2</sup> = 0.884, для RECO MSE = 0.837 гC/m<sup>2</sup>/день и R<sup>2</sup> = 0.883, для NEE MSE = 1.454 гC/m<sup>2</sup>/день и R<sup>2</sup> = 0.773.

Благодарность: исследование выполнено при поддержке Министерства высшего образования и науки РФ (проект NoFEUZ-2023-0023).

Ключевые слова: машинное обучение, потоки углерода, карбоновый цикл, FLUXNET, MODIS

**Источники и литература**

- 1) Zeng J. et al. Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest //Scientific data. – 2020. – Т. 7. – №. 1. – С. 313.
- 2) Yu T., Zhang Q., Sun R. Comparison of machine learning methods to up-scale gross primary production //Remote Sensing. – 2021. – Т. 13. – №. 13. – С. 24-48.
- 3) Zhang C. et al. Mapping CO<sub>2</sub> fluxes of cypress swamp and marshes in the Greater Everglades using eddy covariance measurements and Landsat data //Remote Sensing of Environment. – 2021. – Т. 262. – С. 112-523.

- 4) Chen T., Guestrin C. Xgboost: A scalable tree boosting system //Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. – 2016. – C. 785-794.
- 5) Pastorello, G., Trotta, C., Canfora, E. et al. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data 7, 225 (2020). <http://doi.org/10.1038/s41597-020-0534-3>
- 6) Vermote, E., R. Wolfe. MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006. 2015, distributed by NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MOD09GA.006>.
- 7) Friedl, M., D. Sulla-Menashe. MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006. 2015, distributed by NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MCD12C1.006>
- 8) Loveland T. R., Belward A. S. The international geosphere biosphere programme data and information system global land cover data set (DISCover) //Acta Astronautica. – 1997. – Т. 41. – №. 4-10. – С. 681-689.
- 9) Muñoz Sabater, J. (2019): ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.e2161bac
- 10) Paszke, A. et al., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32. Curran Associates, Inc., pp. 8024–8035. Available at: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.