

Использование технологии машинного обучения для диагностики кризисов в пелагических экосистемах по изотопному составу углерода и кислорода на примере раннего карбона

Журавлев А.В.

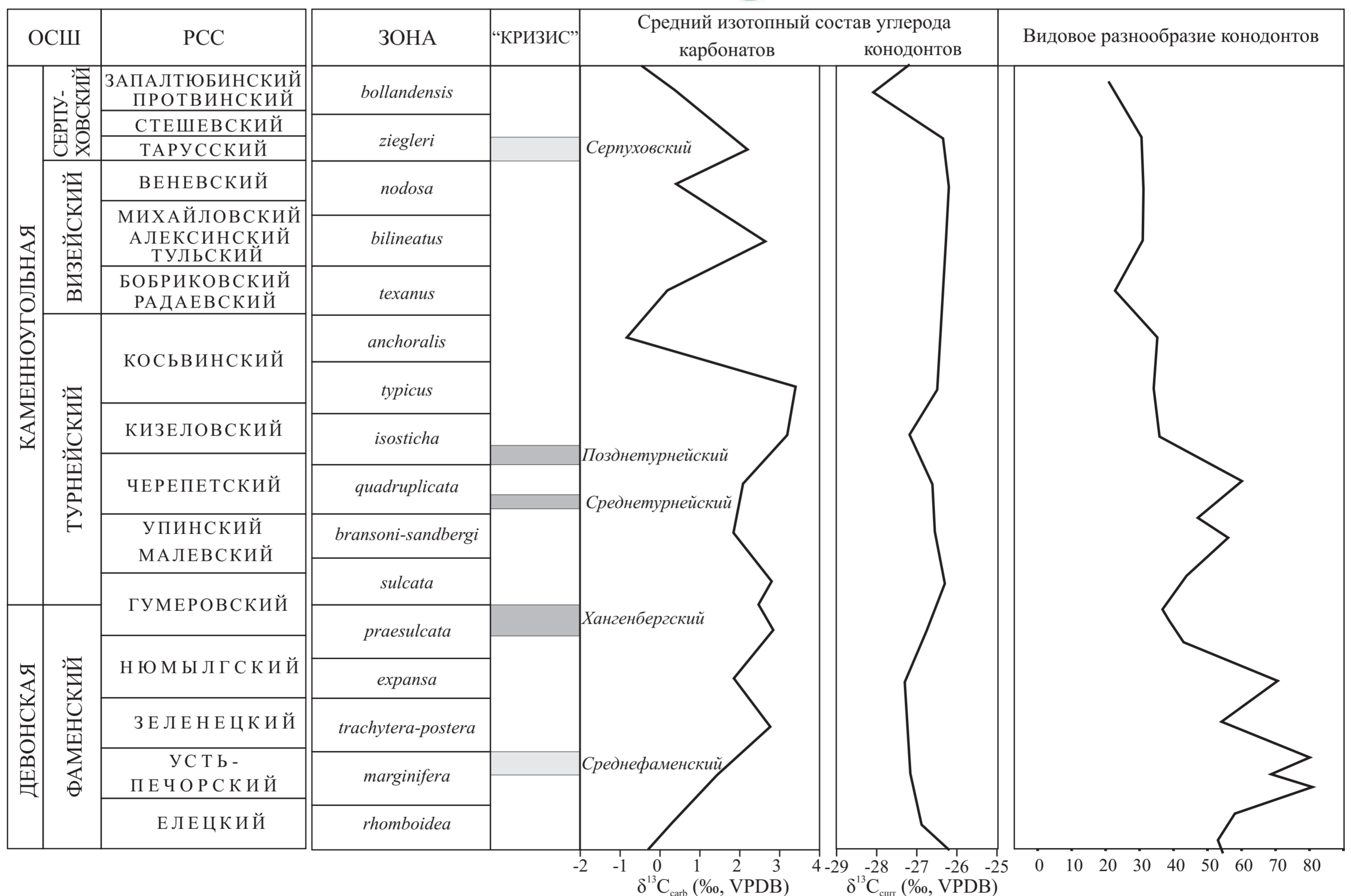
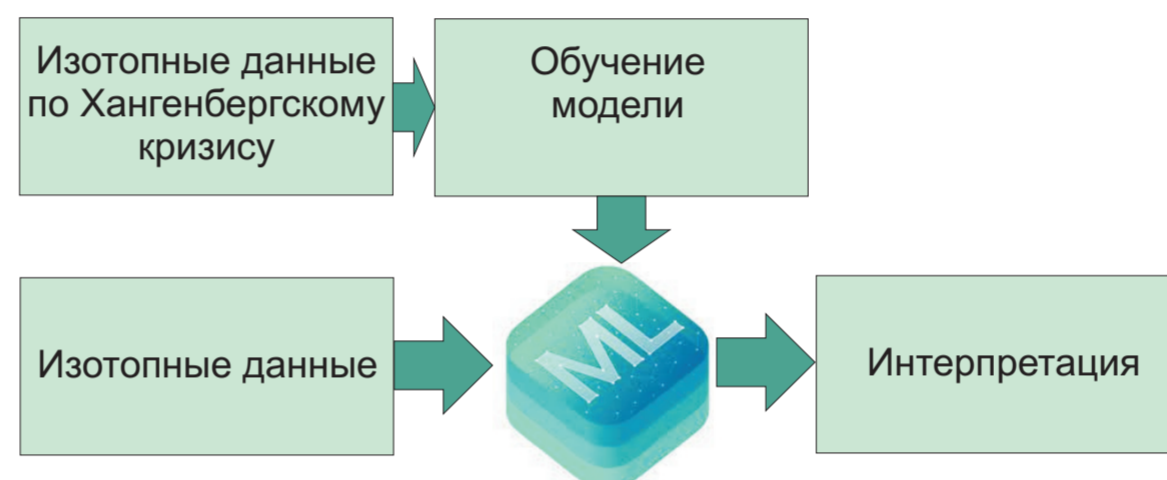
Институт геологии им. акад. Н.П. Юшкина ФИЦ Коми НЦ УрО РАН

Данные об изотопном составе углерода и кислорода осадочных карбонатов и органических остатков широко используются в палеонтологии и стратиграфии. Один из путей содержательного анализа изотопных данных заключается в использовании компьютерных моделей, позволяющих реконструировать локальные или глобальные особенности фракционирования изотопов (Hayes et al., 1999; Hartke et al., 2021; Журавлев, 2022; Zhuravlev, 2023).

Другой путь состоит в получении характерного образа тех или иных состояний древних экосистем в виде набора изотопных данных. Реализация этого подхода возможна на основе технологии машинного обучения.

В качестве модельного интервала была выбрана поздняя фаза Хангенбергского кризиса (по Kaiser et al., 2015) и совпадающая с ним серия изотопных аномалий по углероду (Kaiser et al., 2006, 2008, 2015; Kumpan et al., 2015; Qie et al., 2016; Piszarska et al., 2020; Zhuravlev et al., 2020), а в качестве фонового (не кризисного) интервала – раннее турне. Поздняя фаза Хангенбергского кризиса следует за событием массового вымирания и характеризуется постепенным исчезновением последних девонских представителей некоторых групп организмов и появлением таксонов, характерных для раннекаменноугольной эпохи (Kaiser et al., 2015; 2019).

Обучающая выборка состояла из 80 образцов из пяти разрезов различных фациальных зон северного Приуралья (38 фоновых и 42 кризисных), для каждого из которых определен изотопный состав кислорода карбоната и изотопный состав углерода карбоната и конодонтовых элементов. Обучение модели было проведено с использованием фреймворка Core ML [Apple Inc., 2017-2023] (тип модели Tabular Classifier) и алгоритма Boosted Tree (максимальная глубина 10). Точность модели на обучающей выборке составила 99%. Альтернативная модель была получена на основе той же выборки с использованием библиотеки Scikit-learn [Pedregosa et al., 2011]. Алгоритм обучения модели – Random Forest (максимальная глубина 13). Точность модели на обучающей выборке составила 100%. С помощью полученных моделей был проанализирован массив данных, содержащий информацию об изотопном составе 230 образцов из стратиграфического интервала от среднего фамена до серпухова из разнофациальных разрезов северного Приуралья (исходные данные доступны в качестве приложения к статье Zhuravlev, 2023 и в обновляемом архиве по ссылке <http://doi.org/10.13140/RG.2.2.17585.04961>). Обе модели показали сходные результаты.



Состояния пелагических экосистем, сходные с поздней фазой Хангенбергского кризиса, с невысокой вероятностью предполагаются в поздней *marginifera* (среднефаменское время) и по отдельным образцам, отвечающим фазе *ziegleri* (раннесерпуховское время); с более высокой вероятностью – в фазе *quadruplicata* (начало среднего турне) и ранней *isosticha* (рубеж среднего и позднего турне). Выделенные интервалы, в целом, совпадают с известными биотическими событиями и кризисами в рассматриваемом временном интервале. Начало среднего турне отвечает заключительной фазе события Low Alum Shale; рубеж среднего и позднего турне, известный как событие Mid-Aikuanian, характеризуется вымиранием многих видов конодонтов и существенным снижением их разнообразия в различных палеогеографических областях (Журавлев, 2019; Zhuravlev, Plotitsyn, 2022). Начало серпуховского века также отвечает глобальному экологическому кризису (Yao et al., 2022).

Относительно небольшая обучающая выборка позволяет рассматривать полученные результаты как предварительные. Для более обоснованных интерпретаций состояния древних пелагических экосистем требуется обучение модели на более обширном наборе исходных данных.