

ОПРЕДЕЛЕНИЕ ТОКСИЧЕСКОГО ДЕЙСТВИЯ ТЯЖЁЛЫХ МЕТАЛЛОВ НА ФИТОПЛАНКТОН С ПОМОЩЬЮ КЛАССИФИКАТОРА «СЛУЧАЙНЫЙ ЛЕС»

**Червицов Р.Н., Киселева Д.Г., Дрозденко Т.В.¹, Хрущев С.С., Антал Т.К.¹,
Ризниченко Г.Ю., Плюснина Т.Ю.**

Биологический факультет Московского государственного университета
им. М.В. Ломоносова, Москва, 119234, Россия

¹Лаборатория комплексных экологических исследований Псковского государственного университета, Псков, 180000, Россия

Для анализа большого массива данных необходимо развитие методов их анализа, в связи с этим методы машинного обучения становятся все более востребованными. В работе исследуется действие токсикантов на фотосинтетический аппарат фитопланктона. В задачу работы входило построение и обучение классификатора для выявления действия токсиканта по параметрам кривой индукции флуоресценции хлорофилла *a*, а также поиск методов уменьшения размерности данных. Форма кривой индукции флуоресценции меняется в зависимости от состояния растительного организма и может служить индикатором действия факторов стресса. Изменение состояния фотосинтетического аппарата может быть вызвано наличием в среде выращивания факторов стресса (в том числе токсикантов). В качестве метода определения точности классификатора использовалась кросс-валидация. Также иногда присутствует необходимость уменьшения размерности данных, для этих целей может применяться выбор необходимых параметров из набора или метод главных компонент.

Исследовали действие тяжёлых металлов на пробы природного фитопланктона. Пробы природного фитопланктона из девяти водоёмов Псковской области инкубировали при добавлении Cd^{2+} или $\text{Cr}_2\text{O}_7^{2-}$ (20 и 50 мкМ). На протяжении всего периода эксперимента (несколько суток) регулярно проводилась регистрация кривых индукции флуоресценции проб с фитопланктоном. По кривым индукции флуоресценции были рассчитаны параметры ЛР-теста, которые затем были использованы для построения классификаторов типа «случайный лес». Задача данных классификаторов – выявление наличия или отсутствия токсиканта в пробе. Точность классификатора, основанного на полном наборе параметров ЛР-теста, составила 82%. Самыми важными параметрами для классификации оказались квантовые выходы первичной фотохимии и электронного транспорта. Точность определения в разных водных объектах различается: в отдельных местообитаниях точность составляла менее 50%, а в других точность оказалась близка к 100%, причем точности определения токсикантов в случае разных металлов также различаются. При использовании аналогичных классификаторов, основанных на двух параметрах ЛР-теста, точность сохраняется, если один из этих параметров – квантовый выход первичной фотохимии. При использовании метода главных компонент РСА для уменьшения размерности данных необходимо использовать не менее 4 главных компонент для получения классификатора с такой же точностью, что и классификатор на основе полного набора параметров. Согласно полученным результатам, существуют перспективы применения данного метода при определении наличия токсикантов, в том числе в естественных местообитаниях.

Исследование выполнено в рамках научного проекта государственного задания МГУ №121032500060-0 при частичной поддержке гранта РФФИ № 20-64-46018 (Псковский государственный университет) и гранта РФФИ № 20-04-00465.