

Проблемы интегрального анализа моделей распространения многих видов (на примере мхов России)

Научный руководитель – Сорокин Анатолий Александрович

Орлов Михаил Анатольевич

Выпускник (магистр)

Московский государственный университет имени М.В.Ломоносова, Биологический факультет, Кафедра биофизики, Москва, Россия

E-mail: orlovmikhailanat@gmail.com

Моделирование пространственного распространения видов (Species distribution modelling, SDM) приобретает все большую популярность в среде ботаников, зоологов, эволюционных биологов, специалистов по охране окружающей среды и др. Это эффективный и адаптивное инструмент, позволяющий связать места фактического нахождения организмов с условиями географической среды. В основе такого моделирования лежит концепция экологической ниши, а его реализация использует машинное обучение.

В задачи работы входило разработка алгоритма построения моделей многих видов на едином наборе данных с использованием разных методов машинного обучения, минимизации влияния исследовательской ошибки, связанной с неравномерностью сбора данных, автоматизированного выбора лучшей модели и тестирования накопленной ошибки при интегральном анализе моделей всех видов.

В данной работе выполнено построение SDM ряда видов мохообразных (Musci), локалитеты которых доступны в базе данных "Флора мхов России" [1]. Данная база является наиболее полным источником по разнообразию мхов страны и базовой основой для подготовки одноименной фундаментальной сводки. Исследование выполнено при помощи свободной программной среды R. Предикторы для машинного обучения взяты из баз данных CHELSA, Worldclim. Статистическое смещение (sampling bias) уменьшено с помощью "разбавления" данных (spatial data thinning). Использовано несколько возможно более различных алгоритмов машинного обучения. С помощью которого построен ряд реализаций. Качество моделей оценено по площади под ROC. О вкладе переменных судили по фактору VIF (variable inflation factor).

Итогом работы стали предсказанные ареалы и оценка значимости экологических факторов в модели (что в известной степени отражает экологические запросы самих мхов).

Источники и литература

- 1) Ivanov O.V., Kolesnikova M.A., Afonina O.M., Akatova T.V., Baisheva E.Z., Belkina O.A., Bezgodov A.G., Czerny Adjeva I.V., Dudov S.V., Fedosov V.E., Ignatova E.A., Ivanova E.I., Kozhin M.N., Lapshina E.D., Notov A.A., Pisarenko O.Yu., Popova N.N., Savchenko A.N., Teleganova V.V., Ukrainskaya G.Yu. et al. The database of the moss flora of Russia. Arctoa. 2017. Т. 26. № 1. С. 1-10.
- 2) Orlov M.A., Sheludkov A.V. Quantifying a species ecological niche using SDM (species distribution models) в сборнике XXVII Международная конференция "Математика. Компьютер. Образование" (МКО-2020), серия МАТЕМАТИКА. КОМПЬЮТЕР. ОБРАЗОВАНИЕ, издательство Автономная некоммерческая организация Ижевский институт компьютерных исследований (Ижевск), 2020, том 27, с. 33-33
- 3) Orlov, M., & Sheludkov, A. (2019). Bioclimatic Data Optimization for Spatial Distribution Models. In Springer Proceedings in Earth and Environmental Sciences (pp. 86–95). https://doi.org/10.1007/978-3-030-11720-7_13