

УДК 58.08+519.24+581.9

DOI: 10.24412/2072-8816-2023-17-2-130-144

ОСОБЕННОСТИ МОДЕЛИРОВАНИЯ РАСПРОСТРАНЕНИЯ ВИДОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДА МАКСИМАЛЬНОЙ ЭНТРОПИИ

© 2023 Д.В. Санданов^{1,2}

¹*Институт общей и экспериментальной биологии СО РАН*

ул. Сахьяновой, 6, г. Улан-Удэ, 670047, Россия

²*Национальный парк «Тункинский»*

ул. Ленина, 130, пос. Кырен, 671010, Россия

e-mail: sdenis1178@mail.ru

Аннотация. Экологическое моделирование методом максимальной энтропии на базе пакета Maxent широко используется в различных исследованиях. Однако, для получения валидных результатов и оптимальных моделей важно корректное использование метода, включая подготовку исходных данных и использование различных вариантов настроек в программе. В обзоре обобщены современные публикации по предлагаемой тематике, выявляющие различные нюансы и возможности работы в программе Maxent.

Ключевые слова: моделирование распространения видов, Maxent, максимальная энтропия, преимущества и ограничения метода.

Поступила в редакцию: 27.01.2023. **Принято к публикации:** 15.04.2023.

Для цитирования: Санданов Д.В. 2023. Особенности моделирования распространения видов с использованием метода максимальной энтропии. — Фиторазнообразие Восточной Европы. 17(2): 130–144. DOI: 10.24412/2072-8816-2023-17-2-130-144

Моделирование экологических ареалов видов является одним из современных направлений в экологии и биогеографии. Несмотря на наличие множества публикаций с рекомендациями по стандартам корректного моделирования (Araújo et al., 2019; Zurell et al., 2020) вопросы по подготовке и проведению анализа все еще остаются сложной задачей для начинающих исследователей. В настоящее время существует много методов моделирования распространения видов, основанных на разных алгоритмах (Sandanov, 2019; рис. 1). Среди «моделей присутствия» наиболее популярным является метод максимальной энтропии, реализованный в программе Maxent¹. Возможности и различные детали анализа на основе метода Maxent ранее представлены в обзоре (Lisovsky, Dudov, 2020). В предлагаемой статье представлены особенности работы с методом, основанные на опыте автора, которые позволят лучше понять принципы анализа и получить валидные результаты моделирования. При этом охвачены моменты, которые не обозначены в вышеуказанных обзорах.

В современных биологических исследованиях сбор материала и проведение наблюдений за видами в основном проводится на основе маршрутных или полустационарных исследований. В редких случаях флористические и фаунистические работы полностью охватывают изучаемый контур или ареал вида. Поэтому

¹ Название программы является сокращением от Maximum Entropy (максимальная энтропия). В ранних исследованиях многие авторы сохраняли оригинальное написание программы «MaxEnt». В современных исследованиях чаще используется словосочетание «Maxent», включая и авторов разработанной программы.

исследователи большей частью получают данные о присутствии вида. Достоверные данные об отсутствии вида могут быть получены при оценке территории с использованием регулярной сети отбора проб и наблюдений или же при долговременных исследованиях. При наличии полных данных о присутствии вида и понимании границ его ареала есть возможность ограничить область распространения вида условными точками его отсутствия (так называемое «псевдоотсутствие» вида – pseudo-absence). Такой подход позволяет при моделировании реализовать другие алгоритмы, основанные на данных присутствия и отсутствия вида (рис. 1).

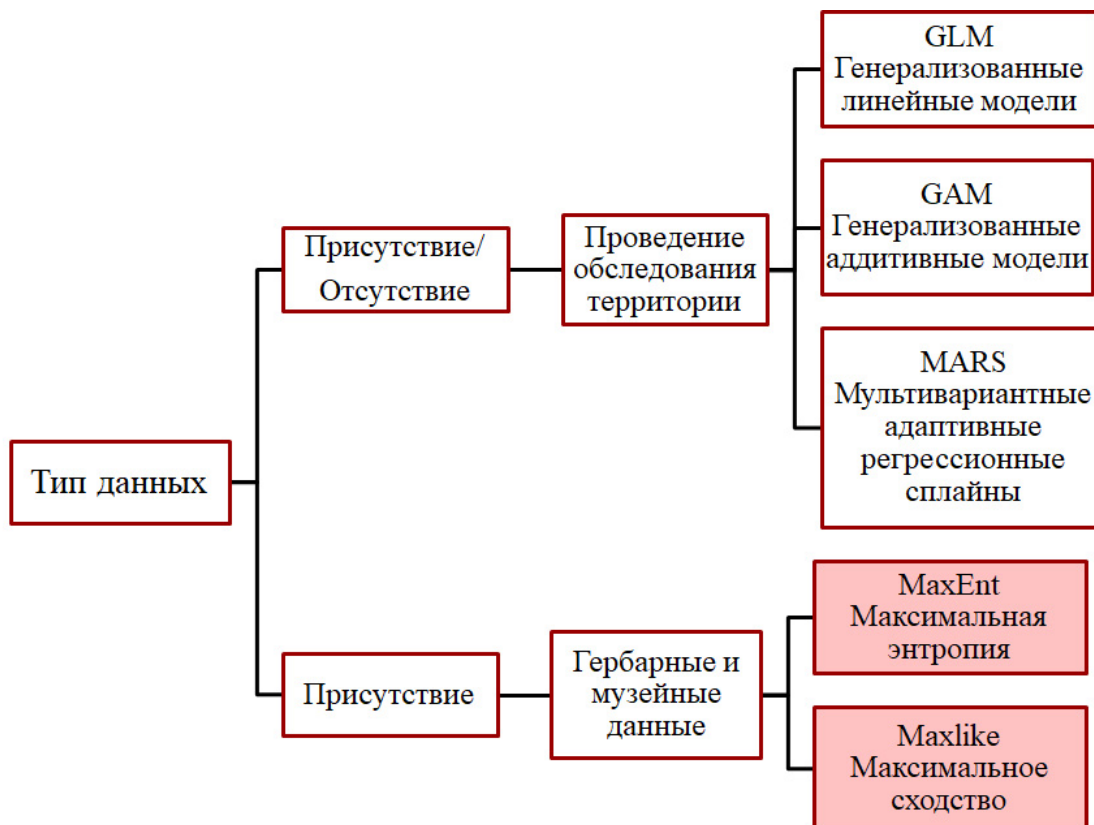


Рис. 1. Основные алгоритмы, использующиеся при экологическом моделировании
Fig. 1. Main algorithms used in species distribution modeling

Метод максимальной энтропии применяется для поиска свойств распределения факторов среды в наборе точек пространства (ячеек изучаемого раstra), в которых вид обитает: по исходной гипотезе в таком наборе свойства распределений факторов среды, влияющих на распространение вида, должны быть максимально близки свойствам распределений этих факторов во всем изучаемом географическом пространстве (Elith et al., 2011). Идея состоит в том, чтобы найти возможное распространение максимальной энтропии (наибольшего распределения) на основе доступной информации по известным местонахождениям вида и экологических условий на всей территории исследования. Модели Maxent представляют собой возможное распределение в виде пикселей на территории исследования без учета того, что пиксели, где вид не обитает не интерпретируются как отсутствие вида в данном местообитании (т.е. не используется «псевдоотсутствие» вида).

Анализ по принципу максимальной энтропии считается одним из самых эффективных методов при моделировании распределения видов на основании данных только о присутствии и широко используется в современных исследованиях (Elith, Leathwick, 2009). Более того, данный метод вошел в пятерку лучших на основе оценки качества и точности моделей (Elith et al., 2006). Алгоритм Maxent представлен в виде

отдельного программного продукта, реализован в пакетах для среды R, имеется в качестве надстройки в ГИС-программах (например, QGIS, SAGA GIS). Как и любой другой метод Maxent имеет свои сильные и слабые стороны (табл. 1).

Таблица 1. Сравнительный анализ различных возможностей программы Maxent с другими методами моделирования

Table 1. Comparative analysis of different possibilities of Maxent program with other methods of species distribution modeling

| Преимущества | Недостатки |
|---|---|
| Оценка взаимоотношений между видом и экологическими переменными, включая и взаимодействие между самими предикторами. | Процесс анализа представляет для исследователя своеобразный «черный ящик», т.е. не так просто понять, как конкретно работает метод. В случае с другими методами есть возможность провести оценку взаимосвязей между предикторами и зависимой переменной на разных этапах. |
| При анализе проводится тестирование с использованием базовых статистических критериев. | Метод не включает много статистики в отличие от генерализованных аддитивных и линейных моделей. |
| Измеряется важность переменных, производится построение кривых отклика на каждую переменную, есть возможность произвести кросс-валидацию и другие настройки параметров моделирования. | Чтобы получить оптимальную модель нужно провести анализ серии моделей с разными настройками. В ряде случаев возникают сложности из-за ограниченности настроек, что снижает возможность более детальной оценки качества результирующей модели. |
| Может быть эффективным при наличии относительно малого числа известных данных по присутствию вида. | Предвзятость выбора образцов является гораздо большей проблемой для методов, оценивающих только присутствие, в отличие от методов присутствия-отсутствия. |
| Возможность использования как непрерывных, так и категориальных экологических переменных (предикторов). | При отсутствии готовых растров с категориальными переменными необходима их предварительная подготовка с одинаковым географическим охватом и размером ячейки. |
| Это относительно современный метод, который по результатам тестов показал лучшие результаты в сравнении с другими методами, основанными на анализе точек присутствия видов. | Современные протоколы рекомендуют проводить сравнительный анализ с использованием различных алгоритмов. |
| Maxent может функционировать в комплексе с различными ГИС программами. В ряде случаев алгоритм уже встроен с эти программы. | Ряд современных пакетов по моделированию также позволяет в дальнейшем анализировать результаты с использованием ГИС. |
| Метод достаточно легко освоить, программа бесплатная и с дружелюбным интерфейсом, есть много форумов и современных публикаций по этой методике. | Простота проведения анализа не должна сказываться на подготовке данных, регулировке настроек и корректной интерпретации полученных результатов. |

Использование метода Maxent может быть успешным в следующих случаях: 1) Если недоступны данные по отсутствию вида (обычно регистрируется только присутствие вида); 2) Когда есть сомнения, что данные по отсутствию вида являются ненадежными; 3) Когда вид может временно отсутствовать на территории; 4) Если вид занимает фрагментированные местообитания.

В программе Maxent имеются различные настройки (от базовых до продвинутых), использование которых может оказать большое влияние на конечную модель потенциального распространения вида. Рассмотрим эти аспекты более подробно.

Вопросы анализа с использованием пакета Maxent довольно подробно отражены в публикации К. Мероу с соавторами (Merow et al., 2013). Р. Хальворсен с соавторами (Halvorsen et al., 2016) предлагают при работе с Maxent не использовать базовые настройки и следовать альтернативной методике Maxent («alternativeMaxEntprocedure» – aMp), которая основана на ручном поэтапном выборе ключевых переменных с применением критерия Фишера для сравнения вложенных моделей.

Нюансы работы с имеющимися входными данными и важность предварительной подготовки точек распространения вида более подробно рассматривались в более ранних публикациях (Sandanov 2019; Lissovsky, Dudov, 2020). Для полноценного анализа в Maxent необходимо равномерно распределить точки распространения вида в пространстве. Для этой процедуры можно использовать возможности различных ГИС-программ или инструмент в пакете SDMtoolbox 2.0, позволяющий распределить точки распространения вида с определенным шагом в зависимости от гетерогенности природных условий. В программном пакете Maxent в установках по умолчанию присутствует параметр (Remove duplicate presence records), исключающий точки, попадающие в одну растровую ячейку.

На конечном этапе важно иметь объем выборки необходимой для проведения корректного анализа. Число точек распространения вида может являться ключевым для получения оптимальных результатов, что показано на серии сравнительных моделей, использующих различные алгоритмы (Zimmermann et al., 2007; Wisz et al., 2008). Также определено, что для широкоареальных видов необходимо использовать в анализе большее число местонахождений (до 150) для большей достоверности полученных результатов (Feeley, Silman, 2011), что было реализовано в наших ранних исследованиях для видов восточноазиатской (Sandanov, Naidanov, 2013) и центральноазиатской приуроченности (Sandanov et al., 2020). В противовес этому отмечается, что при моделировании видов с широкими экологическими предпочтениями (виды-генералисты) можно использовать гораздо меньшее число точек распространения вида (даже всего лишь 10) с большей погрешностью (Soultan, Safi, 2017). Для видов с узкой экологической нишей (виды-специалисты) в анализе с использованием Maxent необходимо использовать не менее 50 местонахождений, при малых выборках валидность моделей заметно снижается. Экспериментально показано, что для алгоритма Maxent необходимо использование не менее 15 точек распространения вида (Papeş, Gaubert, 2007). Тем не менее, основным параметром выборки должен являться максимальный охват всех местообитаний вида, что является основой хорошего моделирования. Нашими исследованиями экспериментально показано, что в гетерогенном ландшафте исключение из анализа даже одной точки распространения вида может оказать большое влияние на результирующую модель (Sandanov, 2022).

Для редких видов или узколокальных эндемиков число точек распространения может быть намного меньше. Обычно при моделировании используется минимум 4 или 5 предикторов, но для редких видов число местонахождений может быть гораздо меньше 40 или 50 точек (исходя из базового соотношения 1 предиктор к 10 точкам регистрации вида). В таком случае лучше использовать меньшее число предикторов, например, только 2 две переменные на 20-29 точек распространения вида. Отметим, что

в Maxent существуют определенные алгоритмы, позволяющие провести моделирование и при небольших выборках (Shcheglovitova, Anderson, 2013). Также отмечается высокая прогностическая ценность моделей, выполненных в Maxent, для эндемичных видов на небольшой территории с использованием спутниковых данных высокого разрешения (Van Gils et al., 2012). В ряде случаев такие исследования позволяют спрогнозировать находки новых популяций редких видов (Williams et al., 2009; Fois et al., 2018). Другим подходом является использование всех возможных комбинаций переменных в бивариантном виде и построение «взвешенного ансамбля» (weighted ensemble). Это позволяет построить множество моделей с двумя разными наборами переменных за один анализ. Потом проводится тестирование полученных моделей по статистическим показателям, на основе которых рассчитывается взвешенное среднее вероятностей (weighted average of probabilities) каждой модели. Моделям, показавшим наилучший результат, присваивается наибольший вес, некорректные модели получают меньший вес (Breiner et al., 2017). Таким образом, моделирование ансамблей (ensemble modeling) позволяет обойти ограничение в объеме выборки для видов с небольшим числом местонахождений.

Maxent поддерживает четыре выходных формата значений модели: исходный (raw), кумулятивный (cumulative), логистический (logistic) и дважды логарифмический (cloglog). В большинстве исследований используется логистический или дважды логарифмический форматы, которые являются экспоненциалом энтропии распределения Maxent и дают оценку пригодности местообитаний в интервале между 0 и 1. Полученные прогнозные карты распространения показывают расчетную вероятность подходящих условий для изучаемого вида: на логистической шкале красным цветом показывается высокая вероятность подходящих условий; зелёным — условия, похожие на те, в которых находится вид; оттенки синего – маловероятные условия. Таким образом, представлена бинарная шкала (наличие вида/отсутствие вида) с плавным переходом между этими двумя состояниями (Philips, Dudik, 2008; Philips et al., 2017). В результатах Maxent формируется растровый файл, отображающий распределение по заданному выходному формату, который можно открыть и анализировать в различных ГИС-пакетах. Оценка полученных выходных карт и выявление различий между ними можно провести с использованием инструмента сравнения карт – Map Comparison Kit <http://mck.riks.nl/>, который позволяет провести сравнение карт полученных моделей с использованием различных алгоритмов (Visser, de Nijs, 2006). Данный пакет также позволяет анализировать и множество других растровых данных. Питер Уилсон (Wilson, 2011) предлагает метод, сочетающий аспекты анализа растровых изображений с ординационным подходом, что позволяет оценивать пространственные различия карт, полученных при моделировании. Оценка смещения ареалов изучаемых видов также можно провести в ГИС-программах при сравнении векторных слоев, сформированных из выходных растров Maxent.

Оценка качества модели в Maxent оценивается на основе расчета площади под ROC-кривой AUC (Area under Curve of the Received Operating Characteristic). По показателю AUC качество моделирования можно разделить на пять категорий: 0.9–1 = «отлично», 0.8–0.9 = «хорошо», 0.7–0.8 = «удовлетворительно», 0.6–0.7 = «плохо», < 0.6 – «очень плохо» (моделирование не удалось) (Araújo et al., 2005). Значение AUC=0.5 соответствует случайному распределению результатов моделирования.

К сожалению, в настройках Maxent нет возможности расчета других статистических критериев точности получаемых моделей, что является определенным ограничением. Рядом авторов отмечается, что анализ AUC может использоваться при оценке полученных моделей, но данный показатель не является тривиальным и может быть заменен другими более надежными критериями (Lobo et al., 2008; Jiménez-Valverde, 2012). Анализ распространения видов с разной размерностью ареалов выявил, что показатель AUC может быть надежным критерием для оценки точности различных

моделей, тогда как показатель каппа Когена (Cohen's kappa) может привести к статистическим артефактам. Также отмечается, что алгоритмы логистической регрессии и дискриминантного анализа очень чувствительны к изменению числа точек распространения вида (McPherson et al., 2004). Отметим, что вопрос относительно надежности того или иного статистического критерия при моделировании распространения видов до сих пор является дискуссионным. Определенные ограничения имеются и в дизайне моделирования с использованием более современных и широко используемых критериев, таких как статистический индекс TSS – True Skill Statistic (Somodi et al., 2017).

Оценка связей экологических переменных (предикторов) с распространением вида в Maxent оценивается на основе анализа трех основных составляющих: 1) Вклад фактора (percent contribution) – рассчитывается на основе коэффициента λ (большее значение λ подразумевает больший вклад предиктора в модель), 2) Важность при пермутации (permutation importance) – вклад предиктора оценивается на сравнении исходной модели с моделью с искаженными значениями одного из факторов, 3) тест jackknife – исключает из анализа по очереди каждую переменную и проводит сравнение моделей с учетом только одной переменной и моделей с ее отсутствием. Для грамотной интерпретации результатов моделирования необходим анализ всех трех параметров.

Оценка структуры вклада факторов среды в полученные модели возможна с применением иерархического кластерного анализа методом попарного внутригруппового невзвешенного среднего (UPGMA) (Lisovsky, Obolenskaya, 2014; Dudov, 2016).

Отмечается, что для получения оптимальной модели необходимо устранить взаимокоррелирующие переменные (Guisan, Zimmermann, 2000; Chapman, 2010). Однако, Дж. Бедия с соавторами (Bedia et al., 2011) отмечает, что при работе в Maxent можно не сокращать число переменных, поскольку данный метод использует взаимодействие между предикторами при построении моделей, а алгоритм регуляризации способствует снижению переобучения модели. Результаты исследований Р. Хальворсена с соавторами (Halvorsen et al., 2016) также не выявляют особого эффекта пространственной автокорреляции на качество получаемых моделей.

В любом случае до начала моделирования необходимо оценить корреляции между используемыми переменными, чтобы в дальнейшем исключить их из анализа или для понимания сходного отклика коррелирующих переменных в результирующей модели. Экспериментально показана необходимость устранения взаимокоррелирующих переменных со значением выше 0.7 (Dormann et al., 2013).

Например, предварительный анализ экологических переменных BIOCLIM и ENVIREM для модельного вида *Scutellaria baicalensis* использованием коэффициента корреляции Пирсона показывает, что имеющиеся данные характеризуются значительными взаимными корреляциями (рис. 2).

Так, переменная BIO 1 (среднегодовая температура) демонстрирует сильные (0.70-0.97) прямые и обратные связи с переменными BIO 2 (средняя суточная амплитуда температуры за каждый месяц), BIO 4 (сезонность температуры), BIO 6 (минимальная температура самого холодного месяца года), BIO 7 (годовая амплитуда температуры), BIO 9 (средняя температура самого сухой четверти года), BIO 11 (средняя температура самой холодной четверти года), а также с годовыми значениями потенциальной эвапотранспирации (annual PET) и ее среднемесячными значениями в самой холодной четверти года (PETColdestQuarter), индексом континентальности (continentality), плювиометрическим фактором Эмбергера (embergerQ), суммой среднемесячных температур для месяцев со средней температурой выше 0°C (growingDegDays0), максимальной температурой самого холодного месяца года (maxTempColdest) и числом месяцев со средней температурой выше 10 градусов (monthCountByTemp10). Важно отметить, что для определенного вида и

модели необходимы индивидуальные расчеты корреляции переменных, т.к. взаимодействие факторов среды в каждом случае является различным.

Ряд авторов, в дополнение используют различные методы ординации (PCA, CVA, CCA, NMDS) и производят выбор переменных, которые лучше объясняются на различных осях (Hirzel et al., 2002, Polce et al., 2013). Однако, в ряде случаев при таком подходе экологическая интерпретация переменных оказывается более сложной, чем простой анализ первичного набора предикторов. Можно также провести сравнительный анализ с полным набором переменных и выделенными переменными. Для этого необходимо выбрать определенное количество случайных точек (число зависит от территории исследования) в изучаемом регионе и на основе расчета вероятностей в каждой из точек провести сравнение моделей классическими статистическими критериями, такими как t-критерий Стьюдента или U-критерий Манна-Уитни.

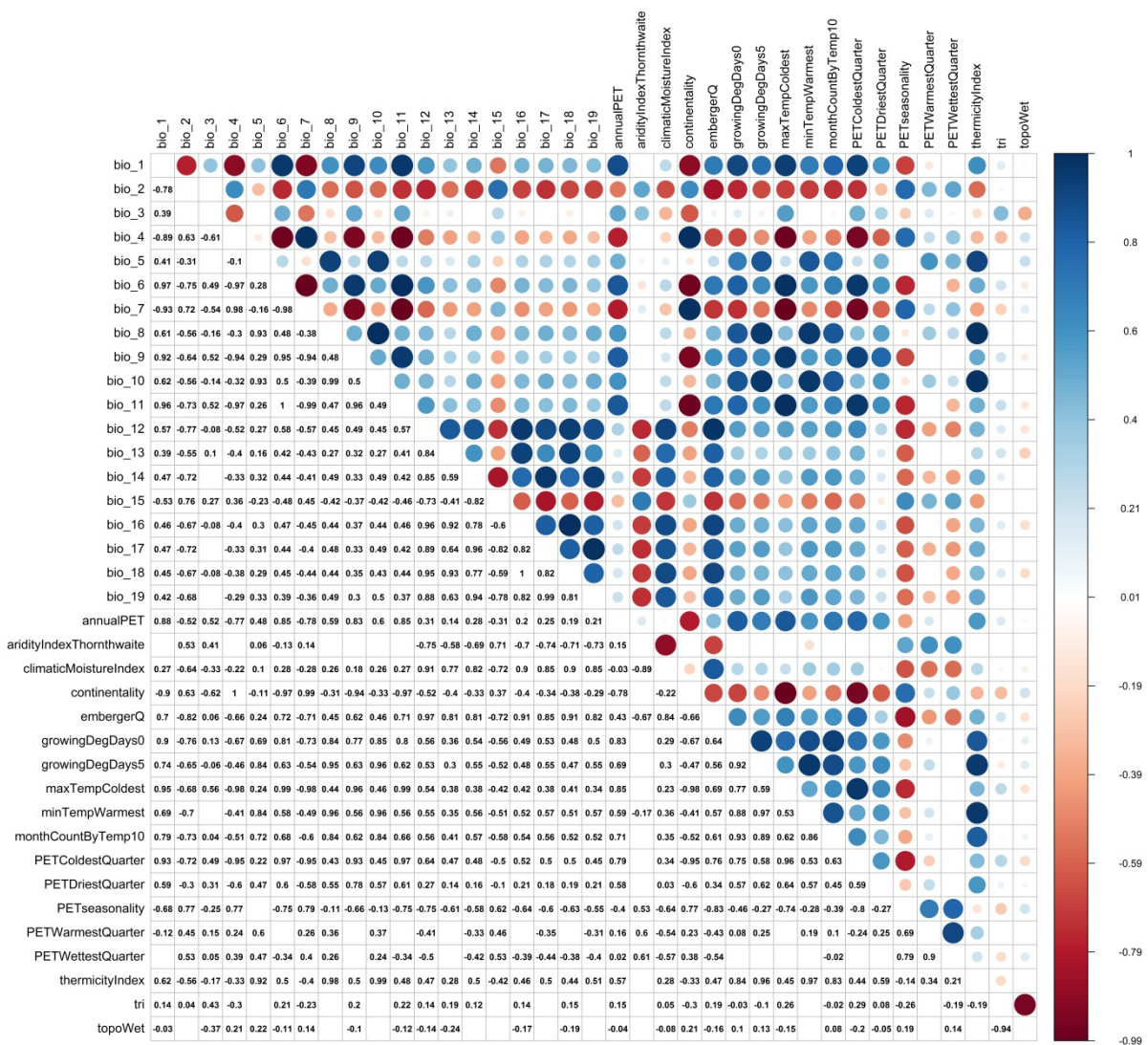


Рис. 2. Корреляционная матрица экологических переменных для местонахождений *Scutellaria baicalensis*

Обозначения: bio_1-bio_19 – биоклиматические переменные (обозначения переменных представлены на сайте <https://www.worldclim.org/>), остальные переменные ENVIREM (обозначения переменных представлены на сайте <https://envirem.github.io/>).

Цифрами обозначены коэффициенты корреляции по Пирсону, эти же показатели дублируются графически в виде кругов (размер круга зависит от силы корреляционных связей), цветовые градации соответствуют шкале справа.

Fig. 2. Correlation matrix of environmental variables for occurrence data of *Scutellaria baicalensis*

Figure captions: bio_1-bio_19 – bioclimatic environmental variables (their meaning is presented on the website <https://www.worldclim.org/>), other variables are ENVIREM (their meaning is presented on the website <https://envirem.github.io/>). Figures in the cells are Pearson's correlation coefficients, which are also shown as circles with diameters proportional to correlation strength. Color grades correspond to the scale on the right.

Возможности Maxent позволяют просмотреть кривые отклика на каждую переменную в выходных файлах и принять решение на этой основе анализа этих показателей. Иногда бывает полезным на первоначальном этапе провести анализ со всем набором переменных с оценкой важности по тесту jackknife, а потом провести сравнительное моделирование после устранения переменных, имеющих маленький вклад в построение модели.

В среде R разработан удобный инструмент, позволяющий провести корректный выбор переменных при моделировании в Maxent (Jueterbock, 2015). Можно также использовать анализ дисперсионного фактора инфляции (Variance Inflation Factor) в пакете BiodiversityR, что позволяет проверить коллинеарность переменных и провести выбор предикторов, необходимых для дальнейшего моделирования (Kindt, 2014).

При использовании различных экологических переменных в Maxent необходимо, чтобы все растровые данные были идентичными по разрешению, размеру и географической проекции. Наиболее быстрым способом подготовки слоев экологических переменных являются готовые пакеты и библиотеки в среде R, такие как raster, rgdal, sf, tidyverse, fasterize. Подготовка растров возможна и с использованием ГИС-программ, но процессы конвертирования немного отличаются в разных ГИС-пакетах.

Можно отметить, что алгоритм Maxent показал наилучшие результаты при переносе исходной модели в другие экологические условия и географически далекие друг от друга регионы (Duque-Lazo et al., 2016). Авторы метода упоминают, что подробный анализ вклада переменных для тестовых данных способствует принятию решения о переносимости модели, в том числе и для оценки распространения вида при условии климатических изменений (Philips, Dudik, 2008). Нашими исследованиями выявлены высокие прогнозные характеристики Maxent при использовании различного набора точек распространения вида с разной степенью детализации (Sandanov, Pisarenko, 2018). Хорошие результаты моделирования в Maxent можно получить при использовании многоступенчатого анализа, который включает последовательную калибровку модели на основе разработки множественных моделей с различными настройками (Ramírez-Gil et al., 2018).

ЗАКЛЮЧЕНИЕ

Таким образом, Maxent является методом моделирования с использованием только данных по присутствию вида. Это современный инструмент, который широко используется в различных исследованиях в глобальном и региональном масштабе. Представленный обзор возможностей метода Maxent показывает, что для получения корректных моделей необходим тщательный подбор всех компонентов моделирования: точки распространения вида, экологические переменные, детализация настроек

анализа. В конечном итоге выбор оптимальной модели проводится на основе множества разработанных моделей.

БЛАГОДАРНОСТИ

Исследования выполнены в рамках государственного задания по темам № 121030900138-8 и № 2-22-128-1.

СПИСОК ЛИТЕРАТУРЫ

Araújo M.B., Pearson R.G., Thuiller W., Erhard M. 2005. Validation of species-climate impact models under climate change. – *Glob Change Biol.* 11 (9): 1504–1513. <https://doi.org/10.1111/j.1365-2486.2005.01000.x>

Araújo M.B., Anderson R.P., Barbosa A.M., Beale C.M., Dormann C.F., Early R., Garcia R.A., Guisan A., Maiorano L., Naimi B., O'Hara R.B., Zimmermann N.E., Rahbek C. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv.* 5 (1): eaat4858. <https://doi.org/10.1126/sciadv.aat4858>

Bedia J., Busqué J., Gutiérrez J.M. 2011. Predicting plant species distribution across an alpine rangeland in northern Spain. A comparison of probabilistic methods. – *Appl Veg Sci.* 14 (3): 415–432. <https://doi.org/10.1111/j.1654-109X.2011.01128.x>

Breiner F.T., Nobis M.P., Bergamini A., Guisan A. 2017. Optimizing ensembles of small models for predicting the distribution of species with few occurrences. – *Methods Ecol. Evol.* 9 (4): 802–808. <https://doi.org/10.1111/2041-210X.12957>

Chapman D.S. 2010. Weak climatic associations among British plant distributions. – *Glob. Ecol. Biogeogr.* 19: 831–841. <https://doi.org/10.1111/j.1466-8238.2010.00561.x>

Dormann C.F., Elith J., Bacher S., Buchmann C., Carl G., Carré G., García Marquéz J.R., Gruber B., Lafourcade B., Leitão P.J., Münkemüller T., McClean C., Osborne P.E., Reineking B., Schröder B., Skidmore A.K., Zurell D., Lautenbach S. 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. – *Ecography.* 36 (1): 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>

[Dudov] Дудов С.В. 2016. География ботанического разнообразия хребта Тукурингра (на примере Зейского государственного природного заповедника): Дисс. ... канд. геогр. наук. Москва. 169 с.

Duque-Lazo J., van Gils T.H., Groen T.A., Navarro-Cerrillo R.M. 2016. Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. – *Ecol. Modell.* 320: 62–70. <https://doi.org/10.1016/j.ecolmodel.2015.09.019>

Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J. McC. M., Townsend Peterson A., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S., Zimmermann N.E. Novel methods improve prediction of species' distributions from occurrence data. – *Ecography* 29: 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Elith J., Leathwick J.R. 2009. Species distribution models: Ecological explanation and prediction across space and time. – *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>

Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E., Yates C.J. 2011. A statistical explanation of MaxEnt for ecologists. – *Divers. Distrib.* 17: 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>

Feeley K.J., Silman M.R. 2011. Keep collecting: accurate species distribution modelling requires more collections than previously thought. – *Divers. Distrib.* 17: 1132–1140. <https://doi.org/10.1111/j.1472-4642.2011.00813.x>

- Fois M., Cuena-Lombrana A., Fenu G., Bacchetta G. 2018. Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions. – *Ecol. Modell.* 385: 124–132. <https://doi.org/10.1016/j.ecolmodel.2018.07.018>
- Guisan A., Zimmermann N.E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Modell.* 135: 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Halvorsen R., Mazzoni S., Dirksen J.W., Næsset E., Gobakken T., Ohlson M. 2016. How important are choice of model selection method and spatial autocorrelation of presence data for distribution modelling by MaxEnt? – *Ecol. Modell.* 328: 108–118. <https://doi.org/10.1016/j.ecolmodel.2016.02.021>
- Hirzel A.H., Hausser J., Chessel D., Perrin N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83 (7): 2027–2036. [https://doi.org/10.1890/0012-9658\(2002\)083\[2027:ENFAHT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2027:ENFAHT]2.0.CO;2)
- Jiménez - Valverde A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modeling. – *Glob. Ecol. Biogeogr.* 21 (4): 498–507. <https://doi.org/10.1111/j.1466-8238.2011.00683.x>
- Jueterbock A. 2015. R package MaxentVariableSelection: selecting the Best Set of Relevant Environmental Variables along with the Optimal Regularization Multiplier for Maxent Niche Modeling. <https://cran.r-project.org/web/packages/MaxentVariableSelection/index.html>
- Kindt R. 2014. BiodiversityR: GUI for Biodiversity, Suitability and Community Ecology Analysis. <https://rdr.io/cran/BiodiversityR/man/BiodiversityRGUI.html>
- [Lissofsky, Obolenskaya] Лисовский А.А., Оболенская Е.В. 2014. Исследование ареалов мелких млекопитающих Юго-Восточного Забайкалья методом моделирования экологической ниши. – *Журн. общ. биологии.* 75 (5): 353–371.
- [Lissofsky, Dudov] Лисовский А.А., Дудов С.В. 2020. Преимущества и ограничения методов экологического моделирования ареалов. 2. MAXENT. – *Журн. общ. биологии.* 81(2): 135–146. <https://doi.org/10.31857/S0044459620020049>
- Lobo J.M., Jiménez - Valverde A., Real R. 2008. AUC: a misleading measure of the performance of predictive distribution models. – *Glob. Ecol. Biogeogr.* 17 (2): 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
- McPherson J. M., Jetz W., Rogers D.J. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? – *J. Appl. Ecol.* 41: 811–823. <https://doi.org/10.1111/j.0021-8901.2004.00943.x>
- Merow C., Smith M.J., Silander J.A. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. – *Ecography.* 36 (10): 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Papeş M., Gaubert P. 2007. Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. – *Divers. Distrib.* 13 (6): 890–902. <https://doi.org/10.1111/j.1472-4642.2007.00392.x>
- Phillips S. J., Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. – *Ecography* 31: 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Phillips S.J., Anderson R.P., Dudík M., Schapire R.E., Blair M.E. 2017. Opening the black box: an open-source release of Maxent. – *Ecography* 40 (7): 887–893. <https://doi.org/10.1111/ecog.03049>
- Polce C., Termansen M., Aguirre-Gutiérrez J., Boatman N.D., Budge G.E., Crowe A., Garratt M.P., Pietravalle S., Potts S.G., Ramirez J.A., Somerwill K.E., Biesmeijer J.C. 2013. Species distribution models for crop pollination: a modelling framework applied to Great Britain. – *PLoS ONE* 8(10): e76308. <https://doi.org/10.1371/journal.pone.0076308>

Ramírez-Gil J.G., Morales J.G., Peterson A.T. 2018. Potential geography and productivity of “Hass” avocado crops in Colombia estimated by ecological niche modeling. *Sci. Hortic.* 237: 287–295. <https://doi.org/10.1016/j.scienta.2018.04.021>

[Sandanov, Naidanov] Санданов Д.В., Найданов Б.Б. 2015. Пространственное моделирование ареалов восточно-азиатских видов растений: современное состояние и динамика под влиянием климатических изменений. – *Растительный мир Азиатской России*. 3 (19): 30–35.

Sandanov D.V., Pisarenko O.Yu. 2018. Bioclimatic modeling of *Crossidium squamiferum* (Viv.) Jur. (Pottiaceae, Bryophyta) distribution. – *Arctoa*. 27: 29–34. <https://doi.org/10.15298/arctoa.27.03>

[Sandanov] Санданов Д.В. 2019. Современные подходы к моделированию разнообразия и пространственному распределению видов растений: перспективы их применения в России. – *Вестник Томского государственного университета. Биология*. 46: 82–114. <https://doi.org/10.17223/19988591/46/5>

[Sandanov et al] Санданов Д.В., Дугарова А.С., Селютина И.Ю. 2020. Моделирование распространения видов секции *Xerobia* Bunge рода *Oxytropis* DC. на территории Центральной Азии при климатических изменениях в прошлом и будущем. – *Вестник Томского государственного университета. Биология*. 52: 85–104. <https://doi.org/10.17223/19988591/52/5>

[Sandanov] Санданов Д.В. 2022. Особенности работы с базами данных по распространению растений и опыт консолидации данных различного формата. – *Природа Внутренней Азии*. 4 (22): 96–104. <https://doi.org/10.18101/2542-0623-2022-4-96-104>

Shcheglovitova M., Anderson R.P. 2013. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. – *Ecol. Modell.* 269: 9–17.

Somodi I., Lepesi N., Botta-Dukát Z. 2017. Prevalence dependence in model goodness measures with special emphasis on true skill statistics. – *Ecol. Evol.* 7 (3): 863–872. <https://doi.org/10.1002/ece3.2654>

Soultan A., Safi K. 2017. The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialization. – *PLoS ONE* 12(11): e0187906. <https://doi.org/10.1371/journal.pone.0187906>

Van Gils H., Conti F., Ciaschetti G., Westinga E. 2012. Fine resolution distribution modelling of endemics in Majella National Park, Central Italy. – *Plant Biosystems*. 146: 276–287. <https://doi.org/10.1080/11263504.2012.685194>

Visser H., de Nijs T. 2006. The Map Comparison Kit. *Environ. Model. Softw.* 21(3): 346–358. <https://doi.org/10.1016/j.envsoft.2004.11.013>

Williams J.N., Seo C., Thorne J., Nelson J.K., Erwin S., O’Brien J.M., Schwartz M.W. 2009. Using species distribution models to predict new occurrences for rare plants. – *Divers. Distrib.* 15 (4): 565–576. <https://doi.org/10.1111/j.1472-4642.2009.00567.x>

Wilson P.D. 2011. Distance - based methods for the analysis of maps produced by species distribution models. – *Methods Ecol. Evol.* 2 (6): 623–633. <https://doi.org/10.1111/j.2041-210X.2011.00115.x>

Wisz M.S., Hijmans R.J., Li J., Peterson A.T., Graham C.H., Guisan A., NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. – *Divers. Distrib.* 14 (5): 763–773. <https://doi.org/10.1111/j.1472-4642.2008.00482.x>

Zimmermann N.E., Edwards Jr T.C., Moisen G.G., Frescino T.S., Blackard J.A. 2007. Remote sensing - based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. – *J. Appl. Ecol.* 44: 1057–1067. <https://doi.org/10.1111/j.1365-2664.2007.01348.x>

Zurell D., Franklin J., König C., Bouchet P.J., Dormann C.F., Elith J., Fandos G., Feng X., Guillera-Arroita G., Guisan A., Lahoz-Monfort J.J., Leitão P.J., Park D. S., Peterson A. T., Rapacciuolo G., Schmatz D.R., Schröder B., Serra-Diaz J. M., Thuiller W., Yates K.L., Zimmermann N.E., Merow C. 2020. A standard protocol for reporting species distribution models. – *Ecography*. 43 (9): 1261–1277. <https://doi.org/10.1111/ecog.04960>

MODELING FEATURES OF SPECIES DISTRIBUTION ON THE BASE OF MAXIMUM ENTROPY METHOD

© 2023 D.V. Sandanov^{1,2}

¹*Institute of General and Experimental Biology SB RAS*

Sakhyanovoi Str., 6, Ulan-Ude, 670047, Russia

²*Tunkinsky National Park*

Lenina Str., 130, Kyren, 671010, Russia

e-mail: sdenis1178@mail.ru

Abstract. Species distribution modeling by maximum entropy methods on the base of Maxent program is widely implemented in different studies. It is important to correctly use this method, including preparation of initial data and setup of various program features in order to get valid results. The modern publications on this topic have been summarized in a review. Different peculiar properties and possibilities of analysis in Maxent have been presented.

Key words: species distribution modeling (SDM), Maxent, maximum entropy, advantages and limitations of the method.

Submitted: 27.01.2023. **Accepted for publication:** 15.04.2023.

For citation: Sandanov D.V. 2023. Modeling features of species distribution on the base of maximum entropy method. — *Phytodiversity of Eastern Europe*. 17(2): 130–144. DOI: 10.24412/2072-8816-2023-17-2-130-144

ACKNOWLEDGEMENTS

This work was supported by the Russian Federal budget (Project Numbers 121030900138-8, № 2-22-128-1).

REFERENCES

Araújo M.B., Pearson R.G., Thuiller W., Erhard M. 2005. Validation of species-climate impact models under climate change. – *Glob Change Biol*. 11 (9): 1504–1513. <https://doi.org/10.1111/j.1365-2486.2005.01000.x>

Araújo M.B., Anderson R.P., Barbosa A.M., Beale C.M., Dormann C.F., Early R., Garcia R.A., Guisan A., Maiorano L., Naimi B., O'Hara R.B., Zimmermann N.E., Rahbek C. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv*. 5 (1): eaat4858. <https://doi.org/10.1126/sciadv.aat4858>

Bedia J., Busqué J., Gutiérrez J.M. 2011. Predicting plant species distribution across an alpine rangeland in northern Spain. A comparison of probabilistic methods. – *Appl Veg Sci*. 14 (3): 415–432. <https://doi.org/10.1111/j.1654-109X.2011.01128.x>

Breiner F.T., Nobis M.P., Bergamini A., Guisan A. 2017. Optimizing ensembles of small models for predicting the distribution of species with few occurrences. – *Methods Ecol. Evol.* 9 (4): 802–808. <https://doi.org/10.1111/2041-210X.12957>

Chapman D.S. 2010. Weak climatic associations among British plant distributions. – *Glob. Ecol. Biogeogr.* 19: 831–841. <https://doi.org/10.1111/j.1466-8238.2010.00561.x>

Dormann C.F., Elith J., Bacher S., Buchmann C., Carl G., Carré G., García Marquéz J.R., Gruber B., Lafourcade B., Leitão P.J., Münkemüller T., McClean C., Osborne P.E., Reineking B., Schröder B., Skidmore A.K., Zurell D., Lautenbach S. 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. – *Ecography*. 36 (1): 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>

Dudov S.V. 2016. Geografiya botanicheskogo raznoobraziya khrebta Tukucingra (na primere Zeiskogo gosudarstvennogo prirodnogo zapovednika) [Geography of botanical diversity of Tukuringra mountain range (on the example of Zeya State Nature Reserve)]: Diss. ... Kand. Sci Moscow. 169 p. (In Russ.).

Duque-Lazo J., van Gils T.H., Groen T.A., Navarro-Cerrillo R.M. 2016. Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. – *Ecol. Modell.* 320: 62–70. <https://doi.org/10.1016/j.ecolmodel.2015.09.019>

Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J. McC. M., Townsend Peterson A., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S., Zimmermann N.E. Novel methods improve prediction of species' distributions from occurrence data. – *Ecography* 29: 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Elith J., Leathwick J.R. 2009. Species distribution models: Ecological explanation and prediction across space and time. – *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>

Elith J., Philips S.J., Hastie T., Dudík M., Chee Y.E., Yates C.J. 2011. A statistical explanation of MaxEnt for ecologists. – *Divers. Distrib.* 17: 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>

Feeley K.J., Silman M.R. 2011. Keep collecting: accurate species distribution modelling requires more collections than previously thought. – *Divers. Distrib.* 17: 1132–1140. <https://doi.org/10.1111/j.1472-4642.2011.00813.x>

Fois M., Cuenca-Lombraña A., Fenu G., Bacchetta G. 2018. Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions. – *Ecol. Modell.* 385: 124–132. <https://doi.org/10.1016/j.ecolmodel.2018.07.018>

Guisan A., Zimmermann N.E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Modell.* 135: 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)

Halvorsen R., Mazzoni S., Dirksen J.W., Næsset E., Gobakken T., Ohlson M. 2016. How important are choice of model selection method and spatial autocorrelation of presence data for distribution modelling by MaxEnt? – *Ecol. Modell.* 328: 108–118. <https://doi.org/10.1016/j.ecolmodel.2016.02.021>

Hirzel A.H., Hausser J., Chessel D., Perrin N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83 (7): 2027–2036. [https://doi.org/10.1890/0012-9658\(2002\)083\[2027:ENFAHT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2027:ENFAHT]2.0.CO;2)

Jiménez - Valverde A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modeling. – *Glob. Ecol. Biogeogr.* 21 (4): 498–507. <https://doi.org/10.1111/j.1466-8238.2011.00683.x>

Jueterbock A. 2015. R package MaxentVariableSelection: selecting the Best Set of Relevant Environmental Variables along with the Optimal Regularization Multiplier for

- Maxent Niche Modeling. <https://cran.r-project.org/web/packages/MaxentVariableSelection/index.html>
- Kindt R. 2014. BiodiversityR: GUI for Biodiversity, Suitability and Community Ecology Analysis. <https://rdr.io/cran/BiodiversityR/man/BiodiversityRGUI.html>
- Lissovsky A.A., Obolenskaya E.V. 2014. Issledovaniye arealov melkikh mlekopitayushikh Yugo-Vostochnogo Zabaikalya metodom modelirovaniya ekologicheskoy nishi [A study of the distribution ranges of small mammals from South-Eastern Transbaikalia with ecological niche modeling methods]. – Zhurn. Obshch. Biologii 75 (5): 353–371. (In Russ.).
- Lissovsky A.A., Dudov S.V. 2020. Preimushchestva i ogranicheniya metodov ekologicheskogo modelirovaniya arealov. 2. MAXENT. [Advantages and limitations of application of the species distribution modeling methods. 2. MAXENT]. – Zhurn. Obshch. Biologii 81(2): 135–146. <https://doi.org/10.31857/S0044459620020049> (In Russ.).
- Lobo J.M., Jiménez - Valverde A., Real R. 2008. AUC: a misleading measure of the performance of predictive distribution models. – Glob. Ecol. Biogeogr. 17 (2): 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
- McPherson J. M., Jetz W., Rogers D.J. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? – J. Appl. Ecol. 41: 811–823. <https://doi.org/10.1111/j.0021-8901.2004.00943.x>
- Merow C., Smith M.J., Silander J.A. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. – Ecography. 36 (10): 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Papeş M., Gaubert P. 2007. Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. – Divers. Distrib. 13 (6): 890–902. <https://doi.org/10.1111/j.1472-4642.2007.00392.x>
- Phillips S. J., Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. – Ecography 31: 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Phillips S.J., Anderson R.P., Dudík M., Schapire R.E., Blair M.E. 2017. Opening the black box: an open-source release of Maxent. – Ecography 40 (7): 887–893. <https://doi.org/10.1111/ecog.03049>
- Polce C., Termansen M., Aguirre-Gutiérrez J., Boatman N.D., Budge G.E., Crowe A., Garratt M.P., Pietravalle S., Potts S.G., Ramirez J.A., Somerwill K.E., Biesmeijer J.C. 2013. Species distribution models for crop pollination: a modelling framework applied to Great Britain. – PLoS ONE 8(10): e76308. <https://doi.org/10.1371/journal.pone.0076308>
- Ramírez-Gil J.G., Morales J.G., Peterson A.T. 2018. Potential geography and productivity of “Hass” avocado crops in Colombia estimated by ecological niche modeling. Sci. Hortic. 237: 287–295. <https://doi.org/10.1016/j.scienta.2018.04.021>
- Sandanov D.V., Naidanov B.B. 2015. Prostranstvennoye modelirovaniye arealov vostochno-aziatskikh vidov rasteniy: sovremennoye sostoyaniye i dinamika pod vliyaniem klimaticheskikh izmeneniy [Spatial modeling of East-Asian plant species distribution: current condition and future dynamic under climatic change]. Rastitel'ny Mir Aziatskoy Rossii. 3 (19): 30–35. (In Russ.).
- Sandanov D.V., Pisarenko O.Yu. 2018. Bioclimatic modeling of *Crossidium squamiferum* (Viv.) Jur. (Pottiaceae, Bryophyta) distribution. – Arctoa. 27: 29–34. <https://doi.org/10.15298/arctoa.27.03>
- Sandanov D.V. 2019. Sovremennyye podkhody k modelirovaniyu raznoobraziya i prostranstvennomu raspdeleniyu vidov rasteniy: perspektivy ikh primeneniya v Rossii [Modern approaches to modeling of plant diversity and spatial distribution of plant species: implication prospects in Russia]. – Vestnik Tomskogo gosudarstvennogo universiteta. Biologiya. 46: 82–114. <https://doi.org/10.17223/19988591/46/5> (In Russ.).

Sandanov D.V., Dugarova A.S., Selyutina I.Yu. 2020. Modelirovaniye rasprostraneniya vidov seksii *Xerobia* Bunge roda *Oxytropis* DC. pri klimaticheskikh izmeneniyakh v proshlom i budushchem [Species distribution modeling for the section *Xerobia* Bunge of the genus *Oxytropis* DC. on the territory of Central Asia under past and future climate change]. – Vestnik Tomskogo gosudarstvennogo universiteta. *Biologiya*. 52: 85–104. <https://doi.org/10.17223/19988591/52/5> (In Russ.).

Sandanov D.V. Д.В. 2022. Osobennosti raboty s bazami dannykh po rasprostraneniyu rasteniy i opyt konsolidatsii danntkh razlichnogo formata [Features of work with databases on plant species distribution and the experience in combining data of various formats]. – Priroda Vnutrenney Azii. 4 (22): 96–104. <https://doi.org/10.18101/2542-0623-2022-4-96-104> (In Russ.).

Shcheglovitova M., Anderson R.P. 2013. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. – *Ecol. Modell.* 269: 9–17.

Somodi I., Lepesi N., Botta-Dukát Z. 2017. Prevalence dependence in model goodness measures with special emphasis on true skill statistics. – *Ecol. Evol.* 7 (3): 863–872. <https://doi.org/10.1002/ece3.2654>

Soultan A., Safi K. 2017. The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialization. – *PLoS ONE* 12(11): e0187906. <https://doi.org/10.1371/journal.pone.0187906>

Van Gils H., Conti F., Ciaschetti G., Westinga E. 2012. Fine resolution distribution modelling of endemics in Majella National Park, Central Italy. – *Plant Biosystems*. 146: 276–287. <https://doi.org/10.1080/11263504.2012.685194>

Visser H., de Nijs T. 2006. The Map Comparison Kit. *Environ. Model. Softw.* 21(3): 346–358. <https://doi.org/10.1016/j.envsoft.2004.11.013>

Williams J.N., Seo C., Thorne J., Nelson J.K., Erwin S., O'Brien J.M., Schwartz M.W. 2009. Using species distribution models to predict new occurrences for rare plants. – *Divers. Distrib.* 15 (4): 565–576. <https://doi.org/10.1111/j.1472-4642.2009.00567.x>

Wilson P.D. 2011. Distance - based methods for the analysis of maps produced by species distribution models. – *Methods Ecol. Evol.* 2 (6): 623–633. <https://doi.org/10.1111/j.2041-210X.2011.00115.x>

Wisz M.S., Hijmans R.J., Li J., Peterson A.T., Graham C.H., Guisan A., NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. – *Divers. Distrib.* 14 (5): 763–773. <https://doi.org/10.1111/j.1472-4642.2008.00482.x>

Zimmermann N.E., Edwards Jr T.C., Moisen G.G., Frescino T.S., Blackard J.A. 2007. Remote sensing - based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. – *J. Appl. Ecol.* 44: 1057–1067. <https://doi.org/10.1111/j.1365-2664.2007.01348.x>

Zurell D., Franklin J., König C., Bouchet P.J., Dormann C.F., Elith J., Fandos G., Feng X., Guillera-Aroita G., Guisan A., Lahoz-Monfort J.J., Leitão P.J., Park D. S., Peterson A. T., Rapacciuolo G., Schmatz D.R., Schröder B., Serra-Diaz J. M., Thuiller W., Yates K.L., Zimmermann N.E., Merow C. 2020. A standard protocol for reporting species distribution models. – *Ecography*. 43 (9): 1261–1277. <https://doi.org/10.1111/ecog.04960>