



SPECIES DISTRIBUTION MODELLING USING MAXENT: OVERVIEW AND PROSPECTS

Yuliia Novoseltseva 

Key words

MaxEnt model, geographic range
analysis, environmental factors

doi

<http://doi.org/10.53452/TU2809>

Article info

submitted 19.12.2024

revised 23.12.2024

accepted 30.12.2024

Language

Ukrainian, English summary

Affiliations

Taras Shevchenko National Uni-
versity of Kyiv (Kyiv, Ukraine)

Correspondence

Yuliia Novoseltseva; Taras
Shevchenko National University
of Kyiv, 60 Volodymyrska Street,
Kyiv, 01033 Ukraine;
Email: julianovoseltse-
va1909@gmail.com;
orcid: 0009-0005-0095-7069

Abstract

Niche modelling of species has become increasingly important in the context of accelerating climate change and anthropogenic impacts on the biosphere. One such tool for predicting the potential distribution of species is the maximum entropy method (MaxEnt). This method is particularly valuable when working with biodiversity data collected from herbaria and museum collections, as such data typically only contain information about where a species has been recorded, rather than where it is absent. It is precisely this feature of MaxEnt that makes it an indispensable tool for biodiversity research based on historical data. This allows for the reconstruction of historical species ranges, the detection of changes in their distribution, and the forecasting of future trends, namely the prediction of potential ranges, the assessment of the impact of climate change and anthropogenic pressure, and the development of effective biodiversity conservation strategies. This article provides a brief overview of the MaxEnt software's operating principle, its capabilities, and limitations. In particular, it analyses the impact of data quality on modelling results and considers various approaches to assessing the importance of ecological factors for species distribution. One of the key issues discussed in the article is the problem of sampling bias. Sampling bias arises because data on the presence of species are often collected non-randomly and depend on the accessibility of the locality, the interests of researchers, and other factors. This can lead to distortions in modelling results. Various methods can be used to correct these biases, such as the bias grid method and the background points method. Another important aspect is the choice of the territory for the background sample. It should be taken into account that when using projections where cells have different areas, MaxEnt may give incorrect results. The article also emphasises the need for cautious interpretation of modelling results. Assessing niche models solely based on AUC (area under ROC curve) can be misleading, therefore, for a more reliable assessment of variable importance, it is worth supplementing it with permutation importance and the jackknife method. Examples of modelling for various groups, including mammals of the Ukrainian fauna, were considered.

Cite as

Novoseltseva, Y. 2024. Species distribution modelling using MaxEnt: overview and prospects. *Theriologia Ukrainica*, **28**: 102–112. [In Ukrainian, with English summary]

Моделювання поширення видів засобами MaxEnt: огляд і перспективи

Юлія Новосельцева

Резюме. Моделювання екологічних ніш видів набуває все більшої актуальності в контексті зростаючих темпів зміни клімату та антропогенного впливу на біосферу. Одним із таких інструментів для прогнозування потенційного ареалу видів є метод максимальної ентропії (MaxEnt). Цей метод особливо цінний під час роботи з даними про біорізноманіття, зібраними в гербаріях та музейних колекціях, оскільки такі дані, як правило, містять інформацію лише про місця, де вид був зареєстрований, а не про місця його відсутності. Саме така особливість MaxEnt — робить його незамінним інструментом для дослідження біорізноманіття на основі історичних даних. Це дає змогу реконструювати історичні ареали видів, виявляти зміни в їхньому поширенні та прогнозувати майбутні тенденції, а саме передбачати потенційні ареали, оцінювати вплив кліматичних змін та антропогенного тиску, а також розробляти ефективні стратегії збереження біорізноманіття. Стаття присвячена короткому розгляду принципу роботи програми MaxEnt, її можливостей та обмежень. Зокрема, аналізується вплив якості даних на результати моделювання, а також розглядаються різні підходи до оцінювання важливості екологічних факторів для розподілу виду. Одним із ключових питань, що розглядаються в статті, є проблема зміщення вибірки. Зміщення вибірки виникає через те, що дані про присутність видів часто збираються не випадковим чином, а залежать від доступності місцевості, інтересів дослідників та інших факторів. Це може призвести до спотворення результатів моделювання. Для корекції цих зміщень можуть бути використані різні методи, такі як метод сітки зміщень та метод фонових точок. Іншим важливим аспектом є вибір території для фонового зразка. Необхідно враховувати, що при використанні проєкцій, де клітинки мають різну площу, MaxEnt може давати некоректні результати. Стаття також наголошує на необхідності обережної інтерпретації результатів моделювання. Оцінка моделей ніш виключно за AUC (площа під ROC-кривою) може бути оманливою, тому для більш надійної оцінки значимості змінних варто доповнювати її показниками важливості перестановки та методом jackknife. Розглянуто приклади моделювання щодо різних груп, зокрема і ссавців фауни України.

Ключові слова: модель Maxent, аналіз географічних ареалів, фактори середовища.

Вступ

Моделювання поширення видів є потужним інструментом у сучасній екології та біогеографії. Воно дає змогу прогнозувати зміни ареалів видів у відповідь на кліматичні зміни, антропогенний вплив та інші фактори. Однак, точність таких моделей значною мірою залежить від якості та повноти даних, що використовуються. Класичні методи моделювання поширення видів зазвичай базуються на даних про присутність та відсутність видів у певних локаціях. Ці дані дають змогу виявити асоціації між поширенням виду та екологічними факторами, такими як клімат, рельєф, тип ґрунту тощо. Однак, такий підхід має певні обмеження. Інформація про відсутність виду в певній локації часто є менш надійною, ніж дані про його присутність. Це пов'язано з тим, що відсутність виду може бути наслідком не лише відсутності відповідних умов, але й недостатньої вибірки, помилок у визначенні тощо [Elith *et al.* 2011].

З огляду на зазначені обмеження, в останні роки з'явилися нові методи моделювання поширення видів, які базуються виключно на даних про присутність видів. Особливе значення це має для музейних колекцій та гербаріїв. Ці установи зберігають унікальні дані про минуле та сучасне біорізноманіття, які можуть бути використані для реконструкції історичних ареалів видів, виявлення змін у їхньому поширенні та прогнозування майбутніх тенденцій [Phillips & Dudík 2008]. Одним із таких методів є метод максимальної ентропії — MaxEnt [Phillips *et al.* 2006]. MaxEnt — це програмне забезпечення, яке займає своє місце серед лідерів в аналізі розподілу видів та їх взаємодії з біо- і геокліматичними градієнтами. Крім цієї програми, існують й інші відомі програми для подібних аналізів, такі як ENMTools, Ecospat, і Biomod2. Кожна з цих програм має свої особливості та переваги, що дозволяє дослідникам обирати найкращий інструмент для конкретного завдання.

Його основна ідея полягає в тому, щоби побудувати модель розподілу виду, яка максимально відповідає наявним даним про місця, де вид був зареєстрований, але водночас не робить жодних додаткових припущень про розподіл виду там, де він не був знайдений. Іншими словами, модель намагається знайти найпростіше пояснення спостережуваного розподілу виду. Однією з найважливіших сфер застосування MaxEnt є прогнозування змін в ареалах видів в умовах зміни клімату, фрагментації середовищ існування та інших антропогенних впливів [Yackulic *et al.* 2013]. Ідея MaxEnt вперше виникла в Центрі біорізноманіття та збереження при Американському музеї природної історії (AMNH) завдяки державно-приватному партнерству між AMNH та AT&T-Research [Phillips 2017].

MaxEnt — це потужний інструмент для моделювання потенційного розподілу видів, особливо корисний, коли відсутні дані про місця, де вид точно не трапляється. Він дає змогу побудувати модель, яка максимально відповідає наявним даним про присутність виду. Однак, як і будь-який інший метод, MaxEnt має свої обмеження [Phillips 2017]. У нашій статті ми зосередимося на тому, яких правил слід дотримуватися для отримання точних результатів моделювання та як працювати з програмою.

Принцип роботи програми

Запуск програми

Програмне забезпечення MaxEnt створене для моделювання видових ніш і розподілу шляхом застосування техніки машинного навчання під назвою моделювання максимальної ентропії та є безплатною. Завантажити програму ви можете на сайті Американського музею природної історії (http://biodiversityinformatics.amnh.org/open_source/maxent/) [Phillips 2017].

MaxEnt оцінює поширення (географічний ареал) виду, знаходячи розподіл, який має максимальну ентропію (тобто є найближчим до географічно рівномірного) з урахуванням обмежень, що впливають з умов навколишнього середовища в зареєстрованих місцях появи. Обмеження визначені в термінах «характеристик» (змінних навколишнього середовища, таких як температура, і простих функцій цих змінних, таких як квадратичні члени), і вимагають, щоб середнє значення кожної характеристики відповідало середньому значенню вибірки [Phillips *et al.* 2017].

Детальний посібник із використання пропонує покрокову інструкцію зі встановлення та запуску програми [Phillips 2017], у нашій статті ми зупинимося на основних моментах.

Для запуску програми вам необхідно надати файл знахідок присутності (знахідок) виду (або кількох видів) та каталог (папку) з файлами, що містять растрові сітки ASCII (у форматі ESRI.asc), кожна з яких описує змінну середовища. Усі сітки мають мати однакові географічні межі та розмір комірки (тобто всі заголовки файлів ASCII мають ідеально відповідати один одному). До прикладу кліматичні зміни середовища, які можна використати для моделювання, доступні на сайті WordClim (<https://www.worldclim.org/>). Важливо зазначити, що перед використанням шарів у дослідженні, їх необхідно кадрувати до меж досліджуваної території за допомогою ГІС-програм.

Модель MaxEnt пропонує чотири різні способи представлення результатів прогнозування: необроблений, кумулятивний, логістичний та cloglog. Кожен із цих форматів має свої особливості та інтерпретацію.

- Необроблений формат є найпростішим форматом, який являє собою безпосередньо значення експоненціальної моделі MaxEnt. Він відображає відносну придатність кожної точки простору для виду, але не має прямої інтерпретації як ймовірність.
- Кумулятивний формат вказує на відсоток площі, де значення моделі не перевищує заданого порогу. Він корисний для оцінки ймовірності відсутності виду в певній зоні.
- Логістичний формат: цей формат перетворює необроблені значення в діапазон від 0 до 1, що дозволяє інтерпретувати результати як ймовірності. Однак, масштабування цих ймовірностей залежить від параметрів моделі.

- Cloglog формат також перетворює необроблені значення в діапазон від 0 до 1, але має більш пряму інтерпретацію як ймовірність присутності виду. Він базується на припущенні, що в середньому на одиницю площі припадає одна особина виду.

Вибір формату виведення залежить від конкретної задачі дослідження та бажаної інтерпретації результатів. Cloglog формат є найбільш поширеним, оскільки він дає змогу отримати інтуїтивно зрозумілі оцінки ймовірності присутності виду.

Статистична обробка

Програма також передбачає статистичну обробку результатів, основним методом є площа під кривою ROC (AUC) [Fielding & Bell 1997]. Площа під ROC-кривою (AUC) є фундаментальною метрикою для оцінки ефективності моделей, які прогнозують наявність або відсутність певного явища (наприклад, виду, захворювання тощо) у певних місцях.

Інтуїтивно, AUC представляє ймовірність того, що випадково обрана точка з позитивним класом (тобто, де явище присутнє) буде оцінена моделлю як більш «позитивна», ніж випадково обрана точка з негативним класом (де явище відсутнє).

Цей показник варіюється від 0.5 (випадкове вгадування) до 1.0 (ідеальна модель) [Elith *et al.* 2011]. Однак, на практиці часто виникають ситуації, коли відсутні дані про негативний клас. У таких випадках використовують «псевдовідсутності» — випадково обрані точки з досліджуваної області, які заповнюють відсутні дані. Застосування псевдовідсутностей дозволяє оцінити AUC, але інтерпретація результатів дещо змінюється. У цьому випадку AUC показує ймовірність того, що випадково обрана точка з позитивним класом буде ранжована вище, ніж випадково обрана фонові точка (псевдовідсутність) [Phillips & Dudík, 2008].

Оцінка моделей ніш виключно за показником AUC може бути оманливою, оскільки високе значення AUC не завжди вказує на якісну модель. Вид, обмежений певними умовами або більш схильний до процесу локального вимирання, може мати низький AUC, навіть якщо модель правильно відображає його екологічні вимоги. Крім того, маніпуляції з кількістю фонових точок можуть штучно підвищити AUC, не покращуючи якість моделі. Таким чином, надмірна реляція на AUC як єдиний критерій оцінки моделей ніш може призводити до некоректних висновків і потребує перегляду [Yackulic *et al.* 2013].

Огляд результатів

Одним із ключових застосувань моделювання розподілу видів є визначення найвпливовіших факторів середовища, які впливають на поширення конкретного виду. MaxEnt, як одна з найпоширеніших моделей для вирішення цього завдання, пропонує кілька підходів до аналізу важливості змінних.

Процес навчання моделі MaxEnt передбачає ітеративне уточнення моделі (процес послідовного вдосконалення моделі шляхом багаторазового повторення певних обчислень або алгоритмів), що описує відповідність виду певним умовам середовища. На кожному кроці алгоритму модель намагається покращити свою здатність прогнозувати присутність виду, змінюючи вагові коефіцієнти, присвоєні різним змінним середовища. Ці коефіцієнти відображають ступінь впливу кожної змінної на розподіл виду. Переводячи у відсотки по закінченню навчання моделі, отримаємо таку таблицю (рис. 1).

Оцінювання внеску змінних у моделі MaxEnt має певні обмеження. Значення відсоткового внеску (другий стовпчик), що визначаються алгоритмом, можуть варіюватися залежно від обраного шляху оптимізації. Крім того, якщо змінні тісно пов'язані між собою, інтерпретація їхніх значень ускладнюється. Наприклад, висока взаємозалежність між річною кількістю опадів та кількістю опадів в окремих місяцях означає, що високий відсотковий внесок для одного з місяців не обов'язково вказує на його більшу значимість для виду.

Для більш надійної оцінки важливості змінних використовують додатковий показник — важливість перестановки (третій стовпчик). Цей показник базується на аналізі зміни якості

моделі при випадковій перестановці значень конкретної змінної. Чим сильніше знижується якість моделі після перестановки, тим важливішою є ця змінна для прогнозування розподілу виду. Цей показник є більш стійким до впливу алгоритму оптимізації та кореляцій між змінними.

Для отримання більш надійної оцінки впливу різних змінних на розподіл виду в моделі MaxEnt часто використовується метод jackknife. Якщо ви хочете використати його при моделюванні, то оберіть відповідний прапорець при запуску програми. Цей метод дозволяє зменшити вплив випадкових флуктуацій даних на оцінку важливості кожної змінної та забезпечує більш стабільні результати.

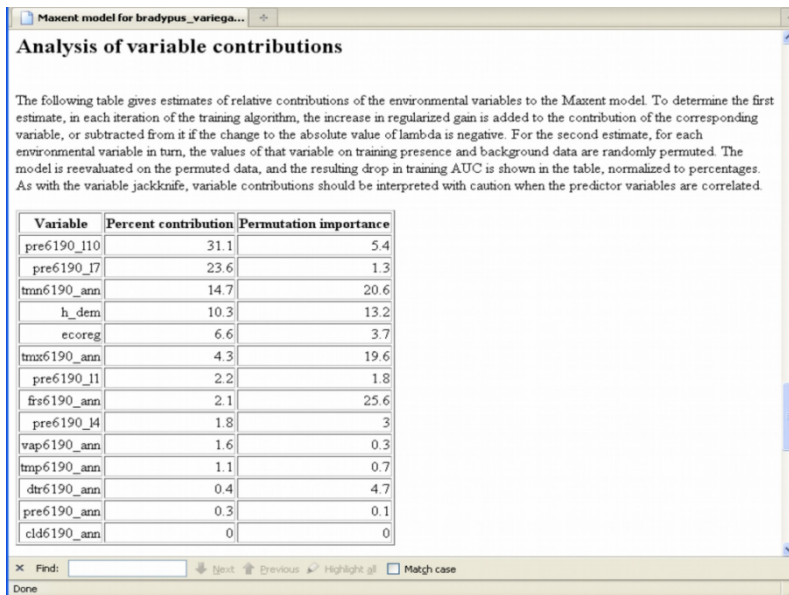


Рис. 1. Відсотковий внесок кожного значення параметра [Phillips 2017].

Fig. 1. Percentage contribution of each parameter value [Phillips 2017].

Суть методу jackknife полягає в послідовному виключенні кожної змінної з набору даних і подальшому побудові моделі на основі решти змінних. Тобто, для кожної змінної створюється окрема модель, у якій ця змінна відсутня. Крім того, будується модель, що включає всі змінні, як і в базовому аналізі. Порівнюючи якість моделей, отриманих із виключенням різних змінних, з моделлю, що включає всі змінні, можна оцінити, наскільки сильно кожна змінна впливає на прогнози моделі.

Результати використання jackknife під час моделювання в MaxEnt відображаються у вигляді трьох гістограм: тест на основі тренувальних даних (тобто ті, що програма використала для навчання моделі), тестувальних даних (дані, які відклала програма для тестування моделі) та на основі AUC-приросту (дані результату статистичної обробки). Стовпці діаграми вказують наскільки кожна змінна є важливою для моделі: сині сектори — при використанні лише однієї змінної, блакитні сектори — при виключенні цієї змінної з використанням усіх інших.

На рис. 2 (тест на основі тренувальних даних) бачимо, що змінна «середня кількість опадів у січні» (pre6190_11) виявилася малоінформативною для прогнозування розподілу виду. Включення цієї змінної в модель не призвело до суттєвого покращення якості прогнозування. На противагу цьому, змінна «кількість опадів у жовтні» (pre6190_110) продемонструвала значно більшу передбачувальну здатність, дозволяючи моделі адекватно відтворювати розподіл виду за тренувальними даними. Можна звернути увагу на те, що деякі світло-блакитні смужки довші за червону смугу, що свідчить про покращення ефективності прогнозування за відсутності відповідних змінних.

Для детальнішого аналізу того, як прогноз залежить від змінних, ми можемо використати опцію створення кривих відгуку моделі (для цього треба вибрати відповідний прапорець на початку запуску програми).

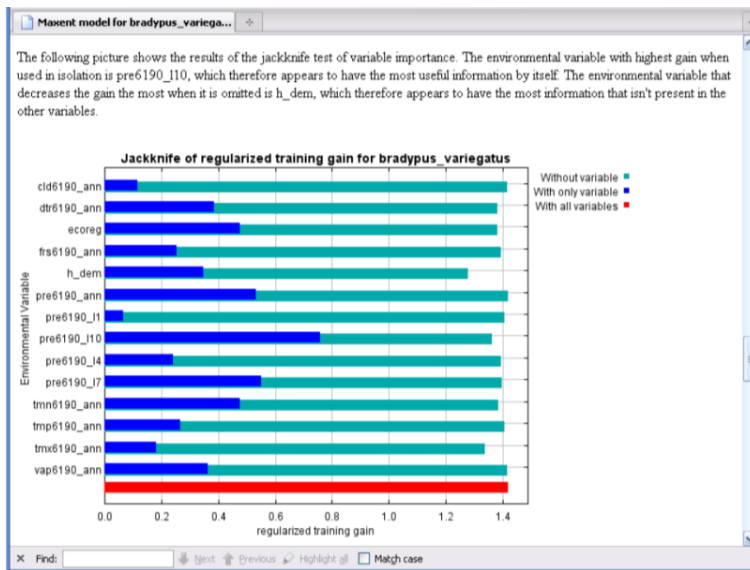


Рис. 2. Тест на основі тренувальних даних [Phillips 2017].

Fig. 2. Test based on training data [Phillips 2017].

- Перший тип кривих показує, як прогнозована ймовірність присутності виду в певній точці змінюється, коли змінюється кожна змінна середовища, зберігаючи всі інші змінні середовища на їх середньому значенні для всієї вибірки.
- Другий тип кривих, на відміну від першого, дозволяє аналізувати, як кожен показник впливає на прогнозовану ймовірність присутності, коли використовується лише він один.

На рис. 3 бачимо приклад таких кривих, значення, відображене на осі ординат, є прогнозованою ймовірністю відповідних умов, як задано форматом логістичного виводу, з усіма іншими змінними, встановленими на їхнє середнє значення для набору місць присутності. По вісі абсцис задано значення окремих показників (вологість, температура і т.д.).

Особливості застосування програми MaxEnt

Оскільки ми вже розглянули частину практичного застосування програми MaxEnt, перейдемо до оцінки її загальних можливостей та обмежень. У цьому розділі ми проаналізуємо особливості, які необхідно пам'ятати при використанні програми. Такий всебічний аналіз дозволить нам краще зрозуміти, коли і як ефективно застосовувати MaxEnt для вирішення конкретних наукових завдань.

Основна та найважливіша перевага програми це моделювання на основі лише даних про присутність. Два найпоширеніші приклади наборів даних лише про наявність — музейні зразки та гербарні записи. Такі дані зазвичай не є результатом випадкової вибірки, а відображають суб'єктивні рішення колекціонерів щодо того, які зразки збирати й де.

Наприклад, колекціонери часто віддають перевагу легкодоступним місцям, таким як узбіччя доріг, або місцям, де вид, що цікавить, уже був помічений [Pearce & Boyce 2006]. Це призводить до того, що в таких даних переважно представлені певні типи місцевостей, що спотворює загальну картину розподілу виду. Наслідки такого зміщення можуть бути серйозними. Якщо не враховувати його при побудові моделей, то отримані результати можуть бути упередженими й не відображати реальний розподіл виду. Наприклад, якщо більшість даних про вид зібрано вздовж доріг, то модель може помилково визначити дороги як оптимальне середовище існування для цього виду, навіть якщо це не так. Саме тому проблема зміщення вибірки є одним із найбільших викликів у моделюванні екологічних ніш видів за допомогою даних про присутність.

Для того, щоб врахувати ці зміщення, Філіпс та його колеги [Phillips *et al.* 2009] запропонували два методи: метод сітки зміщень та метод фонових точок.

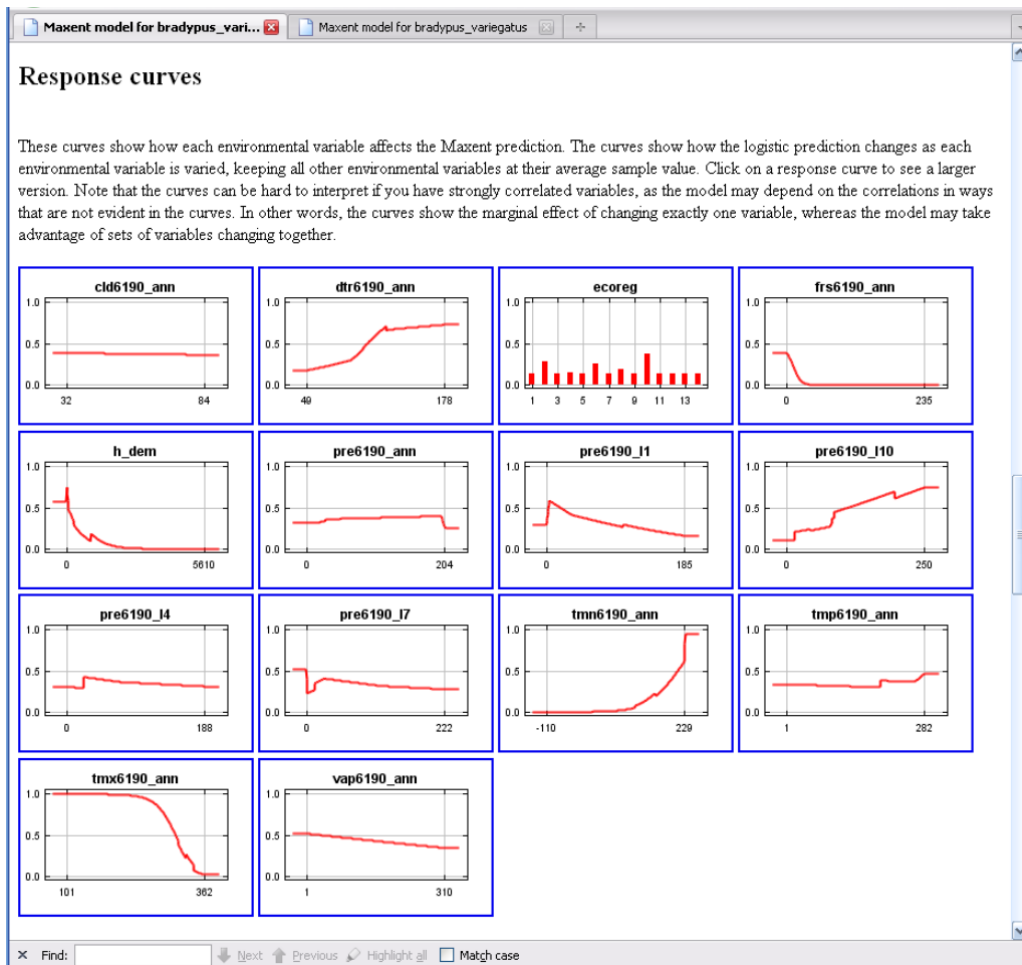


Рис. 3. Графіки відгуку моделі [Phillips 2017].

Fig. 3. Response plots of the model [Phillips 2017].

Метод сітки зміщень передбачає створення сітки, яка відображає відомі зміщення у вибірці. Наприклад, якщо відомо, що більшість зразків було зібрано вздовж доріг, то можна створити сітку, яка буде мати більшу щільність точок вздовж доріг. Ця сітка потім використовується для генерації фонових точок, які будуть враховувати відомі зміщення та скорегувати нашу модель розподілу видів. Основною проблемою цього методу є те, що інформація про зміщення вибірки часто буває неповною або відсутньою. У багатьох випадках, особливо при роботі з історичними даними, важко точно визначити, які фактори вплинули на вибір місць для збору зразків. Тому, застосування методу сітки зміщень може бути обмеженим.

Метод фонових точок передбачає використання даних про інші види, присутні в тому ж регіоні, як фонових точок. Ідея полягає в тому, що якщо інші види також зазнали впливу тих же факторів, які призвели до зміщення вибірки для фокусного виду, то використання цих даних допоможе коригувати це зміщення. Цей метод є більш гнучким, ніж метод сітки зміщень, оскільки він не вимагає детальної інформації про характер зміщення. Однак, у нього є свої обмеження. По-перше, цей метод передбачає, що всі види зазнали впливу однакових факторів, що може бути невірним. По-друге, якщо фокусний вид має дуже вузьку екологічну нішу, то може бути важко знайти інші види, які мають схожий розподіл.

Якулік з кол. [Yackulic *et al.* 2013] проаналізували 108 статей, що опубліковані між 2008 та 2012 рр. і стосувалися моделювання поширення видів засобами MaxEnt. Вони рекомендують дотримуватися таких правил при проведенні дослідження:

1. Оцінюючи наявні дані, необхідно визначити, чи можна їх аналізувати за принципом «присутній-відсутній». Якщо ні, то слід з'ясувати, чи збір даних проводився за чіткою схемою, або ж був хаотичним. У випадку хаотичного збору, існують методи корекції таких похибок, як підвибірка або моделювання. При цьому важливо з'ясувати, чи суперечать ці похибки нашим очікуванням щодо розподілу виду. І, нарешті, варто обміркувати, чи не доцільніше розробити нову схему збору даних, яка б дозволила отримати більш точні результати.

2. Ймовірність виявлення виду під час дослідження завжди менша за одиницю і може варіюватися залежно від умов середовища. Це означає, що навіть якщо вид присутній у певній місцевості, ми можемо його не виявити. Для оцінки ймовірності появи виду, крім даних про його присутність, необхідно враховувати ймовірність виявлення. Якщо дані дозволяють використовувати моделі присутності-відсутності, то за допомогою додаткових спостережень можна оцінити як ймовірність виявлення, так і ймовірність появи виду. Однак, коли такі моделі не застосовні, а ймовірність виявлення залежить від умов середовища, для оцінки ймовірності появи необхідні більш складні моделі та додаткові дані. Якщо ж ймовірність виявлення є сталою, то існують спеціальні методи, які дозволяють оцінити ймовірність появи виду навіть за відсутності даних про відсутність, хоча точність таких оцінок може бути нижчою.

3. Перед початком моделювання необхідно сформулювати попередні припущення (априорні гіпотези) щодо того, як розподіл виду залежить від умов середовища. Якщо ми впевнені, що ймовірність вибору місця для дослідження та ймовірність виявлення виду є стабільними, то ми можемо зосередитися на тому, як кількість особин виду змінюється залежно від різних факторів середовища. Однак, якщо ці ймовірності можуть змінюватися і ми не можемо їх точно контролювати, то нам доведеться враховувати, як ці зміни можуть впливати на нашу оцінку розподілу виду. Важливо уникати ситуації, коли ми підбираємо модель так, щоб вона підтверджувала наші попередні припущення, навіть якщо вони не відповідають дійсності.

4. Отримані в результаті моделювання залежності між даними слід ретельно перевірити на відповідність попереднім припущенням. У разі розбіжностей, необхідно сформулювати нові гіпотези, які пояснюють отримані результати, та розробити план подальших досліджень для їх перевірки.

5. Обов'язковою вимогою до наукової публікації є забезпечення можливості для інших дослідників перевірити та порівняти отримані результати. Для цього необхідно надавати детальну інформацію про використані методи, моделі та отримані дані. Зокрема, окрім візуальної презентації результатів у вигляді карт, слід включати числові характеристики моделей, такі як криві відгуку та значення параметрів. Це значно підвищить прозорість дослідження і дозволить іншим науковцям використовувати отримані результати у своїй роботі. Крім того, бажано надавати відкритий доступ до вихідних даних через онлайн-репозиторії, якщо це дозволено законодавством. Такий підхід сприяє розвитку науки й дозволяє будувати моделі на досягненнях попередніх досліджень.

Чітке окреслення територій для моделювання є таким же важливим, як і вирішення проблеми зміщення вибірки. При моделюванні екологічних ніш за допомогою MaxEnt необхідно ретельно обирати територію для фонового зразка. Необхідно враховувати, що при використанні проєкцій, де клітинки мають різну площу (наприклад, WorldClim), MaxEnt може давати некоректні результати через припущення про однакову площу клітинок. Для вирішення цієї проблеми можна проєктувати сітки на проєкцію рівної площі, створювати сітку варіацій площі комірок або використовувати власний фоновий зразок із відповідними вагами вибірки [Elith *et al.* 2011].

Вибір ландшафту для фонового зразка також критичний [VanDerWal *et al.* 2009]. Він повинен охоплювати весь екологічний діапазон виду і виключати території, які точно не були досліджені. Для місцевих ендеміків фоновий зразок повинен включати території, де вид міг потенційно поширитися. Виключення областей зі зразка фону можна здійснити за допомогою масок. Прогнози для виключених областей можна отримати за допомогою засобів проєкування, але це потребує додаткової обережності [Yackulic *et al.* 2013].

Перспективи застосування

Моделювання нішевих моделей, зокрема з використанням алгоритму MaxEnt, стало потужним інструментом у сучасній екології та біогеографії. Воно дозволяє прогнозувати потенційний ареал видів, оцінювати вплив кліматичних змін та антропогенного тиску на біорізноманіття, а також розробляти стратегії збереження видів та екосистем.

Ми розглянемо потенціал цього методу для вирішення актуальних екологічних проблем, таких як поширення інвазійних видів, аналіз змін клімату, питання збереження біорізноманіття та управління природними ресурсами.

Інвазивні види рослин і тварин є однією з найсерйозніших загроз для глобального біорізноманіття. Швидке поширення цих видів призводить до деградації екосистем, втрати біологічного різноманіття та значних економічних збитків. Тому розробка ефективних стратегій управління інвазіями є актуальним завданням сучасної екології. Одним з ключових аспектів управління інвазіями є раннє виявлення та прогнозування потенційних зон поширення інвазивних видів. Це дозволяє зосередити зусилля на профілактичних заходах та мінімізувати витрати на боротьбу з вже встановленими популяціями.

Дослідження, проведене за допомогою MaxEnt в Австралії [Wilson *et al.* 2009], дозволило виявити різну реакцію двох видів інвазивних трав на зміну клімату. Модель продемонструвала, що зміна клімату може як сприяти, так і обмежувати поширення інвазивних трав. Тропічний вид *Andropogon gayanus* розширює свій ареал, тоді як помірний вид *Nassella neesiana* скорочує його. Це свідчить про необхідність індивідуального підходу до управління кожним інвазивним видом та адаптації стратегій до очікуваних кліматичних змін.

Моделювання є потужним інструментом для виявлення рідкісних та ендемічних видів. Воно дозволяє прогнозувати потенційні місця проживання цих видів на основі даних про їх відомі локалізації та екологічні умови. Застосування MaxEnt значно підвищує ефективність пошукових робіт, допомагає зберегти біорізноманіття та розробити ефективні стратегії охорони рідкісних видів.

Моделюючи потенційний ареал за допомогою MaxEnt науковці з США [Rhoden *et al.* 2017] успішно прогнозували нові місцезнаходження і розподіл середовищ існування для двох рідкісних видів раків (*Fallicambarus harpi* та *Procambarus reimeri*) з незвичайною екологією, що поєднує наземний і водний способи життя. Під час польових досліджень, проведених за прогнозами MaxEnt, дослідники виявили нові популяції обох видів та розширили відомі межі їхнього поширення, що демонструє високу точність моделей.

Окрім того, застосування цієї програми дозволяє детально проаналізувати взаємодію між видами та їхнім середовищем існування, виявити ключові фактори, що впливають на розподіл видів та оцінити вплив різних чинників на природні екосистеми. Наприклад, дослідження, проведене в Зімбабве [Mupfiga *et al.* 2024], мало на меті виявити причини частих пожеж рослинності за допомогою моделі MaxEnt. Аналізуючи дані про пожежі та екологічні фактори, вчені встановили, що висота над рівнем моря, сезонність опадів, температурні коливання та людська діяльність є основними чинниками, які впливають на виникнення пожеж. Модель MaxEnt дозволила точно прогнозувати ймовірність виникнення пожеж, що є важливим для розробки ефективних стратегій боротьби з ними. Результати дослідження підкреслюють необхідність комплексного підходу до управління пожежами, який враховує як природні, так і антропогенні фактори.

Дослідження MaxEnt в українському контексті

Слід зазначити, що алгоритм MaxEnt знаходить все ширше застосування і в українських екологічних та зоологічних дослідженнях. Цей метод активно застосовується для вивчення розподілу видів флори і фауни, прогнозування наслідків зміни клімату та розробки стратегій охорони біорізноманіття.

До прикладу, за останні роки проведено ряд досліджень, серед них — аналіз поширення ліскульки рудої (*Muscardinus avellanarius*) на території України [Novoseltseva 2024], моделю-

вання розподілу собачого кліща (*Ixodes ricinus*) з акцентом на сучасний і майбутній стани клімату на території України й Латвії [Tytar *et al.* 2024], аналіз впливу структура лісових масивів і насаджень на розповсюдження вузькозлатки ясеневіої смарагдової (*Agrilus planipennis*) [Meshkova *et al.* 2024], моделювання поширення мишака жовтогрудого (*Sylvaemus tauricus*) як потенційного переносника збудника хвороби Лайма [Tytar *et al.* 2023], тощо.

Коротко зупинимося на останньому дослідженні. Згаданий вид гризунів є важливим переносником збудників хвороб Лайма та анаплазмозу в Україні. Дослідження за допомогою моделювання показало, що західна і центральна частини України та Крим є найбільш сприятливими для поширення цього виду гризунів. Серед факторів, які впливають на розподіл мишака жовтогрудого, найважливішими є температура, особливості ґрунту та наземного покриття. Виявлена кореляція між поширенням збудника хвороби Лайма та ареалом мишака жовтогрудого підтверджує важливість цього гризуна в епідеміології хвороби Лайма в Україні.

Отже, перспективи використання MaxEnt в екологічних дослідженнях є надзвичайно широкими. Цей метод може стати незамінним інструментом для вивчення впливу кліматичних змін на біорізноманіття, оцінки ефективності природоохоронних заходів та розробки стратегій збереження видів.

Висновки

Моделювання ніш видів за допомогою MaxEnt є потужним інструментом для дослідження біорізноманіття та прогнозування змін ареалів видів. Цей метод дозволяє побудувати модель розподілу виду, яка максимально відповідає наявним даним про місця, де вид був зареєстрований. Однак, для отримання достовірних результатів необхідно враховувати ряд важливих факторів.

По-перше, якість даних про присутність виду має вирішальне значення. Зміщення вибірки, які часто притаманні таким даним, можуть суттєво вплинути на результати моделювання. Для корекції цих зміщень можна використовувати різні методи, такі як метод сітки зміщень та метод фонових точок. По-друге, важливо правильно вибирати територію для фонового зразка. Вона повинна охоплювати весь екологічний діапазон виду і виключати території, де вид точно не зустрічається. Крім того, необхідно враховувати неоднорідність простору та використовувати відповідні проекції для сіток. По-третє, інтерпретація результатів моделювання вимагає обережності. Оцінка моделей ніш виключно за показником AUC може бути оманливою. Для більш надійної оцінки важливості змінних використовується додатковий показник — важливість перестановки. Крім того, для отримання більш стабільних результатів рекомендується використовувати метод jackknife.

Незважаючи на свої обмеження, MaxEnt є потужним інструментом, який широко використовується в сучасній екології. Для отримання достовірних результатів необхідно ретельно підбирати дані, правильно налаштовувати модель і обережно інтерпретувати результати.

Моделювання ніш видів за допомогою MaxEnt є важливим інструментом для дослідження біорізноманіття та прогнозування змін ареалів видів. Однак, для отримання достовірних результатів необхідно ретельно підбирати дані, правильно налаштовувати модель і обережно інтерпретувати результати. Подальший розвиток цього методу дозволить більш точно прогнозувати вплив кліматичних змін та антропогенного впливу на біорізноманіття.

Подяки

Автор висловлює подяку керівнику дослідження Станіславу Мякушку за сприяння у проведенні досліджень.

Декларації

Фінансування. Це дослідження було виконано в рамках дипломного проекту на здобуття ступеня магістра (Київський національний університет імені Тараса Шевченка).

Конфлікт інтересів. Автор не має жодних конфліктів інтересів, які могли б вплинути на зміст цієї статті.

Поводження з матеріалом. Дослідження не передбачало роботу з живим матеріалом.

References

- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and distributions*, **17** (1): 43–57. [CrossRef](#)
- Meshkova, V., O. Zinchenko, V. Us, Y. Skrylnyk. 2024. Emerald ash borer in the park with a long-time history of black ash sawfly defoliation. *Environmental and Earth Sciences Proceedings*, **31** (1): 4. [CrossRef](#)
- Mupfiga, U., O. Mutanga, T. Dube. 2024. Assessing drivers of vegetation fire occurrence in Zimbabwe-insights from Maxent modelling and historical data analysis. *Remote Sensing Applications: Society and Environment*, **37**: 101404 (1–13). [CrossRef](#)
- Novoseltseva, Y. 2024. Analysis of the distribution of *Muscardinus avellanarius* (Rodentia) in Ukraine using the Maxent model. *Theriologia Ukrainica*, **27**: 48–54. [CrossRef](#)
- Pearce, J. L., M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, **43** (3): 405–412. [CrossRef](#)
- Phillips, S. J., R. P. Anderson, R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190** (3–4): 231–259. [CrossRef](#)
- Phillips, S. J., M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31** (2): 161–175. [CrossRef](#)
- Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, [et al.]. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological applications*, **19** (1): 181–197. [CrossRef](#)
- Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, M. E. Blair. 2017. Opening the black box: An open-source release of Maxent. *Ecography*, **40** (7): 887–893. [CrossRef](#)
- Phillips, S. J. 2017. *A Brief Tutorial on Maxent*. [URL](#)
- Rhoden, C. M., W. E. Peterman, C. A. Taylor. 2017. Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. *PeerJ*, **5**: 1–21. [CrossRef](#)
- Tytar, V., O. Lashkova, I. Dzeverin, I. Kozynenko. 2023. Species distribution modelling of the yellow-necked mouse, *Sylvemus tauricus*, with special reference to Ukraine. *Theriologia Ukrainica*, **26**: 147–158. [CrossRef](#)
- Tytar, V., I. Kozynenko, M. Pupins, A. Škute, A. Čeirāns, J. Y. Georges, O. Nekrasova. 2024. Species distribution modeling of *Ixodes ricinus* (Linnaeus, 1758) under current and future climates, with a special focus on Latvia and Ukraine. *Climate*, **12** (11): 184. [CrossRef](#)
- Yackulic, C. B., R. Chandler, E. F. Zipkin, J. A. Royle, J. D. Nichols, [et al.]. 2013. Presence-only modelling using Maxent: when can we trust the inferences? *Methods in Ecology and Evolution*, **4** (3): 236–243. [CrossRef](#)
- VanDerWal, J., L. P. Shoo, C. Graham, S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological modelling*, **220** (4): 589–594. [CrossRef](#)
- Wilson, P. D., P. O. Downey, M. Leishman, R. Gallagher, L. Hughes, J. O'Donnell. 2009. Weeds in a warmer world: predicting the impact of climate change on Australia's alien plant species using MaxEnt. *Plant Protection Quarterly*, **24** (3): 84–87.