Theriologia Ukrainica, **30**: 55–66 (2025)
p-ISSN 2616-7379 • e-ISSN 2617-1120
DOI: 10.53452/TU3007

UNVEILING THE ECOLOGICAL DRIVERS OF THE GREAT JERBOA'S RANGE: A SPECIES DISTRIBUTION MODEL WITH IMPLICATIONS FOR PLAGUE RISK

Sergiy Mezhzherin[1], Volodymyr Tytar[1], Hanna Rashevska[2], Alina Potopa[2]

Key words

Allactaga major, species distribution modelling (SDM), ecological niche, phenological synchrony, zoonotic risk

doi

http://doi.org/10.53452/TU3007

Article info

submitted 04.12.2025
revised 20.12.2025
accepted 30.12.2025

Language

English, Ukrainian summary

Affiliations

[1]I. I. Schmalhausen Institute of Zoology, NAS of Ukraine (Kyiv, Ukraine); [2]Kryvyi Rih State Pedagogical University (Kryvyi Rih, Ukraine)

Correspondence

Dr. Volodymyr Tytar;
I. I. Schmalhausen Institute of Zoology, NAS of Ukraine, 15 Bohdan Khmelnytsky Street, Kyiv, 01054 Ukraine;
Email: vtytar@gmail.com;
orcid: 0000-0002-0864-2548

Abstract

The great jerboa (*Allactaga major*), a keystone rodent of Eurasian deserts and steppes, is of dual conservation and epidemiological concern, being Near Threatened and a natural reservoir of plague. To understand the fundamental drivers of its distribution and identify potential plague reservoir zones, we developed a robust Species Distribution Model (SDM) using a comprehensive set of climatic, soil, and vegetation variables across its Eurasian range. Occurrence data were refined and modelled using the Maxent algorithm within the 'flexsdm' framework, with model interpretation advanced via SHAP (SHapley Additive exPlanations) values. Our model accurately predicted the species' known range from Eastern Europe to Central Asia. SHAP analysis revealed that climate, rather than soil or vegetation biomass, acts as the primary, range-defining filter. The three most influential predictors were Precipitation of the Driest Week (Bio14), Temperature Annual Range (Bio07), and Minimum Temperature of the Coldest Week (Bio06), defining thresholds for aridity tolerance, continentality, and hibernation survival, respectively. Notably, the highest-ranked variable, Bio14, which coincides with the late-winter (February–March) period preceding hibernation emergence, revealed a finely tuned ecological mechanism. The SHAP dependence plot showed a distinct non-linear optimum, where suitability peaks at approximately 6 mm of precipitation. This window likely represents the essential cue for germinating the annual ephemerals that form the critical post-hibernation food pulse, a link supported by a strong correlation ($r = 0.68$) between this precipitation and April vegetation greenness (NDVI). This shifts the understanding of the species' distribution from one of simple physiological tolerance to obligate ecological synchrony. Consequently, areas of high predicted suitability, particularly in southern and eastern Kazakhstan (e.g. Zhambyl, Turkistan, and Almaty oblasts), delineate a continuous ecological corridor representing potential enzootic plague reservoir zones. Our SDM thus transcends a predictive map to diagnose the core abiotic constraints and a key trophic bottleneck defining the species' niche, providing a vital evidence base for both targeted conservation strategies and proactive, risk-based public health surveillance in endemic plague regions.

Cite as

Mezhzherin, S., V. Tytar, H. Rashevska, A. Potopa. 2025. Unveiling the ecological drivers of the great jerboa's range: a species distribution model with implications for plague risk. *Theriologia Ukrainica*, **30**: 55–66. [In English, with Ukrainian summary]

Висвітлення екологічних факторів формування ареалу тушкана великого: модель поширення виду та її значення для оцінки ризиків чуми»

Сергій Межжерін, Володимир Титар, Ганна Рашевська, Аліна Потопа

Резюме. Тушкан великий (*Allactaga major*), ключовий гризун євразійських пустель і степів, має подвійне значення — як з точки зору охорони природи, так і епідеміології. Для розуміння основних чинників його поширення ми розробили за допомогою алгоритму Maxent у рамках програми 'flexsdm' модель поширення виду (SDM), використовуючи набір кліматичних, ґрунтових та рослинних змінних у межах його євразійського ареалу. Інтерпретацію моделі проводили за допомогою значень SHAP (SHapley Additive exPlanations). Модель досить точно описала відомий ареал виду у межах від Східної Європи до Центральної Азії. Аналіз SHAP показав, що клімат, а не ґрунт чи рослинна біомаса, є основним фактором, що визначає ареалу. Головними предикторами виявилися: кількість опадів найпосушливішого тижня, річний діапазон температур та мінімальна температура найхолоднішого тижня, що визначають пороги толерантності до аридності, континентальності та виживання під час сплячки, відповідно. Особливо варто відзначити перший, який у часі збігається з кінцем зими перед виходом тварин зі сплячки. Графік залежності SHAP показав виразний нелінійний оптимум, де придатність середовища за цим предиктором досягає піку приблизно при 6 мм. Це вікно, імовірно, є ключовим сигналом для проростання однорічних ефемерів, які формують критичний харчовий імпульс для тварин після сплячки. Цей зв'язок демонструється значною кореляцією (r = 0,68) між опадами найпосушливішого тижня року та вегетаційною активністю (NDVI) у квітні. Це поглиблює розуміння факторів, що формують поширення виду: крім суто фізіологічної толерантності до навколишніх умов до обов'язкової екологічної синхронізації. Отже, райони з високою передбачуваною придатністю, особливо в південному та східному Казахстані (зокрема, Жамбильська, Туркестанська та Алматинська обл.), окреслюють суцільний екологічний коридор, що представляє потенційні зони ензоотичних осередків чуми. Таким чином, створена SDM — це не лише прогностична карта, а інструмент для діагностики основних абіотичних обмежень та ключового трофічного «вузького місця», що визначає нішу виду. Вона забезпечує життєво важливу доказову базу як для цілеспрямованих стратегій охорони, так і для проактивного ризик-орієнтованого нагляду за громадським здоров'ям в ендемічних по чумі регіонах.

Ключові слова: *Allactaga major*, моделювання поширення видів (SDM), екологічна ніша, фенологічна синхронія, зоонозний ризик.

Introduction

The great jerboa (*Allactaga major*), a charismatic and ecologically significant rodent, inhabits the arid and semi-arid regions of Central Asia and Eastern Europe. The species is native to the steppes and northern deserts of eastern Ukraine and European Russia, through Kazakhstan and northern Uzbekistan to eastern Siberia and western Xinjiang. Its typical habitat is sparse grassland, sloping areas in ravines, road verges and field edges. It is also present in a range of arid and semi-arid habitats, particularly those with some succulent plant growth [Tsytsulina *et al.* 2016].

As one of the largest jerboa species, it is distinguished by its powerful hind limbs for saltatory locomotion, elongated ears, and a long tail ending in a distinctive black-and-white tuft. These morphological adaptations are quintessential for survival in open landscapes such as deserts, semi-deserts, and dry steppes, where resources are patchy and environmental pressures are extreme. Despite its remarkable adaptations, *A. major* faces growing threats from habitat fragmentation and climate change, leading to a classification of Near Threatened on the IUCN Red List and highlighting the urgent need for a deeper understanding of its distributional ecology.

A species' geographic distribution is fundamentally shaped by the interplay of its physiological tolerances and the prevailing environmental conditions [Baselga *et al.* 2012; Bozinovic & Naya 2015]. For species inhabiting extreme environments, climatic variables such as temperature, precipitation, and aridity are often the primary determinants of survival and reproduction, effectively acting as ultimate filters on their potential range [Kindlmann *et al.* 2025].

Species distribution models (SDMs) provide a powerful correlative framework to quantify these relationships by linking known species occurrence records with environmental data to predict the likelihood of suitable habitat across a landscape [Elith & Leathwick 2009; Peterson *et al.* 2011; Rathore & Sharma 2023]. When built upon climatic predictors, these models not only delineate a species' potential contemporary range but also offer critical insights into the fundamental niche—the set of conditions under which a species can persist in the absence of biotic interactions [Hutchinson 1957].

In the context of *A. major*, whose life history is so intimately tied to the challenges of aridity, a climate-based SDM is particularly well-suited. The species' burrowing behaviour, which provides refuge from diurnal heat and winter cold, its diet primarily composed of plant material and seeds, and its water-conserving physiology all suggest that specific climatic thresholds govern its presence. Key factors likely include the temperature regime during its active season, the depth and duration of winter snow cover (which affects both insulation and food accessibility), and the amount and seasonality of precipitation that drives primary productivity. By modelling these relationships, we can move beyond a simple map of occurrence points to a mechanistic understanding of the environmental factors that define the species' home range at a macroecological scale. By projecting these models onto future climate scenarios, we can forecast range shifts, contractions, or expansions, providing critical data for pre-emptive conservation strategies. This study aims to construct a robust SDM for A. major to accurately model its contemporary home range and elucidate the key climatic factors that are the primary determinants of its distribution.

Beyond its value for conservation, understanding the distribution of A. major is crucial due to its significant ecological role. In general, burrowing rodents in arid and semi-arid environments are considered ecosystem engineers due to their activities that modify soil structure and create 'fertile islands' [Mallen-Cooper *et al.* 2019]. These rodents are also highly responsive to climate change and human disturbance, making them valuable environmental indicators in ecosystems where they thrive [Dale & Beyeler 2001]. As a predominantly granivorous and herbivorous species, the greater jerboa influences plant community composition and seed dispersal. Its extensive burrowing activity aerates the soil, enhances water infiltration, and modifies the physical environment, creating microhabitats for other organisms. Furthermore, as an abundant prey item, it forms a critical trophic link for a variety of predators, including foxes, mustelids, and birds of prey (for instance, [Jánossy & Schmidt 1970; Nedyalkov *et al.* 2014]). This position within the food web underscores its importance in maintaining the structural and functional integrity of its ecosystem.

Perhaps of more immediate concern is the role of *A. major* in the epidemiology of zoonotic diseases, most notably plague, caused by the bacterium *Yersinia pestis*. Jerboas, alongside with other wild rodents, are known natural reservoirs for plague in its endemic foci [Rametov *et al.* 2023]. Their burrow systems provide an ideal environment for flea vectors, which act as the primary transmission agents. Epizootics, or outbreaks of the disease within animal populations, can lead to human cases when humans encroach upon these foci or come into contact with infected fleas or tissues. Consequently, the predicted distribution of *A. major* from an SDM can be directly overlain with known plague risk maps. Identifying areas of high climatic suitability for the jerboa can help pinpoint potential zones of sustained zoonotic transmission, thereby informing surveillance efforts and public health initiatives to mitigate the risk of spillover to human populations.

Therefore, the primary objectives of this study are twofold. First, we aim to develop a robust SDM for the jerboa species using global climate data to accurately predict its current potential distribution. Second, we seek to identify and rank the relative importance of key climatic variables in limiting the species' range, thereby elucidating the core environmental drivers of its distribution. The resulting model will serve as a vital tool for conservation planning and identifying potential habitats for future surveys. It also can contribute to a proactive framework for zoonotic disease risk assessment and establishing a baseline for assessing the future impacts of climate change on this unique and vulnerable species. This research underscores the value of SDMs not merely as cartographic exercises, but as essential instruments for unravelling the ecological narratives that shape the species' niche [Soberón & Nakamura 2009; Soberón & Arroyo-Peña 2017].

Materials and Methods

Distributional Data

We obtained presence records for *A. major* through personal systematic field surveys conducted in Ukraine, from existing literature data, and from public databases: the Global Biodiversity Information Facility (URL) and the Ukrainian Biodiversity Information Network (UkrBIN URL). Subsequently we reviewed these datasets by deleting unreliable or ambiguous and excluding those whose geographical location was not precisely defined. The initial dataset compilation contained 628 occurrence points, encompassing locations across Eurasia. In the next step duplicate occurrences were removed by implementing a spatial protocol performed in SAGA GIS using the 'Remove Duplicate Points' module [Conrad *et al.* 2015]. Further, in order to reduce autocorrelation, we adhered to the recommendation implying that data points should be at least 2–3 cells apart (URL). Following Nuñez and Medley [2011], we measured the spatial autocorrelation of occurrences by calculating Moran's I for multiple distance classes using the GeoDa software [Anselin *et al.* 2022]; values < 0.3 were considered acceptable for building meaningful SDMs [Lichstein *et al.* 2002]. Using the 'PointsThinning' module in SAGA GIS, the initial dataset was reduced to 583 record points. The longitude and latitude coordinates (WGS84 datum) of the sample records were stored in an Excel database and converted into CSV format for the establishment of the SDM models.

Modelling approach

The 'flexsdm' R (v. 3.3.3) modelling package [Velazco *et al.* 2022] was applied for projecting the potential geographic distribution of great jerboa across Eurasia. Spatial block partitioning was used to generate pseudo-absence and background points. Filtering the occurrence data was used to reduce sample bias by randomly removing points where they were dense (oversampling) in the environmental and geographical spaces. The 'flexsdm' package offers a wide range of modeling options. Here, we tested out Maximum Entropy (Maxent) [Phillips *et al.* 2006]. This is one of the most popular SDM modelling methods. Maxent can construct simple to highly complex, nonlinear species–environment relationships using various transformations of variables termed features and represented by a number of feature classes (FC) of which we tested linear, quadratic, product and hinge. To reduce overfitting, Maxent uses a regularization procedure to balance model fit with complexity, by penalising models based on the magnitude of their coefficients. Tuned models were built using regularization multiplier values ranging from 0 to 4 with increments of 0.5 and all possible FC combinations.

The model's predictive accuracy was measured using the widely recognised AUC statistic. AUC scores range from 0 to 1, with values closer to 1 reflecting strong discriminatory power in distinguishing habitat suitability [Fielding & Bell 1997] and the true skill statistic (TSS) where the value of >0.4 is considered good, with the range of 0.4–0.8 indicating 'good' performance; a score of >0.8 is considered excellent [Allouche *et al.* 2006; Zhang *et al.* 2015]. But whereas AUC remains a controversial criterion [Lobo 2008], for greater confidence we employed the continuous Boyce index, CBI [Boyce *et al.* 2002], one of the most reliable presence-only evaluation metrics, and also provided by the 'flexsdm' package. It is continuous and varies between −1 and +1. Positive values indicate a model that presents predictions that are consistent with the distribution of presences in the evaluation dataset, values close to zero mean that the model is not different from a random model [Hirzel *et al.* 2006].

Maps of habitat suitability in the GeoTIFF format were processed and visualised in SAGA GIS. Statistical data was analysed using the PAST software package [Hammer *et al.* 2001] and the R environment (https://www.r-project.org).

Environmental variables

Environmental variables focusing primarily on aspects of climate and soil were prepared to summarise important potential drivers of the jerboa's ecological niche. Climate is one of the most influential environmental factors shaping species distributions. This is particularly true in dry regions, where precipitation and temperature directly determine vegetation growth and water availability, and consequently exert indirect effects on food resources and habitat conditions for rodents [Elith & Leathwick 2009]. Derivatives of temperature, precipitation, solar radiation and soil moisture index were obtained from the CliMond archive [Kriticos *et al.* 2012]; http://www.climond.org. The CliMond dataset contains gridded historical climate data used at a 10′ resolution.

Physical and chemical properties of soil (edaphic variation) can play an important role in predicting plant species occurrences [Liu *et al.* 2019; Shabani *et al.* 2025] and thus may indirectly influence jerboa distribution. In addition, it is recommended to incorporate abiotic factors (for example, soil covariates) into predictive frameworks of species responses to build a more accurate picture of its potential home range [Houlahan *et al.* 2016]. Therefore, we incorporated soil explanatory variables in our model from the SoilGrids database (version 2.0; https://soilgrids.org); [Poggio *et al.* 2021]. Initially these were downloaded at a 250 m resolution but resampled for this analysis to match the climatic input layers.

Our initial rationale identifies climate and soil as primary indirect drivers of the jerboa's distribution, although vegetation largely acts as the crucial proximate factor that translates those abiotic conditions into actual suitable or non-suitable habitat. On that account we considered to include such vegetation-related covariates as net primary productivity (NPP) (downloaded from the Socioeconomic Data and Applications Center (SEDAC), http://sedac.ciesin.columbia.edu/es/hanpp.html) and greenness based on the normalised difference vegetation index (NDVI) (downloaded from the EDIT Geoplatform, [Lobo 2007]; URL). The first one directly measures the base of the food web, correlates with overall ecosystem productivity and thus potential food and resource abundance available for the animals [Rodal *et al.* 2025]. The second one is a direct measure of vegetation vigour and cover, has proven extremely useful in predicting herbivore and non-herbivore distribution; previously thought to be most useful in temperate environments, the utility of this satellite-derived index has been demonstrated even in sparsely vegetated areas [Pettorelli *et al.* 2011].

Above these, in order to link the climate data to the ecological reality experienced by the jerboa and its food sources, we decided to employ growing degree days (GDD), one of the key metrics for understanding temperature-plant interactions and fundamental tool that quantifies thermal energy accumulation for predicting the timing of various plant development stages [Baskerville & Emin 1969; Hatfield & Prueger 2015]. GDD is considered crucial for defining the growing season in temperate and arid systems, determines plant growth periods and thus the timing of food availability [Boutin & Lane 2014]. Gridded GDD was downloaded from the Center for Sustainability and the Global Environment site, https://sage.nelson.wisc.edu/).

To ensure consistency across datasets, all considered environmental variables were standardised to a 10′ resolution.

Conditioning factors

Commonly used approaches recommend removing correlated predictor variables before modeling to avoid multicollinearity, which is considered to affect model projections [Zhao *et al.* 2022]. There are several statistical packages offering functions that reduce collinearity in predictors.

In one case we carried out a stepwise selection of environmental variables by calculating the variance inflation factor (VIF) using the library usdm in R [Naimi *et al.* 2014]. VIF values higher than 10 are considered to lead to problematic levels of multi-collinearity [Craney & Surles 2002], therefore factors with VIF values that exceeded 10 were not used in the model.

In the second case we ignored removing correlated predictor variables because the benefits of using all available variables may outweigh the drawbacks of collinearity. Latest research indicates that modeling with correlated climate variables increases accuracy of predictions [Hanberry 2023]. Moreover, complex models such as Maxent take advantage of existing collinearity in finding the best set of parameters [De Marco & Nóbrega 2018].

To further explore the impact of the considered above environmental factors, we employed a SHAP framework from XAI (i.e. eXplainable artificial intelligence) to rank and uncover the most influential drivers [Lundberg *et al.* 2018; Farooq *et al.* 2022]. With a SHAP approach there is no need to consider only uncorrelated environmental drivers [Nikraftar *et al.* 2025]. SHAP (SHapley Additive exPlanations) is a unified framework in explainable artificial intelligence used to interpret the output of any machine learning model by assigning each feature an importance value for a particular prediction. We post-processed the best model results with SHAP by comparing what a model predicts with and without the predictor for all possible combinations of predictors at every single observation. The predictors are then ranked according to their contribution for each observation and averaged across observations. Another useful item are dependence plots. In our case, the R package 'shap-values' (https://github.com/pablo14/; author Pablo Casas) in a modified version was used to perform the SHAP analysis. The application of SHAP for understanding the influence of environmental factors on species distribution has, until today, seen limited exploration, but is now being investigated more widely [e.g. Scavuzzo *et al.* 2022; Song & Estes 2023; Li *et al.* 2025; Buebos-Esteve & Dagamac 2025].

Results and Discussion

Eventually, we considered 46 environmental variables: 35 climate-related, 8 soil covariates, and 3 related to plant biomass. Using this full dataset, we projected the potential geographic distribution of *A. major* across Eurasia (FC hinge, regularisation multiplier 2.0), with means ± standard deviations of AUC and TSS reaching 0.73 ± 0.05 and 0.44 ± 0.06, respectively, and a continuous Boyce index of 0.85±0.19. Also, a selection of environmental variables was carried out by calculating the variance inflation factor (VIF) and factors with VIF values that exceeded 10 were not used in the model. The remaining factors included 8 climate-related, 7 soil covariates, and 2 related to plant biomass.

Repeatedly we projected the potential geographic distribution of *A. major* across Eurasia (FC linear, quadratic, hinge, product, regularisation multiplier 3.0); however, the performance of this particular model was poorer: AUC and TSS reaching only 0.63 ± 0.04 and 0.31 ± 0.03, respectively, and a continuous Boyce index reduced to 0.68 ± 0.32, therefore only the first model accounting for the full set of predictors was taken into further consideration. In general, this model adequately captures the distribution of *A. major* from Ukraine and Kazakhstan through to parts of Xinjiang and Mongolia (Fig. 1).

In order to delineate species suitable and unsuitable habitats, the maximum training sensitivity plus specificity (MaxSS) threshold was used [Bharti *et al.* 2021]. Compared with other thresholds, MaxSS is a promising method for threshold selection when only presence data are available [Liu *et al.* 2013]. Under this threshold the resulting omission rate (OR) will vary but is usually in the 10–20% range. Indeed, the OR of 16.60 ± 0.02% (mean ± SE) can be seen as a sign of good performance [Liu *et al.* 2005], indicating that the used dataset and modelling algorithm effectively simulate *A. major* distribution.

As mentioned above, *A. major* is a concern in the epidemiology of plague, particularly across the arid steppe and desert regions of Central Asia. Within its home range, it serves as a reservoir host, maintaining the bacterium in natural plague foci, where its burrows and flea populations facilitate transmission cycles that pose a risk to humans and other animals. Ukraine, however, is not within the established, natural plague foci associated with the primary reservoir role of *A. major*. Specific rodent communities (primarily ground squirrels *Spermophilus* spp. and voles like *Microtus*) are responsible for maintaining *Yersinia pestis* in Ukraine's steppe regions [Kosminsky & Marinin 1975].

The core distribution of this species' involvement in plague ecology is centred further east in the steppes and deserts of Central Asia (e.g. Kazakhstan). Authors like N. I. Kalabukhov or V. V. Kucheruk explicitly delineate the Volga–Ural, Caspian, and Central Asian foci (with jerboas among the key hosts) from the separate, smaller Black Sea and Azov steppe foci of Ukraine [Kalabukhov 1960; Kucheruk 1965].

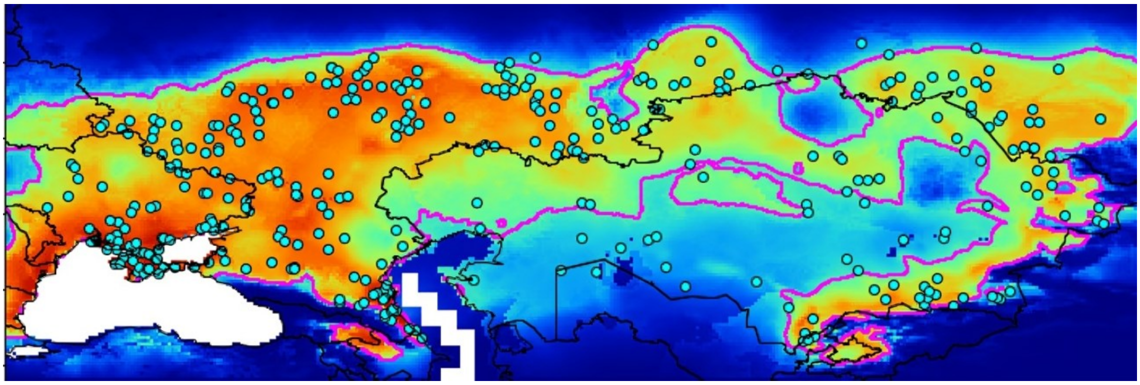

Fig. 1. The model output used to map habitat suitability for *Allactaga major* throughout Eurasia; colours show potential habitat suitability ranging from high (red) to low (blue); blue circles indicate jerboa occurrences (n = 345) filtered by the 'flexsdm' algorithm; the fuchsia-coloured line depicts the maximum training sensitivity plus specificity (MaxSS) threshold (0.44); 1—Ukraine, 2—Russian Federation, 3—Republic of Kazakhstan.

Рис. 1. Результати моделювання на основі яких складена карта придатності середовища для *Allactaga major* на території Євразії; кольори показують потенційну придатність середовища від високої (червоний) до низької (синій); сині кружечки позначають місця знахідок тварин (n=345), відфільтровані алгоритмом 'flexsdm'; фуксієва лінія відображає порогове значення максимальної чутливості навчання плюс специфічність (MaxSS) (0.44); 1 — Україна, 2 — Російська Федерація, 3 — Республіка Казахстан.

In this respect, one of the most extensive natural plague centres, or foci, is located in Central Asia, particularly in Zhambyl Oblast in southern Kazakhstan [Rametov *et al.* 2023]. The cited authors employed a similar methodological approach, using species distribution modelling with the Maxent algorithm, and identified the natural plague foci located in this particular region, notifying from their professional view the high potential risk here of plague outbreak for rural towns and villages. According to our model, in areas of suitable habitat within Zhambyl Oblast the estimated suitability for *A. major* averaged 0.59, maximum 0.77. In continuation of this, our model also points out to other affected regions in Kazakhstan primarily located in the southern, south-eastern, and eastern provinces at times mentioned in previous research reports [Aikimbayev *et al.* 2006; Koshkin *et al.* 2007; Atshabar *et al.* 2015; Abdel *et al.* 2023]. Such key endemic zones include: Turkistan Oblast, formerly South Kazakhstan Oblast (averaged suitability 0.60, maximum 0.80), Almaty Oblast (averaged suitability 0.58, maximum 0.74), Zhetisu Oblast, part of the historical Almaty region (averaged suitability 0.56, maximum 0.74), and Abai Oblast, part of the historical East Kazakhstan region (averaged suitability 0.54, maximum 0.80).

These results reveal a rather consistent pattern, with average suitability values across the identified endemic zones varying by less than 0.06, from 0.54 to 0.60. In epidemiological terms, this homogeneity suggests a continuous, high-risk ecological corridor where plague can be maintained and transmitted enzootically, with minimal geographical barriers to the spread of the bacterium among interconnected jerboa and other rodent populations. The consistency in habitat suitability implies stable and widespread reservoir host availability, which supports persistent plague foci and increases the potential for epizootic events that could spill over to other mammals, including humans. For monitoring and public health strategy, this means that surveillance cannot be geographically limited to a few hotspots; a region-wide, coordinated program across southern and eastern Kazakhstan is essential, as the risk most likely is broadly distributed.

In conclusion, we can state that our model identifies potential plague reservoir zones, as high habitat suitability for *A. major* defines the ecological stage where the plague cycle can be maintained. The final step of human spillover, however, depends on additional actors—such as bridge vectors and hosts [Twigg 1978]—and human behaviour on that stage.

Species distribution modeling serves not only a predictive function but also a diagnostic one, revealing the key environmental relationships that define a species' niche. We employ SHAP (SHapley Additive exPlanations) values to interpret our final models, ranking covariates by their mean absolute SHAP value as a measure of global importance. At the Eurasian scale, the five most influential predictors from the dataset used for model construction were (in ranked order, Fig. 2): Bio14—Precipitation of driest week (mm) [35.4], Bio07—Temperature annual range (°C) [29.7], Bio06—Minimum temperature of coldest week (°C) [19.6], Bio15—Precipitation seasonality (C of V) [7.9], and Bio12—Annual precipitation (mm) [7.4]. [In square brackets: mean absolute SHAP value × 1000].

As previously noted, our set of predictors consisted of environmental variables related to climate, soil and on the whole to plant biomass; however, climate variables (particularly extremes like Bio14 and Bio07) overwhelmingly dominate the model's top predictors, while soil and biomass variables ranked lower. This most obviously means that climate acts as a first-order, macro-scale filter. For a wide-ranging species like *A. major* climate sets across the vast Eurasian continent the absolute, non-negotiable boundaries of its fundamental niche and top variables exactly reflect this fact. In brief, Bio14 (Driest Week Precipitation) defines the absolute aridity tolerance—the physiological and food-resource bottleneck for survival; Bio07 (Temperature Annual Range) captures continentality where a high range is characteristic of the interior deserts and steppes it inhabits, excluding maritime climates with buffered temperatures; and Bio06 (Minimum Temperature of Coldest Week) defines the cold tolerance limit for hibernation survival. These are 'range-defining' constraints and if these macroclimatic conditions remain unmet, the species cannot persist, regardless of local soil or vegetation conditions.

Focusing deeper into the highest-ranked constraint, Bio14 surprisingly reveals a finely tuned ecological mechanism rather than a simple aridity filter. It should be taken into account that across the jerboa's home range, the driest week of the year typically occurs in February–March just prior to the end of hibernation (April–May), with mean precipitation averaging approximately 5.27 mm during this week period. The model's interpretation via SHAP values shows that the biological response to this climate gradient is not a simple, linear increase (Fig. 3).

Instead, the SHAP dependence plot reveals a distinct non-linear response. SHAP values are predominantly negative when Bio14 falls below a critical threshold of ~2–3 mm (indicating insufficient moisture to initiate the food pulse), rise sharply to peak positive values within the optimal of ~6 mm range, and subsequently plateauing or declining as precipitation exceeds ~6 mm (maybe reflecting detrimental effects of excessive moisture).

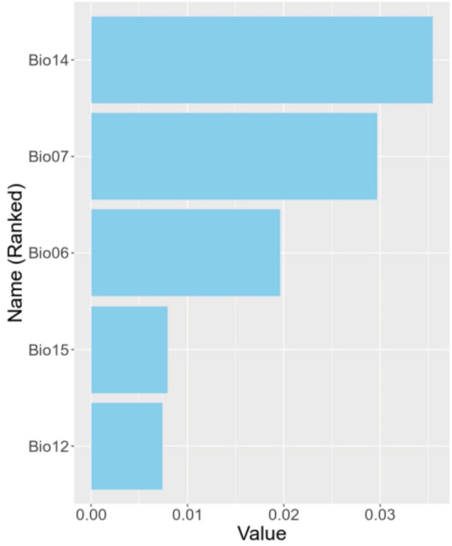

Fig. 2. Absolute summary plot of the complete dataset, where the average absolute value of the SHAP values for each variable is taken in order to obtain a bar chart as a function of the contribution of each variable to the prediction of the model. Top 5 variables are ordered from most (top) to least (bottom) important.
The x-axis represents the SHAP value×1000. The y-axis represents the variables used in the study, which refer to: Bio14—Precipitation of driest week (mm), Bio07—Temperature annual range (°C), Bio06—Minimum temperature of coldest week (°C), Bio15—Precipitation seasonality (C of V), and Bio12—Annual precipitation (mm).

Рис. 2. Зведена діаграма повного набору даних, де береться середнє абсолютне значення SHAP-значень для кожної змінної, щоб отримати стовпчикову діаграму як функцію внеску кожної змінної в прогнозування моделі. Топ-5 змінних упорядковані від найбільш (зверху) до найменш (знизу) важливих.
Вісь X представляє SHAP-значення ×1000. Вісь Y представляє змінні, використані в дослідженні, які відносяться до: Bio14 — опади найпосушливішого тижня (мм), Bio07 — річний діапазон температур (°C), Bio06 — мінімальна температура найхолоднішого тижня (°C), Bio15 — сезонність опадів (коефіцієнт варіації), io12 — річна кількість опадів (мм).

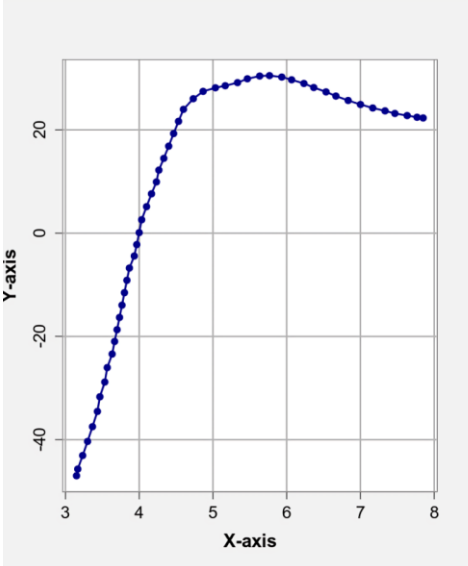

Fig. 3. SHAP Dependence Plot. X-axis: Bio14—Precipitation of Driest Week (mm); Y-axis: SHAP value (impact on habitat suitability).

Рис. 3. Графік залежності SHAP. Вісь X: Bio14 — опади найпосушливішого тижня (мм); Вісь Y: значення SHAP (вплив на показники придатності середовища існування).

The ~6 mm of precipitation in late winter is possibly the essential cue and resource for the germination and initial growth of the annual plants and ephemerals that form the bulk of the jerboa's post-hibernation diet. This mechanistic link is strongly supported by our finding of a significant positive correlation ($r = 0.68$, $p < 0.05$) between precipitation in this driest period and the satellite-derived greenness (NDVI) in April. This small but non-zero moisture window likely initiates plant growth so that fresh, nutrient-rich shoots and seeds are available precisely when jerboas emerge, emaciated and in urgent need of energy for reproduction.

Our finding that habitat suitability for the jerboa depends on precipitation during the late-winter driest week aligns with the established 'pulse-reserve' paradigm in desert ecology [Noy-Meir 1973]. This precipitation pulse subsequently triggers the germination of annual plants, providing the essential food resources required for post-hibernation survival and reproduction, a mechanism documented in other arid-adapted rodent assemblages [e.g. Ernest *et al.* 2000]. In another striking example J.H. Brown and coauthors have demonstrated that population sizes of granivorous rodents are directly driven by winter precipitation, which determines seed production of annual plants; low winter rain leads to a 'seed crash' and population declines [Brown *et al.* 1979]. The variable therefore acts as a climatic proxy for the trophic resource base critical during a key phenological bottleneck.

In conclusion, the dominance of the variable 'Precipitation of the Driest Week' reveals that the jerboa's range is not defined merely by where it can survive, but by where the climate reliably triggers its essential food supply at the most critical annual bottleneck. This shifts the conceptual understanding from one of simple physiological tolerance to a more nuanced ecological synchrony constraint. Consequently, our SDM gains greater power and ecological relevance, as it captures this vital trophic link central to the species' survival strategy.

Conclusions

This study successfully developed a robust Species Distribution Model (SDM) for the great jerboa (*Allactaga major*) across Eurasia, achieving its dual objectives of mapping potential habitat and elucidating the key climatic drivers of its range. The model, built using a comprehensive set of environmental predictors, reliably identified the species' core distribution from the dry steppes of Ukraine to the deserts of Central Asia and Mongolia. The primary outcomes and implications can be summarised as follows:

1. Climate as the Primary Range Determinant: The analysis unequivocally identified climate, rather than soil or vegetation biomass, as the principal first-order filter shaping the continental-scale

distribution of *A. major*. The top predictors—Precipitation of the Driest Week (Bio14), Temperature Annual Range (Bio07), and Minimum Temperature of the Coldest Week (Bio06)—define the non-negotiable macroclimatic thresholds of the species' fundamental niche. These variables represent critical constraints related to aridity tolerance, continentality, and hibernation survival, respectively.

2. Ecological Synchrony Over Simple Tolerance: A key finding was the specific role of Bio14 (Precipitation of the Driest Week). Its importance, coupled with its timing in late winter (February–March), reveals a mechanism more nuanced than simple aridity tolerance. The SHAP dependence plot demonstrated a distinct optimal range (~6 mm), indicating that this modest precipitation pulse acts as a 'critical ecological trigger'. It synchronises the germination of annual plants with the jerboas' post-hibernation emergence, ensuring the availability of essential food resources during a severe energetic bottleneck. This shifts the understanding of the species' distribution from a model of physiological survival to one of obligate ecological synchrony.

3. Implications for Plague Epidemiology: The model's output provides a spatially explicit map of high-probability reservoir habitat for *A. major*. The consistent, high habitat suitability predicted across southern and eastern Kazakhstan delineates a continuous ecological corridor conducive to the maintenance of enzootic plague cycles. This offers a science-based tool for prioritizing surveillance and monitoring efforts in known endemic zones, such as Zhambyl, Turkistan, and Almaty oblasts. While high jerboa suitability defines the stage for plague persistence, the final risk of human spillover remains contingent on additional local factors, including bridge vector dynamics and human activities.

4. Methodological Advance: The application of SHAP (SHapley Additive exPlanations) values proved invaluable for moving beyond prediction to diagnostic ecological insight. This explainable AI (XAI) technique allowed for the clear ranking of variable importance and, crucially, the interpretation of complex, non-linear species-environment relationships, such as the identified optimal precipitation window for Bio14.

In summary, this research demonstrates that a climate-based SDM, interpreted through advanced explanatory frameworks, can transcend cartographic exercise to uncover the core ecological narratives limiting a species' distribution. For the greater jerboa, this narrative is centred on a precise climatic cue that gates annual survival and reproduction. The resulting models serve as vital foundations for both targeted conservation strategies for this Near Threatened species and proactive, risk-based public health planning in regions where it functions as a key plague reservoir. Future work should focus on integrating these suitability maps with data on flea vector distributions and human land use to build comprehensive, dynamic risk assessment systems.

Declarations

Funding. The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of interests. The authors have no conflicts of interest to declare that are relevant to the content of this article.

Handling of materials. All manipulations with living animals were carried out using non-lethal techniques and in compliance with the current legislation of Ukraine.

References

Abdel, Z., B. Abdeliyev, D. Yessimseit, [et al.]. 2023. Natural foci of plague in Kazakhstan in the space-time continuum. *Comparative Immunology, Microbiology and Infectious Diseases*, **100**: 102025. CrossRef

Aikimbayev, A. M., B. B. Atshabar, S. A. Aubakirov, [et al.]. 2006. *Epidemiologic Potential of Natural Plague Foci in Kazakhstan*. Almaty, 1–153. [Russian]

Allouche, O., A. Tsoar, R. Kadmon. 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *The Journal of Applied Ecology*, **43**: 6, 1223–1232. CrossRef

Anselin, L., X. Li, J. Koschinsky. 2022. GeoDa: from the desktop to an ecosystem for exploring spatial data. *Geographical Analysis*, **54**: 439–466. CrossRef

Atshabar, B. B., L. A. Burdelov, U. A. Izbanova, [et al.]. 2015. Passport of regions of Kazakhstan on especially dangerous infections. *Quarantine and zoonotic infections in Kazakhstan*, **1.31**: 5–177. [Russian]

Baselga, A., J. Lobo, J.-C. Svenning, [et al.]. 2012. Global patterns in the shape of species geographic ranges reveal range determinants. *Journal of Biogeography*, **39**: 760–771. CrossRef

Baskerville, G. L. P. Emin. 1969. Rapid estimation of heat accumulation from maximum and minimum temperatures. *Ecology*, **50**: 514–521. CrossRef

Bharti, D. K., G. D. Edgecombe, K. P. Karanth, [et al.]. 2021. Spatial patterns of phylogenetic diversity and endemism in the Western Ghats, India: A case study using ancient predatory arthropods. *Ecology & Evolution*, **11** (23): 16499–16513. CrossRef

Boutin, S., J. E. Lane. 2014. Climate change and mammals: evolutionary versus plastic responses. *Evolutionary Applications*, **7**: 29-41. CrossRef

Boyce, M. S., P. R. Vernier, S. E. Nielsen, [et al.]. 2002. Evaluating resource selection functions. *Ecological Modelling*, **157** (2–3): 281–300. CrossRef

Bozinovic, F., D. Naya. 2015. Linking physiology, climate, and species distributional ranges. *Integrative Organismal Biology*, 277–290. CrossRef

Brown, J. H., O. J. Reichman, D. Davidson. 1979. Granivory in desert ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, **10**: 201–227. CrossRef

Buebos-Esteve, D. E., N. H. A. Dagamac. 2025. Evaluating model-agnostic post-hoc methods in explainable artificial intelligence: augmenting species distribution models. *Biologia futura,* **76** (4): 585–595. CrossRef

Conrad, O., B. Bechtel, M. Bock, [et al.]. 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, **8**: 1991–2007. CrossRef

Craney, T. A., J. G. Surles. 2002. Model-dependent variance Inflation factor cutoff values. *Quality Engineering*, **14** (3): 391–403. CrossRef

Dale, V. H., S. C. Beyeler. 2001. Challenges in the development and use of ecological indicators. *Ecological Indicators*, **1:** 3-10. CrossRef

De Marco, P. Júnior, C. C. Nóbrega. 2018. Evaluating collinearity effects on species distribution models: An approach based on virtual species simulation. *PLoS ONE*, **13** (9): e0202403. CrossRef

Elith, J., J.R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**: 677–697. CrossRef

Ernest, S. M., J. H. Brown, R. R. Parmenter. 2000. Rodents, plants, and precipitation: spatial and temporal dynamics of consumers and resources. *Oikos*, **88** (3): 470–482. CrossRef

Farooq, Z., J. Rocklöv, J. Wallin, [et al.]. 2022. Artificial intelligence to predict West Nile virus outbreaks with ecoclimatic drivers. *The Lancet Regional Health — Europe*, **17**: e100370. CrossRef

Fielding, A. H., J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**: 38–49. CrossRef

Hammer, Ø., D. A. T. Harper, P. D. Ryan. 2001. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, **4** (1): art. 4 (1–9). URL

Hatfield, J. L., J. H. Prueger. 2015. Temperature extremes: effect on plant growth and development. *Weather and Climate Extremes*, **10:** 4–10. CrossRef

Hirzel, A. H., G. Le Lay, V. Helfer, [et al.]. 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199** (2): 142–152. CrossRef

Houlahan, J., S. McKinney, T. Anderson, [et al.]. 2016. The priority of prediction in ecological understanding. *Oikos*, **126** (1): 1–7. CrossRef

Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22:** 415–427. CrossRef

Jánossy, D., E. Schmidt. 1970. Die Nahrung des Uhus (*Bubo bubo*). Regionale und erdzeitliche Änderungen. *Bonner zoologische Beiträge*, **21**: 25–31. URL

Kalabukhov, N. I. 1960. Ecology of Rodents and Natural Focality of Plague. *In: Natural Focality and Epidemiology of Particularly Dangerous Infectious Diseases.* Saratov, 85–95. [Russian]

Kindlmann, P., S. Tsiftsis, L. Buchbauerová, [et al.]. 2025. How will environmental conditions affect species distribution and survival in the coming decades—A Review. *Diversity*, **17:** 11, 793. CrossRef

Koshkin, A.V., L.Y. Nekrasova, V. M. Dubyanskiy. 2007. The great jerboa (Allactaga major Kerr, 1792) in the desert foci of plague in Kazakhstan]. *In: Problems of Particularly Dangerous Infections (Saratov)*, **93**: 44–47. [Russian]

Kosminsky, R. B., V. F. Marinin. 1975. Features of the natural plague focus in the North-West Priazov region]. *In: Questions of the Natural Focality of Diseases (Saratov)*, **7**: 182–189. [Russian]

Kucheruk, V. V. 1965. Rodents as inhabitants of open landscapes of the USSR and Their Epidemiological Significance. *In*: *Fauna and Ecology of Rodents (Moscow)*, **9**: 5–88. [Russian]

Li, Q., R. Hou, X. Zhang, [et al.]. 2025. Integrating explainable machine learning to predict the ecological niche distribution of Cytospora chrysosperma in Xinjiang, China. *Forest Ecology and Management*, **595**: 123031. CrossRef

Lichstein, J. W., T. R. Simons, S. A. Shriner, [et al.]. 2002. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, **72**: 445–463. CrossRef

Liu X.-T., Q. Yuan, J. Ni. 2019. Research advances in modelling plant species distribution in China. *Chinese Journal of Plant Ecology*, **43** (4): 273–283. URL

Liu, C., M. White, G. Newell. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, **40** (4): 778–789. CrossRef

Lobo, J. M. 2007. EDIT Geoplatform. URL

Lobo, J. M. 2008. More complex distribution models or more representative data? *Biodiversity Informatics*, **5**: 15–19. CrossRef

Lundberg, S. M., B. Nair, M. S. Vavilala [et al.]. 2018. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering*, **2** (10): 749e760. CrossRef

Mallen-Cooper, M., S. Nakagawa, D. J. Eldridge. 2019. Global meta-analysis of soil-disturbing vertebrates reveals strong effects on ecosystem patterns and processes. *Global Ecology and Biogeography*, **28**: 661–679. CrossRef

Naimi, B., N. A. S. Hamm, T. A. Groen, [et al.]. 2014. Where is positional uncertainty a problem for species distribution modelling. *Ecography*, **37** (2): 191–203. CrossRef

Nedyalkov, N., A. Levin, A. Dixon, [et al.]. 2014. Diet of saker falcon (Falco cherrug) and Eastern imperial eagle (Aquila heliaca) from Central Kazakhstan. *Ecologia Balkanica*, **6** (1): 25–30.

Nikraftar, Z., E. Parizi, M. Saber, [et al.]. 2025. An interpretable machine learning framework for unraveling the dynamics of surface soil moisture drivers. *Remote Sensing*, **17** (14): 2505. CrossRef

Noy-Meir, I. 1973. Desert ecosystems: environment and producers. *Annual Review of Ecology and Systematics*, **4**: 25–51. CrossRef

Nuñez, M.A., K.A. Medley. 2011. Pine invasions: climate predicts invasion success; something else predicts failure. *Diversity and Distributions*, **17**: 703–713. CrossRef

Peterson, A. T., J. Soberón, R. G. Pearson, [et al.]. 2011. *Ecological Niches and Geographic Distributions.* Princeton University Press, 1–328. CrossRef

Pettorelli, N, S. Ryan, T. Mueller, [et al.]. 2011. The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. *Climate Research*, **46**: 15–27. 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

Poggio, L., L. M. de Sousa, N. H. Batjes [et al.]. 2021. SoilGrids 2.0: Producing Soil Information for the Globe with Quantified Spatial Uncertainty. *The Soil*, **7**: 217–240. CrossRef

Rametov, N. M., M. Steiner, N. A. Bizhanova, [et al.]. 2023. Mapping plague risk using Super Species Distribution Models and forecasts for rodents in the Zhambyl region, Kazakhstan. *GeoHealth*, **7**: e2023GH000853. CrossRef

Rathore, M.K., L.K. Sharma. 2023. Efficacy of species distribution models (SDMs) for ecological realms to ascertain biological conservation and practices. *Biodiversity and Conservation*, **32**: 5393053–3087. CrossRef

Rodal, M., S. Luyssaert, M. Balzarolo [et al.]. 2025. A global database of net primary production of terrestrial ecosystems. *Scientific Data*, **12**: 1534. CrossRef

Scavuzzo, C. M., J. M. Scavuzzo, M. N. Campero, [et al.]. 2022. Feature importance: Opening a soil-transmitted helminth machine learning model via SHAP. *Infectious Disease Modelling*, **7** (1): 262–276. CrossRef

Shabani, F., M. Ahmadi, N. Lorestani, [et al.]. 2025. Pedology and plant provenance can improve species distribution predictions of Australian native flora: A calibrated and validated modeling exercise on 5033 species. *Ecology and Evolution*, **24** (15, 6): e71430. CrossRef

Soberón, J, M. Nakamura. 2009. Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, **17** (106, Suppl. 2): 19644–50. CrossRef

Soberón, J., B. Arroyo-Peña. 2017. Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. *PLOS ONE*, **12**: CrossRef

Song, L., L. Estes. 2023. itsdm: Isolation forest-based presence-only species distribution modelling and explanation in r. *Methods in Ecology and Evolution*, **14**: 831–840. CrossRef

Tsytsulina, K., N. Formozov, I. Zagorodnyuk [et al.]. 2016. *Allactaga major* (errata version published in 2017). *The IUCN Red List of Threatened Species 2016*: e.T857A115052919. CrossRef

Twigg, G. I. 1978. The role of rodents in plague dissemination: a worldwide review. *Mammal Review*, **8**: 77–110. CrossRef

Velazco, S. J. E., M. B. Rose, A. F. A. Andrade, [et al.]. 2022. flexsdm: An R package for supporting a comprehensive and flexible species distribution modelling workflow. *Methods in Ecology and Evolution*, **13** (8): 1661–1669. CrossRef

Zhang, L., S. Liu, P. Sun, [et al.]. 2015. Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PLoS One*, **10** (3): e0120056. CrossRef

Zhao, Z., N. Xiao, M. Shen, [et al.]. 2022. Comparison between optimized MaxEnt and random forest modeling in predicting potential distribution: A case study with *Quasipaa boulengeri* in China. *Science of the Total Environment*, **842**: 156867. CrossRef