Systematic Review

Systematic Review of Variable Selection Bias in Species Distribution Models for *Aedes vexans* (Diptera: Culicidae)

Peter Pothmann [1,2,3,*] , Helge Kampen [4], Doreen Werner [1] and Hans-Hermann Thulke [2]

1. Leibniz Centre for Agricultural Landscape Research, Biodiversity of Aquatic and Semiaquatic Landscape Features, 15374 Müncheberg, Germany
2. Department of Ecological Modelling, Helmholtz-Centre for Environmental Research GmbH, 04318 Leipzig, Germany
3. Department of Forest Science, Technische Universität Dresden, 01069 Tharandt, Germany
4. Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Institute of Infectology, 17493 Greifswald, Germany
* Correspondence: peter.pothmann@zalf.de

Simple Summary

The mosquito *Aedes vexans* is common in Europe and often appears in large numbers after floods. Because it can transmit diseases, scientists use computer models to predict where it might occur. These models rely on environmental factors such as temperature, rainfall, and land use to describe the conditions that make habitats suitable for the species. We reviewed all published studies that used such models for *Aedes vexans* to check whether they include the environmental factors that really matter for this species. We found 28 studies and analysed nearly 500 variables used to describe weather, land use, water, and human activity. Surprisingly, most models did not include information about flooding or temporary water bodies, even though these are essential for the mosquito's reproduction. Instead, many studies focused on urban features, a choice that may reflect where people collect samples rather than where the mosquito actually lives. Our findings highlight that understanding how and where data are collected is as important as the modelling technique itself. Taking flood dynamics into account could make predictions more realistic and help to make management of potential disease risks more effective.

Abstract

We conducted a systematic literature review, following PRISMA guidelines, to assess whether existing species distribution models for *Aedes vexans* reflect its known ecological requirements. This mosquito is closely associated with temporary floodwaters, making hydrological dynamics a critical factor for accurate modelling. From 28 peer-reviewed studies, we extracted 477 environmental and ecological variables and organized them into a hierarchical classification scheme with four main categories: weather, land characteristics, water characteristics, and population. We analysed patterns in variable usage and the reported importance of each variable. Our results show that flood-related variables were largely absent, despite the species' reliance on ephemeral water bodies for reproduction. This may possibly reduce the predictive utility of existing *Aedes vexans* species distribution models. In contrast, urban-landscape variables were frequently used and often ranked as highly predictive, but such results were primarily found in studies that did not account for sampling bias.

Keywords: ecological variable classification; variable selection; model comparison; covariates

1. Introduction

Understanding species–environment relationships is central to ecological research, offering key insights into biodiversity patterns and environmental change [1,2]. Species distribution models, which predict geographic distributions, are widely used tools in this context [3]. They support diverse applications, including conservation planning for endangered species [4,5], forecasting the spread of vector-borne diseases [6] and agricultural pests [7], projecting outbreak dynamics [8,9], and identifying areas of elevated risk [10,11].

Species distribution models are based on two main types of input data: species occurrences and environmental and ecological variables. Occurrence data reflect locations where the species is present or absent and are often derived from field surveys, museum records, or citizen science initiatives. The environmental and ecological variables (hereafter referred to as variables) describe the habitat conditions at occurrence locations. Environmental factors include climate, soil, elevation, hydrology, and land cover, while ecological factors encompass species interactions such as competition, predation, mutualism, and host availability, as well as the distribution of resources like breeding or resting sites. By relating species occurrences to environmental and ecological conditions, species distribution models define environmental and ecological conditions necessary for species persistence and project potential suitable habitat across the landscape, including across unsampled areas [3].

While occurrence data are typically obtained from large databases or the literature [12,13] and methods for the identification of the best-performing model types have been extensively studied [14], selecting an appropriate set of environmental explanatory variables remains a major challenge.

Species distribution models rely on variables to characterize the ecological niche of the species [3,15–18]. The selection of variables typically begins with an initial candidate pool, with the choice of included variables informed by expert knowledge and previous studies. This pool is then often narrowed down using data-driven approaches that assess the individual contribution of each variable to overall model accuracy [19,20]. The success of this refinement depends critically on the ecological completeness of the candidate pool. If essential environmental and ecological constraints are missing, the models may suffer from bias, overlooking key habitat factors that limit species distributions [3,15–18].

We conducted a systematic review of published species distribution models for the floodplain mosquito *Aedes vexans* with two aims: first, to evaluate whether all potentially limiting habitat characteristics are adequately represented in the selection process, thereby revealing possible biases caused by missing or insufficient variables; and second, to identify the variables most frequently used and considered important across existing models. Since these variables have already proven important in diverse modelling contexts, they are also likely to perform well in future applications.

By organizing variables according to the ecological processes or conditions they represent and their reported importance, we provide a structured overview that highlights both the inclusion of key habitat characteristics and potential gaps in current modelling practices.

Understanding the biology and ecology of *Aedes vexans* is crucial for interpreting these findings. As a floodplain mosquito species, *Aedes vexans* exhibits habitat preferences and life-cycle traits that are tightly linked to dynamic water regimes.

Aedes vexans relies on temporary floodwaters for reproduction, which explains the importance of water-level fluctuations in shaping its habitat suitability [21–23]. Females preferentially oviposit near ephemeral ponds formed by river and lake flooding, where drought-resistant eggs remain viable for years until inundation triggers hatching [24,25]. Such flood events can cause rapid population expansions by creating abundant larval

habitats [25]. After hatching, adult mosquitoes typically disperse away from their aquatic breeding sites to seek blood meals.

We specifically chose *Aedes vexans* for this study because it is widely distributed and represents one of the most common floodplain mosquito species, ensuring the availability of many occurrence records and previous modelling efforts. Also, the species occupies a unique ecological niche linked to ephemeral flood habitats, which makes it particularly suitable for assessing how habitat models capture ephemeral environmental dynamics. Finally, *Aedes vexans* is of great public health relevance, as it opportunistically feeds on both humans and animals [25] and serves as a competent vector for pathogens such as Zika, chikungunya, and West Nile viruses [25–27].

2. Materials and Methods

We conducted a systematic literature review following PRISMA standards [28]. The PRISMA checklist and the PRISMA abstract checklist are provided in the Supplementary Materials. The search was performed in the PubMed, Scopus, and Web of Science databases on 17 April 2024. To identify a suitable search term, we first conducted an unstructured search to identify relevant studies, including those on *Aedes vexans* cited in a recent comprehensive review of mosquito species distribution models [29]. After various alternatives had been tested, the term "vexans model" was selected as the most effective. Studies were included if they developed species-specific spatial distribution models of *Aedes vexans*, without restrictions on geography or time frame. Exclusion criteria were limited to language, as only studies published in English were considered, and any studies not meeting the above inclusion criteria were excluded. Our search returned 1250 documents (Figure 1), from which we removed 125 duplicates using a tool in the Zotero software (Version 6.0.23, Vienna, VA, USA) [30]. A single reviewer screened the titles and excluded an additional 1026 documents. The remaining studies were reviewed based on their full content, resulting in the identification of 18 relevant studies. The literature-selection procedure and the subsequent data extraction were validated by a second reviewer.

To capture reports not included in major databases, two reviewers independently performed an unstructured search using Google Scholar, identifying five additional documents. This brought the total to 23 studies.
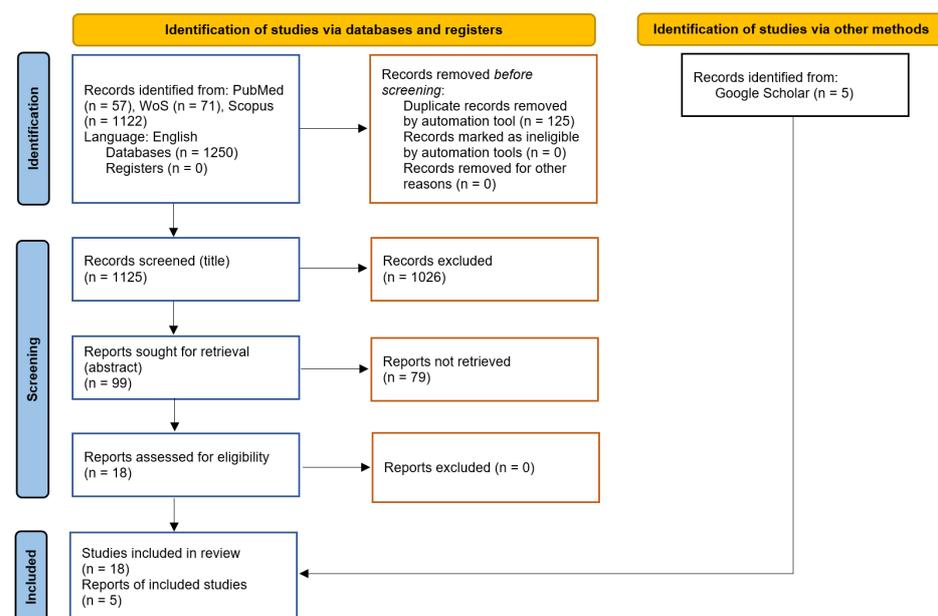

Figure 1. Flow diagram following the PRISMA protocol for databases, registers and other sources. Modified from [28].

We collected all the variables originally used in the models with the aim of assessing whether key ecological requirements of *Aedes vexans* are adequately represented. To facilitate a systematic analysis, we developed a classification scheme, grouping variables into four main categories: 'Land characteristics', 'Water characteristics', 'Population', and 'Weather'. Within each category, we established subcategories to provide a more detailed and precise description of the variables (Table 1). The new grouped variables act as metavariables to organize the original input variables.

Table 1. Categorization and description of the explanatory variables used in the 28 reviewed species distribution models of *Aedes vexans*. The table lists each variable category and subcategory, along with a brief description of their content. If labels differ between the dataset and the figures, the dataset label is shown in parentheses. The original variable names used in the reviewed studies, together with their classifications, are available in the accompanying data repository.

Category	Description	Subcategory	Description
Land characteristic	Data describing the physical features and conditions of the land, often observable from a 'top-down' perspective.	Artificial	Details about human-made structures and landscapes, such as buildings, roads, and urban developments.
		Bare	Details about areas with minimal or no vegetation, such as deserts, rocky regions, or barren soils.
		Infrared	Metrics derived from infrared sensing, including vegetation health, water stress, or thermal properties.
		Terrain	Data on landscape features, such as elevation, slope, or landforms.
		Vegetation	Information about plant life, such as vegetation types, density, or green cover.
		Water	Data on water bodies, including rivers, lakes, ponds, and other forms of standing or flowing water.
		Wetland	Information on marshes, swamps, bogs, or areas seasonally saturated with water.
		Missing data (-)	Missing or unavailable values (e.g., cloud cover).
		Not specified (unspec.)	Variables without detailed classification.
Population	Data on the size or distribution of populations within a given area.	Distance	Measures of distance to the nearest population center.
		Human	Data on human activities, population density, or anthropogenic impacts.
		Other animal	Information about the presence or activities of non-human animal species.
Water characteristic	Data describing the physical and chemical attributes of water bodies.	Amount	Information about the amount of water available to a system or organism.
		Quality	Data on the chemical or biological quality of water.
		Soil water	Metrics related to soil moisture or groundwater levels.
Weather	Data describing atmospheric conditions.	Temperature	Metrics related to temperature and associated variations.
		Precipitation	Data on rainfall and other forms of precipitation.
		Radiation	Details about solar radiation, influencing environmental conditions.
		Humidity	Information on atmospheric moisture levels.
		Not specified (unspec.)	Variables without detailed classification.

Variables often referred to specific time periods, either within a single year or spanning multiple years. The time periods were collected for every model, as were the aggregation methods used to summarize the values within the periods.

Not all variables equally affect model outcomes, as their predictive importance depends on various factors [31]. Importance is typically assessed by evaluating how much model predictions change when a variable is omitted. If omitting a variable leads to substantial changes in predictions, the variable is considered significant. Variables labelled as important in the reviewed studies were recorded as such. However, differences in methods for assessing variable importance introduce some uncertainty. A detailed description of each study's approach to evaluating importance is provided in the supplementary data [32]. To evaluate how well variables predict the distribution of *Aedes vexans*, we looked at how often each variable was used and how many studies considered it important for the model's accuracy. The data were analysed in R (Version 4.3.3) [33]. The collected data are available in a repository for further reference [32]. This review was not registered in a recognised international registry for systematic reviews.

3. Results

3.1. Ecological Completeness of Variables

To assess the ecological completeness of species distribution models for *Aedes vexans*, we examined whether all potentially distribution-limiting variables were represented within the initial variable sets used in existing studies. Ensuring that key habitat factors are included is critical for accurately capturing the species' ecological niche and avoiding bias in model predictions [3,15–18].

Figure 2 provides an overview of the relative representation of different categories and subcategories among the variables used in the 28 model instances derived from the 23 selected studies on *Aedes vexans* habitat [34–56]. Seven variables could not be assigned to subcategories and were therefore excluded from the analysis.

The category 'Land characteristics' accounts for the largest share of explanatory variables, primarily as a result of the variables in the subcategory 'Vegetation'. In addition to 'Vegetation', we introduced 'Wetland' as a separate subcategory because of the known association of *Aedes vexans* with these habitats [25]. However, 'Wetland' variables are used less frequently than those related to 'Water' or artificial surfaces ('Artificial'). None of the variables in the 'Water' subcategory explicitly represent areas with fluctuating water levels, although these are a critical aspect of *Aedes vexans* biology. Such regions are represented only indirectly, for instance by 'Wetland' variables or elevation data classified under the 'Terrain' subcategory, which may serve as proxies for flood-prone areas.

The second-largest category is 'Weather', which primarily comprises 'Temperature' and 'Precipitation' variables, with temperature being more prevalent. Together with 'Vegetation', the subcategories 'Temperature' and 'Precipitation' constitute the dominant group of explanatory variables. These variables capture ecologically relevant factors identified across multiple studies and settings, suggesting they are likely to influence *Aedes vexans* distribution in future models and across different regions or conditions. For example, temperature is a known determinant of larval development and vegetation affects site selection for oviposition by female mosquitoes [21–23,25].

Some less frequently used variables also demonstrated strong predictive performance and may offer underexplored insights into the species' ecological niche. Soil moisture, for instance, prevents egg desiccation and supports oviposition, yet it is rarely included despite its relevance [25]. The categories 'Population' and 'Water characteristic' were considered far less frequently.
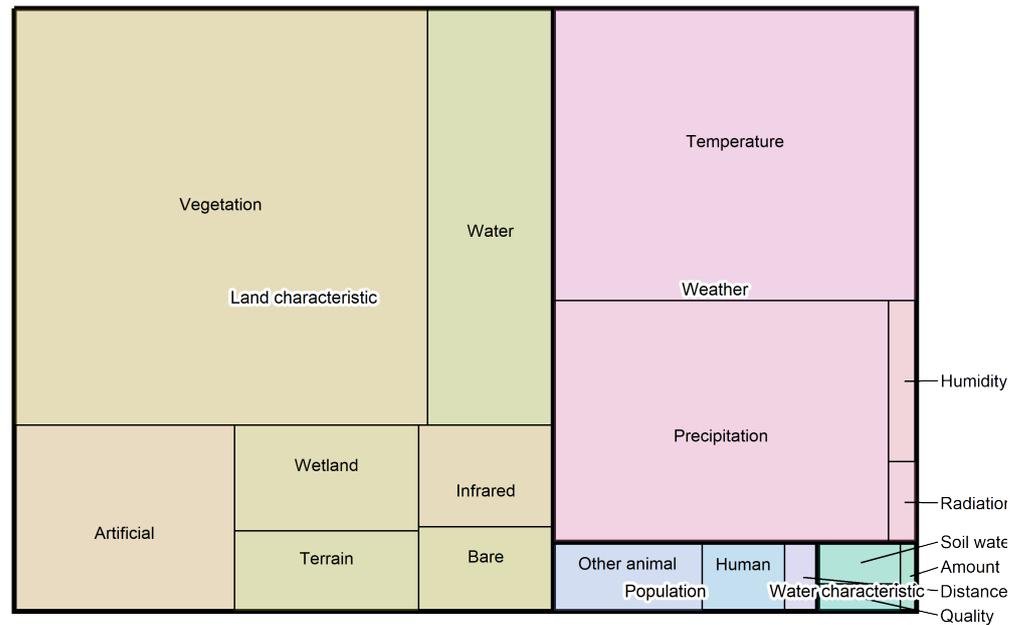
Figure 2. Treemap of explanatory variables used in the reviewed studies. Each rectangle represents a category of explanatory variables, with its area proportional to the number of variables used across the studies. Colour-coded boxes with white-bordered text indicate the overarching categories, while smaller boxes in matching hues represent the corresponding subcategories. If one category occupies twice the area of another, variables within that category were used twice as frequently. In total, 477 variables were included. Seven entries assigned to the subcategories 'Missing Data' and 'Not specified' were excluded from the figure and subsequent analysis.

Most variables originate from global datasets such as CORINE (20%), WORLDCLIM (20%), MODIS (13%), and GLOBCOVER (9%), despite the majority of studies being conducted at local or national scales [57–60]. Specifically, over half of the studies (58.6%) were performed at national (20.7%) or sub-national (37.9%) levels, with only a small fraction (3.5%) conducted at the global scale and 37.9% covering multiple countries without reaching a global scope. This highlights a mismatch between the spatial resolution of available environmental data and the geographic focus of most research efforts.

The high percentage of studies employing these commonly used data products reveals that projections of *Aedes vexans* distributions under future scenarios are rarely undertaken. To date, only a single study explored the species' potential distribution under climate change, and this analysis was conducted at the global scale. No study has yet incorporated regionally downscaled climate projections to assess future distributions at the local level.

The variables used in the studies are aggregated over multiple years in 67% of cases. Nine percent are considered on an annual basis, and another nine percent are considered on a quarterly basis. The remaining variables are included in the models with greater temporal resolution.

3.2. Essential Variables Used in Existing Models

We analysed the correspondence between the frequency of use of variables and their recognised importance in characterizing *Aedes vexans* habitat (Figure 3). The variables used in the modelling studies vary in their influence on the models' outcomes. If a variable is removed and the accuracy of the model drops significantly, that variable is of great importance as a descriptor of the potential *Aedes vexans* distribution. This metric is known as variable importance.

Five studies had to be excluded from the analysis because they used methods which did not allow for the determination of the predictive power of variables [35,36,38,44,48]. These are primarily customised models, the use of which limits the ability to assess predictive quality between alternative model versions.

Approximately 39% of the studies reported the importance of individual variables. Among these, the most commonly used method was permutation importance or variable contribution (38%), while other approaches such as jackknife tests (25%), fitting individual models with selected variables (25%), or SHAP values (12.5%) were applied less frequently.

The yellow top-right square of Figure 3 highlights variables that are frequently used, and when applied, were often considered important for model performance. Main subcategories of variables such as 'Vegetation', 'Temperature', and 'Precipitation' fall into this group. In contrast, the blue top-left section includes variables that are also frequently used but rarely demonstrate a substantial influence on model results. The lower half of the figure features variables that are rarely considered. In particular, the green bottom-right box comprises variables that are seldom used but consistently flagged as important when they are included. The variables in this box are classified in the category 'Water characteristics'.
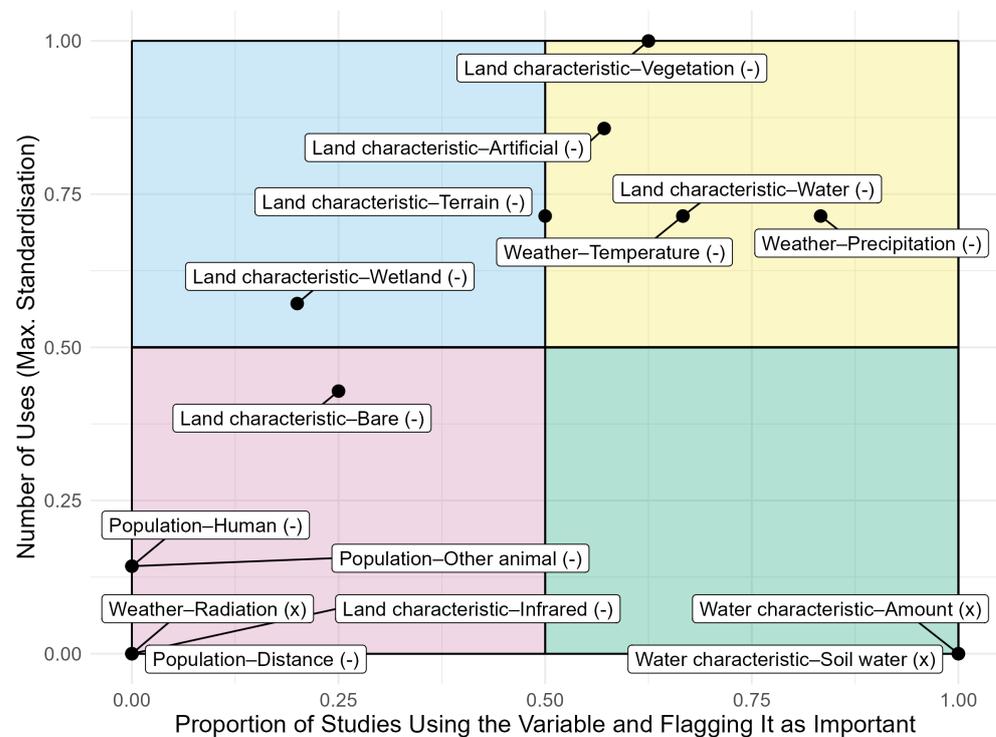

Figure 3. Relationship between the frequency of each explanatory variable used at the subcategory level and its perceived importance in species distribution models. The *y*-axis represents the number of times each variable was used across studies, standardised to a maximum value of 1. The *x*-axis shows the proportion of studies that not only used the variable but also identified it as important for model performance. When multiple variables from the same subcategory were included in a study, they were merged and the subcategory was considered important if any of its constituent variables significantly influenced the model outcome. The seven values of the subcategories 'Missing data' and 'Not specified' were excluded. Symbols next to variable labels indicate whether spatial sampling bias in the occurrence data was addressed in the corresponding studies. The symbol '-' indicates a mixture of studies that did and did not account for sampling bias, and the symbol 'x' indicates none of the studies accounted for sampling bias. None of the variables was exclusively used in studies that accounted for sampling bias in occurrence data. Six studies did not use occurrence data and were excluded from this analysis.

None of the variables was exclusively used in studies that fully accounted for sampling bias in occurrence data, and therefore no '+' can be found in the figure. For most variables, usage was distributed across studies that either addressed or neglected sampling bias. Notably, three variables were used exclusively in studies that did not apply any methods for sampling-bias correction.

4. Discussion

Species distribution models depend on the careful selection of variables to accurately represent a species' ecological niche [3,15–18]. This study pursued two primary objectives: (1) to identify variables that are frequently used and reported as important in existing *Aedes vexans* species distribution models and are therefore likely to be relevant in future models; and (2) to detect and critically evaluate potential biases in current variable-selection practices. Our review of 23 studies and 28 model instances yielded 472 reported variables. We classified them based on their frequency of use and reported predictive importance.

Our review suggests that, with the exception of the subcategory 'Artificial', most variables shown in the right half of Figure 3 should be included in the initial variable selection for future species distribution models for *Aedes vexans*. In particular, variables in the upper-right quadrant were frequently used, demonstrated robust predictive performance, and represent essential dimensions of the species' ecological niche, such as temperature, precipitation, and vegetation cover [25]. However, even these variables warrant critical ecological evaluation to ensure that key aspects of the species' biology are not overlooked.

A prominent example of such an omission concerns the lack of variables capturing flood dynamics, despite their central importance to *Aedes vexans* ecology. The species relies on temporary floodwaters for egg hatching, with breeding sites often located in areas that flood regularly. Exceptional flood events can substantially expand breeding habitats, leading to mass emergence events [25]. Yet, none of the reviewed models explicitly accounted for flooding dynamics. Moreover, long-term averaged variables commonly used in species distribution models tend to obscure such short-term fluctuations. As a result, current models likely fail to represent the full complexity of *Aedes vexans* habitats. Including flood-related variables could significantly improve model realism and ecological completeness. Nevertheless, such data are not universally accessible, and their predictive power may appear limited in models focused on adult occurrences, as most available records pertain to adult mosquitoes. The frequent use and high importance of temperature, precipitation, and land-cover variables—such as water and vegetation—in species distribution models for *Aedes vexans* align with the findings of a broader review study covering multiple mosquito species [29].

Another key finding concerns the frequent use of variables related to human settlements, particularly those in the 'Artificial' subcategory. While *Aedes vexans* may indeed benefit from blood-meal availability in densely populated areas, the apparent importance of these variables in many species distribution models likely results from sampling bias. Most occurrence records originate from easily accessible, urbanized areas [61,62]. Among the eight studies that included 'Artificial' variables, only two corrected for spatial sampling bias—and neither identified 'artificial' as an important predictor. In contrast, five of the six studies without bias correction reported 'Artificial' as highly relevant. A similar pattern was found for the related 'Human' subcategory, underscoring the often-addressed substantial effect of uncorrected sampling biases on model outcomes [63–71].

Current models also neglect the full host range of *Aedes vexans*, focusing solely on humans and livestock, despite evidence that the species feeds on a wide variety of hosts, including birds and rodents [25,72–74]. Incorporating variables that represent all rele-

vant host groups could markedly improve the accuracy of species distribution models, particularly in regions where non-human hosts are abundant.

Only 2 of the 23 reviewed models explicitly specify which of the four recognized *Aedes vexans* subspecies they represent. Subspecies of *Aedes vexans* exhibit distinct geographic distributions [75], a factor rarely accounted for in current models. Consequently, it is often unclear whether the different subspecies occupy partially distinct ecological niches. To date, subspecies-specific modelling has been performed only for *Aedes vexans vexans* [41,42], and no tailored distribution models exist for the remaining subspecies, limiting our understanding of their potential differentiation.

Our review highlights a frequent mismatch between the spatial scales of the variables and those of the distribution models. While over half of the reviewed studies were conducted at national or sub-national scales, many relied on global or continental datasets such as CORINE, WORLDCLIM, MODIS, or GLOBCOVER [57–60]. Datasets covering global and continental scales are particularly valuable for large-scale analyses, as they provide key insights into historical dynamics and potential future global changes. However, conducting research on local or regional scales using global or continental datasets can obscure local ecological gradients and lead to inaccurate habitat-suitability predictions, as global or continental datasets often fail to capture fine-scale environmental heterogeneity [76–78]. Additionally, they are often biased [79,80], which can introduce artefacts into the distribution model for species with specific microhabitat requirements, such as *Aedes vexans*. However, various methods and recent advances exist to improve the quality of global datasets used at regional or local scales. Approaches that integrate regional or local data with global datasets appear promising for enhancing input data, even for areas lacking gridded datasets specifically compiled for that location [76,81]. Other methods focus on correcting biases in existing data products [82–84]. Our findings align with concerns raised by Lippi et al. [29], who noted that many mosquito species distribution models lack biological justification in variable selection and often neglect to address collinearity. This reflects a general deficit in ecological hypothesis testing, which may lead to oversimplified or misleading predictions. Our review confirms this issue for *Aedes vexans*, highlighting the need for more biologically grounded approaches in future modelling efforts.

Limitations and Methodological Considerations

Several challenges arose during the review process, and these may have influenced the interpretation and generalisability of our findings. One key limitation was the inconsistent reporting and classification of explanatory input variables across studies. Often, variables were described vaguely or referenced only by the data product from which they were derived, without further specification. This limited the precision with which variables could be compared or interpreted.

To manage this heterogeneity, we developed a hierarchical classification scheme with major categories and subcategories. While this structure allowed us to organise and analyse a large number of variables effectively, it also introduced a trade-off: broader groupings may overgeneralise ecologically distinct variables, potentially masking differences in their relevance for *Aedes vexans* habitat.

Another major limitation was the frequent lack of reporting on variable importance. Sixty percent of the reviewed studies did not quantify or describe the contribution of individual variables to model performance. This omission creates uncertainty in evaluating which variables truly influence model predictions. Notably, Lippi et al. [29] also reported that nearly 20% of distribution-modelling studies on mosquitoes failed to report variable importance, indicating that this issue persists across the broader field, albeit to a lesser extent. Moreover, retrospective assessment of variable importance was usually not possible,

as only one study partially shared its code and just 17% provided access to datasets or model outputs. Without standardised reporting practices and the open sharing of data and model results, it remains challenging to trace back model inputs. Initiatives advocating for greater transparency and reproducibility in ecological modelling represent promising steps toward improving the reliability and comparability of future studies [85].

5. Conclusions

Our review highlights a broader tendency in species distribution modelling toward a data-driven, convenience-based "all-in, goods-out" variable-selection approach [3,29]. While such models may yield acceptable results for generalist species with broad ecological niches and high detectability, they are poorly suited to species like *Aedes vexans*, whose distribution is driven by dynamic, temporally variable processes such as flooding. Modelling such species requires variables that are specifically adapted to these dynamic conditions and directly capture the spatial variation.

To improve model quality, a more targeted, hypothesis-driven selection of ecologically meaningful variables is essential [86]. Our classification framework, which ranks variables based on their usage frequency and reported predictive performance, offers a structured starting point for variable selection.

The grouped variables introduced in this study serve as metavariables that provide a structured approach for organizing the large set of individual input variables. These metavariables capture the main environmental and ecological dimensions relevant to the life cycle and habitat requirements of *Aedes vexans*. Their utility extends beyond the specific dataset used here: on a general level, they allow for the identification of key dimensions that should be considered when constructing species distribution models. At a more detailed level, they facilitate the selection of specific datasets that accurately capture these dimensions or help to identify gaps where relevant data are missing, such as absence of flood-frequency data from the *Aedes vexans* models. This dual applicability highlights the value of metavariables for both guiding the overall modelling strategy and informing the choice of input variables.

Beyond its application to *Aedes vexans*, this systematic approach is transferable to other species, provided that sufficient modelling literature exists. It can facilitate more transparent, reproducible, and ecologically grounded model development by identifying both consistently used variables and overlooked but potentially relevant predictors. Nevertheless, variable relevance remains inherently species- and context-specific. Our classification should therefore be adapted to the ecological traits of the target species and the objectives of each modelling study; however, variable selection in future models should not be based on prior use alone, but must also reflect the variable's ecological plausibility as a factor limiting the spatial distribution of the species.

Data Availability Statement: The data presented in this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.14603893.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chauvier, Y.; Descombes, P.; Guéguen, M.; Boulangeat, L.; Thuiller, W.; Zimmermann, N.E. Resolution in species distribution models shapes spatial patterns of plant multifaceted diversity. *Ecography* **2022**, *2022*, e05973. [CrossRef]
2. Díaz, S.; Malhi, Y. Biodiversity: Concepts, Patterns, Trends, and Perspectives. *Annu. Rev. Environ. Resour.* **2022**, *47*, 31–63. [CrossRef]
3. Elith, J.; Leathwick, J.R. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [CrossRef]
4. Wilson, C.D.; Roberts, D.; Reid, N. Applying species distribution modelling to identify areas of high conservation value for endangered species: A case study using *Margaritifera margaritifera* (L.). *Biol. Conserv.* **2011**, *144*, 821–829. [CrossRef]
5. Guisan, A.; Tingley, R.; Baumgartner, J.B.; Naujokaitis-Lewis, I.; Sutcliffe, P.R.; Tulloch, A.I.T.; Regan, T.J.; Brotons, L.; McDonald-Madden, E.; Mantyka-Pringle, C.; et al. Predicting species distributions for conservation decisions. *Ecol. Lett.* **2013**, *16*, 1424–1435. [CrossRef]
6. Burrows, H.; Slatculescu, A.M.; Feng, C.X.; Clow, K.M.; Guillot, C.; Jardine, C.M.; Leighton, P.A.; Krause, P.J.; Kulkarni, M.A. The utility of a maximum entropy species distribution model for *Ixodes scapularis* in predicting the public health risk of Lyme disease in Ontario, Canada. *Ticks Tick-Borne Dis.* **2022**, *13*, 101969. [CrossRef]
7. Godefroid, M.; Cruaud, A.; Streito, J.C.; Rasplus, J.Y.; Rossi, J.P. *Xylella fastidiosa*: Climate suitability of European continent. *Sci. Rep.* **2019**, *9*, 8844. [CrossRef]
8. Messina, J.P.; Kraemer, M.U.; Brady, O.J.; Pigott, D.M.; Shearer, F.M.; Weiss, D.J.; Golding, N.; Ruktanonchai, C.W.; Gething, P.W.; Cohn, E.; et al. Mapping global environmental suitability for Zika virus. *eLife* **2016**, *5*, e15272. [CrossRef] [PubMed]
9. Cao, B.; Bai, C.; Wu, K.; La, T.; Su, Y.; Che, L.; Zhang, M.; Lu, Y.; Gao, P.; Yang, J.; et al. Tracing the future of epidemics: Coincident niche distribution of host animals and disease incidence revealed climate-correlated risk shifts of main zoonotic diseases in China. *Glob. Change Biol.* **2023**, *29*, 3723–3746. [CrossRef]
10. Bosso, L.; Luchi, N.; Maresi, G.; Cristinzio, G.; Smeraldo, S.; Russo, D. Predicting current and future disease outbreaks of *Diplodia sapinea* shoot blight in Italy: Species distribution models as a tool for forest management planning. *For. Ecol. Manag.* **2017**, *400*, 655–664. [CrossRef]
11. Ikegami, M.; Jenkins, T.A. Estimate global risks of a forest disease under current and future climates using species distribution model and simple thermal model—Pine Wilt disease as a model case. *For. Ecol. Manag.* **2018**, *409*, 343–352. [CrossRef]
12. Pasanisi, E.; Pace, D.; Orasi, A.; Vitale, M.; Arcangeli, A. A global systematic review of species distribution modelling approaches for cetaceans and sea turtles. *Ecol. Inform.* **2024**, *82*, 102700. [CrossRef]
13. Benavides Rios, E.; Sadler, J.; Graham, L.; Matthews, T.J. Species distribution models and island biogeography: Challenges and prospects. *Glob. Ecol. Conserv.* **2024**, *51*, e02943. [CrossRef]
14. Valavi, R.; Guillera-Arroita, G.; Lahoz-Monfort, J.J.; Elith, J. Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecol. Monogr.* **2022**, *92*, e01486. [CrossRef]
15. Leathwick, J.R.; Whitehead, D. Soil and atmospheric water deficits and the distribution of New Zealand's indigenous tree species. *Funct. Ecol.* **2001**, *15*, 233–242. [CrossRef]
16. Austin, M.P.; Van Niel, K.P. Improving species distribution models for climate change studies: Variable selection and scale. *J. Biogeogr.* **2011**, *38*, 1–8. [CrossRef]
17. Jarnevich, C.S.; Stohlgren, T.J.; Kumar, S.; Morisette, J.T.; Holcombe, T.R. Caveats for correlative species distribution modeling. *Ecol. Inform.* **2015**, *29*, 6–15. [CrossRef]
18. Mod, H.K.; Scherrer, D.; Luoto, M.; Guisan, A. What we use is not what we know: Environmental predictors in plant distribution models. *J. Veg. Sci.* **2016**, *27*, 1308–1322. [CrossRef]
19. Vignali, S.; Barras, A.G.; Arlettaz, R.; Braunisch, V. SDMtune: An R package to tune and evaluate species distribution models. *Ecol. Evol.* **2020**, *10*, 11488–11506. [CrossRef]
20. Adde, A.; Rey, P.L.; Fopp, F.; Petitpierre, B.; Schweiger, A.K.; Broennimann, O.; Lehmann, A.; Zimmermann, N.E.; Altermatt, F.; Pellissier, L.; et al. Too many candidates: Embedded covariate selection procedure for species distribution modelling with the covsel R package. *Ecol. Inform.* **2023**, *75*, 102080. [CrossRef]
21. Berec, L.; Gelbič, I.; Šebesta, O. Worthy of Their Name: How Floods Drive Outbreaks of Two Major Floodwater Mosquitoes (Diptera: Culicidae). *J. Med. Entomol.* **2014**, *51*, 76–88. [CrossRef]

22. Vaux, A.; Watts, D.; Findlay-Wilson, S.; Johnston, C.; Dallimore, T.; Drage, P.; Medlock, J. Identification, surveillance and management of *Aedes vexans* in a flooded river valley in Nottinghamshire, United Kingdom. *J. Eur. Mosq. Control Assoc.* **2021**, *39*, 15–25. [CrossRef]

23. Hesson, J.; Schäfer, M.; Lundström, J. Increased abundance of floodwater mosquito species over three seasons in Forshaga municipality, western Sweden. *J. Eur. Mosq. Control Assoc.* **2023**, *41*, 25–34. [CrossRef]

24. Gjullin, C.M.; Yates, W.W.; Stage, H.H. Studies on *Aedes xexans* (Meig.) and *Aedes Sticticus* (Meig.), Flood-Water Mosquitoes, in the Lower Columbia River Valley. *Ann. Entomol. Soc. Am.* **1950**, *43*, 262–275. [CrossRef]

25. Becker, N.; Petrić, D.; Zgomba, M.; Boase, C.; Madon, M.B.; Dahl, C.; Kaiser, A. Biology of Mosquitoes. In *Mosquitoes: Identification, Ecology and Control*; Springer International Publishing: Cham, Switzerland, 2020; pp. 11–27. [CrossRef]

26. Kampen, H.; Walther, D. Vector Potential of Mosquito Species (Diptera: Culicidae) Occurring in Central Europe. In *Mosquito-Borne Diseases: Implications for Public Health*; Benelli, G., Mehlhorn, H., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 41–68. [CrossRef]

27. Martinet, J.P.; Ferté, H.; Failloux, A.B.; Schaffner, F.; Depaquit, J. Mosquitoes of North-Western Europe as Potential Vectors of Arboviruses: A Review. *Viruses* **2019**, *11*, 1059. [CrossRef]

28. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]

29. Lippi, C.A.; Mundis, S.J.; Sippy, R.; Flenniken, J.M.; Chaudhary, A.; Hecht, G.; Carlson, C.J.; Ryan, S.J. Trends in mosquito species distribution modeling: Insights for vector surveillance and disease control. *Parasit. Vectors* **2023**, *16*, 302. [CrossRef]

30. Corporation for Digital Scholarship. *Zotero*, Version 6.0.23; Corporation for Digital Scholarship: Vienna, VA, USA, 2024. Available online: https://www.zotero.org/ (accessed on 12 October 2025).

31. Smith, A.B.; Santos, M.J. Testing the ability of species distribution models to infer variable importance. *Ecography* **2020**, *43*, 1801–1813. [CrossRef]

32. Pothmann, P.; Kampen, H.; Werner, D.; Thulke, H.H. Data from: Systematic Review of Variable Selection Bias in Species Distribution Models for *Aedes vexans* (Diptera: Culicidae). Zenodo. 2025. Available online: https://doi.org/10.5281/zenodo.14603894 (accessed on 12 October 2025).

33. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2024.

34. Diuk-Wasser, M.A.; Brown, H.E.; Andreadis, T.G.; Fish, D. Modeling the Spatial Distribution of Mosquito Vectors for West Nile Virus in Connecticut, USA. *Vector-Borne Zoonotic Dis.* **2006**, *6*, 283–295. [CrossRef] [PubMed]

35. Tourre, Y.; Lacaux, J.P.; Vignolles, C.; Ndione, J.A.; Lafaye, M. Mapping of Zones Potentially Occupied by *Aedes vexans* and Culex poicilipes Mosquitoes, the Main Vectors of Rift Valley Fever in Senegal. *Geospat. Health* **2008**, *3*, 69–79. [CrossRef] [PubMed]

36. Vignolles, C.; Lacaux, J.P.; Tourre, Y.M.; Bigeard, G.; Ndione, J.A.; Lafaye, M. Rift Valley fever in a zone potentially occupied by *Aedes vexans* in Senegal: Dynamics and risk mapping. *Geospat. Health* **2009**, *3*, 211–220. [CrossRef]

37. Larson, S.R.; DeGroote, J.P.; Bartholomay, L.C.; Sugumaran, R. Ecological Niche Modeling of Potential West Nile Virus Vector Mosquito Species in Iowa. *J. Insect Sci.* **2010**, *10*, 110. [CrossRef]

38. Caminade, C.; Ndione, J.A.; Kebe, C.M.F.; Jones, A.E.; Danuor, S.; Tay, S.; Tourre, Y.M.; Lacaux, J.P.; Vignolles, C.; Duchemin, J.B.; et al. Mapping Rift Valley fever and malaria risk over West Africa using climatic indicators. *Atmos. Sci. Lett.* **2011**, *12*, 96–103. [CrossRef]

39. Chuang, T.W.; Henebry, G.M.; Kimball, J.S.; VanRoekel-Patton, D.L.; Hildreth, M.B.; Wimberly, M.C. Satellite microwave remote sensing for environmental modeling of mosquito population dynamics. *Remote Sens. Environ.* **2012**, *125*, 147–156. [CrossRef]

40. Ducheyne, E.; Versteirt, V.; Hendrickx, G. Abundance of Rift Valley Fever vectors in Europe and the Mediterranean Basin. *EFSA Support. Publ.* **2013**, *10*, 420E. [CrossRef]

41. Tran, A.; Ippoliti, C.; Balenghien, T.; Conte, A.; Gely, M.; Calistri, P.; Goffredo, M.; Baldet, T.; Chevalier, V. A Geographical Information System-Based Multicriteria Evaluation to Map Areas at Risk for Rift Valley Fever Vector-Borne Transmission in Italy. *Transbound. Emerg. Dis.* **2013**, *60*, 14–23. [CrossRef] [PubMed]

42. Versteirt, V.; Ducheyne, E.; Schaffner, F.; Hendrickx, G. Systematic literature review on the geographic distribution of rift valley fever vectors in Europe and the neighbouring countries of the Mediterranean Basin. *EFSA Support. Publ.* **2013**, *10*, 412E. [CrossRef]

43. Mughini-Gras, L.; Mulatti, P.; Severini, F.; Boccolini, D.; Romi, R.; Bongiorno, G.; Khoury, C.; Bianchi, R.; Montarsi, F.; Patregnani, T.; et al. Ecological Niche Modelling of Potential West Nile Virus Vector Mosquito Species and Their Geographical Association with Equine Epizootics in Italy. *EcoHealth* **2014**, *11*, 120–132. [CrossRef] [PubMed]

44. Talla, C.; Diallo, D.; Dia, I.; Ba, Y.; Ndione, J.A.; Sall, A.A.; Morse, A.; Diop, A.; Diallo, M. Statistical Modeling of the Abundance of Vectors of West African Rift Valley Fever in Barkédji, Senegal. *PLoS ONE* **2014**, *9*, e114047. [CrossRef]

45. Arsevska, E.; Hellal, J.; Mejri, S.; Hammami, S.; Marianneau, P.; Calavas, D.; Hénaux, V. Identifying Areas Suitable for the Occurrence of Rift Valley Fever in North Africa: Implications for Surveillance. *Transbound. Emerg. Dis.* **2016**, *63*, 658–674. [CrossRef] [PubMed]

46. Schaffner, F.; Versteirt, V.; Bortel, W.V.; Zeller, H.; Wint, W.; Alexander, N.S. VBORNET gap analysis: Mosquito vector distribution models utilised to identify areas of potential species distribution in areas lacking records. *Open Health Data* **2016**, *4*, e6. [CrossRef]

47. Giordano, B.V.; Turner, K.W.; Hunter, F.F. Geospatial Analysis and Seasonal Distribution of West Nile Virus Vectors (Diptera: Culicidae) in Southern Ontario, Canada. *Int. J. Environ. Res. Public Health* **2018**, *15*, 614. [CrossRef]

48. Tran, A.; Fall, A.G.; Biteye, B.; Ciss, M.; Gimonneau, G.; Castets, M.; Seck, M.T.; Chevalier, V. Spatial Modeling of Mosquito Vectors for Rift Valley Fever Virus in Northern Senegal: Integrating Satellite-Derived Meteorological Estimates in Population Dynamics Models. *Remote Sens.* **2019**, *11*, 1024. [CrossRef]

49. Wint, W.; Petric, D.; Bortel, W.; Alexander, N.; Schaffner, F. RVF vector spatial distribution models: Vector abundance. *EFSA Support. Publ.* **2020**, *17*, 1847E. [CrossRef]

50. Abdelkrim, O.; Samia, B.; Said, Z.; Souad, L. Modeling and mapping the habitat suitability and the potential distribution of Arboviruses vectors in Morocco. *Parasite* **2021**, *28*, 37. [CrossRef]

51. Liu, C.; White, M.; Newell, G. Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography* **2011**, *34*, 232–243. [CrossRef]

52. Wieland, R.; Kuhls, K.; Lentz, H.H.K.; Conraths, F.; Kampen, H.; Werner, D. Combined climate and regional mosquito habitat model based on machine learning. *Ecol. Model.* **2021**, *452*, 109594. [CrossRef]

53. Outammassine, A.; Zouhair, S.; Loqman, S. Global potential distribution of three underappreciated arboviruses vectors (*Aedes japonicus*, *Aedes vexans* and *Aedes vittatus*) under current and future climate conditions. *Transbound. Emerg. Dis.* **2022**, *69*, e1160–e1171. [CrossRef]

54. Li, Y.; An, Q.; Sun, Z.; Gao, X.; Wang, H. Distribution areas and monthly dynamic distribution changes of three *Aedes* species in China: *Aedes aegypti*, *Aedes albopictus* and *Aedes vexans*. *Parasit. Vectors* **2023**, *16*, 297. [CrossRef]

55. Ludwig, A.; Rousseu, F.; Kotchi, S.O.; Allostry, J.; Fournier, R.A. Mapping the abundance of endemic mosquito-borne diseases vectors in southern Quebec. *BMC Public Health* **2023**, *23*, 924. [CrossRef]

56. Talbot, B.; Kulkarni, M.A.; Rioux-Rousseau, M.; Siebels, K.; Kotchi, S.O.; Ogden, N.H.; Ludwig, A. Ecological Niche and Positive Clusters of Two West Nile Virus Vectors in Ontario, Canada. *EcoHealth* **2023**, *20*, 249–262. [CrossRef]

57. Leroy, M.; Bicheron, P.; Brockmann, C.; Krämer, U.; Miras, B.; Huc, M.; Ninô, F.; Defourny, P.; Vancutsem, C.; Petit, D.; et al. GlobCover: A 300 m global land cover product for 2005 using ENVISAT MERIS time series. *Remote Sens. Environ.* **2008**, *114*, 2091–2107.

58. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]

59. European Environment Agency. CORINE Land Cover. 2021. Available online: https://land.copernicus.eu/en/products/corine-land-cover?tab=datasets (accessed on 12 October 2025).

60. Didan, K. MODIS/Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid V061 [Data Set]. 2021. Available online: https://doi.org/10.5067/MODIS/MOD13A1.061 (accessed on 27 January 2025).

61. Araújo, M.B.; Guisan, A. Five (or so) challenges for species distribution modelling. *J. Biogeogr.* **2006**, *33*, 1677–1688. [CrossRef]

62. Bowler, D.E.; Callaghan, C.T.; Bhandari, N.; Henle, K.; Benjamin Barth, M.; Koppitz, C.; Klenke, R.; Winter, M.; Jansen, F.; Bruelheide, H.; et al. Temporal trends in the spatial bias of species occurrence records. *Ecography* **2022**, *2022*, e06219. [CrossRef]

63. Reddy, S.; Dávalos, L.M. Geographical sampling bias and its implications for conservation priorities in Africa. *J. Biogeogr.* **2003**, *30*, 1719–1727. [CrossRef]

64. Kadmon, R.; Farber, O.; Danin, A. Effect of Roadside Bias on the Accuracy of Predictive Maps Produced by Bioclimatic Models. *Ecol. Appl.* **2004**, *14*, 401–413. [CrossRef]

65. Bystriakova, N.; Peregrym, M.; Erkens, R.H.J.; Bezsmertna, O.; Schneider, H. Sampling bias in geographic and environmental space and its effect on the predictive power of species distribution models. *Syst. Biodivers.* **2012**, *10*, 305–315. [CrossRef]

66. Syfert, M.M.; Smith, M.J.; Coomes, D.A. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS ONE* **2013**, *8*, e55158. [CrossRef]

67. Kramer-Schadt, S.; Niedballa, J.; Pilgrim, J.D.; Schröder, B.; Lindenborn, J.; Reinfelder, V.; Stillfried, M.; Heckmann, I.; Scharf, A.K.; Augeri, D.M.; et al. The importance of correcting for sampling bias in MaxEnt species distribution models. *Divers. Distrib.* **2013**, *19*, 1366–1379. [CrossRef]

68. Troudet, J.; Grandcolas, P.; Blin, A.; Vignes-Lebbe, R.; Legendre, F. Taxonomic bias in biodiversity data and societal preferences. *Sci. Rep.* **2017**, *7*, 9132. [CrossRef]

69. Titley, M.A.; Snaddon, J.L.; Turner, E.C. Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. *PLoS ONE* **2017**, *12*, e0189577. [CrossRef]

70. Hughes, A.C.; Orr, M.C.; Ma, K.; Costello, M.J.; Waller, J.; Provoost, P.; Yang, Q.; Zhu, C.; Qiao, H. Sampling biases shape our view of the natural world. *Ecography* **2021**, *44*, 1259–1269. [CrossRef]

71. García-Roselló, E.; González-Dacosta, J.; Lobo, J.M. The biased distribution of existing information on biodiversity hinders its use in conservation, and we need an integrative approach to act urgently. *Biol. Conserv.* **2023**, *283*, 110118. [CrossRef]

72. Nasci, R.S. Variations in the Blood-Feeding Patterns of *Aedes vexans* and *Aedes trivittatus* (Diptera: Culicidae)1. *J. Med. Entomol.* **1984**, *21*, 95–99. [CrossRef] [PubMed]

73. Molaei, G.; Andreadis, T.G. Identification of Avian- and Mammalian-Derived Bloodmeals in *Aedes vexans* and *Culiseta melanura* (Diptera: Culicidae) and Its Implication for West Nile Virus Transmission in Connecticut, U.S.A. *J. Med. Entomol.* **2006**, *43*, 1088–1093. [CrossRef] [PubMed]

74. Biteye, B.; Fall, A.G.; Seck, M.T.; Ciss, M.; Diop, M.; Gimonneau, G. Host-feeding patterns of Aedes (Aedimorphus) vexans arabiensis, a Rift Valley Fever virus vector in the Ferlo pastoral ecosystem of Senegal. *PLoS ONE* **2019**, *14*, e0215194. [CrossRef]

75. Wilkerson, R.C.; Linton, Y.M.; Strickman, D. *Mosquitoes of the World*, 1st ed.; Hardcover; 199 Line Drawings, 168 Color Plates; Johns Hopkins University Press: Baltimore, MD, USA, 2021; Volume 1–2, p. 1332. [CrossRef]

76. Poggio, L.; Simonetti, E.; Gimona, A. Enhancing the WorldClim data set for national and regional applications. *Sci. Total Environ.* **2018**, *625*, 1628–1643. [CrossRef]

77. Tulbure, M.G.; Hostert, P.; Kuemmerle, T.; Broich, M. Regional matters: On the usefulness of regional land-cover datasets in times of global change. *Remote Sens. Ecol. Conserv.* **2022**, *8*, 272–283. [CrossRef]

78. Hemp, A.; Hemp, J. Weather or not—Global climate databases: Reliable on tropical mountains? *PLoS ONE* **2024**, *19*, e0299363. [CrossRef]

79. Svoma, B.M.; Cerveny, R. Analyzing bias in prominent climatic data sets. *Prog. Phys. Geogr. Earth Environ.* **2012**, *36*, 333–347. [CrossRef]

80. Roberts, D.R.; Wood, W.H.; Marshall, S.J. Assessments of downscaled climate data with a high-resolution weather station network reveal consistent but predictable bias. *Int. J. Climatol.* **2019**, *39*, 3091–3103. [CrossRef]

81. Wang, T.; Hamann, A.; Spittlehouse, D.; Carroll, C. Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLoS ONE* **2016**, *11*, e0156720. [CrossRef] [PubMed]

82. Räty, O.; Räisänen, J.; Ylhäisi, J.S. Evaluation of delta change and bias correction methods for future daily precipitation: Intermodel cross-validation using ENSEMBLES simulations. *Clim. Dyn.* **2014**, *42*, 2287–2303. [CrossRef]

83. Miao, C.; Su, L.; Sun, Q.; Duan, Q. A nonstationary bias-correction technique to remove bias in GCM simulations. *J. Geophys. Res. Atmos.* **2016**, *121*, 5718–5735. [CrossRef]

84. Beyer, R.; Krapp, M.; Manica, A. An empirical evaluation of bias correction methods for palaeoclimate simulations. *Clim. Past* **2020**, *16*, 1493–1508. [CrossRef]

85. Zurell, D.; Franklin, J.; König, C.; Bouchet, P.J.; Dormann, C.F.; Elith, J.; Fandos, G.; Feng, X.; Guillera-Arroita, G.; Guisan, A.; et al. A standard protocol for reporting species distribution models. *Ecography* **2020**, *43*, 1261–1277. [CrossRef]

86. Qiao, H.; Soberón, J.; Peterson, A.T. No silver bullets in correlative ecological niche modelling: Insights from testing among many potential algorithms for niche estimation. *Methods Ecol. Evol.* **2015**, *6*, 1126–1136. [CrossRef]