Species distribution models: Administrative boundary centroid occurrences require careful interpretation

Justin R. Barker [a,*], Hugh J. MacIsaac [b,a]

[a] *Great Lakes Institute for Environmental Research, University of Windsor, Windsor, Ontario N9B 3P4, Canada*
[b] *School of Ecology and Environmental Science, Yunnan University, Kunming 650091, China*

ABSTRACT

Describing and understanding species distributions and the factors driving them is fundamental to ecology and biogeography. Species distribution models (SDMs) allow one to investigate objectives of identifying ecologically important factors to the distribution, estimating species-environment responses, predicting the probability of species occurrence, and predicting species presence or absence. Mosquito occurrence records used in SDMs are often imprecise and represented as a centroid of a geopolitical/administrative boundary. Using a virtual species, we investigated the effect of centroids on SDMs and determined which methodology was best suited to provide accurate and applicable conclusions for each of the objectives. We compared 12 distinct algorithms, four levels of pseudo-absences, and three predictor sets to determine the optimal SDM methodology for each objective. The ability of methodology considerations to account for the effects of centroids varied for each objective. Ecologically important predictors were misidentified but could be best approximated by generalized additive models with 10,000 pseudo-absences. Response curves only captured the expected positive or negative trends. Centroids limited SDMs' ability to differentiate expected probabilities, resulting in overprediction of high probability areas. Response curves and occurrence probabilities were best estimated by generalized boosting regression models. Species presence was largely over-estimated within southern regions, but underpredicted in northern regions, and was best estimated by weighted mean ensembles. Overall, generalized boosting regression methods and (weighted) mean ensembles provided the most reliable conclusions across all four objectives. Further, the most reliable conclusions were consistently observed with equal pseudo-absences when considered with the removal of low-contributing predictors, except for predictor identification.

1. Introduction

Species distribution models (SDMs), also known as ecological niche models, are powerful tools to correlate the occurrence of a species with environmental predictors to explain and/or predict a distribution (Guisan and Zimmermann, 2000). Explanations of distributions include testing hypotheses related to identifying ecologically important conditions which determine the distribution (Bradie and Leung, 2017), and/or estimating species-environment response curves (e.g. Ikegami and Jenkins, 2018). Meanwhile, SDMs predict the probability of occurrence (e.g. Jarnevich and Reynolds, 2011) and/or the expected presence or absence of a species (e.g. Johnson et al., 2017) in geographic and temporal space.

The ability of SDMs to explain and/or predict a distribution depends on, but not limited to, the algorithm, number and quality of response and predictor variables considered (Wisz et al., 2008; Synes and Osborne, 2011; Heikkinen et al., 2012). Previous comparisons have indicated that different algorithms are better suited for other objectives and input data (Heikkinen et al., 2012; Aguirre-Gutiérrez et al., 2013). The response variable consists of presences, absences, abundance, or species richness of target species or subjects (Elith et al., 2006; Wisz et al., 2008). Absence records are rarely available or reliable, with pseudo-absences applied instead (Barbet-Massin et al., 2012). Pseudo-absences represent a location where the species has not been observed, whether sampled or not and is presumed to be absent without a validating absence record (Grimmett et al., 2020). Predictors reflect the abiotic or biotic conditions for which responses are considered (Guisan and Zimmermann, 2000). Though only predictors with solid evidence indicating a causal relationship with the species' distribution should be considered when possible, and generally higher number of

---

response variables is recommended (Austin, 2002; Wisz et al., 2008; but see Boria and Blois, 2018). Scale is one of the qualities of response and predictor variables that affects SDM performance (Lechner et al., 2012; Lecours et al., 2015). Depending on the context, scale relates to the spatial relationship at which a species interacts with the environmental influences (i.e. dispersal range) (Jackson and Fahrig, 2015), the spatial characteristics of the occurrence records (i.e. sampling unit, area of observation, positional accuracy) or predictors (i.e. cell size, grain, resolution), or the spatial and temporal accuracy applied to the SDM to analyze the distribution (Lecours et al., 2015). For this paper, we use scale to describe the level of spatial detail for each form of data.

Failure to explicitly consider the scale can limit SDMs ability to provide reliable conclusions, which are applicable and accurate to aid management actions, are directly affected (Moudrý et al., 2019). The applied scale must satisfy SDMs' assumptions of i) the occurrence records contain no error; and ii) the provided environmental predictors are representative of the physiological tolerances or resource requirements of the species' niche (Austin, 2002; Osborne and Leitão, 2009). However, response quality and predictor variable availability limit the choice of scale, specifically irregular lattice occurrence records (i.e. administrative regions centroids) with high geographic uncertainty (Cheng et al., 2021; Connor et al., 2018) Response quality limits the geographic precision, while predictor scale limits the level of environmental and temporal detail available (Connor et al., 2018; Lawler et al., 2006). Previous investigations suggested that half or more of repository occurrence records (i.e. presence/absence) within the United States and select species in Europe are geographically inaccurate and represent centroids of administrative/geopolitical boundaries, hereafter centroids (Collins et al., 2017; Park and Davis, 2017; Cheng et al., 2021). Consideration of centroids in SDMs alters variable contribution assessment, interpretation of responses, and an overall decrease in predictive ability (Johnson and Gillingham, 2008; Osborne and Leitão, 2009; Naimi et al., 2011).

Many methodological approaches have been suggested but cannot entirely correct the imprecision of centroids (e.g. Araújo and New, 2007; Pacifici et al., 2019). Considering the mean predictor value within each boundary may limit the effect of centroids (Park and Davis, 2017). However, aggregation causes a loss of fine-scale variation, which may be an important determinant of species distributions. Therefore is not practical for large boundaries (Collins et al., 2017; Cheng et al., 2021). Additionally, aggregation increases potential artifact effects through the modifiable areal unit problem (MAUP; Openshaw 1984). MAUP is a source of statistical bias where correlations can vary from positive to negative depending on the aggregation scale (i.e. county, state) (Goodchild, 2011). Accordingly, the resulting aggregated scale satisfies assumption (i) but potentially violates (ii) through obscuring fine-scale variation and MAUP effects (Pearson and Dawson, 2003; Moudrý and Šímová, 2012; Manzoor et al., 2018). Further, SDMs calibrated with centroids are often misleading (Nelson, 2001; Lecours et al., 2015).

SDM applications are limited, owing to the abundance of centroids for some species and geographic regions to produce sophisticated SDMs (Duputié et al., 2014). Reliable SDMs conclusions can only be established if the applied scale captures important environmental characteristics of the species' distribution (Vergara et al., 2016). Issues related to centroids remain impossible to solve without accurate occurrence data (Josselin and Louvet, 2019). Regardless, the application of centroids is common practice in SDMs of epidemiologically relevant species in North America (e.g. Escobar et al., 2013; Johnson et al., 2017). For example, occurrence records of the arbovirus vector mosquitoes *Aedes aegypti* and *Aedes albopictus* are predominately available as county centroids within the U.S. (Hahn et al., 2017) *A. aegeypti* and *A. albopictus* can transmit over 20 pathogens, including zika, chikungunya, dengue, and yellow fever viruses (Leta et al., 2018). Aided by human-mediated dispersal and high eco-plasticity, both species have rapidly expanded their range in recent years, causing increased introduction and incidence of pathogens within novel regions (Weaver, 2014; Ibañez-Justicia,

2020). Johnson et al. (2017) predicted the distribution of *A. aegpyti* and *A. albopictus* within the contiguous U.S. from county centroids. They indicated the unlikely occurrence of both species within Wayne County, MI (Johnson et al., 2017). However, both species were recorded across the border in the Windsor-Essex region of south-western Ontario, Canada, from 2017 to 2019 (Giordano et al., 2019), with *A. albopictus* continuing to be present through June 2021 (Windsor-Essex Health Unit, 2021). Therefore, concern exists that SDMs trained with centroids may not provide reliable conclusions.

There is increasing awareness that SDMs must be tested and designed for the desired objective using available data (Aguirre-Gutiérrez et al., 2013; Guillera-Arroita et al., 2015; Araújo et al., 2019). The effect of centroids on different SDM algorithms, quantity of response and predictor variables is unknown. Previous investigations of SDMs calibrated with centroids are limited to the probability of occurrence predictions by Maximum Entropy (MaxEnt; Collins et al., 2017; Park and Davis 2017; Cheng et al., 2021). Given the availability of multiple algorithms, it is therefore vital to assess the effect of administrative centroids on different algorithms to achieve common SDM objectives. Common SDM objectives include identifying ecologically important predictors, estimating species-environment response curves, predicting the probability of occurrence, and classifying presence-absence maps (Aguirre-Gutiérrez et al., 2013; Araújo et al., 2019). Here, we provide an in-depth analysis of SDMs' ability to achieve these common objectives for epidemiologically relevant species of mosquitoes based on centroid occurrences of a virtual species in North America. To simulate reality, we created a virtual species to resemble epidemiologically relevant species with known administrative region centroid occurrences, *A. aegypti* and *A. albopictus*. We focus on assessing SDM's ability to explain a species' niche by identifying the correct ecologically important predictors and estimating appropriate species-environment responses. Additionally, we determine the ability of SDMs to predict the probability of occurrence and binary presence-absence maps from centroids. Lastly, we investigated which SDM methodology is best suited to interpret centroids across all described objectives. This work will guide the use of centroids in SDMs to improve the reliability of conclusions for public health applications.

2. Methods

We constructed a realistic virtual species to simulate real-world application of centroids, with the environmental suitability traits based on *A. aegypti* and *A. albopictus*. Environmental suitability of both species has been primarily related to temperature for development and survival (Brady et al., 2013; Eisen et al., 2014). Additionally, both species are container-breeding species that rely on precipitation for suitable reproduction conditions in natural or human-made containers (i.e. tires, flower plots) (Gama and Islamiyah, 2013; Dhimal et al., 2015). *A. albopictus* has successfully established in more northern areas, attributed to its ability to diapause and survive cold winters (Denlinger and Armbruster, 2014). *A. aegypti* has only demonstrated diapause in some populations (Lima et al., 2016). Both species occur beyond environmentally suitable areas by employing micro-niches as refugia when the macro-climate is unsuitable (Hayden et al., 2010; Murdock et al., 2017). *A. aegypti* is mainly found in urban areas and has an anthropophilic feeding habit. *A. albopictus* inhabits urban, peri-urban, rural, or forested habitats and is an opportunistic feeder (Yang et al., 2021). When species distributions overlap, *A. albopictus* is a superior larval competitor, and *A. albopictus* males will cross-mate with *A. aegypti* females causing reproductive losses (Braks et al., 2004; Lounibos et al., 2016).

2.1. Predictor variables

Based on the literature, we considered predictors that have direct and indirect effects on *A. albopictus* and *A. aegypti*, including climate,

topography, land cover and use, vegetation indices, and socioeconomic predictors (Larson et al., 2010; Mughini-Gras et al., 2014; Alahmed et al., 2015). We calculated the aggregate means of each predictor to represent mean administrative boundary predictor values with Zonal Statistics in ArcGIS 10.6 (ESRI, 2008) Boundary land cover and use categories were represented by the mean percent land cover and use per category within a boundary. We assessed multi-collinearity using Pearson's correlation matrix followed by the variance inflation factor (VIF) (Leroy et al., 2016; Naimi et al., 2014). Multi-collinearity was evaluated at the original 1 km² and aggregated scale to address all potential collinearity. We excluded highly correlated variables ($|r|>0.7$, VIF >10) to prevent analysis errors (Dormann et al., 2013). We manually selected a single predictor from each highly correlated group (Table 1; S1).

We calculated some predictors before aggregation. Growing degree days (GDD) represented the magnitude of mean monthly temperatures above a baseline temperature of 5 °C, below which development or survival cannot occur (McMaster and Wilhelm, 1997). We calculated the selected bioclimatic variables (annual mean temperature (BIO1), mean diurnal temperature (BIO2), precipitation of the wettest month (BIO13), and precipitation of the driest month (BIO14)) and GDD from Daymet version 4 monthly climate summaries (Thornton et al., 1997, 2017) in dismo and envirem R packages, respectively (Hijmans et al., 2017; Bemmels, 2018). Additionally, we estimated human population density from the most recent census data for each administrative boundary as a proxy for rural and urban areas, as well as anthropophilic nature of *A. aegpyti* (National Institute of Statistics and Geography, 2018; Statistics Canada, 2018; U.S. Census Bureau, 2017; Yang et al., 2021). We applied the normalized difference vegetation index (NDVI) to quantify vegetation greenness and provide an understanding of vegetation density (Pettorelli et al., 2005). We restricted the geographic extent to North America with administrative regions represented by Canadian health

Table 1
Predictor variables considered. Asterisks indicate predictors which were calculated from source data.

Predictor	Abbreviation	Scale	Source
Annual mean temperature*	BIO1	1 km²	Thornton et al., 1997, 2017
Mean diurnal temperature*	BIO2		
Precipitation of the wettest month*	BIO13		
Precipitation of the driest month*	BIO14		
Growing degree days*	GDD		
Snow cover	SC	24 km²	Brown and Brasnett, 2010
Elevation	EV	30 ′onds	Earth Resources Observation and Science Center, 2017
Deciduous broadleaf forests	L_DBF	0.05°	Friedl et al., 2010
Evergreen needle leaf forests	L_ENF		
Mixed forests	L_MF		
Urban settlements	L_UB		
Water	L_WT		
Woody savannas	L_WS		
Normalized difference vegetation index	NDVI	0.1°	Allen and Stockli, 2018
Human population density	PD	Canadian health regions	Statistics Canada, 2018
		U.S. counties	U.S. Census Bureau, 2017
		Mexican States	National Institute of Statistics and Geography, 2018

regions, U.S. counties, and Mexican states, resulting in a mean scale of 6424 km² (Table S2).

2.2. Virtual species design

We designed a virtual species through the virtualspecies R package (Leroy et al., 2016). We designed the habitat suitability to reflect common characteristics of *A. aegypti* and *A. albopictus* based on annual mean temperature, precipitation of the wettest month, and elevation at the most detailed available scale of 1 km². This scale was appropriate as both species have mean flight ranges of ≤1 km (Verdonschot and Besse-Lototskaya, 2014). However, this scale would not reflect appropriate micro-niches or refugia. Annual mean temperature related to survival and development (Reinhold et al., 2018). Precipitation of the wettest month indicated the availability and quality of oviposition sites for reproduction (Becker et al., 2010). Lower elevation related to increased human disturbance, creating more breeding habitats through artificial water-holding containers (Gama and Islamiyah, 2013; Dhimal et al., 2015). Accordingly, we determined the virtual species' habitat suitability from defined responses of each predictor within North America (Fig. 1a). We modeled the virtual species' habitat suitability by a multiplicative suitability index of annual mean temperature, precipitation of the wettest month, and elevation as β-function, logistic, and linear responses, respectively (Fig. 1a-c; Supplementary methods). A β-function pattern allowed for specific upper and lower thresholds with different slopes, representing temperature thresholds for survival and development (Cunze et al., 2016; Koch et al., 2016). The logistic function represented a gradual increase in suitability to a maximum asymptote, reflecting the availability of oviposition sites as precipitation increased (Alaniz et al., 2017). We expressed elevation as a piecewise linear decrease representing lower host population and pooling water availability at higher elevations (Alaniz et al., 2017; Santos and Meneses, 2017). We interpreted habitat suitability as the probability of occurrence for comparison of SDMs. The final probability map was visually examined to ensure similarity to published predictions of *A. aegypti* and *A. albopictus* for ecological realism (Fig. 1b).

2.3. Modeling building

To train SDMs, we created a geographically-structured response by converting the suitability index into presence-absence records (Fig. 1b, c). We randomly generated 3500 unique, 1 km spatially thinned and corrected by suitability presence-only records within the contiguous U. S. and Mexico. These records were then converted into centroids, such that all occurrences within a boundary were represented by one centroid, resulting in 1414 centroids.

To assess SDM methodology for centroid occurrences, we investigated common model building choices of the: algorithm; the number of pseudo-absences generated; and predictor variable selection (Fig. 1d, e). These choices represent how the species-environment relationship is interpreted, characterized, and assessed and predicted against (Elith et al., 2006; Wisz et al., 2008; Aguirre-Gutiérrez et al., 2013).

2.3.1. Algorithms

We applied twelve different distinct algorithms in either their original software, denoted by "_O", or as part of ensemble SDM software, biomod2 (Thuiller et al., 2020), denoted by "_B" (Table 2). The software or platform applied potentially impacts SDM performance as each has its unique features and considerations (Aguirre-Gutiérrez et al., 2013; Guillera-Arroita et al., 2015). The ensemble SDM framework provided by biomod2 is the most widely used within the SDM literature (Hao et al., 2019). However, the black-box nature of the biomod2 framework results in only rough model calibrations. It inflates model evaluations by rescaling predictions by a binomial generalized linear model (Thuiller et al., 2009) not present in the original software. Thus, a comparison of original and biomod2 methods assesses the consideration of software
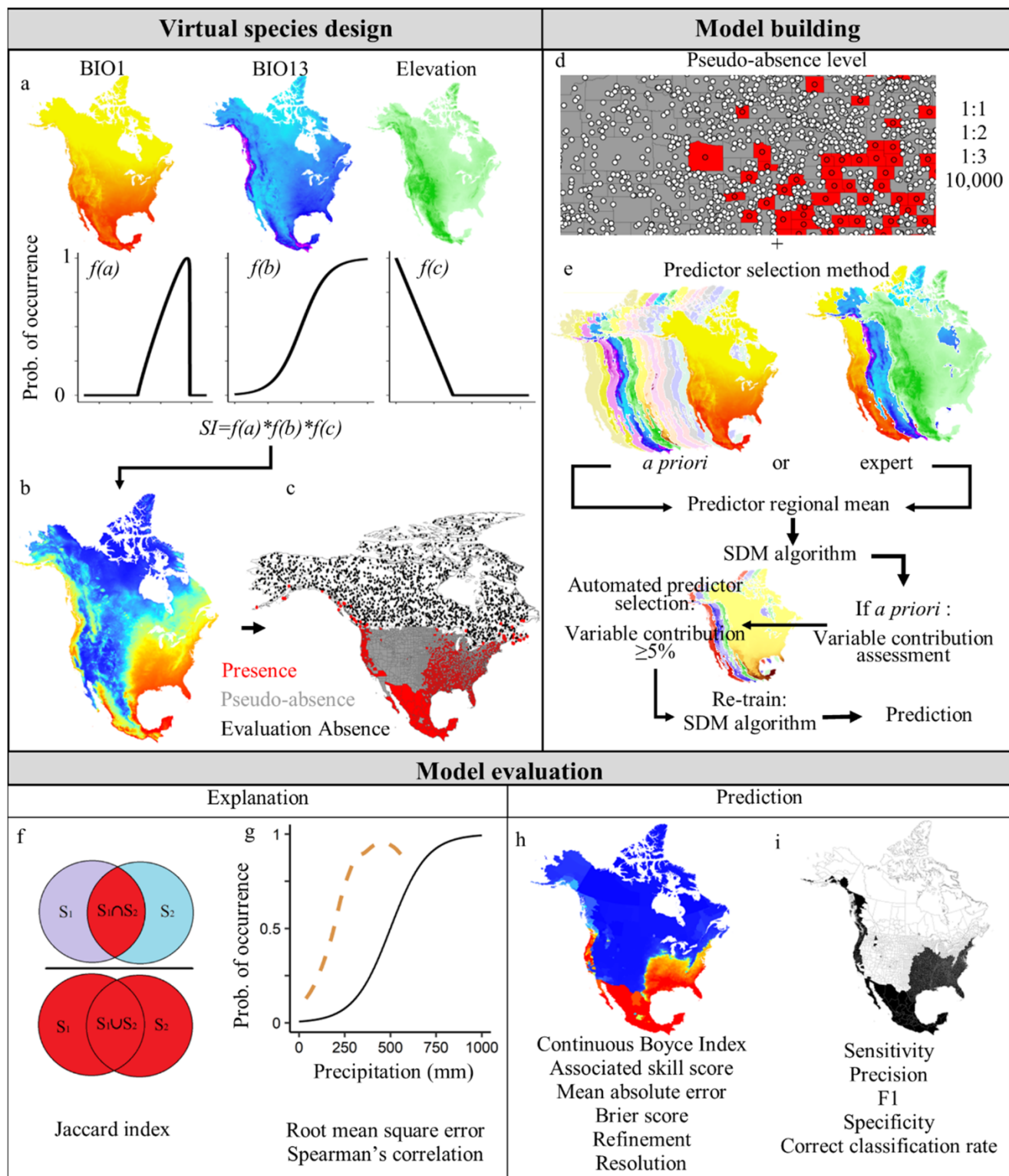
Fig. 1. Methodology workflow of virtual species design, model building, and evaluation. Virtual species were designed according to response curves from real environmental values of annual mean temperature (BIO1), precipitation of the wettest month (BIO13), and elevation (a), calculation and mapping of virtual species suitability index (b) and converted suitability index to binary training centroid occurrences (red polygon), presumed absences (gray polygon), testing occurrences (red dots) and absences (black dots) (c). Model building consisted of determining random pseudo-absences (white dots) within regions without an occurrence (red dots) at four levels (d), then one of three predictor selection methods, *a priori,* expert, or re-trained with automated predictor selection derived from *a priori* measure of variable contribution (e). Model evaluation was divided into SDMs' ability to explain or predict a distribution. SDMs' ability to explain was evaluated through their ability to identify the ecologically important predictors by Jaccard index, J, where S_1 is expert predictors, S_2 is the automated predictors, and red areas indicate the values considered to calculate J (f) and estimated response curve (orange-dashed line) accuracy relative to expected species-environment response curve (solid black line) by root mean square error and Spearman's correlation (g). Evaluation of predictions was determined by predicted probability of occurrence predictions (h) compared to true probability (b) and classification of binary presence-absences maps at a minimal training presence threshold (i).

Table 2
SDM algorithms implemented and corresponding R package.

Class of model	Algorithm	Abbreviation	Software	Reference
Machine learning	Artificial neural network	ANN_O	nnet	Venables and Ripley, 2002
		ANN_B	biomod2	Thuiller et al., 2020
	Maximum entropy	MaxEnt_O	dismo	Hijmans et al., 2017
		MaxEnt_B	biomod2	Thuiller et al., 2020
	MaxLike	MXL_O	maxlike	Royle et al., 2012
	Random forest	RF_O	randomForest	Liaw and Wiener, 2001
		RF_B	biomod2	Thuiller et al., 2020
Regression	Classification tree analysis	CTA_O	rpart	Therneau and Atkinson, 2019
		CTA_B	biomod2	Thuiller et al., 2020
	Generalized additive model	GAM_O	gam	Hastie, 2020
		GAM_B	biomod2	Thuiller et al., 2020
	Generalized linear model	GLM_O	Base R	R Core Team 2019
		GLM_B	biomod2	Thuiller et al., 2020
	Multiple adaptive regression splines	MARS_O	earth	Milborrow et al., 2019
		MARS_B	biomod2	
	Flexible discriminant analysis	FDA_O	mda	Hastie et al., 2020
		FDA_B	biomod2	Thuiller et al., 2020
Regression, boosting	Generalized boosted regression method	GBM_O	gbm	Greenwell et al., 2020
		GBM_B	biomod2	Thuiller et al., 2020
Envelope	Surface range envelopes	SRE_O	dismo	Hijmans et al., 2017
		SRE_B	biomod2	Thuiller et al., 2020
Ensemble	Committee average	EMca_B	biomod2	Thuiller et al., 2020
	Mean	EMmean_B		
	Median	EMmedian_B		
	Weighted mean	EMwmean_B		

selection on model reliability. We constructed the original regression and random forest models using additive formulas. Also, we calibrated the original generalized linear models with first and second-order polynomial variables of all predictors to reflect corresponding bio-mod2 formulas. Otherwise, all algorithms were run with default settings for simplicity and comparison (Table S3). For descriptions of each algorithm, see Elith et al. (2006), Araújo and New (2007), Reiss et al. (2011), and Fitzpatrick et al. (2013).

2.3.2. Pseudo-absences

We considered pseudo-absences for model training to simulate common practice instead of known absence. Pseudo-absences must be derived from areas accessible to the species (Barve et al., 2011). *A. aegypti* and *A. albopictus* are primarily distributed by human activity (Eritja et al., 2017). Thus, all administrative boundaries were considered accessible. Accordingly, we generated 100,000 pseudo-absences at least 1 km apart within the 1761 boundaries without a centroid occurrence in

the contiguous U.S. and Mexico. We considered pseudo-absence numbers equal, double, or triple the number of centroids (1:1, 1:2, 1:3, respectively) (e.g. Tiffin et al., 2019) or a total of 10,000. We determined 100 unique training sets per pseudo-absence level consisting of all centroid occurrences and a random subset without replacement of all pseudo-absences to ensure consistent training (Fig. 1d).

2.3.3. Predictor selection

We considered predictor selection by *a priori*, automated, and expert selection. *A priori* represented all predictors with suspected direct, indirect, or resource effects (e.g. Low et al., 2021). Alternatively, ecologically important predictors are estimated from *a priori* by the percent reduction in model fit when each predictor is randomly permuted, hereafter automated (Harisena et al., 2021). To demonstrate each predictor selection, we first trained SDMs with all predictors to represent *a priori* (Table 1). For each *a priori* training set, we calculated predictor importance per algorithm and pseudo-absence level. Predictor importance was assessed by randomizing a single predictor variable position, then making a new prediction with the new randomized predictor variables. Predictor importance reflects the difference between Pearson's correlation of the original and randomized predictors' prediction from one (Thuiller et al., 2020). We repeated this process 1500 times per predictor to allow for adequate convergence. Predictor importance was determined by functions in biomod2 or equivalent functions in caret R packages for original algorithms (Kuhn, 2020; Liaw et al., 2019; Thuiller et al., 2020). We converted the predictor contribution measures to percent contribution such that all values for a single repetition summed to one. Automated predictors represented predictors with greater than 5% mean contribution across all training sets per algorithm and pseudo-absence level. We demonstrated expert predictors by considering only the driving predictors of annual mean temperature, precipitation of the wettest month, and elevation (Fig. 1e). Therefore, we generated and compared 30,000 SDMs (25 algorithms x 4 pseudo-absence levels x 3 predictor selections x 100 training sets). We trained SDMs with an 80% random subset of training sets. Algorithms included in ensembles achieved a sensitivity of ≥95% (see Model evaluation) of omitted training presence records.

2.4. Model evaluation

We evaluated centroid explanation ability by identifying expected important predictors and response curve estimation (Fig. 1f, g). We evaluated predictive ability by the accuracy of the predicted probability of occurrence and binary discrimination (Fig. 1h, i). Each algorithm provided a different range of probabilities for interpretation. Consequently, we normalized all predictions before evaluation. Details of all evaluations are available in Supplementary methods.

2.4.1. Identification of ecologically important predictors

The Jaccard index (J) quantified the ability to identify ecologically important predictors (Fig. 1f). J is a measure of similarity between two datasets. J ranges from zero to one, such that one and zero indicate identical and no similarities, respectively (Jaccard 1908). We determined J between automated and expert predictors. For each training set and SDM, we determined how many of the three predictors used to create the virtual species were selected during automated selection and divided by the total number of unique automated and expert predictors (e.g. Inman et al., 2021) (Fig. 1f).

2.4.2. Response curve estimation

We calculated response curves for each predictor used to create the virtual species by extracting the predicted probability and corresponding aggregated predictor values. We smoothed the resulting responses by generalized additive smoothing to ensure a single probability per environmental value. Expected responses were calculated for each environmental value and compared by root mean square error (RMSE)

and Spearman's correlation (ρ) (Fig. 1 g). RMSE provided a measure of variation between the expected and estimated response values. Larger RMSE values indicated a greater error in estimated responses. ρ indicated if the estimated response curve determined the appropriate pattern relative to the expected one. ρ ranged from −1 to 1, with negative values indicating negative correlation, zero no association, and positive values indicating a positive correlation.

2.4.3. Probability of occurrence

We evaluated the predicted probability of occurrences by a full diagnostic verification by calibration, bias, skill, accuracy, refinement, and resolution of training and testing regions (Murphy and Winkler 1992). Calibration described the degree to which relative suitability of a presence correlated with predicted probability. Bias indicated the degree to which the predicted probability differed from the known probability of occurrence. Skill measured the accuracy of the predicted probability relative to an expected binary prediction. Forecast accuracy reflected the overall degree to which binned predicted probability corresponded to the expected binned probability of occurrence. Refinement indicated the mean square difference of binned predicted and expected probability values. Resolution described the ability of SDMs could separate different probabilities relative to expected separation (Murphy and Winkler 1992). Training evaluations represented comparisons of the predicted and expected probability of occurrence within the contiguous U.S. and Mexico, and testing evaluation was within Canada and Alaska unless otherwise stated.

We assessed calibration by the continuous Boyce index (CBI) in the `ecospat` R package (Broennimann et al., 2020). CBI measures the prediction accuracy of occurrence events by determining the Spearman rank correlation coefficient of the predicted-to-expected ratio. CBI values of one, zero, and negative indicated predictions consistent with occurrences, equal to random, and inconsistent with occurrences, respectively (Boyce et al., 2002; Hirzel et al., 2006). Training CBI was determined by 20% of withheld occurrences. Testing CBI was evaluated against 2000 generated unique, spatially thinned by 1 km, and corrected by suitability presence-absence records within Canada and Alaska (Fig. 1c).

We determined the unconditional bias and skill by mean absolute error (MAE) and the associated skill score (SS), respectively (Murphy ,1988; Roebber, 1998). MAE was the mean difference between predicted and expected probability values. An MAE value of zero indicated accurate predictions, and greater values indicated a higher error in predictions. SS was a mean square error measure between predicted probabilities and expected binary outcomes. A SS value of one indicated perfect skill, with greater than zero indicating better than random, and less than zero indicated worse than random. MAE and SS were determined by extracting corresponding predicted and expected probability values into the appropriate formulas (Supplementary methods).

Brier score (BS) represented forecast accuracy (Murphy and Winkler, 1992). BS assessed the mean squared error between predicted and expected binned probabilities. A value of zero indicated accurate predictions, 0.25 indicated predictions are equal to random, and greater than 0.25 indicated predictions are inaccurate and worse than random (Brier, 1950). Refinement and resolution were quantified as part of BS and examined qualitatively by attribute figures (Hsu and Murphy, 1986). Higher refinement indicated a greater difference in predicted and expected binned probabilities, and zero indicated no difference. Higher resolution values indicated a greater ability to separate different probabilities, with a minimum of zero, which indicated no separation (Murphy and Winkler, 1992). We interpreted attribute diagrams according to the categorization of reliability based on the slope of the resulting reliability line; perfect, useful, marginally useful, not useful, or dangerously useless predictions (Weisheimer and Palmer, 2014). BS, refinement, and resolution values were determined through the `verification` R package (NCAR, 2015).

2.4.4. Presence-absence map

We evaluated the binary presence-absence predictions by discrimination. Discrimination determined the threshold-dependent ability of an SDM to classify presence or absence (Fielding and Bell, 1997). We used a minimal presence threshold of omitted training centroids to create binary maps. We defined discrimination by sensitivity, specificity, precision, F1 score, and correct classification rate (CCR) (Fielding and Bell, 1997). Sensitivity and specificity were the probability that known presences or absences were predicted correctly, respectively. Precision was the probability that a predicted occurrence was an observed occurrence. F1 score was the harmonic means of sensitivity and precision. CCR is the conditional probability that presence and absences were correctly classified (Fielding and Bell, 1997). All discrimination metrics ranged from zero to one, such that one indicated perfect discrimination while those ≤ 0.5 indicated random discrimination. This range of discrimination metrics provided a complete perspective of SDMs' ability to classify presence and absences.

Training discrimination was evaluated against the 20% of presence and pseudo-absences withheld. Meanwhile, discrimination was tested against the 2000 presence-absences within Canada and Alaska described previously (Fig. 1c). Accounting for the black-box nature of `biomod2`, SDMs were trained and projected with a pre-determined 80% presence and pseudo-absence subset. Subsequent thresholds were determined by evaluating prediction against the 80% training data by threshold and evaluation functions in `dismo` R package (Hijmans et al., 2017). Lastly, the fit of each probability and binary evaluation was assessed by the minimum difference between training and testing evaluations. The minimal difference is based on the logic that overfit models will predict the training data well, but poorly on test data. Positive values indicate over-fit, while negative values indicate under-fit models (Warren and Seifert, 2011).

2.5. Analysis

Variation in reliability among model building choices was determined through Type II Wald χ^2 tests fit by linear mixed-effects models in the `car` and `lmerTest` R packages (Kuznetsova et al., 2017; Fox and Weisberg, 2019). Mixed-effects models allowed for the examination of each model building consideration while accounting for repeated measures on training sets. We determined a single mixed-effect model for each validation and evaluation metric. Mixed-effects models for J included fixed effects of algorithm and pseudo-absences interaction with random effects of the training set. We considered algorithm, pseudo-absence number, and predictor selection method as fixed effects and training set as a random effect in the mixed-effects model for response curve estimation and prediction evaluations. We transformed all evaluations by order quantile normalization before statistical analysis to normalize mixed-effect model residuals (Peterson and Cavanaugh, 2020). We examined minimal differences according to their absolute value to demonstrate deviation from fit.

We determined the mean relative performances of all associated evaluations to determine which methodology was best suited to identify ecologically important predictors, estimate species-environment response curves, predict probability of occurrence, and classify presence-absence maps. First, posthoc comparisons by estimated marginal means with Dunn–Šidák correction for pairwise comparisons were conducted for each validation and evaluation (Length 2020). We conducted posthoc tests on pseudo-absences and predictor selection variation per algorithm, and between algorithms, pseudo-absence levels, and predictor selections to determine algorithm-specific, and overall performance patterns, respectively. Second, we assigned the resulting marginal means a normalized score from zero to one based on their significance group classification for each posthoc test, such that SDMs with the same statistical group classification received the same score. We assigned the relative performance of one to the highest or lowest mean if the target value was one (highest) or zero, respectively

(Table S4). We interpreted these ranks to represent poor, fair, average, good, or excellent relative performance (Fig. 5). Third, the relative performance of each model building consideration combination was determined by calculating the relative performance mean across each evaluation per objective. We determined the relative performance of response curve estimation by the normalized mean of RMSE and ρ relative performances for annual mean temperature, precipitation of the wettest month, and elevation. Similarly, we calculated the relative performance of overall predictive ability by determining the normalized mean across all measures of training, testing, and predictive performance fit for probability of occurrence and presence-absence map classification. Overall explanation or prediction relative performance was determined by the normalized mean across each respective group of evaluation metrics. Lastly, total relative performance was determined by the normalized mean of explanation and prediction scores (Table S8). All SDM computations and analysis were completed in R v.3.6.0 (R Core Team, 2019).

3. Results

The ability of model building considerations to account for the effects of centroids varied for identification of ecologically important predictors, estimation species-environment response curves, predicting probability of occurrence, and classification presence-absence maps objectives ($p < 0.05$; Table S6). Evaluating the relative performance per objective indicated that at least one SDM successfully limited centroid effects to provide appropriate SDM conclusions (Table S7). However, model building considerations to optimize each evaluation and objective were inconsistent between evaluations, objectives, model type, software, and within algorithms (Fig. S3–4). Only two algorithms, multiple adaptive regression splines, and surface range envelopes optimized all objectives under a single methodology.

3.1. Identification of ecologically important predictors

The ability to identify ecologically important predictors from centroids was poor ($J = 0.56 \pm 0.25$; mean ± standard deviation (SD)) (Fig. 2). Only 15% of SDMs identified the ecologically important predictors. Instead, SDMs typically identified ecologically important and non-important predictors (57%) or identified two of three ecologically important predictors with (26%) or without (2%) non-important

predictors. GDD, NDVI, and precipitation of the driest month were among the most misidentified by 62%, 39%, and 32% of SDMs, respectively (Fig. S1). The effect of pseudo-absences was algorithm-dependent and not consistent within model classes, except envelope methods. Pseudo-absences did not affect predictor identification for 12 of 25 algorithms. We observed that eight algorithms improved identification with equal numbers of pseudo-absences, four for 10,000 pseudo-absences, and one improved at double pseudo-absences (Fig. 2). Overall, original generalized additive models with 10,000 pseudo-absences provided the best identification. Generalized linear models from `biomod2` with balanced datasets, and ensemble methods excluding committee average across pseudo-absences, identified all ecologically important predictors but less consistently. Poor identification was exhibited by envelope and machine learning methods (Fig. S3a).

3.2. Response curve estimation

Estimated response curves determined appropriate positive or negative trends (ρ: 0.71 ± 0.17; mean ± SD) but corresponding probabilities were miscalculated (RMSE: 0.42 ± 0.15; Fig. 3). Across ecologically important predictors, only 1% of response trends were poorly estimated ($\rho \leq 0$). On the other hand, only 1% of responses closely approximated the appropriate response (RMSE < 0.1), and 28% exhibited high miscalculation (RMSE ≥ 0.5) (Table S7). Altogether, only 0.1% of SDMs resulted in poor trend and response estimation, particularly by artificial neural networks and committee average ensembles.

Estimated response patterns varied between ecologically important predictors and based on the platform used (Fig. 3). Annual temperature response estimates did not capture the expected β-function but indicated a monotonic increase or a negative unimodal response which over-predicted at freezing temperatures by `biomod2` and original algorithms, respectively (Fig. 3a). Conversely, precipitation response estimates captured the expected logistic curve but were accurate only at minimal and maximal thresholds. Intermediate precipitation values generally over-estimated corresponding probability, specifically by `biomod2` algorithms (Fig. 3b). Responses of annual temperature and precipitation of the wettest month were further limited by the loss of conditions greater than 27 °C and 700 mm, respectively, causing truncation (Fig. 3a, b). Lastly, elevation exhibited the most consistent estimates of approximately linear monotonic decrease though trends varied (Fig. 3c). Original algorithms under predicted occurrence probability at low to
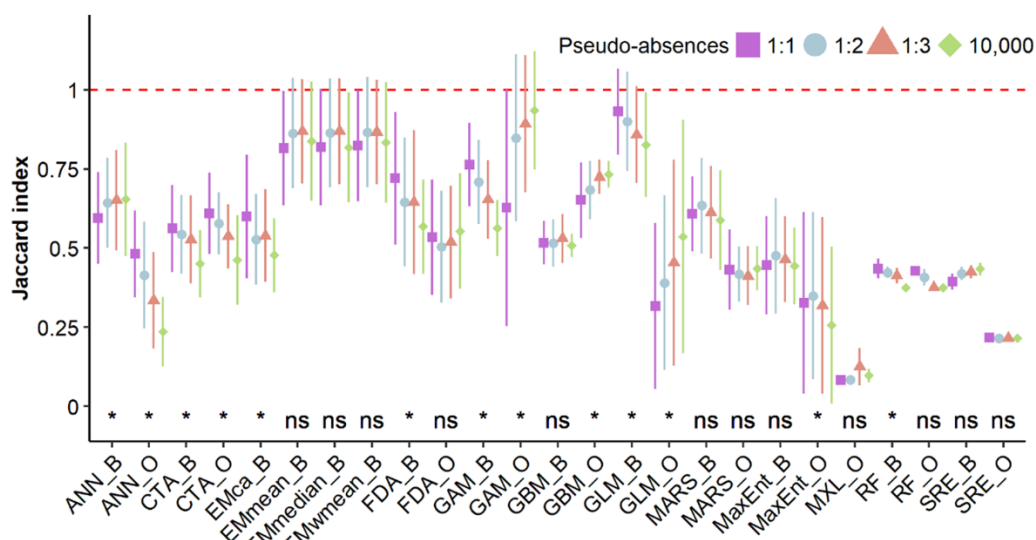
Fig. 2. Mean Jaccard index ± standard deviation per algorithm and pseudo-absences. Algorithm-specific effect of pseudo-absences are shown along x-axis indicating if algorithms observed an effect (*) or not (ns) among pseudo-absences per algorithm. Dashed line indicates perfect identification of ecologically important predictors. Algorithm abbreviations are provided in Table 2.
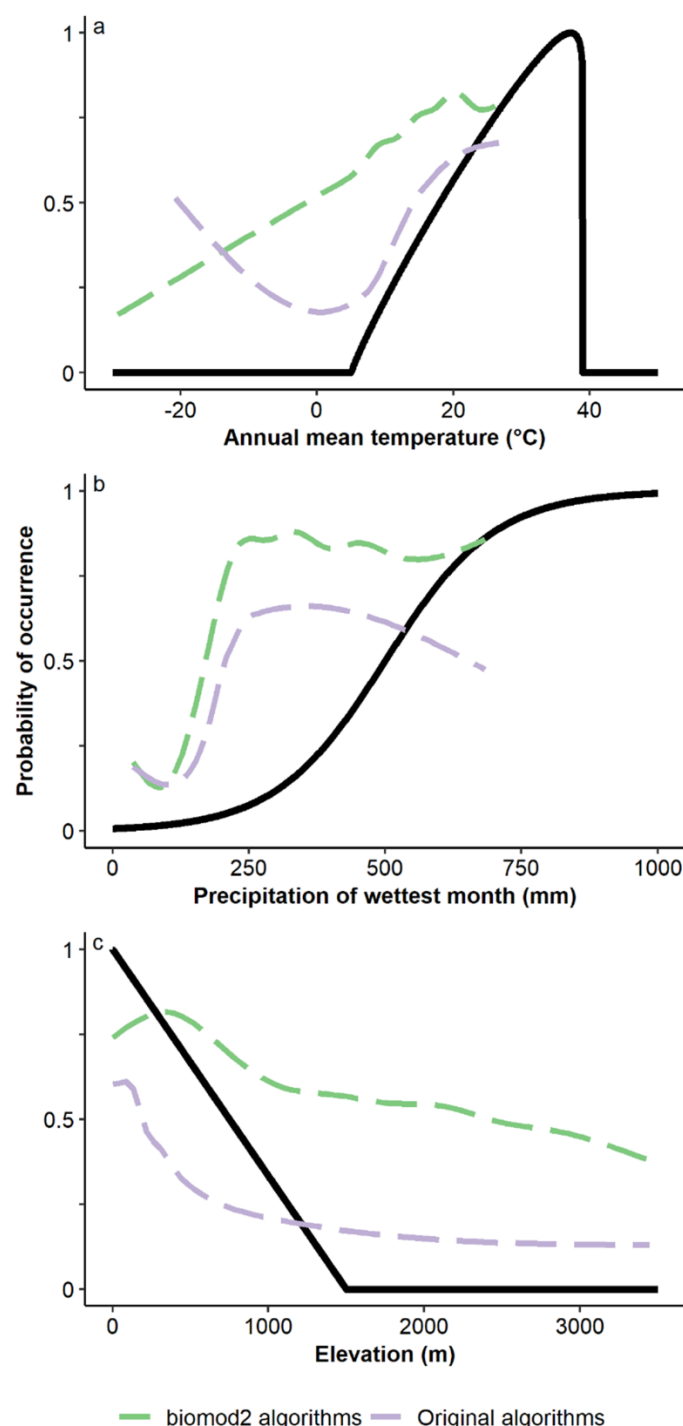
Fig. 3. Algorithm mean responses of annual mean temperature (a), precipitation of the wettest month (b), and elevation (c). Response curves shown represent generalized additive smoothing. The expected response is shown by the solid black line. Accuracies of independent models are available in Supplementary material: Table S7.

moderate elevations and over probability at high elevations. Algorithms from `biomod2` only slightly under predicted occurrence at low elevation, while over predicting occurrence at moderate to higher elevations (Table S7).

Relative performance evaluations indicated that 72% of algorithms required training with only the ecologically important predictors and triple or 10,000 pseudo-absences to improve response estimates (Fig. S3b). However, five algorithms achieved the highest response

estimation when considered with equal centroids and pseudo-absences and automated predictor selection. These five algorithms included the overall best responses estimated by original generalized boosting regression methods with balanced training and automated predictor selection. Excellent response estimation was also exhibited by generalized linear models from `biomod2` with automated and 10,000 pseudo-absences (Table 3). The most inaccurate responses were estimated by envelope and committee average ensembles models (Table S7; Fig. S3b).

3.3. Probability of occurrence

Centroids generally reflected expected presence or absences but introduced high amounts of error into occurrence probability predictions owing to an inability to differentiate probabilities (Fig. 4). Most predictions exhibited little variation with higher-than-expected proportions of low probable areas, supported by training and testing resolutions of 0.04±0.01 and 0.02±0.01 (mean ± SD), respectively (Fig. 4c, e, Fig. S2). Alternatively, areas of high probability were overpredicted across most of North America, suggesting low probability only within non-coastal areas of western U.S. and Canada (Fig. 4 g). Low variability of probabilities provided good calibration and skill. Probabilities were consistent (CBI>0) with centroids and testing occurrences for 88% and 82% of SDMs, respectively (Table S7). Further, 58% and 62% of predicted probabilities reflected expected presence or absences in the training and testing region, respectively (SS>0). Lower skill compared to calibration resulted from overprediction of occurrence in expected low probability areas (Fig. 4a, g). Observed probability trends resulted in the overall bias of 0.19±0.09 and 0.22±0.10 (mean ± SD) in training and testing regions, respectively. This observed error was comparable to that indicated by forecast accuracy and refinement (Table S7). Overall, probability predictions were marginally useful at best (Fig. S2). Probability predictions were overfit to the training data with minimal differences exhibiting improved calibration, less error, and improved separation of values in the training region. However, the testing region indicated more skill owing to a higher proportion of species absence (Table S7).

Probability of occurrences was best estimated from centroids with equal or double pseudo-absences by 76% of algorithms (Fig. S4a). Predictor selection required to account for centroids was less consistent, with 40%, 36%, and 24% of algorithms requiring automated, *a priori*, and expert predictors, respectively. Across all SDMs, occurrence probability was best estimated by original generalized regression boosting methods with equal pseudo-absences and automated predictors (Table 3). Excellent occurrence probabilities were also generated by mean and weighted mean ensembles with double pseudo-absences and automated predictors. The least reliable probabilities were provided by neural networks and envelope models (Fig. S4a).

Table 3
Recommended methodology per objective. Overall objectives represent the relative performance for all corresponding objectives. Full relative performances are available in Supplementary material: Table S7.

| Objective	Algorithm	Pseudo-absences	Predictor selection
Explanation			
Identification of ecologically important predictors	GAM_O	10,000	–
Response curve estimation	GBM_O	1:1	Automated
Overall	GLM_B	1:1	Automated or expert
Prediction			
Probability of occurrence	GBM_O	1:1	Automated
Presence-absence map	EMwmean_B	1:1	Automated
Overall	GBM_O	1:1	Automated
Explanation and prediction			
Overall	GBM_O	1:1	Automated
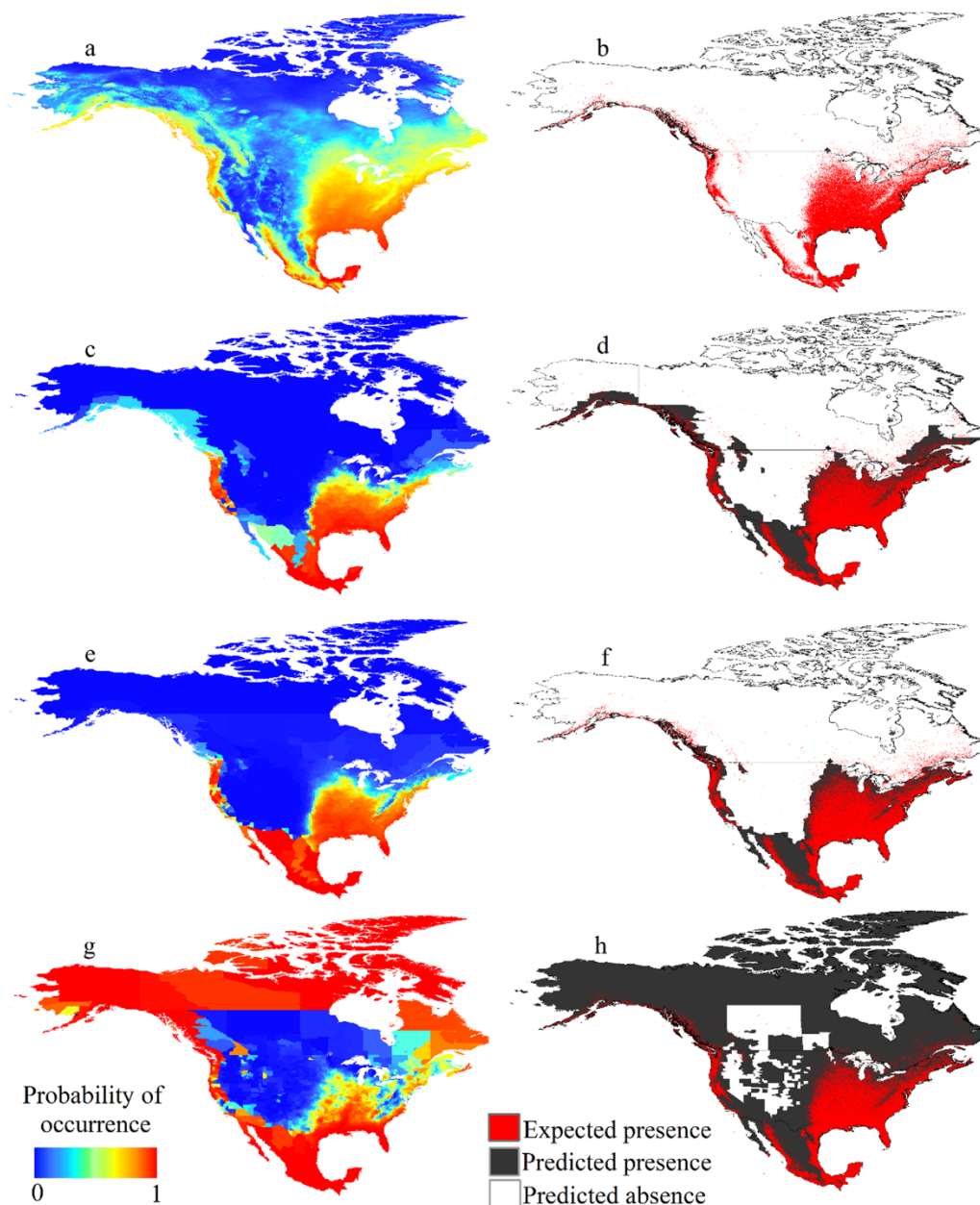
Fig. 4. Predicted probability of occurrence and binary presence-absence maps compared to expected. Maps correspond to expected probability of occurrence (a), true presence-absence (b). An example of excellent relative performance by GBM_O at 1:1 with automated predictors' predicted probability of occurrence(c) and binary prediction (d). An example of an average relative performance by GLM_B with 10,000 pseudo-absences and *a priori* predictors' probability of occurrence (e) and binary prediction (f). An example of poor relative performance predictions by ANN_O with 10,000 pseudo-absences and automated predictors' probability of occurrence (g) and binary prediction (h). See Table 2 for algorithm abbreviations.

3.4. Presence-absence map

Discrimination of presence-absence from centroids exhibited over- and under-prediction of species' presence, depending on region and model building considerations (Fig. 4b, d, f, h). For example, boundaries with a high-density of species presence were predominantly correctly classified but also overpredicted within Mexico and along the U.S.-Mexico border (Precision: 0.5 ± 0.22; F1: 0.62±0.18; mean ± SD, Fig. 4d, F). Extending classification outside of the training area observed either low detection of species presence or high overprediction (Precision: 0.19±0.13; F1: 0.21±0.1; mean ± SD, Fig. 4f, h). Boundaries with a low density of species presence were under-predicted and misclassified as absent within both regions (Fig. 4d, F) unless presence was vastly over-predicted across arctic and coastal boundaries (Fig. 4h). As a result, extrapolation of centroid-trained SDMs generally observed improved classification of species absence (Specificity: 0.69±0.14; mean ± SD) over presence (Sensitivity: 0.55±0.12; mean ± SD). Overall SDMs could only provide moderately accurate discriminations of presence or absence within both training and testing regions (CCR: 0.68±0.15 and 0.69±0.14, respectively, mean ± SD). Further, 86% of SDMs provided better than random discrimination. Accordingly, minimal difference between regions indicated SDMs were overfitted to the centroids, with lower discrimination of presences in the testing region compared to absences (Table S7).

Discrimination from centroids improved when considering equal or double pseudo-absences to centroids for 87% of algorithms (Fig. S4b). Prediction selection considerations were less consistent, with 37%, 34%, and 29% of algorithms requiring automated, *a priori*, and expert predictors, respectively. The relative performance indicated the most reliable discrimination was obtained with weighted mean ensembles with equal pseudo-absences to centroids, and automated predictors (Table 3). Non-weighted mean ensembles and original generalized additive models under the same model building considerations also provided excellent discrimination. Poorest discrimination ability was provided by envelope, committee average ensembles, and neural networks models (Fig. S4b).

3.5. Explanation and prediction

The ability of centroids to reliably provide explanations was best achieved by generalized linear models from `biomod2` with equal pseudo-absences to centroids to identify predictors and then estimate response curves. Envelope and machine learning methods tended to provide the least reliable explanations (Fig. S3c). Overall predictive ability improved relative to other methodologies when boosting or non-committee average ensemble methods with automated predictors and equal pseudo-absences were considered. Neural networks and envelope models provided the poorest predictions (Fig. S4c).

Overall, optimization of all SDM objectives from centroids required automated predictor selection regardless of pseudo-absences for 44% of algorithms (Fig. 5). Alternatively, application of expert predictors or *a priori* with triple or fewer pseudo-absences to centroids was required for 38% and 10% of algorithms, respectively (Fig. S5). Relative performance across all metrics indicated that original generalized boosting regression methods followed by mean, median, or weighted mean ensembles with equal centroids and pseudo-absences and automated

Fig. 5. Overall relative performance between all considered algorithm, pseudo-absences, and predictor selection considering niche explanation and prediction accuracy. A breakdown of relative performance is available in Supplementary materials: Figs. S3–5. See Table 2 for SDM abbreviations.

predictor selection provided the most reliable conclusions to explain and predict a distribution (Fig. 5). Meanwhile, remaining regression methods provided excellent or good relative performance. Machine learning methods exhibited average to poor relative performance. Finally, poor performance was provided by both envelope methods and committee average ensembles (Fig. S5).

4. Discussion

SDMs are the most common method to aid management of specific species. Models can be built from different algorithms, predictors, and quantities of response variables, whose conclusions vary in reliability depending on the objective and precision of the response variable. In this study, we observed that it is possible to determine appropriate conclusions from centroids if one constructs an SDM carefully considering the methods used, particularly algorithms. Specifically, generalized boosting regression methods (GBMs) followed by mean or weighted mean ensembles provided the most accurate conclusions across objectives. Previous research highlighted GBMs as high-performing (Elith et al., 2006; Wisz et al., 2008; Heikkinen et al., 2012) or moderately-performing (Aguirre-Gutiérrez et al., 2013; Breiner et al., 2018) choices. Their strength comes from using an iterative mean ensemble of the boosting and regression-tree algorithms to emphasize previously misidentified training responses (Elith et al., 2008; Shirley et al., 2013). Consequently, GBMs are adept at interpreting non-linear responses, removing of non-contributing predictors, fitting multiple interactions, accounting for interactions, outliers, and collinearity, and analyzing and interpreting complex responses (Elith et al., 2008; Yu et al., 2020). Notably, GBMs can interpret imprecise occurrences (Graham et al., 2008; Naimi et al., 2011; Bombi and D'Amen, 2012) and are less affected by coarsening of scale (Aguirre-Gutiérrez et al., 2013). Though GBMs have been criticized for their tendency to overfit and produce an unreasonable probability of occurrences (Becker et al., 2020), we did not observe that here (Table S7).

Similarly, mean or weighted mean ensembles also tended to provide high relative performance across all objectives (Fig. 5). These methods could predict appropriate presence-absence maps while also identifying ecologically important predictors and maintaining excellent probability of occurrence predictions. Ensembles benefit from considering multiple algorithms to highlight areas of agreement. The exceptional predictive ability of ensembles stems from presence predictions that are strictly limited to cells for which the majority of SDMs agree (Aguirre-Gutiérrez et al., 2013). Additionally, as all SDMs were able to identify at least two of three ecologically important predictors, ensembles were able to assign lower contributions to non- ecologically important predictors. Ensemble methods have risen in popularity to account for variation among algorithms (Araújo and New, 2007). Hao et al. (2019) compared ensemble to singular methods and found the former were the best or nearly best-performing in most studies. Though ensemble methods can be improved with more careful fine-tuning and consideration of algorithms included as opposed to a sensitivity threshold applied here.

The recommended use of GBMs or ensembles contrasts previous SDM mosquito comparisons, which suggested that MaxEnt, generalized linear models (GLMs), or random forests provided the highest performance in the Bermuda Islands and global predictions (Khatchikian et al., 2011; Ding et al., 2018). However, Khatchikian et al. (2011) evaluated against precise occurrences and did not consider GBMs or ensembles. Meanwhile, Ding et al. (2018) considered GBMs but evaluated SDMs without independent data. More mosquito distribution publications to date have relied on MaxEnt owing to its perceived flexibility and high performance (Merow et al., 2013). However, we observed only average relative performance of MaxEnt overall (Fig. 5). MaxEnt estimates the probability of occurrence by determining the stable equilibrium state of parameters with the highest entropy (Phillips et al., 2006; Phillips and Dudík, 2008; Booth et al., 2014). MaxEnt's consideration of centroids reduced environmental heterogeneity and thus incorrectly indicated
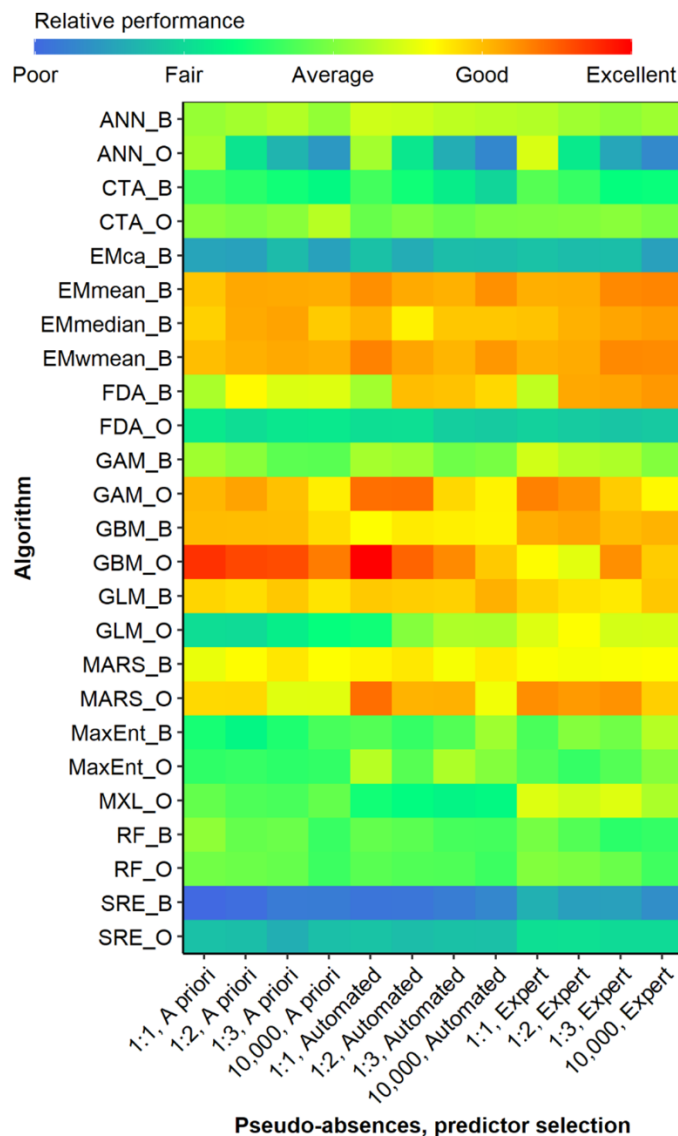
stable equilibrium from MAUP and degraded performance (Bombi and D'Amen, 2012). As a result, the performance of MaxEnt was reduced at coarser scales, while the GBMs were not. GBMs have consistently indicated comparable or better performance than MaxEnt with coarse scales (Guisan et al., 2007b; Graham et al., 2008). Bombi and D'Amen (2012) indicated no hindrance to GBM predictive performance until a 24-fold change from the suspected true ecological scale. Guisan et al. (2007a) demonstrated MaxEnt performance decreased compared to GBM at a minimal 10-fold scale increase. Further, other machine learning SDMs are generally considered to provide more reliable results (Elith et al., 2006; Lawler et al., 2006). The assumption of machine learning superiority may be driven by interpolative evaluation, instead of the full diagnostic evaluation conducted here. Machine learning methods may require more attention to fine-tuning hyper-parameterizations to ensure reliable conclusions than was considered here (Araújo et al., 2019). This suggests that despite MaxEnt and other machine learning methods' hold on the SDM literature, its use in applications where coarse scales and MAUP are of concern is ill-advised without fine-tuning and independent evaluation.

Further model building considerations indicated algorithm- and objective-specific effects, with no pattern observed across algorithms, which is consistent with previous studies (Synes and Osborne, 2011; Barbet-Massin et al., 2012; Heikkinen et al., 2012). These results further support the need to determine SDM methodology for the available data and objective, instead of on commonly applied algorithms (Aguirre-Gutiérrez et al., 2013).

4.1. Identification of ecologically important predictors

Overall, predictor identification was poor. These results corroborate the findings of Inman et al. (2021). They tested MaxEnt's ability to identify the ecologically important predictors and observed correct identification by only 3% of SDMs with aggregation corrected bias (Inman et al., 2021). The scale changing inhibited predictor identification as drivers of distribution change with scale (Hortal et al., 2010). The difference in scale alters the spatial autocorrelation and heterogeneity within each predictor. Accordingly, SDMs appeared to only select predictors with higher spatial autocorrelation or heterogeneity (Fig. S6). Increased heterogeneity and spatial autocorrelation inflate predictor importance, particularly when a distribution is driven by multiple predictors (Connor et al., 2018; Smith and Santos, 2020). This suggests ecologically important predictors identified by SDMs trained with centroids can approximate appropriate ecologically important predictors by indicating predictors demonstrating spatial autocorrelation and heterogeneity at the provided scale. These findings may be limited to the predictor identification method applied. Harisena et al. (2021) detailed the predictor importance assessment used here, which is sensitive to spatial autocorrelation, pseudo-replication, and truncated responses. Another predictor importance assessment or selection threshold may improve predictor identification with centroids (Synes and Osborne, 2011; Harisena et al., 2021).

Consistent predictor identification required generalized additive models (GAMs) with 10,000 pseudo-absences. Predictor identification was driven more by the algorithm than the pseudo-absence level. Smith and Santos (2020) evaluated the effects of sample size, scale, and collinearity on GAMs and MaxEnt to determine predictor contribution. They concluded that small sample sizes and coarse scales inhibited predictor identification, but the algorithm was the greater indicator of identification success. Specifically, large sample sizes provided the best identification, but MaxEnt was more affected by coarse scales (Smith and Santos, 2020). Therefore, our findings partially agree with Smith and Santos (2020). The smallest sample size considered here was larger than Smith and Santos (2020)'s largest, thus potentially explaining the variation in the pseudo-absence number effect. This suggests that though larger sample sizes are better, some algorithms have upper and lower sample size thresholds for predictor identification (Fig. 2).

Further, Aguirre-Gutiérrez et al. (2013) indicated only high variability of GAM's variable contribution measures. This is further supported by a decreased predictor selection ability of GAM at lower pseudo-absence levels (Fig. 2). Therefore, GAM's predictor identification may not be reflected when fewer centroids are available (Smith and Santos, 2020).

4.2. Response curve estimation

Aggregation caused truncated species response curves, thus increasing niche width only capturing general trends. This finding supports previous observations that mismatched scales inhibit species response detection (Azaele et al., 2012; Araújo and Rozenfeld, 2014). Change in scale is known to alter the shape of responses from linear to asymmetric or skewed curves owing to MAUP (Rydgren et al., 2003; Lechner et al., 2012). As the scale coarsens, heterogeneity is increasingly masked, thus limited the estimated response (Wiens, 1989). However, regardless of response imprecision, Guisan and Zimmermann (2000) described that SDMs are not expected to provide realistic responses, nor to inform about their underlying mechanisms. Instead, one should consider if the response curve trend was biologically possible and not overly complex (Jarnevich et al., 2015). Accordingly, correlative SDMs should only be expected to capture the general positive or negative response validated against physiological studies. Here, SDMs captured the expected positive or negative trends, indicating appropriate realism for an SDM. However, overestimating niche breadth suggests that centroids are unsuitable for transfering beyond the training extent (Thuiller et al., 2004; Manzoor et al., 2018). Future studies may improve niche estimation by considering accounting for uncertainty within predictors, such as standard error or considering additional aggregates (i.e. min, max, median), considering pseudo-absences from a more restrictive range, or only investigating univariate model responses (Thuiller et al., 2004; Winters et al., 2008; Santika and Hutchinson, 2009; Stoklosa et al., 2015).

Across SDMs, we identified GLMs with expert or successful automated identification and equal pseudo-absences to centroids provided the best explanation. This finding is contrary to previous studies, which suggested that GLMs were inferior to other algorithms such as GAM or GBM for response estimation (Santika and Hutchinson, 2009). This inconsistency may be due to previous GLMs considering only linear responses while we considered a higher-order approach. Higher-order approaches are recommended to provide increased flexibility to fit complex responses (Segurado et al., 2006; Dormann et al., 2007). Further, explanation was best estimated for most algorithms when considering expert predictors and high numbers of pseudo-absences. The requirement of expert predictors for proper explanation re-enforces the recommendation building SDMs with only predictors that are well supported (Araújo et al., 2019).

4.3. Probability of occurrence

Predictions of virtual species generally improved when considered with equal pseudo-absences and automated predictor selection. Coarse scales increase the probability of false absences. Thus more pseudo-absences bias SDMs to consider more potentially suitable habitats as unsuitable (McPherson et al., 2006). In this study, the 1414 centroids limited pseudo-absences to a maximum of 1761 unique environmental conditions. Therefore, consideration of equal pseudo-absences to centroids provided the least bias training relative to other pseudo-absence levels investigated, supported by previous studies (Moffett et al., 2007; Phillips and Dudík, 2008; Liu et al., 2018). For example, Johnson et al. (2017) indicated that discrimination of *A. albopictus* centroids improved when considering equal background points to centroids. However, *A. aegypti,* which had fewer occurrence records, improved with double pseudo-absences (Johnson et al., 2017). This suggests the number of pseudo-absences required varies based on the number of responses available and unique environmental conditions.

Automated selection of predictors generally allowed for more accurate spatial conclusions. Accounting for the imprecision of centroids fundamentally alters how the responses and resulting distributions are interpreted (Levin, 1992). While the effect of one factor may be prominent at a fine scale, its effect may be negligible at another (Schweiger and Beierkuhnlein, 2016). Hortal et al. (2010) described how as scale coarsens, the impact of biotic factors on insect distribution decreases, and abiotic factors' effects increase. Mouton et al. (2009) suggested that data-driven SDMs, such as automated predictors, outperformed expert models as predictors were selected based on the available scale, thus providing a more reliable interpretation. Therefore, automated predictor selection estimates ecologically important predictors at the provided scale to better represent the responses at the observed over expected responses.

The overall probability of occurrence predictions was only moderately accurate with low generality suggesting marginal usefulness in application. Some studies suggest that lower degrees of scale coarsening preserved environmental characteristics (Guisan et al., 2007b; Trivedi et al., 2008; Bombi and D'Amen, 2012), but generally, scale coarsening decreased SDM predictive performance (Rahbek and Graves, 2001; Thompson and McGarigal, 2002; Guisan et al., 2007a; Seo et al., 2009; Mertes and Jetz, 2018). The degree of scale coarsening observed in this study (Table S2) did not reflect those of the lower degree previously described (≤ 10 km^2) and therefore supported decreased performance at coarse scales. Our observed performance reflected a greater decrease in predictive ability than in previous centroid applications. Johnson et al. (2017) applied centroids of *A. aegypti* and *A. albopictus* across the contiguous U.S. counties and observed good area under the receiver operating characteristic curve scores but without independent validation. Similarly, Collins et al. (2017) compared SDMs trained with precise or centroids occurrences to investigate the degree of bias introduced for butterfly, dragonfly, and damselfly species in the contiguous U.S. counties. Niche similarity metrics indicated centroids only somewhat compromised predictions, with the effect more pronounced in larger and environmentally heterogenous counties which could not be accounted for by boundary scale (Collins et al., 2017). One possible explanation for decreased degradation in our study is that all previous studies investigated administrative regions within a single country. In contrast, we applied centroid occurrences across administrative regions of three countries. Consequently, across-country considerations resulted in a more considerable variation in size and shape of boundaries, thereby leading to greater uncertainty and instability of predictors by MAUP (Openshaw, 1984). Specifically, larger and more heterogeneous administrative regions introduced greater predictor uncertainty, namely boundaries in the western U.S. or Canada (Collins et al., 2017; Cheng et al., 2021).

4.4. Presence-absence map

Prediction of species' presence or absence tended to over-predict presence with some cases of under-prediction, depending on the methodology. Methodology considerations for species presence or absence predictions followed the patterns described previously. Presence-absence maps require further review of the threshold considered. Threshold selection is one of many possible sources of bias in SDMs (Bean et al., 2012). In this study, we applied a minimum presence threshold (Fig. 4b). Yet, in practice, threshold selection should be considered relative to the desired SDM application and the importance of omission and commission errors (Liu et al., 2013). Alternatively, thresholds derived from the maximization of the sum between sensitivity and specificity or minimization of the difference between sensitivity and specificity are generally seen as superior to others, or authors may examine an overlay of different thresholds to determine suitable areas (Liu et al., 2005; Jiménez-Valverde and Lobo, 2007). To our knowledge, no studies have investigated the effect of binary threshold choice associated with MAUP or administrative/geopolitical boundary centroids.

4.5. Implications

Management of mosquito populations often focuses on chemical or biological agents to limit of reproduction and spread. Yet, these management practices require fine-scale and precise predictions to be effective, which are not captured using centroids (Fouet and Kamdem, 2019; Pascoe et al., 2019). Our results indicated the prediction of only administrative regions at the highest risk of establishment while misclassifying lower density of true occurrence sites. This suggested that with very careful consideration of SDM methods, centroids can be used to predict boundaries at the highest risk of arthropod vector establishment, if sufficient propagules are introduced in new boundaries. These corresponding boundaries would be encouraged to enact public education campaigns and increase vector surveillance. However, SDMs consistently missed lower risk areas which may indicate early detection. This highlighted the loss of environmental detail with the use of centroids. Therefore, the corresponding SDMs are unsuitable for estimating regions with isolated risks such as micro-niches. Further investigations on a local scale or with more precise requirements would be required to provide management applications. Additional or alternative methods may accomplish this may be necessary to provide reliable conclusions from centroids by applying mixed-effect models (Hamil et al., 2016), Bayesian (Velásquez-Tibatá et al., 2016), integration methods (Collins et al., 2017; Pacifici et al., 2019) or considering movement and biotic factors in addition to abiotic (Soberón and Peterson, 2005).

Centroid responses and resulting ecologically insignificant scales in SDMs are not limited to *A. aegypti* and *A. albopictus*. Approximately half of the occurrences from species of all major taxonomic groups are limited to centroids in occurrence repositories in the U.S. (Park and Davis, 2017). The results of this study provide guidance to improve centroid application in SDMs, but not conclusive guidelines. Virtual species ensure SDM assumptions are met, including that occurrence records are precise, entire geographic and environmental extent are sampled without bias, species are at equilibrium, and they provide a known truth for evaluation (Guisan and Zimmermann, 2000).

Conversely, using a virtual species may bias the results, as it gives a simplified distribution based on limited environmental conditions without movement or biotic impacts. The results here reflect the ability of SDMs to determine appropriate ecologically important predictors, estimate responses, predict probability and presence only if the species responses are based on a combination of β, logistic, and linear responses. We selected these responses to ensure ecological realism given ecological and physiological studies of *A. aegypti* and *A. albopictus*. Considering linear and non-linear responses limited potentially bias towards a particular analysis (Hirzel et al., 2001). Other responses may not observe the same results, thus re-enforcing the need to test models for the given data and objective. For example, many regression and tree-based SDMs poorly model linear responses, as the logit-link and grouping strategies are non-linear and represent a threshold response (Meynard and Quinn, 2007).

Additionally, we defined the virtual species from a multiplicative formula. Yet, we trained SDMs with additive formulas. This allowed for the interaction of environmental conditions on species occurrence while maintaining common practice in SDMs, as shown by the default settings of an additive formula for most SDMs (Elith et al., 2006; Meynard and Quinn, 2007). Meynard and Quinn (2007) examined SDM discrimination ability of virtual species determined by additive and multiplicative habitat suitability indexes. They found that additive and multiplicative training followed similar performance patterns but varied depending on the evaluation metric. Future SDM investigations of centroids may contrast our findings if trained with multiplicative rather than additive formulas. Therefore, one major limitation was considering a single virtual species across SDMs, as additional virtual species would allow more variation in the known habitat suitability. Additionally, it is important

to note we examined all SDMs under default settings. Default settings may have improved the performance of one SDM over another. For example, the default interaction depth of GBMs and multiple adaptive regression splines have a default interaction depth of seven and zero, respectively. Further, default settings allowed GBMs to have greater maximum number of trees and minimal number of observations per node than random forest and classification tree (Table S3). These would allow GBMs to resemble the original multiplicative suitability index more closely, as well as a greater ability to determine the differentiate between environmental condition importance, presence, and absence. If all algorithms were fine-tuned to the provided data, we may have observed different results (Merow et al., 2013; Araújo et al., 2019). Future work is required to determine if our results hold across taxonomic groups.

CRediT authorship contribution statement

Justin R Barker: conceived the study, collected, analyzed the data, and drafted the manuscript. **Hugh J. MacIsaac**: conceived the study, revised and approved the manuscript.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2022.110107.

References

Aguirre-Gutiérrez, J., Carvalheiro, L.G., Polce, C., van Loon, E.E., Raes, N., Reemer, M., Biesmeijer, J.C., 2013. Fit-for-purpose: species distribution model performance depends on evaluation criteria – Dutch Hoverflies as a case study. PLoS ONE 8, e63708. https://doi.org/10.1371/journal.pone.0063708.

Alahmed, A.M., Naeem, M., Kheir, S.M., Sallam, M.F., 2015. Ecological distribution modeling of two malaria mosquito vectors using geographical information system in Al-Baha Province, Kingdom of Saudi Arabia. Pak. J. Zool. 47, 1797–1806.

Alaniz, A.J., Bacigalupo, A., Cattan, P.E., 2017. Spatial quantification of the world population potentially exposed to Zika virus. Int. J. Epidemiol. 46, 966–975. https://doi.org/10.1093/ije/dyw366.

Allen, J., Stockli, R., 2018. Vegetation Index (1 month - Terra/MODIS) | Vegetation Index (1 month - Terra/MODIS) | NASA. https://neo.gsfc.nasa.gov/view.php?datasetId=MOD_NDVI_M (accessed July 26 2018).

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

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22, 42–47. https://doi.org/10.1016/j.tree.2006.09.010.

Araújo, M.B., Rozenfeld, A., 2014. The geographic scaling of biotic interactions. Ecography 37, 406–415. https://doi.org/10.1111/j.1600-0587.2013.00643.x.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Model. 157, 101–118. https://doi.org/10.1016/S0304-3800(02)00205-3.

Azaele, S., Cornell, S.J., Kunin, W.E., 2012. Downscaling species occupancy from coarse spatial scales. Ecol. Appl. 22, 1004–1014.

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many?: How to use pseudo-absences in niche modelling? Methods Ecol. Evol. 3, 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol. Model. 222, 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011.

Bean, W.T., Stafford, R., Brashares, J.S., 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35, 250–258. https://doi.org/10.1111/j.1600-0587.2011.06545.x.

Becker, E.A., Carretta, J., Forney, K.A., Barlow, J., Brodie, S., Hoopes, R., Jacox, M.G., Maxwell, S.M., Redfern, J., Sisson, N.B., Welch, H., Hazen, E.L., 2020. Performance evaluation of cetacean species distribution models developed using generalized additive models and boosted regression trees. Ecol. Evol. 10, 5759–5784. https://doi.org/10.1002/ece3.6316.

Becker, N., Petric, D., Zgomba, M., Boase, C., Madon, M., Dahl, C., Kaiser, A., 2010. Mosquitoes and their control. Springer, Berlin Heidelberg.

Bemmels, J., 2018. ENVIREM: An expanded set of bioclimatic and topographical variables increases flexibility and improves performance of ecological niche modeling. Ecography 41, 291–307.

Bombi, P., D'Amen, M., 2012. Scaling down distribution maps from atlas data: a test of different approaches with virtual species. J. Biogeogr. 39, 640–651. https://doi.org/10.1111/j.1365-2699.2011.02627.x.

Booth, T.H., Nix, H.A., Busby, J.R., Hutchinson, M.F., 2014. bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. Divers. Distrib. 20, 1–9. https://doi.org/10.1111/ddi.12144.

Boria, R.A., Blois, J.L., 2018. The effect of large sample sizes on ecological niche models: Analysis using a North American rodent, Peromyscus maniculatus. Ecol. Model. 386, 83–88. https://doi.org/10.1016/j.ecolmodel.2018.08.013.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. Ecol. Model. 157, 281–300. https://doi.org/10.1016/S0304-3800(02)00200-4.

Bradie, J., Leung, B., 2017. A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. J. Biogeogr. 44, 1344–1361. https://doi.org/10.1111/jbi.12894.

Brady, O.J., Johansson, M.A., Guerra, C.A., Bhatt, S., Golding, N., Pigott, D.M., Delatte, H., Grech, M.G., Leisnham, P.T., Maciel-de-Freitas, R., Styer, L.M., Smith, D.L., Scott, T.W., Gething, P.W., Hay, S.I., 2013. Modelling adult Aedes aegypti and Aedes albopictus survival at different temperatures in laboratory and field settings. Parasi. Vectors. 6, 351. https://doi.org/10.1186/1756-3305-6-351.

Braks, M.A.H., Honório, N.A., Lounibos, L.P., Lourenço-De-Oliveira, R., Juliano, S.A., 2004. Interspecific competition between two invasive species of container mosquitoes, Aedes aegypti and Ae. albopictus (Diptera: Culicidae), in Brazil. Ann. Entomol. Soc. Am. 97, 130–139. https://doi.org/10.1603/0013-8746(2004)097[0130:ICBTIS]2.0.CO;2.

Breiner, F.T., Nobis, M.P., Bergamini, A., Guisan, A., 2018. Optimizing ensembles of small models for predicting the distribution of species with few occurrences. Methods Ecol. Evol. 9, 802–808. https://doi.org/10.1111/2041-210X.12957.

Brier, G., 1950. Verification of forecasts expressed in terms of probability. Mon. Weather. Rev. 78, 1–3. https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Broennimann, O., Di Cola, V., Guisan, A., 2020. ecospat: Spatial ecology miscellaneous methods. Ecography 40, 774–787. https://doi.org/10.1111/ecog.02671.

Brown, R.D., Brasnett, B., 2010. Canadian meteorological centre (CMC) daily snow depth analysis data. https://nsidc.org/data/nsidc-0447/versions/1 (accessed July 26 2018).

Cheng, Y., Tjaden, N.B., Jaeschke, A., Thomas, S.M., Beierkuhnlein, C., 2021. Using centroids of spatial units in ecological niche modelling: Effects on model performance in the context of environmental data grain size. Glob. Ecol. Biogeogr. 30, 611–621. https://doi.org/10.1111/geb.13240.

Collins, S.D., Abbott, J.C., McIntyre, N.E., 2017. Quantifying the degree of bias from using county-scale data in species distribution modeling: Can increasing sample size or using county-averaged environmental data reduce distributional overprediction? Ecol. Evol. 7, 6012–6022. https://doi.org/10.1002/ece3.3115.

Connor, T., Hull, V., Viña, A., Shortridge, A., Tang, Y., Zhang, J., Wang, F., Liu, J., 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41, 1270–1282. https://doi.org/10.1111/ecog.03416.

Cunze, S., Koch, L.K., Kochmann, J., Klimpel, S., 2016. Aedes albopictus and Aedes japonicus - two invasive mosquito species with different temperature niches in Europe. Parasit. Vectors. 9, 573. https://doi.org/10.1186/s13071-016-1853-2.

Denlinger, D.L., Armbruster, P.A., 2014. Mosquito diapause. Annu. Rev. Entomol. 59, 73–93. https://doi.org/10.1146/annurev-ento-011613-162023.

Dhimal, M., Gautam, I., Joshi, H.D., O'Hara, R.B., Ahrens, B., Kuch, U., 2015. Risk factors for the presence of chikungunya and dengue vectors (Aedes aegypti and Aedes albopictus), their altitudinal distribution and climatic determinants of their abundance in central Nepal. PLOS Negl. Trop. Dis. 9, e0003545 https://doi.org/10.1371/journal.pntd.0003545.

Ding, F., Fu, J., Jiang, D., Hao, M., Lin, G., 2018. Mapping the spatial distribution of Aedes aegypti and Aedes albopictus. Acta Trop 178, 155–162. https://doi.org/10.1016/j.actatropica.2017.11.020.

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

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kuehn, I., Ohlemueller, R., Peres-

Neto, P.R., Reineking, B., Schroeder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30, 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x.

Duputié, A., Zimmermann, N.E., Chuine, I., 2014. Where are the wild things? Why we need better data on species distribution. Glob. Ecol. Biogeogr. 23, 457–467. https://doi.org/10.1111/geb.12118.

Earth Resources Observation and Science Center, 2017. Global 30 arc-second elevation (GTOPO30). 10.5066/f7df6pqs (accessed July 26 2018).

Eisen, L., Monaghan, A.J., Lozano-Fuentes, S., Steinhoff, D.F., Hayden, M.H., Bieringer, P.E., 2014. The Impact of Temperature on the Bionomics of Aedes (Stegomyia) Aegypti, with Special Reference to the Cool Geographic Range Margins. J. Med. Entomol. 51, 496–516. https://doi.org/10.1603/ME13214.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.M., Peterson, A. T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

Environmental Systems Research Institute (Esri), 2018. ArcGIS desktop: Release 10.8.1. Redlands, CA.

Eritja, R., Palmer, J.R.B., Roiz, D., Sanpera-Calbet, I., Bartumeus, F., 2017. Direct evidence of adult Aedes albopictus dispersal by car. Sci. Rep. 7, 14399. https://doi.org/10.1038/s41598-017-12652-5.

Escobar, L.E., Peterson, A.T., Favi, M., Yung, V., Pons, D.J., Medina-Vogel, G., 2013. Ecology and geography of transmission of two bat-borne rabies lineages in Chile. PLOS Neglect. Trop. Dis. 7, e2577. https://doi.org/10.1371/journal.pntd.0002577.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38–49.

Fitzpatrick, M.C., Gotelli, N.J., Ellison, A.M., 2013. MaxEnt versus MaxLike: empirical comparisons with ant species distributions. Ecosphere 4, art55. https://doi.org/10.1890/ES13-00066.1.

Fouet, C., Kamdem, C., 2019. Integrated mosquito management: Is precision control a luxury or necessity? Trends Parasitol 35, 85–95. https://doi.org/10.1016/j.pt.2018.10.004.

Fox, J., Weisberg, S., 2019. An R companion to applied regression. Sage, Thousand Oaks, CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. https://doi.org/10.1016/j.rse.2009.08.016.

Gama, Z.P., Islamiyah, M., 2013. Distribution patterns and relationship between elevation and the abundance of Aedes aegypti in Mojokerto city 2012. Open J. Anim. Sci. 3, 11.

Giordano, B.V., Gasparotto, A., Liang, P., Nelder, M.P., Russell, C., Hunter, F.F., 2019. Discovery of an Aedes (Stegomyia) albopictus population and first records of Aedes (Stegomyia) aegypti in Canada. Med. Vet. Entomol. 34, 10–16. https://doi.org/10.1111/mve.12408.

Goodchild, M.F., 2011. Scale in GIS: An overview. Geomorphology, Scale Issues in Geomorphology 130, 5–9. https://doi.org/10.1016/j.geomorph.2010.10.004.

Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Peterson, A.T., Loiselle, B.A., 2008. The influence of spatial errors in species occurrence data used in distribution models. J Appl Ecol 45, 239–247. https://doi.org/10.1111/j.1365-2664.2007.01408.x.

Greenwell, B., Cunningham, J., GBM Developers, 2020. gbm: Generalized boosted regression models. (R package version 2.1.8). https://CRAN.R-project.org/package=gbm.

Grimmett, L., Whitsed, R., Horta, A., 2020. Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics. Ecol. Model. 431, 109194 https://doi.org/10.1016/j.ecolmodel.2020.109194.

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R., Wintle, B.A., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Glob. Ecol. Biogeogr. 24, 276–292. https://doi.org/10.1111/geb.12268.

Guisan, A., Graham, C.H., Elith, J., Huettmann, F., 2007a. Sensitivity of predictive species distribution models to change in grain size. Divers. Distrib. 13, 332–340. https://doi.org/10.1111/j.1472-4642.2007.00342.x.

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

Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S., Peterson, A.T., 2007b. What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? Ecol. Monogr. 77, 615–630. https://doi.org/10.1890/06-1060.1.

Hahn, M.B., Eisen, L., McAllister, J., Savage, H.M., Mutebi, J.-P., Eisen, R.J., 2017. Updated reported distribution of Aedes (Stegomyia) aegypti and Aedes (Stegomyia) albopictus (Diptera: Culicidae) in the United States, 1995–2016. J. Med. Entomol. 54, 1420–1424. https://doi.org/10.1093/jme/tjx088.

Hamil, K.-A.D., Iannon, B.V.I., Huang, W.K., Fei, S., Zhang, H., 2016. Cross-scale contradictions in ecological relationships. Landscape Ecol 31, 7–18. https://doi.org/10.1007/s10980-015-0288-z.

Hao, T., Elith, J., Guillera-Arroita, G., Lahoz-Monfort, J.J., 2019. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Divers. Distrib. 25, 839–852. https://doi.org/10.1111/ddi.12892.

Harisena, N.V., Groen, T.A., Toxopeus, A.G., Naimi, B., 2021. When is variable importance estimation in species distribution modelling affected by spatial correlation? Ecography 44, 778–788. https://doi.org/10.1111/ecog.05534.

Hastie, T., 2020. gam: Generalized Additive Models (R package version 1.20.1). https://CRAN.R-project.org/package=gam.

Hastie, T., Tibshirani, R., Leisch, F., Hornik, K., Ripley, B.D., 2020. mda: Mixture and flexible discriminant analysis. (R package version 0.5-2). https://CRAN.R-project.org/package=mda.

Hayden, M.H., Uejio, C.K., Walker, K., Ramberg, F., Moreno, R., Rosales, C., Gameros, M., Mearns, L.O., Zielinski-Gutierrez, E., Janes, C.R., 2010. Microclimate and human factors in the divergent ecology of Aedes aegypti along the Arizona, U.S./Sonora, MX Border. EcoHealth 7, 64–77. https://doi.org/10.1007/s10393-010-0288-z.

Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography 35, 276–288. https://doi.org/10.1111/j.1600-0587.2011.06999.x.

Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2017. dismo: species distribution modeling. (R package version 1.3-5). https://CRAN.R-project.org/package=dismo.

Hirzel, A.H., Helfer, V., Metral, F., 2001. Assessing habitat-suitability models with a virtual species. Ecol. Model. 145, 111–121. https://doi.org/10.1016/S0304-3800(01)00396-9.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecol .Model 199, 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017.

Hortal, J., Roura-Pascual, N., Sanders, N.J., Rahbek, C., 2010. Understanding (insect) species distributions across spatial scales. Ecography 33, 51–53. https://doi.org/10.1111/j.1600-0587.2009.06428.x.

Hsu, W., Murphy, A.H., 1986. The attributes diagram a geometrical framework for assessing the quality of probability forecasts. Int. J. Forecast. 2, 285–293. https://doi.org/10.1016/0169-2070(86)90048-8.

Ibañez-Justicia, A., 2020. Pathways for introduction and dispersal of invasive Aedes mosquito species in Europe: a review. JEMCA 38, 1–10.

Ikegami, M., Jenkins, T.A.R., 2018. 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. Ecol. Manag. 409, 343–352. https://doi.org/10.1016/j.foreco.2017.11.005.

Inman, R., Franklin, J., Esque, T., Nussear, K., 2021. Comparing sample bias correction methods for species distribution modeling using virtual species. Ecosphere 12, e03422. https://doi.org/10.1002/ecs2.3422.

Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44, 223–270.

Jackson, H.B., Fahrig, L., 2015. Are ecologists conducting research at the optimal scale? Glob. Ecol. Biogeogr. 24, 52–63. https://doi.org/10.1111/geb.12233.

Jarnevich, C.S., Reynolds, L.V., 2011. Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. Biol. Invasions. https://doi.org/10.1007/s10530-010-9798-4.

Jarnevich, C.S., Stohlgren, T.J., Kumar, S., Morisette, J.T., Holcombe, T.R., 2015. Caveats for correlative species distribution modeling. Ecol. Inform. 29, 6–15. https://doi.org/10.1016/j.ecoinf.2015.06.007.

Jiménez-Valverde, A., Lobo, J.M., 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. Acta. Oecol. 31, 361–369. https://doi.org/10.1016/j.actao.2007.02.001.

Johnson, C.J., Gillingham, M.P., 2008. Sensitivity of species-distribution models like error, bias, and model design: An application to resource selection functions for woodland caribou. Ecol. Model. 213, 143–155. https://doi.org/10.1016/j.ecolmodel.2007.11.013.

Johnson, T.L., Haque, U., Monaghan, A.J., Eisen, L., Hahn, M.B., Hayden, M.H., Savage, H.M., McAllister, J., Mutebi, J.-P., Eisen, R.J., 2017. Modeling the environmental suitability for Aedes (Stegomyia) aegypti and Aedes (Stegomyia) albopictus (Diptera: Culicidae) in the contiguous United States. J. Med. Entomol. 54, 1605–1614. https://doi.org/10.1093/jme/tjx163.

Josselin, D., Louvet, R., 2019. Impact of the scale on several metrics used in geographical object-based image analysis: Does GEOBIA mitigate the modifiable areal unit problem (MAUP)? ISPRS Int. J. Geo-Inf. 8, 156. https://doi.org/10.3390/ijgi8030156.

Khatchikian, C., Sangermano, F., Kendell, D., Livdahl, T., 2011. Evaluation of species distribution model algorithms for fine-scale container-breeding mosquito risk prediction. Med. Vet. Entomol. 25, 268–275. https://doi.org/10.1111/j.1365-2915.2010.00935.x.

Koch, L.K., Cunze, S., Werblow, A., Kochmann, J., Dörge, D.D., Mehlhorn, H., Klimpel, S., 2016. Modeling the habitat suitability for the arbovirus vector Aedes albopictus (Diptera: Culicidae) in Germany. Parasitol. Res. 115, 957–964. https://doi.org/10.1007/s00436-015-4822-3.

Kuhn, M., 2020. caret: Classification and regression training. (R package version 6.0-91). https://CRAN.R-project.org/package=caret.

Kuznetsova, A., Brockhoff, P., Christensen, R., 2017. lmerTest Package: Tests in linear mixed effects models. J. Stat. Softw. 82, 1–26. https://doi.org/10.18637/JSS.V082.I13.

Larson, S.R., Degroot, J.P., Bartholomay, L.C., Sugumaran, R., 2010. Ecological niche modeling of potential West Nile virus vector mosquito species in Iowa. J. Insect. Sci. 10, 110. https://doi.org/10.1673/031.010.11001.

Lawler, J.J., White, D., Neilson, R.P., Blaustein, A.R., 2006. Predicting climate-induced range shifts: model differences and model reliability. Glob. Change Biol. 12, 1568–1584. https://doi.org/10.1111/j.1365-2486.2006.01191.x.

Lechner, A.M., Langford, W.T., Jones, S.D., Bekessy, S.A., Gordon, A., 2012. Investigating species–environment relationships at multiple scales: Differentiating between intrinsic scale and the modifiable areal unit problem. Ecol. Complex. 11, 91–102. https://doi.org/10.1016/j.ecocom.2012.04.002.

Lecours, V., Devillers, R., Schneider, D.C., Lucieer, V.L., Brown, C.J., Edinger, E.N., 2015. Spatial scale and geographic context in benthic habitat mapping: review and future directions. Mar. Ecol. Prog. Ser. 535, 259–284. https://doi.org/10.3354/meps11378.

Length, R., 2020. emmeans: Estimated marginal means, aka least-squares means. (R package version 1.7.4). https://CRAN.R-project.org/package=emmeans.

Leroy, B., Meynard, C.N., Bellard, C., Courchamp, F., 2016. virtualspecies, an R package to generate virtual species distributions. Ecography 39, 599–607. https://doi.org/10.1111/ecog.01388.

Leta, S., Beyene, T.J., De Clercq, E.M., Amenu, K., Kraemer, M.U.G., Revie, C.W., 2018. Global risk mapping for major diseases transmitted by Aedes aegypti and Aedes albopictus. Int. J. Infect. Dis. 67, 25–35. https://doi.org/10.1016/j.ijid.2017.11.026.

Levin, S.A., 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. Ecology 73, 1943–1967. https://doi.org/10.2307/1941447.

Liaw, A., Wiener, M., 2001. Classification and Regression by randomForest. R News 2, 18–22. https://CRAN.R-project.org/doc/Rnews/.

Liaw, Y., Weis, G., Wait, K., Graham, E.M., Woolford, W., Pyke, R., 2019. BCCVL/org.bccvl.compute.Githubrepository.

Lima, A., Lovin, D.D., Hickner, P.V., Severson, D.W., 2016. Evidence for an overwintering population of Aedes aegypti in Capitol Hill neighborhood, Washington, DC. Am. J. Trop. Med. Hyg. 94, 231–235. https://doi.org/10.4269/ajtmh.15-0351.

Liu, C., Berry, P., M., Dawson, T.P., Pearson, R., G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28, 385–393.

Liu, C., Newell, G., White, M., 2018. The effect of sample size on the accuracy of species distribution models: considering both presences and pseudo-absences or background sites. Ecography 4, 535–548. https://doi.org/10.1111/ecog.03188.

Liu, C., White, M., Newell, G., 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. J. Biogeogr. 40, 778–789. https://doi.org/10.1111/jbi.12058.

Lounibos, L.P., Bargielowski, I., Carrasquilla, M.C., Nishimura, N., 2016. Coexistence of Aedes aegypti and Ae. albopictus (Diptera: Culicidae) in peninsular Florida two decades after competitive displacements. J. Med. Entomol. 53, 1385–1390. https://doi.org/10.1093/jme/tjw122.

Low, B.W., Zeng, Y., Tan, H.H., Yeo, D.C.J., 2021. Predictor complexity and feature selection affect Maxent model transferability: Evidence from global freshwater invasive species. Divers. Distrib. 27, 497–511. https://doi.org/10.1111/ddi.13211.

Manzoor, S.A., Griffiths, G., Lukac, M., 2018. Species distribution model transferability and model grain size – finer may not always be better. Sci Rep 8, 7168. https://doi.org/10.1038/s41598-018-25437-1.

McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87, 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0.

McPherson, J.M., Jetz, W., Rogers, D.J., 2006. Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. Ecol. Model. 192, 499–522. https://doi.org/10.1016/j.ecolmodel.2005.08.007.

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x.

Mertes, K., Jetz, W., 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41, 1604–1615. https://doi.org/10.1111/ecog.02871.

Meynard, C.N., Quinn, J.F., 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. J. Biogeogr. 34, 1455–1469. https://doi.org/10.1111/j.1365-2699.2007.01720.x.

Milborrow, S., Hastie, T., Tibshirani, R., Miller, A., Lumley, T., 2019. earth: Multivariate adaptive regression splines. (R package version 5.3.0). https://CRAN.R-project.org/package=earth.

Moffett, A., Shackelford, N., Sarkar, S., 2007. Malaria in Africa: Vector species' niche models and relative risk maps. PLoS ONE 2, e824. https://doi.org/10.1371/journal.pone.0000824.

Moudrý, V., Lecours, V., Malavasi, M., Misiuk, B., Gábor, L., Gdulová, K., Šímová, P., Wild, J., 2019. Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. Ecol. Inform. 54, 100987 https://doi.org/10.1016/j.ecoinf.2019.100987.

Moudrý, V., Šímová, P., 2012. Influence of positional accuracy, sample size and scale on modelling species distributions: a review. Int. J. Geogr. Inf. Syst. 26, 2083–2095. https://doi.org/10.1080/13658816.2012.721553.

Mouton, A.M., De Baets, B., Goethals, P.L.M., 2009. Knowledge-based versus data-driven fuzzy habitat suitability models for river management. Environ. Model. Softw. 24, 982–993. https://doi.org/10.1016/j.envsoft.2009.02.005.

Mughini-Gras, L., Mulatti, P., Severini, F., Boccolini, D., Romi, R., Bongiorno, G., Khoury, C., Bianchi, R., Montarsi, F., Patregnani, T., Bonfanti, L., Rezza, G., Capelli, G., Busani, L., 2014. Ecological niche modelling of potential West Nile Virus vector mosquito species and their geographical association with equine epizootics in Italy. EcoHealth 11, 120–132. https://doi.org/10.1007/s10393-013-0878-7.

Murdock, C.C., Evans, M.V., McClanahan, T.D., Miazgowicz, K.L., Tesla, B., 2017. Fine-scale variation in microclimate across an urban landscape shapes variation in mosquito population dynamics and the potential of Aedes albopictus to transmit arboviral disease. PLOS Negl. Trop. Dis. 11, e0005640 https://doi.org/10.1371/journal.pntd.0005640.

Murphy, A.H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Weather Rev. 116, 2417–2424. https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

Murphy, A.H., Winkler, R.L., 1992. Diagnostic verification of probability forecasts. Int. J. Forecast. 7, 435–455. https://doi.org/10.1016/0169-2070(92)90028-8.

Naimi, B., Hamm, N., Groen, T.A., Skidmore, A.K., Toxopeus, A.G., 2014. Where is positional uncertainty a problem for species distribution modelling. Ecography 37, 191–203.

Naimi, B., Skidmore, A.K., Groen, T.A., Hamm, N.A.S., 2011. Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. J. Biogeogr. 38, 1497–1509. https://doi.org/10.1111/j.1365-2699.2011.02523.x.

National Institute of Statistics and Geography, 2018. Densidad de población por entidad federativa, 1990 a 2015. https://en.www.inegi.org.mx/app/tabulados/interactivos/?pxq=Poblacion_Poblacion_07_fb7d5132-39f0-4a6c-b6f6-4cbe440e048d (accessed 12.10.18).

NCAR - Research Applications Laboratory, 2015. verification: Weather forecast verification utilities. (R package version 1.42). https://CRAN.R-project.org/package=verification.

Nelson, A., 2001. Analysing data across geographic scales in Honduras: detecting levels of organisation within systems. Agric. Ecosyst. Environ. 85, 107–131. https://doi.org/10.1016/S0167-8809(01)00191-8.

Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data. Environ. Plan. A. 16, 17–31. https://doi.org/10.1068/a160017.

Osborne, P.E., Leitão, P.J., 2009. Effects of species and habitat positional errors on the performance and interpretation of species distribution models. Divers. Distrib. 15, 671–681. https://doi.org/10.1111/j.1472-4642.2009.00572.x.

Pacifici, K., Reich, B.J., Miller, D.A.W., Pease, B.S., 2019. Resolving misaligned spatial data with integrated species distribution models. Ecology 100, e02709. https://doi.org/10.1002/ecy.2709.

Park, D.S., Davis, C.C., 2017. Implications and alternatives of assigning climate data to geographical centroids. J. Biogeogr. 44, 2188–2198. https://doi.org/10.1111/jbi.13029.

Pascoe, E.L., Pareeth, S., Rocchini, D., Marcantonio, M., 2019. A lack of "environmental earth data" at the microhabitat scale impacts efforts to control invasive arthropods that vector pathogens. Data. 4, 133. https://doi.org/10.3390/data4040133.

Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Glob. Chang. Biol. 12, 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x.

Peterson, R.A., Cavanaugh, J.E., 2020. Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. J. Appl. Stat. 47, 2312–2327. https://doi.org/10.1080/02664763.2019.1630372.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol. Evol. 20, 503–510. https://doi.org/10.1016/j.tree.2005.05.011.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Rahbek, C., Graves, G.R., 2001. Multiscale assessment of patterns of avian species richness. PNAS 98, 4534–4539. https://doi.org/10.1073/pnas.071034898.

Reinhold, J.M., Lazzari, C.R., Lahondère, C., 2018. Effects of the environmental temperature on Aedes aegypti and Aedes albopictus mosquitoes: A review. Insects 9, 158. https://doi.org/10.3390/insects9040158.

Reiss, H., Cunze, S., Koenig, K., Neumann, H., Kroencke, I., 2011. Species distribution modelling of marine benthos: a north sea case study. Mar. Ecol. Prog. Se. 442, 71–86. https://doi.org/10.3354/meps09391.

Roebber, P.J., 1998. The regime dependence of degree day forecast technique, skill, and value. WAF. 13, 783–794. https://doi.org/10.1175/1520-0434(1998)013<0783:TRDODD>2.0.CO;2.

Royle, J.A., Chandler, R.B., Yackulic, C., Nichols, J.D., 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods Ecol. Evol. 3, 545–554. https://doi.org/10.1111/j.2041-210X.2011.00182.x.

Rydgren, K., Økland, R.H., Økland, T., 2003. Species response curves along environmental gradients. A case study from SE Norwegian swamp forests. J Veg. Sci. 14, 869–880. https://doi.org/10.1111/j.1654-1103.2003.tb02220.x.

Santika, T., Hutchinson, M.F., 2009. The effect of species response form on species distribution model prediction and inference. Ecol. Model. 220, 2365–2379. https://doi.org/10.1016/j.ecolmodel.2009.06.004.

Santos, J., Meneses, B.M., 2017. An integrated approach for the assessment of the Aedes aegypti and Aedes albopictus global spatial distribution, and determination of the zones susceptible to the development of Zika virus. Acta. Trop. 168, 80–90. https://doi.org/10.1016/j.actatropica.2017.01.015.

Schweiger, A.H., Beierkuhnlein, C., 2016. Scale dependence of temperature as an abiotic driver of species' distributions. Glob. Ecol. Biogeogr. 25, 1013–1021. https://doi.org/10.1111/geb.12463.

Segurado, P., Araujo, M.B., Kunin, W.E., 2006. Consequences of spatial autocorrelation for niche-based models. J. Appl. Ecol. 43, 433–444. https://doi.org/10.1111/j.1365-2664.2006.01162.x.

Seo, C., Thorne, J.H., Hannah, L., Thuiller, W., 2009. Scale effects in species distribution models: implications for conservation planning under climate change. Biol. Lett. 5, 39–43. https://doi.org/10.1098/rsbl.2008.0476.

Shirley, S.M., Yang, Z., Hutchinson, R.A., Alexander, J.D., McGarigal, K., Betts, M.G., 2013. Species distribution modelling for the people: unclassified landsat TM imagery predicts bird occurrence at fine resolutions. Divers. Distrib. 19, 855–866. https://doi.org/10.1111/ddi.12093.

Smith, A.B., Santos, M.J., 2020. Testing the ability of species distribution models to infer variable importance. Ecography 43, 1801–1813. https://doi.org/10.1111/ecog.05317.

Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distribution areas. Biodivers. Inform. 2, 1–10. https://doi.org/10.17161/bi.v2i0.4.

Statistics Canada, 2018. Census profile - age, sex, type of dwelling, families, households, marital status, language, income, immigration and ethnocultural diversity, housing, aboriginal peoples, education, labour, journey to work, mobility and migration, and language of work for Canada, provinces, territories and health regions, 2016 Census (No. 98-401-X2016058) https://www150.statcan.gc.ca/n1/en/catalogue/98-401-X2016058 (accessed December 10 2018).

Stoklosa, J., Daly, C., Foster, S.D., Ashcroft, M.B., Warton, D.I., 2015. A climate of uncertainty: accounting for error in climate variables for species distribution models. Methods Ecol. Evol. 6, 412–423. https://doi.org/10.1111/2041-210X.12217.

Synes, N.W., Osborne, P.E., 2011. Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. Glob. Ecol. Biogeogr. 20, 904–914. https://doi.org/10.1111/j.1466-8238.2010.00635.x.

Therneau, T., Atkinson, B., 2019. rpart: Recursive partitioning and regression trees. (R package version 4.1-15). https://CRAN.R-project.org/package=rpart.

Thompson, C.M., McGarigal, K., 2002. The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). Landscape Ecol 17, 569–586. https://doi.org/10.1023/A:1021501231182.

Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. J. Hydrol. 190, 214–251. https://doi.org/10.1016/S0022-1694(96)03128-9.

Thornton, P.E., Rupesh Shrestha, M.M., Thornton, S.C., Kao, Y.W., Wei, B.E., 2017. Daymet: Monthly climate summaries on a 1-km grid for North America, version 4. ORNL DAAC, Oak Ridge, Tennessee, USA.

Thuiller, W., Brotons, L., Araújo, M.B., Lavorel, S., 2004. Effects of restricting environmental range of data to project current and future species distributions. Ecography 27, 165–172. https://doi.org/10.1111/j.0906-7590.2004.03673.x.

Thuiller, W., Georges, D., Engler, R., Breiner, F., 2020. biomod2: Ensemble platform for species distribution modeling. (R package version 3.5.0). https://CRAN.R-project.org/package=biomod2.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD – a platform for ensemble forecasting of species distributions. Ecography 32, 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x.

Tiffin, H.S., Peper, S.T., Wilson-Fallon, A.N., Haydett, K.M., Cao, G., Presley, S.M., 2019. The influence of new surveillance data on predictive species distribution modeling of Aedes aegypti and Aedes albopictus in the United States. Insects 10, 400. https://doi.org/10.3390/insects10110400.

Trivedi, M.R., Berry, P.M., Morecroft, M.D., Dawson, T.P., 2008. Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. Glob. Chang. Biol. 14, 1089–1103. https://doi.org/10.1111/j.1365-2486.2008.01553.x.

U.S. Census Bureau, 2017. Annual estimates of the resident population: April 1, 2010 to July 1, 2017 https://www.census.gov/geo/maps-data/data/tiger-data.html (accessed December 10 2018).

Velásquez-Tibatá, J., Graham, C.H., Munch, S.B., 2016. Using measurement error models to account for georeferencing error in species distribution models. Ecography 39, 305–316. https://doi.org/10.1111/ecog.01205.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S-PLUS. Springer Science & Business Media.

Verdonschot, P.F.M., Besse-Lototskaya, A.A., 2014. Flight distance of mosquitoes (Culicidae): A metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. Limnologica 45, 69–79. https://doi.org/10.1016/j.limno.2013.11.002.

Vergara, M., Cushman, S.A., Urra, F., Ruiz-González, A., 2016. Shaken but not stirred: multiscale habitat suitability modeling of sympatric marten species (Martes martes and Martes foina) in the northern Iberian Peninsula. Landscape Ecol 31, 1241–1260. https://doi.org/10.1007/s10980-015-0307-0.

Warren, D.L., Seifert, S.N., 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol. Appl. 21, 335–342. https://doi.org/10.1890/10-1171.1.

Weaver, S.C., 2014. Arrival of chikungunya virus in the new world: prospects for spread and impact on public health. PLOS Negl. Trop. Dis. 8, e2921. https://doi.org/10.1371/journal.pntd.0002921.

Weisheimer, A., Palmer, T.N., 2014. On the reliability of seasonal climate forecasts. J. R. Soc. Interface. 11, 20131162 https://doi.org/10.1098/rsif.2013.1162.

Wiens, J., 1989. Spatial scaling in ecology. Ecol 3, 385–397.

Windsor-Essex Health Unit, 2021. News release: Mosquito trap reveals one Aedes albopictus mosquito in the area https://www.wechu.org/newsroom/news-release-mosquito-trap-reveals-one-aedes-albopictus-mosquito-area (accessed January 16 2022).

Winters, A.M., Bolling, B.G., Beaty, B.J., Blair, C.D., Eisen, R.J., Meyer, A.M., Pape, W.J., 2008. Combining mosquito vector and human disease data for improved assessment of spatial West Nile virus disease risk. Am. J. Trop. Med. Hyg. 78, 654–665.

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

Yang, B., Borgert, B.A., Alto, B.W., Boohene, C.K., Brew, J., Deutsch, K., DeValerio, J.T., Dinglasan, R.R., Dixon, D., Faella, J.M., Fisher-Grainger, S.L., Glass, G.E., Hayes, R., Hoel, D.F., Horton, A., Janusauskaite, A., Kellner, B., Kraemer, M.U.G., Lucas, K.J., Medina, J., Morreale, R., Petrie, W., Reiner, R.C., Riles, M.T., Salje, H., Smith, D.L., Smith, J.P., Solis, A., Stuck, J., Vasquez, C., Williams, K.F., Xue, R.-D., Cummings, D.A.T., 2021. Modelling distributions of Aedes aegypti and Aedes albopictus using climate, host density and interspecies competition. PLOS Neglect. Trop. Dis. 15, e0009063 https://doi.org/10.1371/journal.pntd.0009063.

Yu, H., Cooper, A.R., Infante, D.M., 2020. Improving species distribution model predictive accuracy using species abundance: Application with boosted regression trees. Ecol. Model. 432, 109202 https://doi.org/10.1016/j.ecolmodel.2020.109202.