



OPEN

Local environmental factors drive distributions of ecologically-contrasting mosquito species (Diptera: Culicidae)

Roel M. Wouters^{1,2,3✉}, Wouter Beukema^{1,4}, Maarten Schrama², Koos Biesmeijer^{1,2}, Marieta A. H. Braks^{2,5}, Pepijn Helleman², Francis Schaffner⁶, Joey van Slobbe⁷, Arjan Stroo⁸ & Jordy G. van der Beek^{1,2,9}

Mosquitoes are important vectors of disease pathogens and multiple species are undergoing geographical shifts due to global changes. As such, there is a growing need for accurate distribution predictions. Ecological niche modelling (ENM) is an effective tool to assess mosquito distribution patterns and link these to underlying environmental preferences. Typically, macroclimatic variables are used as primary predictors of mosquito distributions. However, they likely undervalue local conditions and intraspecific variation in environmental preferences. This is problematic, as mosquito control takes place at the local scale. Utilising high-resolution (10 × 10 m) Maxent ENMs on the island of Bonaire as model system, we explore the influence of local environmental variables on mosquito distributions. Our results show a distinct set of environmental variables shape distribution patterns across ecologically-distinct species, with urban variables strongly associated with introduced species like *Aedes aegypti* and *Culex quinquefasciatus*, while native species show habitat preferences for either mangroves, forests, or ephemeral water habitats. These findings underscore the importance of distinct local environmental factors in shaping distributions of different mosquitoes, even on a small island. As such, these findings warrant further studies aimed at predicting high-resolution mosquito distributions, opening avenues for preventative management of vector-borne disease risks amidst ongoing global change and ecosystem degradation.

Mosquitoes belong to the family of Culicidae with approximately 3700 species¹. A subset of these species plays an important role in global public health as vector of infectious diseases for humans, livestock and wildlife. According to recent estimations, by 2050, half of the world's population will be at risk of vector-borne diseases². As a result of global change, many mosquito species are on the move³⁻⁶. Hence, there is a growing need to accurately predict mosquito distributions. As such, it is essential to understand the underlying drivers of mosquito distributions.

Large scale spatial projections of mosquito distributions rely on the predictive value of large-scale climatic factors (i.e. temperature and precipitation), resulting in relatively coarse-scale maps of species distributions^{2,4,7-11}. However, at the local scale (< 10 km²), these climatic factors are largely indistinctive and local scale environmental

¹NL Biodiversity and Society Research Group, Naturalis Biodiversity Center, 2333 CR Leiden, The Netherlands. ²Institute of Environmental Sciences, Leiden University, 2333 CC Leiden, The Netherlands. ³Department of Ecology, Faculty of Science, Charles University, 12844 Prague, Czechia. ⁴RAVON, Reptile, Amphibian and Fish Conservation Netherlands, 6501 BK Nijmegen, The Netherlands. ⁵Centre for Zoonoses and Environmental Microbiology, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, 3721 MA Bilthoven, The Netherlands. ⁶Francis Schaffner Consultancy, 4125 Riehen, Switzerland. ⁷Bonaire Public Health Department, Public Body Bonaire, 4PXG+GH4, Kralendijk, Dutch Caribbean, The Netherlands. ⁸Centre for Monitoring of Vectors (CMV), Netherlands Food and Consumer Product Safety Authority (NVWA), 6706 EA Wageningen, The Netherlands. ⁹Pandemic and Disaster Preparedness Center, Delft, Rotterdam, The Netherlands. ✉email: WoutersRoel1@gmail.com

parameters are thought to dictate the heterogeneity in mosquito distribution and local distribution maps of mosquitoes are generally lacking^{12–16}. This greatly hampers the use of distribution maps to guide policy, e.g., to guide vector-control and other management/mitigation strategies that operate at similar local levels¹⁷. Hence, there is a growing need to include (local) environmental conditions, other than climatic variables in understanding mosquito distributions^{12,13,17}.

Species distribution modelling, also known as ecological niche modelling (ENM), allows for fine-scale predictions of species occurrences^{18–20}. ENMs comprise a family of statistical or machine-learning models, which may be used to predict the potential distribution of a given species in a given area by relating occurrence data to ecologically-relevant environmental conditions²¹. While ENMs have previously been used to generate global distribution maps largely relying on abovementioned climatic variables^{2,4,7–11}, the models also allow for the inclusion of more local-scale environmental variables thus contributing to the understanding of the fine scale distribution of a given species. As such, ENM holds the potential to be applied for vector-borne disease management efforts, improving traditional disease risk maps and increasing knowledge on the ecology of a given species¹⁸.

Indeed, studies applying fine-scale ENM highlight the potential importance of local-scale variables^{12,17}. For example, the occurrence of the (introduced) arbovirus vectors *Aedes aegypti* and *Aedes albopictus* in a human-dominated landscape in the tropics and subtropics can be explained by a number of key urban and nature-related factors (including population density, house density and distance from vegetation or water), and to a lesser extent also by local climatic factors such as temperature and precipitation^{20,22–25}. However, there is a major bias in ENM-based studies on mosquitoes: the vast majority is geared towards a limited number of species involved in disease transmission, while precious little is known about the distribution of other species²⁶. This greatly limits our understanding of the importance of environmental variables in shaping the distribution of ecologically divergent mosquito species and in particular how introduced invasive species differ from the established native species.

Here, we use the island of Bonaire (288 km²), one of the islands of the Lesser Antilles in the Caribbean region (Fig. 1H), as a model to show how a confined set of key environmental factors at high resolution can be used to explain the local distributions of mosquito species with contrasting ecological strategies.

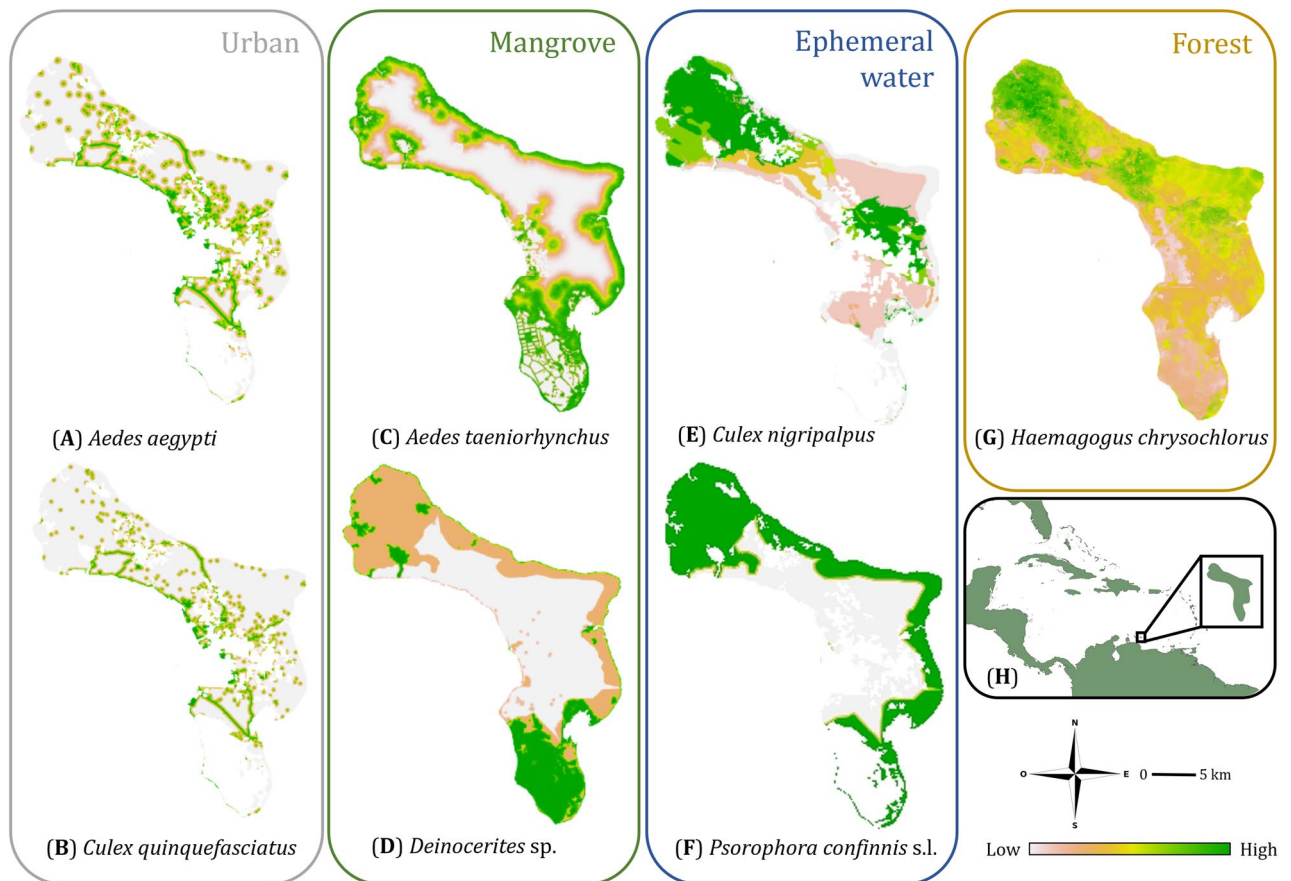


Figure 1. Habitat suitability maps for the seven modelled mosquito species: (A) *Aedes aegypti*, (B) *Culex quinquefasciatus*, (C) *Aedes taeniorhynchus*, (D) *Deinocerites* sp., (E) *Culex nigripalpus*, (F) *Psorophora confinnis* s.l., (G) *Haemagogus chrysochlorus*. (H) A map of the island of Bonaire as part of the Lesser Antilles, located near coast of Venezuela. Habitat suitability (relative likelihood) is illustrated with a gradient fill, where brown indicates low habitat suitability and green indicates high suitability. Species are grouped by corresponding habitat type.

Results and discussion

Habitat suitability (relative likelihood) for seven out of twelve mosquito species collected on Bonaire was estimated using a high-resolution ENM approach involving variable selection procedures, null model validation, and model selection to obtain an optimally parameterised model for each species. For five out of the twelve species, models could not be generated because the number of observations was insufficient to run the models (see Fig. S1 and Table S1). The most important contributing environmental variables for each species and the model performance (i.e., AUC and Boyce index) varied between mosquito species (Table 1), resulting in strikingly different predicted distribution patterns of the species on the island (Fig. 1). Below, the key findings concerning overall model performance and the important environmental factors contributing to the habitat suitability for the seven species are reported. See Table S2 and Fig. S2–S8 for complete model selection, performance and variable response curves per species and see Table S3 for an overview of all spatial data included in the models.

Aedes aegypti and *Culex quinquefasciatus*

Both *Ae. aegypti* and *Cx. quinquefasciatus* show a relatively similar predicted occurrence over the island (Fig. 1A,B). Models performed moderately (AUC = 0.65 and 0.64, and Boyce index = 0.41 and 0.31 respectively) and the most important variable contributing to the habitat suitability of both species is the distance to urban habitat (100% and 79% respectively, Table 1), which is negatively correlated thus implying that habitat suitability decreases with increasing distance to urbanisation. For *Cx. quinquefasciatus* also land use (mangrove habitat, 21%) contributes to the habitat suitability. The predicted habitat suitability maps for these two species (Fig. 1A,B) reveal a wide distribution over the island with urban environments including the cities of Kralendijk (central west) and Rincon (northwest), as well as the main roads and other urban elements as important predictors.

Aedes taeniorhynchus

In the model for *Ae. taeniorhynchus*, the most important contributing environmental variable is the distance to mangrove habitat (58.4%). The negative correlation between this variable and the occurrence of this species implies that the likelihood of observing *Ae. taeniorhynchus* decreases further away from mangrove habitat. Distance to nature is the second important variable (18.8%). Elevation (negative correlation) and the categorical variable mangrove habitat contribute respectively 12.9% and 9.8% (Table 1). The selected model has a high performance with AUC = 0.89 and Boyce index = 0.74. The predicted habitat suitability for *Ae. taeniorhynchus* is particularly high at lower elevations, including the salt plains and mangrove habitats in the south (Fig. 1C).

Deinocerites sp.

The model for the yet undescribed species of *Deinocerites* sp. performs well with an AUC of 0.78 and a Boyce index of 0.47. The most important variable contributing to the habitat suitability of *Deinocerites* sp. is the distance to protected natural area (54.4%). The second contributing variable is the distance to mangrove habitat (46%),

Species	Contributing environmental variables				Known ecology	AUC value	Boyce index
	Variable 1	Variable 2	Variable 3	Variable 4			
<i>Aedes aegypti</i>	Distance to urban habitat (100%)				Introduced; container breeder; urban ^{28,35}	0.65	0.41
<i>Culex quinquefasciatus</i>	Distance to urban habitat (79%)	Land use Mangrove habitat (21%)			Introduced; nutrient-rich, dirty standing water breeder; urban ²⁸	0.64	0.31
<i>Aedes taeniorhynchus</i>	Distance to mangrove habitat (58.4%)	Distance to natural habitat (18.8%)	Elevation—negative correlation (12.9%)	Land use mangrove habitat (9.8%)	Breeds in fresh, brackish and saline waterbodies; mangrove ^{27,28}	0.89	0.74
<i>Deinocerites</i> sp.	Distance to protected nature area (54.4%)	Distance to mangrove habitat (46%)			Land crab-hole breeder; brackish and saline water tolerant; restricted to land crab holes of the family Geacarididae ^{29,38}	0.78	0.47
<i>Culex nigripalpus</i>	Distance to temporary waterbody (84.5%)	Distance to urban vegetation (15.5%)			Breeds in flooded ditches and small ponds; related to rainfall and ephemeral waterbodies ^{27,28}	0.50	0.11
<i>Psorophora confinnis</i> s.l.	Distance to protected nature area (85.1%)	Distance to temporary waterbody (14.9%)			Breeds in flooded plains and ponds; related to rainfall and ephemeral waterbodies ^{28,39}	0.56	0.63
<i>Haemagogus chrys-ochlorus</i>	Distance to natural habitat (48%)	Elevation—positive correlation (18.6%)	Land use forest and high scrub habitat (12.4%)	Geology high terrace (7.2%)	Tree-hole breeder ^{29,37}	0.56	−0.11

Table 1. Variable importance (%), ecology known from the literature and model performance (AUC values) for each modelled species. Cumulative contribution of the variables is at least 80%. The *distance to* variables are all negative correlations where increasing distance means lower suitability.

Table 1). Both variables show negative correlations with the respective environmental variables, implying that the probability of occurrence decreases further away from these habitats. A very high predicted suitability was observed in the southern part of Bonaire, where the protected salt plains are located (Fig. 1D). The vegetation on these salt plains consists predominantly of (low shrub) mangrove, and the waters are brackish or (hyper) saline; this is also the location with highest observed densities of land crabs which are closely associated with this species. Several suitable hotspots can also be found in the north in protected mangrove habitats.

Culex nigripalpus

The model for *Cx. nigripalpus* showed that the two most important contributing variables are (1) the distance to temporary waterbodies (84.5%) and (2) the distance to urban vegetation (15.5%), which includes city parks and other green areas within the urbanised areas (Table 1). Both variables show no clear correlation (see Fig. S7) and the model performance is poor with an AUC of 0.5 and a Boyce index of 0.11). The habitat suitability map predicts high suitability in the north and northwest (i.e. in a protected nature area, agricultural habitat and around the city of Rincon) and east of Kralendijk (Fig. 1E), both areas consisting primarily of agricultural and unprotected nature areas. Several suitable habitats can be found in the southeast, where protected mangrove vegetation is dominant.

Psorophora confinnis s.l.

For *Ps. confinnis* s.l., 85.1% of the modelled habitat suitability can be explained by the distance to protected nature area and the remaining 14.9% is explained by the distance to temporary waterbodies (Table 1). The model predicts high habitat suitability for this species in the central north and central south, which are both protected nature areas, as well as along the east coast (Fig. 1F). The model has a poor to moderate performance with an AUC of 0.56 and a Boyce index of 0.63.

Haemagogus chrysochlorus

Haemagogus chrysochlorus was found at 23 unique sampling locations, the largest sample size of all modelled species. Nevertheless, the selected model had a poor performance (AUC = 0.56 and Boyce index = -0.11). The four most contributing environmental variables were respectively: (1) distance to natural habitat (48%), (2) elevation (18.6%), (3) land use (forest and high scrub habitat, 12.4%) and (4) geology (higher terrace, 7.2%) (Table 1). For *Hg. chrysochlorus*, the model predicts high suitability in the higher elevated areas which are covered with forests containing trees of low height. Those areas are located in the north and centre of the island, both in protected natural areas. Several smaller suitable spots on the map are located central east and central south (Fig. 1G).

Model performance and validation

We observe the lowest AUC validation scores for *Cx. nigripalpus* and *Hg. chrysochlorus* (AUC of 0.50 and 0.56, and Boyce index of 0.11 and -0.11 respectively). For *Ps. confinnis* s.l. we observe a low AUC score of 0.56 but a better Boyce index score of 0.63. As we applied random stratified sampling based on the available habitat types, we do not expect a collection bias is the cause of these low validation scores. We argue instead that lacking environmental information at local to micro scales affects model performances in these species. Specifically, information (in the form of GIS layers) in the form of ephemeral waterbodies and tree hole availability, which are used as breeding sites, is currently not available at spatial broad scales. Inclusion of such small-scale landscape elements could likely improve model performance in the future. These relatively poor model performances (Table 1) imply that these ENM outputs need to be interpreted with caution and further ecological validation is necessary to affirm model results. Nevertheless, the contributing environmental variables explaining their distribution are largely in line with the known ecology of these species throughout their geographic distributions^{27–32} (Table 1). This underscores the notion that, despite the suboptimal performance, these models may still be useful to generate good local projections of the occurrence of the species of interest.

In the sampling data set, larval, pupal and adult stages of the modelled species were grouped to ensure a large enough sample size for model input. However, adult mosquitoes have the ability to disperse in contrast to the strictly water-bound larvae and pupae. Adults of some species, notably *Ae. taeniorhynchus*, are known to be able to cover large distances^{33,34}. Interestingly, our findings show a *distance to* variable has the highest explanatory contribution for all seven species, suggesting sample sizes were robust enough to account for the dispersal of adult stages. This stresses the value of including *distance to* matrices when modelling flying insects such as mosquitoes. Using data from only a single life stage could potentially underestimate (larval) or overestimate (adult) the potential distribution and spatial invasion risk.

Distinct environmental variables shape distributions

There was a remarkable difference between the predicted distribution of the introduced species (*Ae. aegypti* and *Cx. quinquefasciatus*; most likely introduced on Bonaire in the fifteenth th/sixteenth century^{35,36}) on the one hand and mosquitoes that are likely of native origin (*Ae. taeniorhynchus*, *Deinocerites* sp., *Cx. nigripalpus*, *Ps. confinnis* s.l. and *Hg. chrysochlorus*) on the other hand. The introduced mosquito species show a distribution strictly centred around urban environments which matches the known ecology and lifecycle of these species along their distribution globally and in the Caribbean region^{2,28,35}. Inclusion of more high-resolution urban spatial layers, for instance dealing with urban structure or demography, might reveal more detailed environmental information about the local distribution of these species that play an important role in public health.

In marked contrast to the urban-dwelling species discussed above, the distribution of the various native mosquito species is explained by different elements of the natural environment. Based on the predicted habitat suitability outcomes of the models, their distribution appears to reflect three separate habitat preferences: (1)

mangrove-dwelling species, (2) forest-dwelling species and (3) ephemeral water-dwelling species (Fig. 1). The models for the native mangrove-dwelling mosquitoes (i.e. *Ae. taeniorhynchus* and *Deinocerites* sp.) show a distribution closely aligned with natural mangrove habitats, which are predominantly located on the coastal areas. This coincides with the known ecology and breeding habits of these salt tolerant, mangrove breeding species^{27,28}. The habitat suitability maps for these species suggest possible co-occurrence which was confirmed in our field survey. Two other species show a clear habitat preference towards ephemeral fresh waterbodies restricted to natural areas (*Ps. confinnis* s.l.) or vegetation-rich areas in urban environments (*Cx. nigripalpus*). This matches previous observations as both species are known to depend on ephemeral fresh waterbodies like natural and artificial ponds and floodplains^{27,28}. To further increase model performance for this specific group, more fine-scale spatial information on different kinds and stages of ephemeral waterbodies would likely be needed. The forest-dwelling species group is represented by a single species: *Hg. chrysochlorus*. This species typically breeds in tree holes^{29,37}, a behaviour consistent with our observed field data and the predicted distribution in forests at slightly elevated elevations on the high terrace of Bonaire. Notably, this region has many older trees, primarily *Guaiacum officinale* L., which frequently harbour tree holes. Additionally, larvae of the species were recorded in rockpools with abundant leaf litter within the same area. Overall our predicted maps closely match the known ecology of the species in this study, thus highlighting the strength of this approach to generate fine-scale distribution maps for mosquito species. However, this strongly relies on the availability of information on the local ecological factors at a given place.

Conclusion

The results of this study show that, on a small but ecologically diverse island used as a model system, contrasting distributions of mosquito species can be explained by a set of ecologically-relevant environmental variables using the ENM approach Maxent. We observed a clear distinction between introduced species, of which the distributions were primarily explained the distance to urbanisation, and the native mosquitoes, of which the distribution is primarily explained by environmental predictors that are characteristic of natural habitats. Within the native mosquito species pool, predicted distributions were associated with three distinct habitat associations: mangrove-dwelling species, forest-dwelling species, and ephemeral water-dwelling species, which is largely in line with the known ecology of the modelled species despite poor model performances for three of seven species. Overall, our data implies that high-resolution ENM modelling reveals crucial insights in local habitat preferences that may ultimately benefit vector management and control. As such it can be a highly valuable tool to generate predictions on local mosquito distributions, but this relies strongly on the availability of high resolution data on local environmental factors. With increasing availability of remote sensing techniques leading to broad-scale, high-resolution spatial data, the application of ENM approaches for modelling vectors and their relatives will likely become more important in predicting distributions. This paves the way for generating fine-scale distribution maps that may benefit local policy and management (e.g. vector-control strategies) and other management practice related to (wildlife) veterinary and public health.

Methods

Study site and mosquito diversity survey

The island of Bonaire, located in the Lesser Antilles, is part of the Dutch Caribbean and with an area of 288 km², it is one of the smaller inhabited islands in the Caribbean region. The island is located 80 km off the coast of Venezuela and has a semi-arid tropical savanna climate. Bonaire is primarily covered by limestone with basalt lava rock formations breaking through on the higher elevations⁴⁰. The vegetation mainly consists of low shrub cacti and acacias with some ruminant low forest on the higher elevations. Both in the south and the north, protected mangrove vegetation surrounds the numerous salt lakes, locally called salinas⁴⁰. We conducted a thorough mosquito diversity survey on the island of Bonaire in December 2022 as part of a larger inventory of the mosquito diversity on the Dutch ABC islands (Aruba, Bonaire and Curaçao). The survey was conducted for a period of fourteen days (from 30 November until 14 December 2022). We sampled during the end of the rainy season (September through December), which is generally a favourable time for mosquito activity due to the presence of different types of waterbodies. The rainy season of 2022 was incredibly wet, with almost double the amount of precipitation compared to the annual average⁴¹. We conducted random stratified sampling based on the diversity of habitat types present on Bonaire. We selected the different habitat types based on existing vegetation and land use maps^{40,42} (Table S3) and we sampled all (aquatic) habitat types in Bonaire (excluding deep sea and shallow coastal waters, as they do not form suitable habitat for mosquitoes in general). Habitats were reached either by car or by foot and places suitable for trapping were identified on sight and with expertise from local collaborators.

Mosquito collection

Mosquitoes are holometabolic insects, meaning they undergo complete metamorphosis and therefore have a larval and pupal stage before reaching their imago (adult) stage. To get a complete overview of the diversity, the used sampling methods were aimed to trap all larval, pupal and adult stages. Larvae and pupae are strictly bound to water and therefore, they were sampled from waterbodies using traditional dipping methods⁴³. Adult specimens were trapped using traps baited with CO₂ produced by yeast (Biogents Pro trap⁴⁴), as well as using aspirators and netting techniques. CO₂-baited traps were placed for a period of 24 h, covering diurnal, nocturnal and crepuscular species. From each trapping site corresponding coordinates and habitat type were registered, along with other environmental parameters (including pH and salinity of the water). Coordinates and corresponding environmental information were later attributed to the identified species. Collected samples were morphologically identified on-site to species level using a compilation of identification keys^{27–29,45–49}. We omitted

specimens that could not be identified to species level from further modelling. Table S1 shows all species identified from Bonaire.

Data preparation, modelling pipeline and model selection

Occurrence data cleaning, preparation of environmental variables used to characterise species distributions, and ENM building were conducted in the R (version 4.2.2) open source statistical programming language (R Core Team 2022). Packages used include *data.table* (v.1.14.8)⁵⁰ and *dplyr* (v.1.1.2)⁵¹ (general functions), *ncdf4* (v.1.22)⁵², *raster* (v.3.6.26)⁵³, *rgdal* (v.1.6.7)⁵⁴, *rgeos* (v.0.6.4)⁵⁵, *sf* (v.1.0.14)⁵⁶, *sp* (v.2.1.2)⁵⁷, and *vegan* (v.2.6.4)⁵⁸ (spatial functions), and *ade4* (v.1.7.22)⁵⁹, *dismo* (v.1.3.14)⁶⁰, *ecospat* (v.3.5.1)⁶¹, *ENMeval* (v.2.0.4)⁶², *ENMTools* (v.1.0.6)⁶³, and *spatialEco* (v.2.0.1)⁶⁴ (ENM-related functions).

We exclusively used occurrence data of specimens we collected during our field survey and did not work with open databases or formerly collected data. See Fig. S1 for a spatial rendering of the occurrence data for all modelled species. For the occurrence data, no distinction was made between the life stages (i.e. larval, pupal and adult stages) and only one occurrence point per unique location where a given species was present was used. This means only presence was taken into account, not abundance. Species with less than four unique collection sites were not modelled. Subsequently, we prepared a variety of biotic and abiotic environmental spatial layers. We started with publicly available layers which included layers that are known to be important for explaining mosquito distributions (e.g. urban habitat). Spatial layers include land use, vegetation, geology, soil type, elevation, protected nature areas, information on riverbeds and larger semi-temporary waterbodies and urban variables including urban areas, neighbourhoods, roads and footpaths. We did not provide a priori environment-species interactions and therefore initially included all spatial layers in each species model (see Table S3 for all base maps used for the environmental spatial layers). Because of its small size, Bonaire has just one macroclimatic zone. Therefore, no macroclimatic or weather variables were added to the models. Notably, a recent study examining the impact of climate change on Bonaire, did not distinguish more than one macroclimatic region on the island in their modelling approaches⁶⁵. No microclimatic variables (e.g. higher temperature or lower humidity because of differences in urbanisation and tree cover) were available for Bonaire, and as such these could not be included. We rescaled the spatial layers with an original resolution of 2 × 2 m or 5 × 5 m to a resolution of 10 × 10 m in R and QGIS and we changed the coordinate system to WGS84. For most of the categorical layers, we created a *distance to* matrix as well. This enables numerical analyses of these variables and the importance of the proximity to a specific category.

Our chosen presence-background ENM approach needs a dataset of ‘background sites’ that are compared with the environmental variation at locations where mosquitoes were collected. Therefore, we created such a background dataset by randomly selecting a total of 10,000 occurrence points within the well-sampled areas to prevent overprediction of the model^{62,66,67}. As a threshold, only background areas were selected when at least two or three species were collected.

The machine-learning ENM algorithm Maxent⁶⁶ was used to generate suitability predictions via the R package *ENMeval*⁶². Maxent is one of the most widely used SDM algorithms, often outperforms other ENMs, and generally achieves good results when using small datasets (e.g., Wisz et al.⁶⁷). Applying Maxent in *ENMeval* permits splitting occurrence data into ‘spatial blocks’, which among other benefits decreases the chance of high spatial autocorrelation between training and testing sites, and thereby counters model overfitting⁶⁸. *ENMeval* also includes model selection procedures⁶⁹.

The ENM pipeline consisted of the following steps. Occurrence- and background data were first divided into spatial blocks to decrease chances for high spatial autocorrelation between training and testing sites. Then, a set of initial models was created for these subsets, each with different combinations of parameter settings, which together represent all combinations of regularization multiplier (1–5) and feature class settings [‘linear’ (L), ‘linear & quadratic’ (LC), hinge (H), and linear, quadratic and hinge (LQH)]. The optimal model was defined as the model with the lowest mean percentage of test sites falling outside the predicted range (i.e., lowest mean omission error), and the highest mean evaluation (AUC) value^{70,71}. To check whether this model performed significantly better than models based on randomly selected data, a series of 100 null models was created⁷² based on Raes and Ter Steege²¹. Optimal models of *Aedes aegypti*, *Aedes taeniorhynchus*, *Culex quinquefasciatus*, *Culex nigripalpus*, *Deinocerites* sp., *Haemagogus chrysochlorus* and *Psorophora confinnis* s.l. were found to perform significantly better than models based on randomly selected data. See Table 1 for more information on their ecology and habitat preferences. All environmental variables were used in the initial models; no initial correlation assessments were performed. Because Maxent is a machine learning algorithm, (strongly) correlating variables have no effect on performance when predictions are only made for the current time period⁷³. However, there is a risk of overparameterisation, and suitability maps based on ecologically relevant variables are more accurate than maps based on arbitrary environmental variables^{74–76}. The percentage contribution of each environmental variable was therefore retrieved from the results of the initial models and used to select the most influential variables. If within this selection there were still sets of variables with a correlation of less than –0.7 or more than 0.7 Pearson’s *r*, the least important variable was removed (based on Jueterbock⁷⁷). The final set of environmental variables was used to generate a new set of models, again for all different combinations of parameter settings. The optimal model from this series was selected as the final model. From this model, plots of the occurrence data divided into spatial blocks, metadata, species-environment relationship plots, and the spatial predictions (suitability maps) were exported as separate files. Taking into account the growing criticism, the evaluation of model performance continues to predominantly rely on the validation of Area Under the Curve (AUC) scores^{78,79}. In the context of our investigation, we adopt a nuanced approach to interpreting AUC values: an AUC value of ≤ 0.5 is considered not performing better than random. Those within the range of AUC > 0.5 but < 0.7 are classified as demonstrating poor to moderate performance; AUC values > 0.7 but < 0.9 signify models exhibiting

moderate to high performance; and notably, an $AUC > 0.9$ is indicative of excellent performance^{80–82}. On top of AUC values we also evaluated model performance based on the Boyce index⁸³. The index ranges from -1 to 1 . A value closer to 1 indicates model predictions align with presences' distribution while a value closer to -1 indicates the model predictions do not align with the presence's distribution. Values around zero indicate the model does not differ from a random model. Variable response curves were also evaluated to understand the correlations between variables and predicted distributions (Figure S2). The entire pipeline can be found under Data Availability Statement.

Data availability

All data, including used scripts, can be found on our GitHub repository, link: https://github.com/wouterbeukema/Wouters_et_al_MosquitosBonaire.

Received: 19 February 2024; Accepted: 14 June 2024

Published online: 20 August 2024

References

1. Wilkerson, R. C., Linton, Y.-M. & Strickman, D. *Mosquitoes of the World* (John Hopkins University Press, 2020).
2. Kraemer, M. U. G. *et al.* Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* **4**, 854–863 (2019).
3. Caminade, C. *et al.* Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 119–124 (2017).
4. Kraemer, M. U. G. *et al.* The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife* **4**, 1–18 (2015).
5. Medley, K. A. Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models. *Glob. Ecol. Biogeogr.* **19**, 122–133 (2010).
6. Lounibos, L. P. Invasions by insects vectors of human disease. *Annu. Rev. Entomol.* **47**, 233–266 (2002).
7. Ding, F., Fu, J., Jiang, D., Hao, M. & Lin, G. Mapping the spatial distribution of *Aedes aegypti* and *Aedes albopictus*. *Acta Trop.* **178**, 155–162 (2018).
8. Brady, O. J. *et al.* Global temperature constraints on *Aedes aegypti* and *Ae. albopictus* persistence and competence for dengue virus transmission. *Parasit. Vectors* <https://doi.org/10.1186/1756-3305-7-338> (2014).
9. Campbell, L. P. *et al.* Climate change influences on global distributions of dengue and chikungunya virus vectors. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 1–9 (2015).
10. Lafferty, K. D. The ecology of climate change and infectious diseases. *Ecology* **90**, 888–900 (2009).
11. Reinhold, J. M., Lazzari, C. R. & Lahondère, C. Effects of the environmental temperature on *Aedes aegypti* and *Aedes albopictus* mosquitoes: A review. *Insects* <https://doi.org/10.3390/insects9040158> (2018).
12. Stresman, G. H. Beyond temperature and precipitation: Ecological risk factors that modify malaria transmission. *Acta Trop.* **116**, 167–172 (2010).
13. Schrama, M. *et al.* Human practices promote presence and abundance of disease-transmitting mosquito species. *Sci. Rep.* **10**, 1–6 (2020).
14. Zhang, Y., Bi, P. & Hiller, J. E. Climate change and the transmission of vector-borne diseases: A review. *Asia. Pac. J. Public Health* **20**, 64–76 (2008).
15. Reiskind, M. H., Griffin, R. H., Janairo, M. S. & Hopperstad, K. A. Mosquitoes of field and forest: the scale of habitat segregation in a diverse mosquito assemblage. *Med. Vet. Entomol.* **31**, 44–54 (2017).
16. O'Brien, V. A. & Reiskind, M. H. Host-seeking mosquito distribution in habitat mosaics of southern great plains cross-timbers. *J. Med. Entomol.* **50**, 1231–1239 (2013).
17. Chandrasegaran, K., Lahondère, C., Escobar, L. E. & Vinauger, C. Linking mosquito ecology, traits, behavior, and disease transmission. *Trends Parasitol.* **36**, 393–403 (2020).
18. Barker, J. R. & MacIsaac, H. J. Species distribution models applied to mosquitoes: Use, quality assessment, and recommendations for best practice. *Ecol. Model.* **472**, 110073 (2022).
19. Cianci, D., Hartemink, N. & Ibáñez-Justicia, A. Modelling the potential spatial distribution of mosquito species using three different techniques. *Int. J. Health Geogr.* **14**, 1–10 (2015).
20. Khatchikian, C., Sangermano, F., Kendall, D. & Livdahl, T. Evaluation of species distribution model algorithms for fine-scale container-breeding mosquito risk prediction. *Med. Vet. Entomol.* **25**, 268–275 (2011).
21. Raes, N. & Ter Steege, H. A null-model for significance testing of presence-only species distribution models. *Ecography (Cop.)* **30**, 727–736 (2007).
22. Estallo, E. L. *et al.* Modelling the distribution of the vector *Aedes aegypti* in a central Argentine city. *Med. Vet. Entomol.* **32**, 451–461 (2018).
23. Johnson, T. L. *et al.* 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 (2017).
24. Hopperstad, K. A., Sallam, M. F. & Reiskind, M. H. Estimations of fine-scale species distributions of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in Eastern Florida. *J. Med. Entomol.* **58**, 699–707 (2021).
25. Li, Q. *et al.* Ecological niche modeling identifies fine-scale areas at high risk of dengue fever in the Pearl River delta, China. *Int. J. Environ. Res. Public Health* **14**, 1–13 (2017).
26. Lippi, C. A. *et al.* Trends in mosquito species distribution modeling: Insights for vector surveillance and disease control. *Parasites Vectors* **16**, 1–17 (2023).
27. Belkin, J. N., Heinemann, S. J. & Page, W. A. The Culicidae of Jamaica (Mosquito Studies. XXI). *Contrib. Am. Entomol. Inst.* **6**, 1–458 (1970).
28. Van der Beek, J. G. *et al.* Taxonomy, ecology and distribution of the mosquitoes (Diptera: Culicidae) of the Dutch Leeward Islands, with a key to the adults and fourth instar larvae. *Contrib. Zool.* **89**, 373–392 (2020).
29. Van der Kuyp, E. Mosquitoes of the Netherlands Antilles and their hygienic importance. *Stud. Fauna Curaçao Other Caribb. Islands* **5**, 37–114 (1954).
30. Bonne, C. & Bonne-Wepster, J. *Mosquitoes of Surinam. A Study of Neotropical Mosquitoes* (Het Instituut, 1925).
31. Cova-García, P., Sutil, E. & Rausseo, J. A. *Mosquitos de Venezuela* Vol. II (Ministerio de Sanidad y Asistencia Social, 1966).
32. Cova-García, P., Sutil, E. & Rausseo, J. A. *Mosquitos de Venezuela* Vol. I (Ministerio de Sanidad y Asistencia Social, 1966).
33. Verdonschot, P. F. M. & Besse-Lototskaya, A. A. Flight distance of mosquitoes (Culicidae): A metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologia* **45**, 69–79 (2014).
34. Elmore, C. M. Jr. & Schoof, H. F. Dispersal of *Aedes taeniorhynchus* Weidemann near Savannah. *Georgia. Mosquito News* **23**, 1–7 (1963).

35. Powell, J. R. & Tabachnick, W. J. History of domestication and spread of *Aedes aegypti*—A review. *Mem. Inst. Oswaldo Cruz* **108**, 11–17 (2013).
36. Belkin, J. N. *The Mosquitoes of the South Pacific (Diptera, Culicidae)* Vol. 1 and 2 (University of California Press, 1962).
37. Arnell, J. H. Mosquito studies (Diptera, Culicidae). A revision of the genus *Haemagogus*. *Contrib. Am. Entomol. Inst.* **10**, 1–174 (1973).
38. Chase, F. A. & Holthuis, L. B. Land and fresh water decapod Crustacea from The Leeward Group and Northern South America. *Stud. Fauna Curaçao Other Caribb. Islands* **3**, 21–28 (1948).
39. Lanzaro, G. C., Collier, T. C. & Lee, Y. Defining genetic, taxonomic, and geographic boundaries among species of the *Psorophora fonninis* (Diptera: Culicidae) complex in North and South America. *J. Med. Entomol.* **52**, 907–917 (2015).
40. Freitas, J. a De, Nijhof, B. S. J., Rojer, a C. & Debrot, a O. *Landscape Ecological Vegetation Map of Bonaire*. <http://www.dcbd.nl/sites/www.dcbd.nl/files/documents/Freitasetal2005LandscapeecologicalvegetationmapBonaire.pdf> (2005).
41. KNMI. KNMI: hour data Caribbean area. https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens_Caribisch_gebied (2023).
42. Mûcher, C. A. & Verweij, P. J. F. M. *Land Cover Classification Bonaire*. (2020).
43. Brisco, K. K., Cornel, A. J., Lee, Y., Mouatcho, J. & Braack, L. Comparing efficacy of a sweep net and a dip method for collection of mosquito larvae in large bodies of water in South Africa. *F1000Research* **5**, 1–9 (2016).
44. Van der Beek, J. G. *et al.* First record of a phlebotomine sand fly (Diptera: Psychodidae) on the Dutch Caribbean island of Curaçao. *Deinsea* **22**, 23–30 (2024).
45. Lane, J. *Neotropical Culicidae* Vol. 1 (University of São Paulo, 1953).
46. Lane, J. *Neotropical Culicidae* Vol. 2 (University of São Paulo, 1953).
47. Pratt, H. D. *Key to the Culicidae of Puerto Rico*. (1969).
48. Darsie, R. F., Taylor, D. S., Prusak, Z. A. & Verna, T. N. Checklist of the mosquitoes of the Bahamas with three additions to its fauna and keys to the adult females and fourth instars. *J. Am. Mosq. Control Assoc.* **26**, 127–134 (2010).
49. Burkett-Cadena, N. D. *Mosquitoes of the Southeastern United States* (The University of Alabama Press, 2013).
50. Dowle, M. & Srinivasan, A. data.table: Extension of 'data.frame'. R package version 1.14.8.
51. Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. dplyr: A Grammar of Data Manipulation. R package version 1.1.2.
52. Pierce, D. ncd4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R package version 1.22.
53. Hijmans, R. J. raster: Geographic Data Analysis and Modeling. R package version 3.6-26.
54. Bivand, R., Keitt, T. & Rowlingson, B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.6-7.
55. Bivand, R. & Rundel, C. rgeos: Interface to Geometry Engine—Open Source ('GEOS'). R package version 0.6-4.
56. Pebesma, E. Simple features for R: Standardized support for spatial vector data. *R J.* **10**, 439–446 (2018).
57. Bivand, R., Pebesma, E. & Gomez-Rubio, V. *Applied Spatial Data Analysis with R* 2nd edn. (Springer, 2013).
58. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.6-4.
59. Dray, S., Dufour, A. & Chessel, D. The ade4 package - II: Two-table and K-table methods. *R News* **7**, 47–52 (2007).
60. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species Distribution Modeling. R package version 1.3-14.
61. Broennimann, O., Di Cola, V. & Guisan, A. ecospat: Spatial Ecology Miscellaneous Methods. R package version 3.5.1.
62. Kass, J. M. *et al.* ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods Ecol. Evol.* **12**, 1602–1608 (2021).
63. Warren, D. & Dinnage, R. ENMTools: Analysis of Niche Evolution using Niche and Distribution Models.
64. Evans, J. S. & Murphy, M. A. spatialEco. R package version 2.0-1.
65. Van Oosterhout, L. *et al.* An integrated assessment of climate change impacts and implications on Bonaire. *Econ. Disasters Clim. Chang.* **7**, 147–178 (2023).
66. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model* **190**, 231–259 (2006).
67. Wisz, M. S. *et al.* Effects of sample size on the performance of species distribution models. *Divers. Distrib.* **14**, 763–773 (2008).
68. Valavi, R. *et al.* Modelling climate change effects on Zagros forests in Iran using individual and ensemble forecasting approaches. *Theor. Appl. Climatol.* **137**, 1015–1025 (2019).
69. Morales, N. S., Fernández, I. C. & Baca-González, V. MaxEnt's parameter configuration and small samples: Are we paying attention to recommendations? A systematic review. *PeerJ* **2017**, 1–16 (2017).
70. Low, B. W., Zeng, Y., Tan, H. H. & Yeo, D. C. J. Predictor complexity and feature selection affect Maxent model transferability: Evidence from global freshwater invasive species. *Divers. Distrib.* **27**, 497–511 (2021).
71. Velasco, J. A. & González-Salazar, C. Akaike information criterion should not be a “test” of geographical prediction accuracy in ecological niche modelling. *Ecol. Inform.* **51**, 25–32 (2019).
72. Bohl, C. L., Kass, J. M. & Anderson, R. P. A new null model approach to quantify performance and significance for ecological niche models of species distributions. *J. Biogeogr.* **46**, 1101–1111 (2019).
73. Feng, X., Park, D. S., Liang, Y., Pandey, R. & Papeš, M. Collinearity in ecological niche modeling: Confusions and challenges. *Ecol. Evol.* **9**, 10365–10376 (2019).
74. Rödder, D., Schmidlein, S., Veith, M. & Lötters, S. Alien invasive slider turtle in unpredicted habitat: A matter of niche shift or of predictors studied?. *PLoS One* <https://doi.org/10.1371/journal.pone.0007843> (2009).
75. Zeng, Y., Low, B. W. & Yeo, D. C. J. Novel methods to select environmental variables in MaxEnt: A case study using invasive crayfish. *Ecol. Modell.* **341**, 5–13 (2016).
76. Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C. & Guisan, A. Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions. *Glob. Ecol. Biogeogr.* **26**, 275–287 (2017).
77. Jueterbock, A. 'MaxentVariableSelection' vignette. <https://cran.r-project.org/web/packages/MaxentVariableSelection/vignettes/MaxentVariableSelection.pdf> (2016).
78. Shabani, F., Kumar, L. & Ahmadi, M. Assessing accuracy methods of species distribution models: AUC, Specificity, Sensitivity and the True Skill Statistic. *Glob. J. Hum. Soc. Sci.* **18**, 6–18 (2018).
79. Konowalik, K. & Nosol, A. Evaluation metrics and validation of presence-only species distribution models based on distributional maps with varying coverage. *Sci. Rep.* **11**, 1–15 (2021).
80. Fielding, A. H. & Bell, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **24**, 38–49 (1997).
81. Yackulic, C. B. *et al.* Presence-only modelling using MAXENT: When can we trust the inferences?. *Methods Ecol. Evol.* **4**, 236–243 (2013).
82. Merow, C., Smith, M. J. & Silander, J. A. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography (Cop.)* **36**, 1058–1069 (2013).
83. Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Modell.* **199**, 142–152 (2006).

Acknowledgements

We would like to thank the team of the Vector Control Unit from the Bonaire Public Health Department for their assistance and expertise in the sample location decisions and field sampling. We would like to thank *Stichting*

Nationale Parken (STINAPA) Bonaire, in particular Jilly Sarpong and Monique Grol, for granting access to the parks. Furthermore, we would like to acknowledge Merijn Moens for troubleshooting during the modelling process and Jory van Thiel for reviewing and commenting the manuscript.

Author contributions

Conceptualisation: R.M.W., W.B., J.G. van der B., M.A.H.B., M.S. Funding acquisition: M.A.H.B., J.G.v.d.B., K.B. Field work, data collection and identification: R.M.W., J.G.v.d.B., P.H., F.S., J. van S., A.S., M.A.H.B., M.S. Modelling and analyses: R.M.W., W.B. Visualisation: R.M.W. Writing—original draft: R.M.W. All authors have read, contributed and agreed to the published version of the manuscript.

Funding

The mosquito diversity expedition to the Dutch Caribbean was coordinated by the MOBOCON program (Capacity building program for MOsquitoBORne disease CONtrol through sustainable integrated mosquito management in Dutch Caribbean) from the RIVM, National Institute for Public Health and Environment which was funded by the Ministry of Health, Welfare and Sport, The Netherlands. Additionally, this project was supported the Pandemics and Disaster Preparedness Center (PDPC) as part of the frontrunner project 'Climate Change and Vectorborne Virus Outbreaks.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-64948-y>.

Correspondence and requests for materials should be addressed to R.M.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024