



Evaluation of species distribution model algorithms for fine-scale container-breeding mosquito risk prediction

C. KHATCHIKIAN^{1*}, F. SANGERMANO^{2†}, D. KENDELL³
and T. LIVDAHL¹

¹Department of Biology, Clark University, Worcester, MA, U.S.A., ²Clark Laboratories, Clark University, Worcester, MA, U.S.A. and ³Bermuda Ministry of Health, Paget, Bermuda, U.K.

Abstract. The present work evaluates the use of species distribution model (SDM) algorithms to classify high densities of small container-breeding *Aedes* mosquitoes (Diptera: Culicidae) on a fine scale in the Bermuda Islands. Weekly ovitrap data collected by the Department of Health, Bermuda for the years 2006 and 2007 were used for the models. The models evaluated included the algorithms Bioclim, Domain, GARP (genetic algorithm for rule-set prediction), logistic regression and MaxEnt (maximum entropy). Models were evaluated according to performance and robustness. The area under the receiver operating characteristic curve was used to evaluate each model's performance, and robustness was assessed according to the spatial correlation between classification risks for the two datasets. Relative to the other algorithms, logistic regression was the best and MaxEnt the second best model for classifying high-risk areas. We describe the importance of covariables for these two models and discuss the utility of SDMs in vector control efforts and the potential for the development of scripts that automate the task of creating risk assessment maps.

Key words. *Aedes*, risk prediction, species distribution models, SDMs, Bermuda Islands.

Introduction

Because *Aedes* mosquitoes are important vectors of human arboviruses (Beatty & Aitken, 1979; Dohm *et al.*, 1995; Mitchell, 1995), the accurate determination, at multiple scales, of areas that are prone to breed high densities of mosquitoes is critical to the development of control and mitigation strategies. Public health officers in many countries conduct surveys and sampling programmes that allow them to direct their resources efficiently to protect the public. Different types of traps and collection devices, designed to collect mosquitoes at different life stages, provide information on species presence and densities. Additional methods that detect the presence of specific human pathogens in mosquito populations, such as polymerase chain reaction-based technologies (Porter *et al.*,

1993; Hadfield *et al.*, 2001; Shi *et al.*, 2001), could be combined to provide critical public health information.

Ovitrap sampling allows the collection of eggs from small container-breeding mosquitoes and has been commonly used to monitor *Aedes* populations (e.g. Evans & Bevier, 1969; Lee, 1992; Dhang *et al.*, 2005; Morato *et al.*, 2005; Kaplan *et al.*, 2010). Data obtained with this sampling scheme are often used as proxy estimators for the presence, activity or size of mosquito populations in the vicinity of traps. This methodological approach allows for the collection of large amounts of data with relatively low effort. However, several sources of error can remain, including the inhibition of oviposition by the presence of chemical clues in the water, habitat selection (including attraction to preferable habitats near the trap), or multiple ovitrap oviposition events. The data

Correspondence: Todd Livdahl, Department of Biology, Clark University, 950 Main Street, Worcester, MA 01610, U.S.A. Tel.: +1 508 793 7514; Fax: +1 508 793 7174; E-mail: tlivdahl@clarku.edu

*Present address: Department of Biology, University of Pennsylvania, Philadelphia, PA, U.S.A.

†These authors contributed equally to this paper.

obtained with ovitraps, as with any other sampling procedure, are subject to stochastic variation and may make the production of a comprehensive risk assessment map challenging.

Several different species distribution models (SDMs) are commonly used to produce coarse-scale risk prediction maps (i.e. on continental, national and state or provincial scales) and an extensive list of examples can be found in the literature. For example, Benedict *et al.* (2007) used an ecological modelling algorithm [genetic algorithm for rule-set prediction (GARP)] to predict worldwide *Aedes albopictus* risk and Foley *et al.* (2010) used two different algorithms to predict the distribution of *Anopheles* spp. (Diptera: Culicidae) in the Republic of Korea. Species distribution models use presence, presence/absence or quantitative data for the species of interest to produce predictions based on a combination of geographically referenced climatic, biological, demographic and/or physical data. Similar procedures are seldom used to produce comparable predictions on a very fine scale (i.e. with pixel sizes of <100 m) for small study areas (i.e. areas of <100 km²), perhaps because the availability of environmental variables at such fine scales is scarce, their degree of variability over short distances is small and species information at a very fine scale that could successfully account for microhabitat variation is lacking.

Two of the most commonly used presence data-only algorithms, MaxEnt (maximum entropy) (Phillip *et al.*, 2004) and GARP (Stockwell & Peters, 1999), have often shown accurate prediction capabilities in simulations and evaluations, outperforming classical modelling approaches, such as those of Domain, Bioclim and logistic regression (Phillip *et al.*, 2004; Hijmans & Graham, 2006; Phillip & Dudík, 2008; see also Stockman *et al.*, 2006). These algorithms differ in their rationales and procedures. Briefly, MaxEnt finds the maximum entropy probability distribution that agrees with the provided presence data based on environmental data; a large literature describes in detail the underlying MaxEnt algorithms and rationale (e.g. Phillip *et al.*, 2004, 2006). By contrast, GARP includes multiple, non-deterministic iterative procedures that incorporate various model distribution methods such as logistic regression and range envelopes, producing with each run predicted binary maps of presences and absences. Multiple optimal models are produced for each dataset, which can be converted into presence likelihoods. The three additional algorithms mentioned above, Domain (multivariate distance metric algorithm), Bioclim (envelop algorithm) and logistic regression (which considers both presences and absences) are less complex than the former two, and often perform poorly in simulations (Phillip *et al.*, 2004; Wisz *et al.*, 2008). Different algorithms produce different outputs, but in general convey presence probabilities or some arbitrary value that can be interpreted in a similar fashion.

The Bermuda Islands, an archipelago located in the Atlantic Ocean off the east coast of the U.S.A. (32°14'–32°24' N, 64°39'–64°53' W), have a subtropical climate, with mild winters and hot, humid summers, which provides suitable conditions for *Aedes* mosquitoes. Historically, these mosquitoes have been responsible for extensive outbreaks of yellow and dengue fevers. No vector-borne diseases have been recently reported, despite the presence of *Aedes* mosquitoes in the

area. However, the Health Department of Bermuda maintains an aggressive mosquito monitoring and control programme to prevent potential health risks found in similar locations and to reduce the biting nuisance that may affect the local economy.

In this study, we aim to assess the performance of commonly used SDM algorithms to classify areas that are prone to supporting high densities of small container-breeding mosquitoes. Specifically, we aim initially to evaluate the performance of the Bioclim, Domain, GARP, logistic regression and MaxEnt algorithms at a very fine scale in the Bermuda Islands. Subsequently, we aim to identify environmental covariates that contribute to high mosquito prevalence at this fine scale.

Materials and methods

The small size of the Bermuda Islands (<54 km²) makes the application of SDMs for the generation of risk maps a challenging task, as typical bioclimatic variables cannot be used as a result of the lack of spatial variability in the region. However, the region seems suitable for the empirical determination of infestation risk, given: (a) the presence of an extensive weekly ovitrap programme [see Kaplan (2006) for a comprehensive description of the sampling programme]; (b) the almost exclusive presence of a single small container-breeding mosquito species, *Ae. albopictus* (Skuse), with the marginal and occasional presence of *Aedes aegypti* (L.) (Kaplan *et al.*, 2010), and (c) the availability of mosquito records for two consecutive years in which the population appears to remain constant (see below), which allows area classifications to be compared between datasets as a measure of the robustness of the algorithms.

The selection of environmental data layers was based on availability and *a priori* expectation of influences in the mosquito population. Distance to buildings, distance to roads and human population densities were selected as proxies of human influence on the mosquito population because human activities provide both breeding habitats (artificial containers and other breeding grounds) and dispersal opportunities (through the movement of containers colonized by eggs or larvae). Elevation and slope were selected in consideration of their influence on water accumulation. Slope was also presumed to influence the access of cleaning crews to steep areas. Distance to shore was selected to consider seawater effects, such as salt spray, and wind exposure. Aspect was selected to consider effects of solar irradiation, wind incidence and their potential effects on egg desiccation. Some variables that are commonly used in SDMs, including primary productivity, temperature, precipitation and moisture, were not used in our study. There is low variation in these variables because of the small size of the islands (see above).

The same ovitrap data and environmental variables were used in all models (elevation, slope, aspect, distance to buildings, distance to shore, distance to roads, human population). The Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) with a spatial resolution of 90 m was used for elevation data; the DEM was resampled based on a bilinear interpolation in order to match the 45-m resolution

of the remaining layers. Slope and aspect layers were derived from the DEM. Aspect was rescaled to an index of southwestness, using the cosine function (aspect degrees, 225°) (Franklin *et al.*, 2000), in which values range from 1 (representing southwest) to -1 (representing northeast). This measure was considered appropriate in view of the islands' general orientation along a southwest–northeast axis. Layers representing distance to buildings, distance to roads and distance to shore, as well as human population according to public vote registration records, were also included as variables. All geographic information system (GIS) processing to obtain the layers described above was performed using IDRIS Version 16.05 Taiga edition (Eastman, 2010).

Weekly presence data for 292 ovitraps from 2006 to 2007 were used in all models. The prevalence values [frequency of positive ovitraps: 2006, 0.12 ± 0.02 standard error (SE); 2007, 0.11 ± 0.02 SE] of *Aedes* eggs for these years appeared stable and were found not to differ significantly (Kaplan *et al.* 2010). In order to classify ovitraps as positive for high density, we selected those that presented values higher than the mean (i.e. disturbance criterion). Ovitrap were scored as positive (presence data point for the SDMs) for high mosquito density if eggs were detected in ≥ 6 weeks in each year, which represents values higher than the mean for both datasets (mean number of positive weeks: 2006, 5.89 ± 0.03 SE; 2007, 5.79 ± 0.03 SE). Ovitrap with lower values were considered negative and included as absences for logistic regression and receiver operating characteristic (ROC) calculations (see below).

The 292 ovitraps used in this study were deployed by the Bermuda Ministry of Health. In order to perform statistical comparisons and validations, data were partitioned into training sets consisting of 75% of the observations (used to develop the prediction models) and testing sets consisting of 25% of the data (used to evaluate the accuracy of the results). The process was performed in DIVA-GIS Version 5.2.0.2 (Hijmans *et al.*, 2001) with 30 repetitions in order to obtain 30 training–testing subsamples for each dataset. Each modelling algorithm was run independently with each of the 30 subsamples. The algorithms used were Bioclim (Nix & Busby, 1986), as implemented in DIVA-GIS, Domain, as implemented in DIVA-GIS, GARP Version 1.1.6 (Stockwell & Peters, 1999), logistic regression, as implemented in IDRIS Version 16.05 module LOGREG (Eastman, 2010), and MaxEnt Version 3.3.1 (Phillip *et al.*, 2004).

MaxEnt was run using linear, quadratic, product, threshold and hinge features, with output set to logistic (Phillip & Dude, 2008). GARP was run with optimization parameters set to 100 runs with a convergence limit of 0.005 and a maximum of 1000 iterations. All four rule types (atomic, range, negated range, logistic regression) were used. The best subset option was enabled to select the 10 best models (omission threshold = 20%, commission threshold = 50%, 20 total models under hard omission threshold). The 10 best models were imported into IDRIS Version 16.05 and converted into probabilities. Bioclim and Domain were run using the default 0.025 percentile cut-off level. The logistic regression algorithm also requires absence values (i.e. negative presence points) and these were included in its input file.

Because many algorithms produce dimensionally different outputs, the outputs from each run were rescaled to values of 0–1 by dividing each pixel by the maximum pixel value. The performance of each model predicting yearly presences was evaluated using the area under the ROC curve (AUC), using the online tool JLABROC4 (www.jrocf.it.org). The AUC can range from 0 to 1; a value of 1 indicates perfect model agreement, a value of 0.5 indicates agreement equal to chance, and 0 indicates complete disagreement. Differences among AUC values obtained for each subsample with each algorithm were tested by ANOVA. A subsequent Tukey–Kramer honestly significant differences (HSD) procedure was used to assess pairwise differences. The AUC values obtained for the 2007 dataset failed a normality goodness-of-fit test (Shapiro–Wilk test, $W = 0.98$, $P < 0.05$) and were transformed ($y' = -1/y$) to achieve normality. In order to check the robustness of algorithms, a spatial correlation between the classification predictions for 2006 and 2007 for each subsample and each algorithm was performed using IDRIS (REGRESS module). The obtained statistic failed the normality goodness-of-fit test (Shapiro–Wilk test, $W = 0.88$, $P < 0.05$) and no satisfactory transformation for normality was found. For pairwise comparisons, bootstrap procedures were used to produce 95% confidence intervals (CIs) after 10 000 replicates using S-plus Version 8.0.4 (Insightful Corp, 2007). Non-overlapping CIs were considered statistically different. Unless stated otherwise, statistical analyses were performed using JMP Version 7.0 (SAS Institute, 2007).

To obtain classification risk maps, we used the complete datasets for both years and the AUC values as goodness-of-fit indicators for the two models that performed best according to the AUC criterion described above. We compare the overall performance of the algorithms based on classification risk maps and AUCs.

Results

Model performance

The five models performed differently with both datasets based on the 30 subsamples (ANOVA: 2006 dataset, $F_{149} = 49.9$, $P < 0.001$; 2007 dataset, $F_{149} = 75.5$, $P < 0.001$). In terms of the overall performance of the algorithms in each dataset, two models surpassed the rest: logistic regression and MaxEnt. For the 2006 dataset, logistic regression performed best, followed by MaxEnt. For the 2007 dataset, both logistic regression and MaxEnt performed better than any other model. For both datasets, Bioclim presented an intermediate performance, whereas GARP and Domain presented consistently low performances. Figure 1 shows the overall performance and pairwise statistical differences for all models for both datasets. Examination of the best performance among the 30 subsamples for each algorithm presented a similar pattern: logistic regression and MaxEnt showed the highest AUC values. Table 1 summarizes the performance of each model.

Using both complete datasets, the goodness-of-fit for each of the two best models, logistic regression and MaxEnt, indicates acceptable performances for each dataset (Fig. 2).

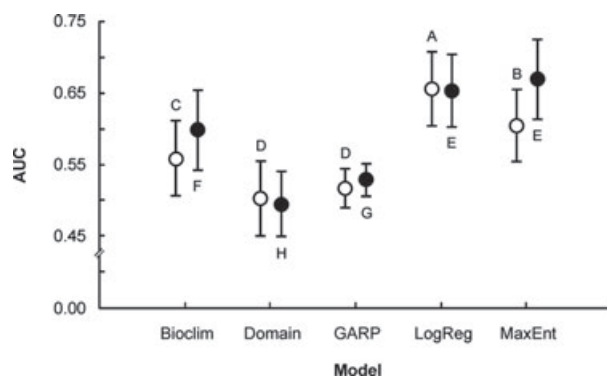


Fig. 1. Performance [area under the curve (AUC)] of the five algorithms evaluated (MaxEnt, logistic regression, GARP, Domain, Bioclim) with the two different datasets based on the 30 training–testing subsamples. ○, mean values for 2006; ●, mean values for 2007; error bars, 1 standard deviation. Markers not connected by same letter are significantly different [Tukey–Kramer honestly significant differences (HSD) test, $\alpha = 0.05$].

Table 1. Algorithm performance with the 30 training–testing subsamples for the two datasets, based on the area under the receiver operating characteristic curve (AUC).

Model	Mean AUC		Median AUC		Best model AUC	
	2006	2007	2006	2007	2006	2007
Bioclim	0.56	0.60	0.57	0.60	0.68	0.69
Domain	0.50	0.49	0.51	0.49	0.58	0.58
GARP	0.52	0.53	0.53	0.53	0.56	0.60
Logistic regression	0.66	0.65	0.76	0.65	0.81	0.75
MaxEnt	0.61	0.67	0.76	0.68	0.75	0.76

AUC values: 1 indicates a perfect model agreement; 0.5 indicates agreement equal to chance; 0 indicates complete disagreement.

Model robustness

The algorithms showed high robustness (all algorithms presented mean coefficients of correlation of $>82\%$), which is consistent with high spatial correlation between the predicted risk for the two consecutive years. Bioclim and Domain presented the highest agreement between risk classifications across the years, with coefficient of regression values between the two datasets of >0.96 . Interestingly, the two algorithms that performed best according to the AUC criterion, MaxEnt and the logistic regression approach, showed lower robustness values than Bioclim and Domain, and the model with the lowest robustness across dates was GARP (Fig. 3). This suggests that the predicted risk surfaces produced with these algorithms were slightly different for each yearly dataset.

Importance of variables

For the logistic regression algorithm, four variables stand out as most important in both datasets: distance to shore; distance to roads; distance to buildings, and elevation.

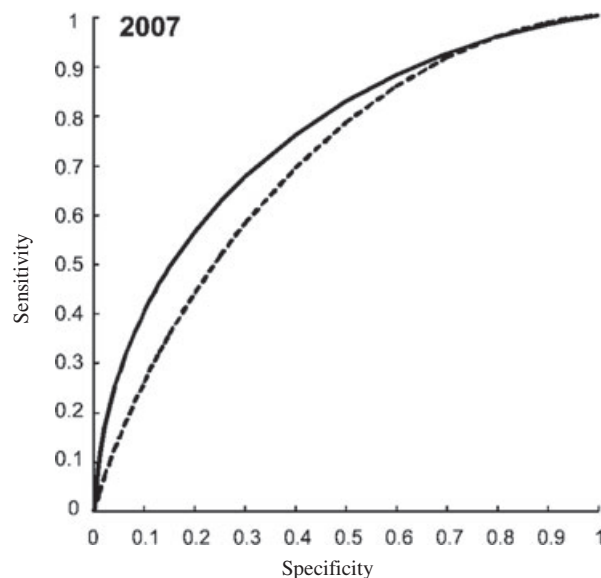
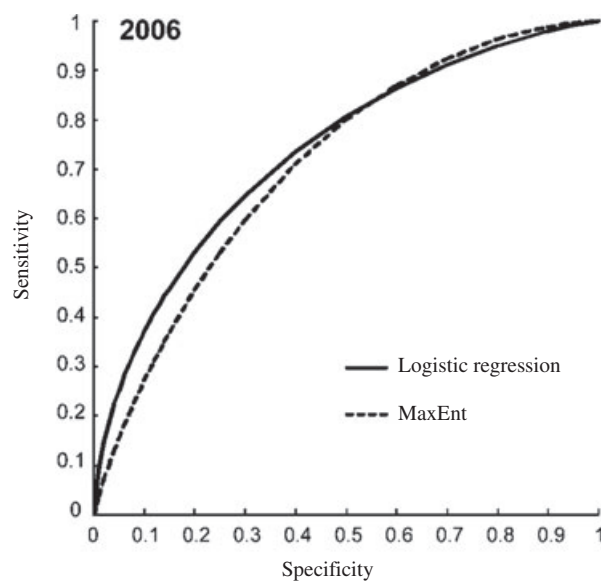


Fig. 2. Goodness-of-fit of the two best performing models using the complete datasets for 2006 and 2007: logistic regression [area under the curve (AUC): 2006, 0.71; 2007, 0.70] and MaxEnt (AUC: 2006, 0.74; 2007, 0.76).

However, predictions across the different subsamples vary widely (Fig. 4A, C). For MaxEnt, identifying the importance of the respective variables is more difficult. The variables that contribute most to the model are distance to shore for the 2006 dataset (Fig. 4B) and slope for the 2007 dataset (Fig. 4D). Calculations of the contributions of variables in MaxEnt are sensitive to correlations between variables. If two variables are important, MaxEnt will assign a large contribution to one of them and a low contribution to the other. This characteristic of MaxEnt may be responsible for the variability observed between contributions in 2006 and 2007. Although the correlation between distance to shore and

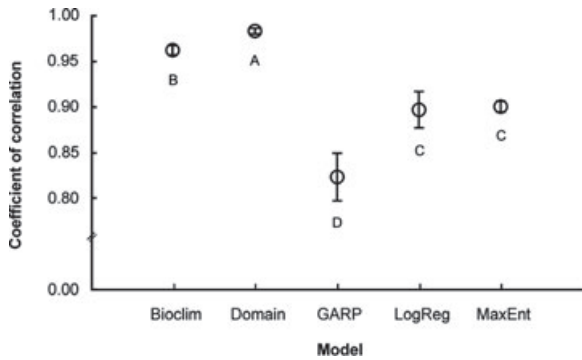


Fig. 3. Coefficients of correlation between the risk classifications of the five algorithms evaluated (MaxEnt, logistic regression, GARP, Domain, Bioclim) with the two datasets based on the 30 training-testing subsamples. ○, mean values for 10 000 replicates; bars indicate 95% bootstrap confidence intervals (CIs) for 10 000 replicates. Symbols marked by same letter indicate overlapping CIs.

slope is low ($r = -0.18$), it seems large enough to affect MaxEnt determination of weights. This can be corroborated by observing the individual importance of each variable to the

model (MaxEnt jack-knife approach; Table 2); the variables that contribute the most alone (when run with that variable in isolation) are elevation, distance to shore, slope and distance to buildings. However, they all have redundant information (model gain does not vary when the variable is excluded) and therefore contribute quite diversely across models. Distance to roads has the least redundant information and, if excluded, is the variable that affects model performance the most. However, this variable is not important when used alone, which suggests the presence of interactions between distance to roads and other variables that improve model performance.

Overall, three variables (distance to shore, elevation and slope) vary in their contributions across the different subsamples, with high importance values in some, but not in others. By contrast, aspect and population consistently have very low importance (Fig. 4).

Risk assessment

The risk maps produced from all the points with the two better-performing models, logistic regression and MaxEnt, differ in their classification of risk areas. Both models agreed

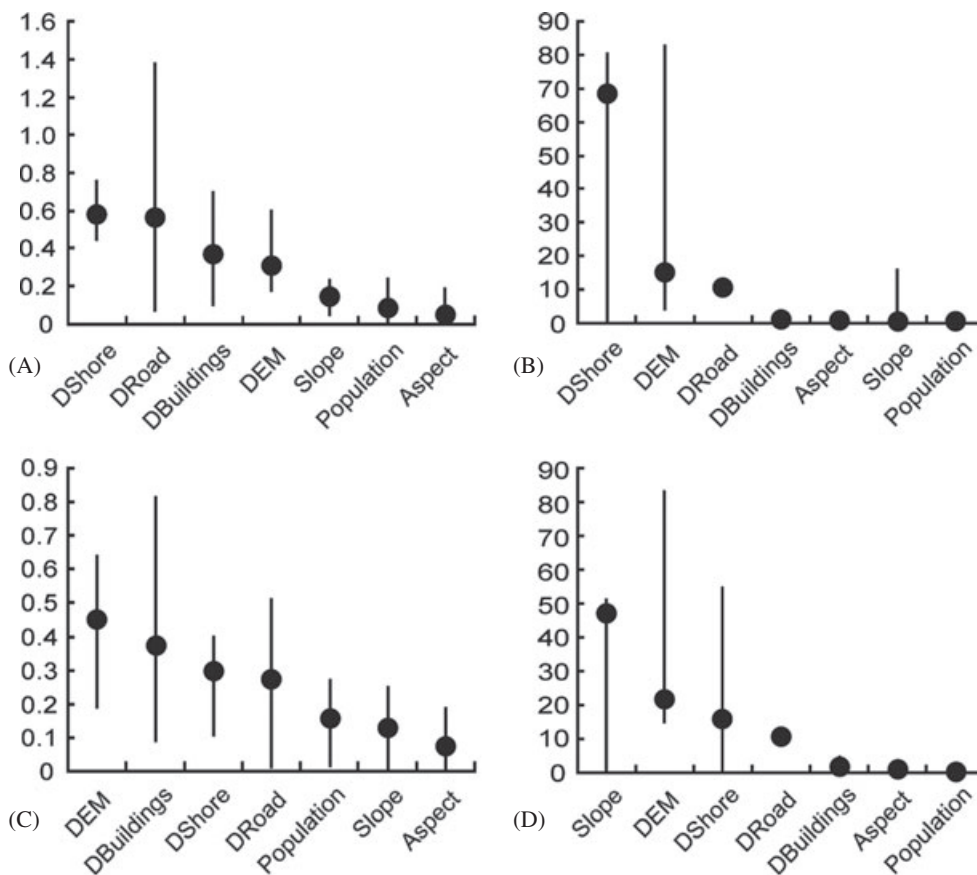


Fig. 4. Importance of environmental variables (covariates) for the two best models: logistic regression and MaxEnt. (A) Logistic regression for the 2006 dataset. (B) MaxEnt for the 2006 dataset. (C) Logistic regression for the 2007 dataset. (D) MaxEnt for the 2007 dataset. Importance is expressed in terms of regression coefficients for logistic regression and as a percentage for MaxEnt. ●, median values; lines represent range values. DEM, digital elevation model; DBuildings, distance to buildings; DShore, distance to shore; DRoad, distance to road.

Table 2. MaxEnt jack-knife of variable importance.

	2006		2007	
	With only	Without	With only	Without
DShore	2.00	2.39	1.91	2.46
DEM	1.99	2.42	2.00	2.44
Slope	1.82	2.45	1.86	2.44
DBuildings	1.62	2.43	1.81	2.44
Aspect	1.43	2.44	1.44	2.45
Population	0.50	2.45	0.56	2.47
DRoad	0.01	2.16	0.01	2.17

The values represent the training gains when the variable is used in isolation ('with only') and when the variable is excluded ('without'). Gain represents the fit between MaxEnt's probability distribution and the distribution of the sample observation data. A variable that contributes useful information to the model will have a high gain when used in isolation. A variable that contributes unique information will have the gain reduced when it is excluded from the model. Gain with all variables equals 2.58 for 2006 and 2.61 for 2007.

DShore, distance to shore; DEM, digital elevation model; DBuildings, distance to buildings; DRoad, distance to road.

on the classification of some areas as low risk (dark areas in Fig. 5); these areas included the airport, a large nature reserve and small islands. Interestingly, these areas share some general characteristics, including closeness to shore, low elevation, low slope and low population, in addition to a low prevalence of human-related features such as roads and buildings.

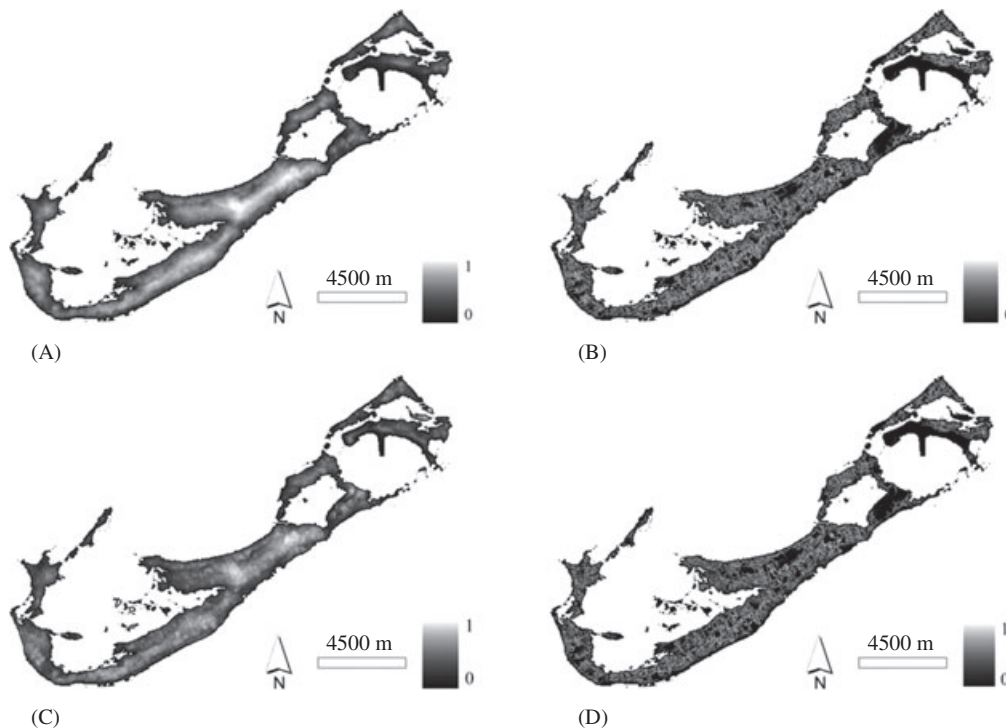


Fig. 5. Relative risk prediction maps produced by logistic regression (A, C) and MaxEnt (B, D) for 2006 (A, B) and 2007 (C, D). Lighter shading indicates areas at higher risk.

Discussion

Several modelling strategies are often used to produce vector risk maps, but generally these approaches are implemented on a much coarser scale than that used here (e.g. Benedict *et al.*, 2007; Moffett *et al.*, 2007; Foley *et al.*, 2010). In this study, we explored a similar approach (i.e. the probability of mosquito presence) on a very fine scale, which allowed the modelling of risk in specific, restricted local areas. This modelling approach can produce especially advantageous results in the case of mosquitoes breeding in small containers, given that breeding habitats are indeed very small and are thus well below the resolution of satellite images that can be used in risk modelling exercises for other vector mosquitoes, such as marsh breeders.

Most algorithms analysed here produced better-than-random classifications (Table 1, Fig. 1), and some algorithms performed better than others, especially logistic regression. However, the overall classification performance of the algorithms could be considered as low, which may stem from the ecological situation. The ovitrap data reflect the stochasticity in females' oviposition events in discrete time intervals. Remarkably, two of the algorithms that produced low predictions (Bioclim and Domain) presented extremely high robustness values (>96%). It is possible that these algorithms closely follow independent variables that only partially explain the dependent variable and thus provide highly correlated predictions that fail to accurately predict the risk areas for container-breeding mosquitoes.

The differences in classification performance across dates detected in most models is puzzling and unexpected; no *a priori* observation or descriptive statistic suggested any significant difference between these two datasets. Nevertheless, the difference allows two interesting conclusions: firstly, that any modelling approach should use multiple datasets when possible, and, secondly, that the logistic regression approach was much less sensitive to this variation, presenting similar classification performance values according to the AUC criterion (Table 1, Fig. 1). However, some authors (Peterson *et al.*, 2007) have recently raised concerns about the extensively used AUC to assess the accuracy of ecological models, as two very different ROC curves can produce similar areas and, consequently, two very spatially different models may result in similar AUCs. Nevertheless, it has been pointed out that the criterion remains extremely useful for comparing the relative performances of different models (Wisz *et al.*, 2008); further evaluations, beyond the scope of this work, are necessary to address this issue.

All factors considered, the logistic regression algorithm, as implemented in IDRISI, provides the best modelling approach for mapping areas at risk for mosquito infestation in Bermuda. Logistic regression presented the higher AUC values for both datasets, with high consistency between years in terms of both the visual examination (Fig. 5B, D) and the correlation between the classification areas (Fig. 3). Moreover, a comparison of the areas classified as prone to high densities of mosquitoes with those of the closest competing model, MaxEnt, suggests that the logistic regression algorithm produces a more realistic classification. The MaxEnt classification appears to be biased by the presence of roads. Interestingly, distance to roads does not have high covariance in MaxEnt runs, nor an ample range among replicates (Fig. 4) (i.e. distance to roads does not make a high contribution in MaxEnt models, although the jack-knife analysis of variable importance suggests the possible presence of interactions between distance to roads and other variables, as mentioned above).

The algorithms evaluated in this study were selected for their simplicity and feasibility of use by mosquito control officials; most of these algorithms are available in standalone packages that require minimal data preparation once the environmental variables have been produced. Some recent or more elaborate algorithms, such as GAM (generalized additive models), GDM (generalised dissimilarity modelling) and MARS (multivariate adaptive regression splines), have not been evaluated here, either because they require more complex scripting and the use of advanced statistical languages, or because they had not been released for public use at the time our evaluations were performed, as in the cases of OM-GARP and LIVES (Elith *et al.*, 2006; Wisz *et al.*, 2008). It seems feasible to implement an automatic script within IDRISI (i.e. by using the macro functionality) that could take data obtained by the monitoring programmes of control agencies at any point in the season and immediately produce risk maps. Such a procedure may allow for quicker responses and a more efficient use of vector control and prevention resources.

From a critical point of view, none of the algorithms presented here are able to produce extremely accurate classifications (Table 1), but this evaluation may help to promote

the development or improvement of specific algorithms suitable for this task. The only dynamic variables (i.e. those that change over time) used in the models were distance to roads, distance to buildings and population. Because these variables contributed the least to the models, we can consider that our results are static representations of risk and are therefore of limited use in analyses of future risk. Other dynamic variables that might influence risk on this fine scale, such as canopy density and evapotranspiration, micro-climate conditions and abundance of discarded small containers, were not included in the analyses because information on these variables at the spatial scale of this research was lacking. If spatially detailed micro habitat variables were available for the Bermuda Islands, the empirical models of best accuracy (logistic regression in this work) could be used to model future mosquito risk. There exist, however, limitations in the application of empirical species distribution models to predict future risk. Because the main purpose of the mathematical formulations of empirical models is to describe the distribution of the observations and not the underlying 'cause-effect' (Guisan & Zimmermann, 2000), these models may decrease in performance when projecting species distributions under future conditions. Decreased prediction accuracy can arise as a result of the dependence of inferred relationships on current conditions and sample data as a model may be trained under a combination of environmental conditions that may not exist in the future, or the species may adapt to combinations of variables that currently do not exist (Hijmans & Graham, 2006). Moreover, if the sample data used are spatially biased (e.g. as a result of accessibility), the relationships may be taken on only part of the species niche and therefore the accuracy of both current and future projections may be affected. Mechanistic models are considered superior for understanding the effect of climate on species distribution under the assumption of universal dispersal and absence of competition (Hijmans & Graham, 2006). This is because mechanistic models are based on knowledge of the physiology of the species and thus do not depend on sampling schemes. Although detailed physiological data are required to parameterize mechanistic models, they may produce better representations of future risk. Their applicability in the context of fine-scale production of risk for container-breeding mosquitoes should be evaluated in future research as they may represent useful tools for controlling and eradicating *Aedes* populations.

Acknowledgements

We thank Ross Furbert and other personnel of the Vector Control Department at the Bermuda Ministry of Health for their efforts in monitoring the traps that produced the data for this study. The National Institutes of Health (grant R15 AI062712-01 to TPL) and the Keck Foundation provided support for this project. We thank two anonymous reviewers for helpful comments.

References

- Beatty, J. & Aitken, T.H.G. (1979) *In vitro* transmission of yellow fever virus by geographic strains of *Aedes aegypti*. *Mosquito News*, **39**, 232–238.

- Benedict, M.Q., Levine, R.S., Hawley, W.A. & Lounibos, L.P. (2007) Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. *Vector-Borne Zoonotic Diseases*, **7**, 76–85.
- Dhang, C.C., Benjamin, S., Saranum, M.M., Fook, C.Y., Lim, L.H., Ahmad, N.W. & Sofian-Azirun, M. (2005) Dengue vector surveillance in urban residential and settlement areas in Selangor, Malaysia. *Tropical Biomedicine*, **22**, 39–43.
- Dohm, D.J., Logan, T.M., Barth, J.F. & Turell, M.J. (1995) Laboratory transmission of Sindbis virus by *Aedes albopictus*, *Ae. aegypti*, and *Culex pipiens* (Diptera: Culicidae). *Journal of Medical Entomology*, **32**, 818–821.
- Eastman, R. (2010) IDRISI Taiga Version 16.05. Clark Laboratories, Worcester, MA.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distribution from occurrence data. *Ecography*, **29**, 129–151.
- Eng, J. (2009) ROC analysis: web-based calculator for ROC curves. www.jrocf.it.org [Accessed December 2009].
- Evans, B.R. & Bevier, G.A. (1969) Measurement of field populations of *Aedes aegypti* with the ovitrap in 1968. *Mosquito News*, **29**, 347–353.
- Foley, D.H., Klein, T.A., Kim, H.C., Brown, T., Wilkerson, R.C. & Rueda, L.M. (2010) Validation of ecological niche models for potential malaria vectors in the Republic of Korea. *Journal of the American Mosquito Control Association*, **26**, 210–213.
- Franklin, J., McCullough, T. & Gray, C. (2000) Terrain variables used for predictive mapping of vegetation communities in Southern California. *Terrain Analysis: Principles and Applications* (ed. by J. Wilson & J. Gallant), pp. 331–353. John Wiley & Sons, New York, NY.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hadfield, T.L., Turell, M., Dempsey, M.P., David, J. & Park, E.J. (2001) Detection of West Nile virus in mosquitoes by RT-PCR. *Molecular Cellular Probes*, **15**, 147–150.
- Hijmans, R.J. & Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 1–10.
- Hijmans, R.J., Guarino, L., Cruz, M. & Rojas, E. (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Research News*, **127**, 15–19.
- Insightful Corporation (2007) *S-plus Version 8.0.4*. Insightful Corp., Palo Alto, CA.
- Kaplan, L. (2006) *Aedes aegypti* and *Aedes albopictus* in Bermuda: the spatial and temporal distribution from 2000–2005. MS Thesis. Clark University, Worcester, MA.
- Kaplan, L., Kendell, D., Robertson, D., Livdahl, T. & Khatchikian, C. (2010) *Aedes aegypti* and *Aedes albopictus* in Bermuda: extinction, invasion, invasion and extinction. *Biological Invasions*, **12**, 3277–3288.
- Lee, H.L. (1992) *Aedes* ovitrap and larval survey in several suburban communities in Selangor, Malaysia. *Mosquito-Borne Diseases Bulletin*, **9**, 9–15.
- Mitchell, C.J. (1995) The role of *Aedes albopictus* as an arbovirus vector. *Parasitology*, **37**, 109–113.
- Moffett, A., Shackelford, N. & Sarkar, S. (2007) Malaria in Africa: vector species' niche models and relative risk maps. *PLoS ONE*, doi:10.1371/journal.pone.0000824.
- Morato, V.C.G., Teixeira, M.G., Gomes, A.C., Bergamaschi, D.P. & Barreto, M.L. (2005) Infestation of *Aedes aegypti* estimated by oviposition traps in Brazil. *Revista Saúde Pública*, **39**, 553–558.
- Nix, H.A. & Busby, J.R. (1986) BIOCLIM: a bioclimate analysis and prediction system. Research Report No. 1983–85. Division of Water and Land Resources, Canberra, ACT.
- Peterson, A.T., Papes, M. & Eaton, M. (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and MaxEnt. *Ecography*, **30**, 550–560.
- Phillip, S.J. & Dudík, M. (2008) Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Phillip, S.J., Dudík, M. & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference of Machine Learning, Banff, AB, ACM International Conference Proceedings Series*, **69**, p. 8. New York, NY.
- Phillip, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distribution. *Ecological Modelling*, **190**, 231–259.
- Porter, K.R., Summers, P.L., Dubois, D. *et al.* (1993) Detection of West Nile virus by the polymerase chain reaction and analysis of nucleotide sequence variation. *American Journal of Tropical Medicine and Hygiene*, **48**, 440–446.
- SAS Institute (2007) *JMP Version 7.0*. SAS Institute, Inc., Cary, NC.
- Shi, P., Kauffman, E.B., Ren, P. *et al.* (2001) High-throughput detection of West Nile virus RNA. *Journal of Clinical Microbiology*, **39**, 1264–1271.
- Stockman, A.K., Beamer, D.A. & Bond, J.E. (2006) An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distributions*, **12**, 81–89.
- Stockwell, D. & Peters, D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Sciences*, **13**, 143–158.
- Wishart, E. (1999) Adult mosquito (Diptera: Culicidae) and virus survey in metropolitan Melbourne and surrounding areas. *Australian Journal of Entomology*, **38**, 310–313.
- Wisz, M.S., Hijman, R.J., Li, J. *et al.* (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Accepted 12 October 2010