



OPEN

Validating species distribution models to illuminate coastal fireflies in the South Pacific (Coleoptera: Lampyridae)

Laura N. Sutherland^{1,2✉}, Gareth S. Powell² & Seth M. Bybee^{2,3}

The coastal areas of Vanuatu are under a multitude of threats stemming from commercialization, human development, and climate change. *Atyphella* Olliff is a genus of firefly that includes species endemic to these coastal areas and will need protection. The research that has already been conducted was affected by accessibility due to the remote nature of the islands which left numerous knowledge gaps caused by a lack of distributional data (e.g., Wallacean shortfall). Species distribution models (SDM) are a powerful tool that allow for the modeling of the broader distribution of a taxon, even with limited distributional data available. SDMs assist in filling the knowledge gap by predicting potential areas that could contain the species of interest, making targeted collecting and conservation efforts more feasible when time, resources, and accessibility are major limiting factors. Here a MaxEnt prediction was used to direct field collecting and we now provide an updated predictive distribution for this endemic firefly genus. The original model was validated with additional fieldwork, ultimately expanding the known range with additional locations first identified using MaxEnt. A bias analysis was also conducted, providing insight into the effect that developments such as roads and settlements have on collecting and therefore the SDM, ultimately allowing for a more critical assessment of the overall model. After demonstrating the accuracy of the original model, this new updated SDM can be used to identify specific areas that will need to be the target of future conservation efforts by local government officials.

The results of rapid industrialization (e.g., climate change, deforestation and light pollution) have a significant impact worldwide. An ability to quickly document and assess biodiversity in neglected and/or unknown areas of the world is paramount to biodiversity sciences and conservation. This need is compounded for biodiversity in isolated or unique areas of the world, where the habitat of endemics can be threatened^{1,2}. A major challenge to biodiversity discovery, study, and potential conservation efforts is often the amount of area to cover, obtaining proper permits, limited funding, and time. There is an urgent need for tools that can focus biodiversity studies with little prior knowledge and/or input data.

Islands represent some of the most unique and unknown areas of the world while also being among the most, if not the most, fragile and threatened environments due to climate change, sea level rise, plastic pollution, and general human commercialization³. Islands also provide unique opportunities to investigate many biogeographical questions including fragmentation and biodiversity dynamics^{4–6}. The South Pacific is an excellent region for biogeography, biodiversity, and conservation studies due to the number of islands and amount of variation within and between the archipelagos^{7–9}. Specifically, Vanuatu is of interest because it is much younger (~2 million years old) than the surrounding island chains^{5,9}. Vanuatu is comprised of 80+ islands, mainly of volcanic origin, with an approximate area of 12,190 km² spanning 1300 km^{9,10}. Vanuatu has a relatively lower documented species diversity due to the young age, size, and isolation, but is home to several unique and rare endemics¹¹. These endemics are subject to potential dispersal barriers within the archipelago. The main barrier is the hypothesized biogeographic break, known as Cheesman's line, between the islands of Efate and Erromango, separating the flora and fauna of the northern and southern islands (Fig. 1)^{7,9,12}.

Vanuatu is an area of special interest to study coastline limited species due to the size, variation, and number of islands with varying coastline habitat. Across the four most diverse insect orders (Coleoptera, Hymenoptera,

¹Department of Entomology, Purdue University, 901 West State Street, West Lafayette, IN, USA. ²Department of Biology, Brigham Young University, 4102 LSB, Provo, UT, USA. ³Monte L. Bean Museum, Brigham Young University, Provo, UT, USA. ✉email: sutherll@purdue.edu

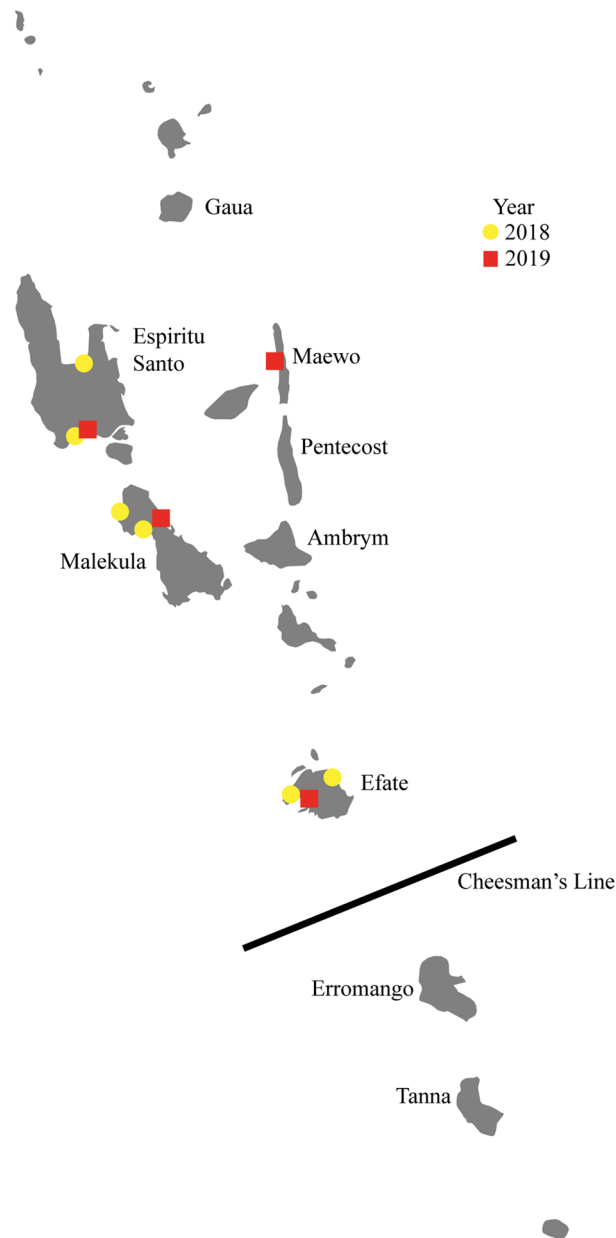


Figure 1. Distribution map showing occurrence data for *Atyphella* across Vanuatu based on field expeditions in 2018 and new locations from 2019. The islands explored over both years are labeled. Black line denotes the biogeographic break known as Cheesman's Line (Vemaps.com, Adobe, San Jose, CA, USA).

Lepidoptera, Diptera), accounting for almost 775,000 described species globally, species that are dependent on coast habitats are rare¹³. Specifically, within Coleoptera there are approximately 52 species from 13 families (<0.01%) that are considered marine^{14–16}. Doyen¹⁵ defined marine ranging from spending some time submerged by high tides to fully aquatic in the ocean. While many coastal species fall under this definition, we are defining coastal by restricting the definition of “marine” to those species that inhabit intertidal zone areas that are fully submerged during high tide.

Fireflies (Coleoptera: Lampyridae) are famous for their bioluminescence and there are ~2,250 described species¹⁷. Interestingly, there are several coastal species of firefly. *Micronaspis* Green has been collected from the Bahamas, Brazil, and Florida and there is a potential for a further intertidal species in Jamaica that remains unnamed^{14,18,19}. *Atyphella* Olliff included a single coastal species, *A. aphrogeneia* (Ballantyne) that was found in Papua New Guinea and specimens were later collected in Vanuatu expanding their range and information known about their habitat requirements^{14,20}.

In the S.E. Asia and Australopacific region there are ~222 species which are split across 28 genera in Luciolinae²¹, with only three species now known to be coastal and they are all members of *Atyphella*. Until 2018, *A. aphrogeneia* was the only coastal species within this genus. Two additional species endemic to Vanuatu

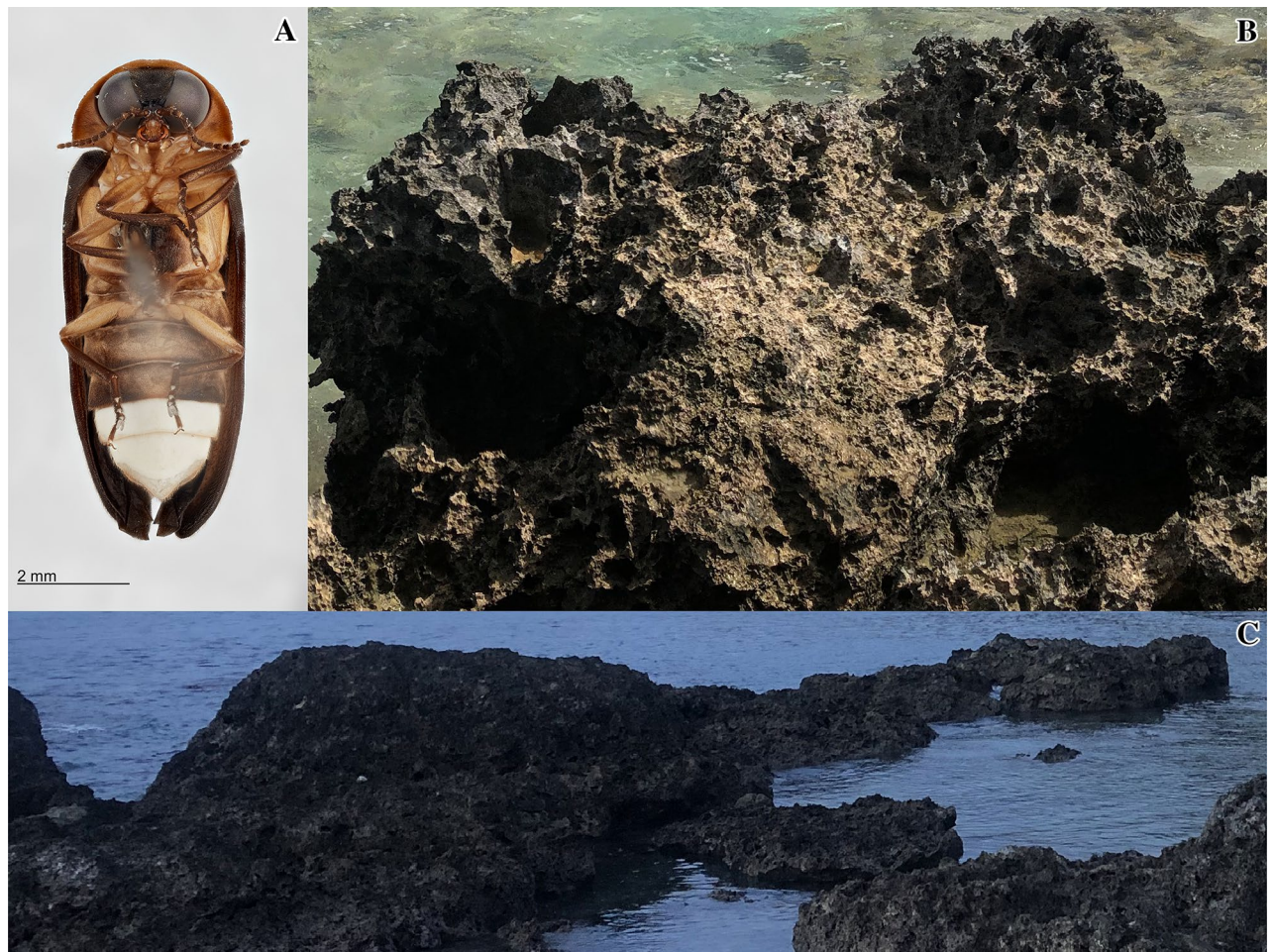


Figure 2. (A) Holotype of *A. maritimus*. (B,C) Photos of habitat, volcanic rock with extremely sharp and jagged edges lined the coast where *Atyphella* was collected. At high tide, the rocks were either mostly submerged or completely covered (Photo credit: (A,B) Colin Jensen and Natalie Saxton, (C) Saxton et al.²⁴).

were discovered: *A. maritimus* Saxton and Powell and *A. marigenous* Saxton and Bybee²². *Atyphella maritimus* has only been collected from islands in the Malampa and Loyalty Basin bioregion while *A. marigenous* has only been collected from Efate which sits in the East Shefa bioregion, both species are exclusively found north of Cheesman's line^{22–24}. Both *A. maritimus* and *A. marigenous* can be observed at twilight with adult males flying over the intertidal zone and the larvae and females being found in the pockets and pocks of the sharp volcanic rock (Fig. 2)^{14,24}. Larvae are often found below the water surface.

Species distribution models (SDM) of rare and endemic groups (i.e., *Atyphella*) are useful when there are logistic and/or financial constraints to survey specific areas for a species overall biodiversity (e.g., across Vanuatu). These models can function as powerful tools to direct field efforts to maximize the chance of locating, defining a range, and documenting biologically important data²⁵. MaxEnt is one such program that uses environmental and climate data to determine criteria for a species niche and extrapolates it to produce a geographic range with an associated probability for each area^{26–28}. MaxEnt is typically used for species level predictions, but it can be appropriate to use at a higher taxonomic level (e.g., genus) if all the species within that genus require the same or similar habitat conditions²⁹. It has become a popular tool because it is a presence/pseudo-absence program and is relatively accurate with small (> 5 observations) data sets^{30–33}. Despite the popularity of the tool, it is still rare to be able to subsequently validate a model with additional fieldwork, especially in an area as isolated as Vanuatu. Here, we provide results from two consecutive years of fieldwork, with the second heavily guided by the predictive model resulting from the first expedition.

Collecting efforts in both 2018 and 2019 resulted in potentially spatially biased localities because collecting areas were largely influenced by accessibility, cost, and infrastructure. When considering any sort of practical conservation, an important variable to consider is distributional data, especially with rare or endemic taxa³⁴. The Wallacean shortfall makes conservation efforts difficult in general, but the problem is exacerbated with under described taxa or by accessibility to remote areas^{35–38}. However, SDMs have been successful in conservation efforts within mammals^{39–41}, plants^{42–44}, and even insects^{34,45}.

MaxEnt has been shown to be successful with species that possess both aquatic and terrestrial characteristics²⁵. However, there is a large amount of variation within the coastal habitats of Vanuatu. Here we used additional

field work to test the predictive power of a previously produced SDM for a coastal endemic taxon. Additionally, an updated model is presented, and results of a subsequent bias analysis are provided to add further context to the models' interpretation.

Methods

Collection methods. Field expeditions were conducted in 2018 and 2019 on the islands of Ambrym, Efate, Erromango, Espiritu Santo, Gaua, Maewo, Malekula, Pentecost, and Tanna in search of *Atyphella* (Fig. 1). In 2018, specimens were found at six locations across three islands. These six presence locations were used for a preliminary MaxEnt analysis to predict areas to search for this endemic genus in the future²⁴. Islands selected to be sampled in 2019 were chosen based on the Saxton et al.²⁴ original model. The 2019 trip resulted in four additional localities from Efate, Espiritu Santo, Maewo, and Malekula being added to the original 2018 dataset. Only four of the islands were checked both years (Efate, Espiritu Santo, Malekula, and Tanna) which allowed for three new islands: Ambrym, Maewo, and Pentecost to be sampled in 2019. Three to five locations were checked on each of these new islands. These sites amounted to approximately five to seven kilometers of coastline on each island.

Ecological niche model. The 30 s bioclimatic variable files accessed from WorldClim^{46,47} and the base map were trimmed and converted to ASC files using DIVA⁴⁸. MaxEnt version 3.3.3k⁴⁹ was run with default settings and auto features. All of the bioclimatic variables available on WorldClim were included, and a jackknife was run to help handle and visualize the correlation between the variables. The jackknife was also used to measure the variable importance. A random selection from 10,000 background grid cells was used. The default prevalence was set to 0.5, the convergence threshold was set to 0.00001, the regularization multiplier was set to one, and cross validation was used in the repeated runs. The output format was cloglog and the model was evaluated using the threshold independent area under the receiver operating characteristic curve (AUC)^{26,50}. The AUC is a summary statistic to show how well a classifier, in this case a MaxEnt model, is at correctly distinguishing between a true presence and pseudo-absence location²⁶. Since MaxEnt is a presence only program, it uses a variation on this technique. Instead of using known absence locations, it uses randomly selected background points—any point within the study area not marked as present—also known as pseudo-absences^{26,51}. Therefore, the interpretation of the AUC is “the probability that a randomly chosen presence location is ranked higher than a randomly chosen background point”⁵⁰. Because it uses pseudo-absences the maximum AUC is always less than one²⁶. An AUC of 0.5 shows that it is no better at ranking the presence locations above the background locations^{26,51}. The higher the AUC value the better it is at minimizing either the false positives or negatives.

The WorldClim raster layers of the top contributing factors for the 2018 and 2019 models (Fig. 4) were viewed in QGIS Zanzibar 3.8⁵². The raster layers were then overlaid with occurrence points. The values associated with a known occurrence location were extracted using the identify features tool to show the variation within a factor across these locations in an attempt to explain the differences in the AUC values. All three (2018, 2019, combined) MaxEnt predictions were uploaded into QGIS to visualize how the areas predicted as suitable changed depending on different cut off values and the overlap between the predicted area and biasing factors (Fig. 7). In the cloglog output format the relative probability equates to the projected habitat suitability⁵³. The probabilities were binned into ten sections with the upper probability included in that bin (e.g., 0.2–0.3). The area considered suitable in each bin was then estimated using the raster calculator. These new binary raster layers were created from just the values within each of the bins. The unique values report was then used to provide a summary of all the presence and absence pixel counts for each bin. These pixel counts were used to find the total area and a proportion within the entire archipelago for each bin (Fig. 5, Supplementary Table S2).

Investigation of collecting bias. Shapefiles of potential biasing factors (e.g., roads and settlements) in Vanuatu were obtained from NextGIS (<https://data.nextgis.com/en/region/VU/>, data accessed March 2020). The bias estimation was run in RStudio 4.02⁵⁴ through the calculate bias function from the *sambias* package version 1.0.4⁵⁵. The package created a grid from the border we supplied (shapefile of Vanuatu) and the distance from each cell to the nearest biasing factor was found⁵⁵. We increased the precision of the distance estimation by setting the resolution to 0.1 (11 km²). The function used a Poisson sampling process to estimate the number of expected occurrences and a Bayesian approach to estimate the weight of each biasing factor⁵⁵.

Results

Additional records. During 2019 the known range for *Atyphella* was expanded. One additional location was added on each of the following islands: Espiritu Santo, Malekula, and Efate (Fig. 1, Supplemental Table S1). The majority of known presence locations on Efate, Espiritu Santo, and Malekula from 2018 were checked in 2019 and all still contained *Atyphella*. *Atyphella* was also collected on Maewo which is the first record from the eastern islands in the Malampa and Loyalty Basin bioregion.

Updated ecological niche model. Three MaxEnt outputs were generated (Fig. 3) and the highest contributing variables used in model training are reported (Table 1). These factors are shown to be much more variable in 2019 than 2018 (Fig. 4). In 2018, the range within the December average temperature and March minimum temperature values across the occurrence sites is 2 °C*10. There is effectively no variation within these values at the occurrence sites. In 2019, there is a difference of 24 °C*10 within the September minimum temperature values and a difference of 599 °C*10 within the temperature seasonality values. This stark increase in variation is shown to weaken the overall model and explains the observed result which is a prediction that lacks precision or high support (Fig. 3).

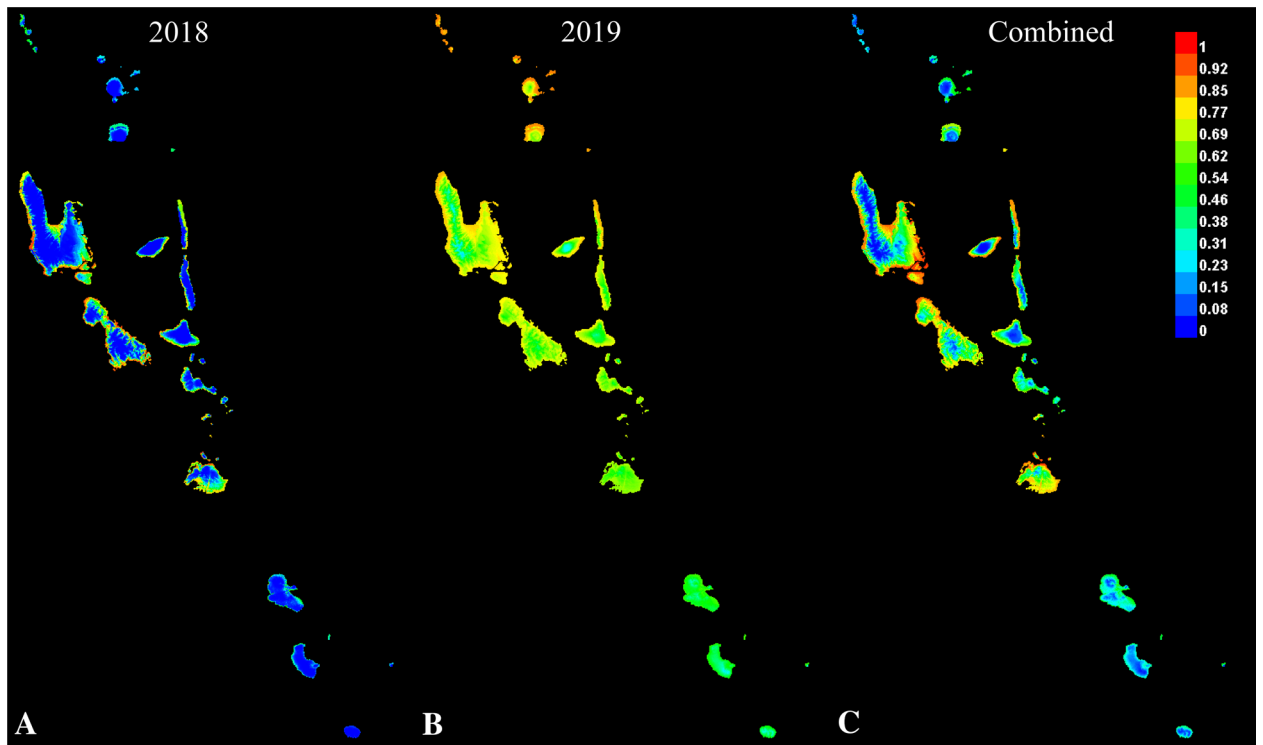


Figure 3. Results from MaxEnt analysis based on collection records from (A) only 2018 collecting, (B) new locations from 2019, (C) combined collecting events (MaxEnt version 3.3.3 k, Adobe, San Jose, CA, USA).

Variable	Contribution (%)
2018	
December average temperature	31.9
Mean temperature of wettest quarter	20.8
May precipitation	9.4
March minimum temperature	5.8
2019	
September minimum temperature	41.3
Temperature seasonality	32.9
Minimum temperature of coldest month	18.2
October minimum temperature	4.1
Combined	
August maximum temperature	21.1
Temperature seasonality	16.5
September minimum temperature	7.8
March average temperature	6.4

Table 1. Comparison of major contributing factors to each MaxEnt model. Temperature and precipitation are frequently used in model training in both aquatic and terrestrial systems, we found this to be the case as well. Since both coastal species of *Atyphella* from Vanuatu spend the majority of the time in the intertidal zone submerged by water, precipitation is not as influential if they were a completely terrestrial species.

The first model based on the six locations from 2018 had an AUC of 0.93 and highlighted coastal areas within the Malampa and Loyalty Basin and East Shefa bioregions²³. The model using the four locations from 2019 had an AUC of 0.69 and the output did not show extreme probabilities (<0.2 or >0.8). The combined model using ten locations resulted in an AUC of 0.86 and the areas with the highest probability are along the coasts of the islands north of Chessman's line. The amount of area within each probability is provided (Fig. 5, Supplementary Table S2). The lines for the 2018 and combined models share the same general trend which is as the probability increases the area decreases at a constant rate. For the 2019 model, as the probability increases, the area increases until 0.7, when it switches and begins to quickly decrease.

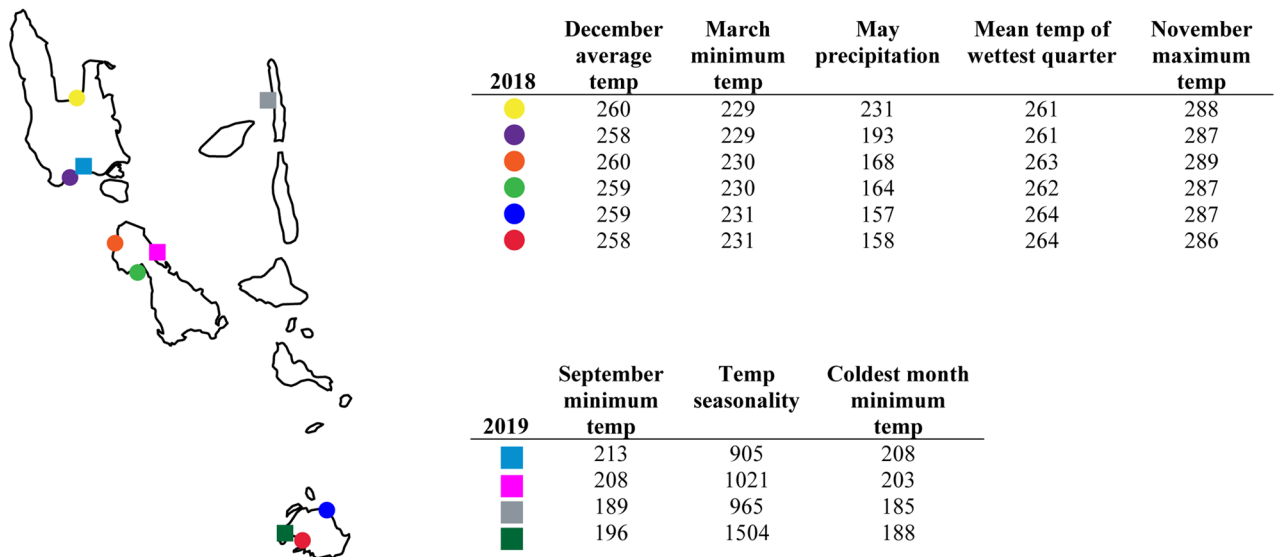


Figure 4. Values for the top contributing factors at each location (temperature values given in °C*10 and precipitation values give in mm). The range of values within 2018 is minimal for each factor. There is a wider range in the values for each factor in 2019 (Vemaps.com, Adobe, San Jose, CA, USA).

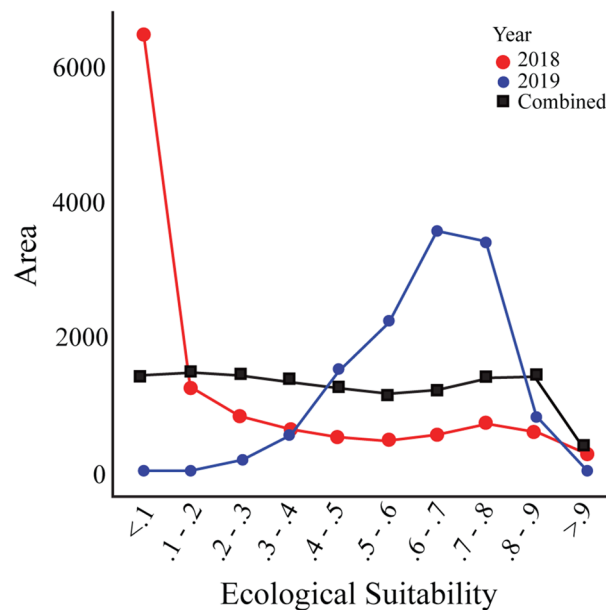


Figure 5. Line graph showing area of islands (in km²) predicted to be suitable for the presence of *Atyphella* at 10% intervals (upper probability included in interval).

Collecting bias. The output from the *sambias* package showed that both roads and settlements affected our sampling rate because they each have a negative slope (Fig. 6). As the distance from these factors increased the less likely that area was to be sampled. This can be seen when overlaying these biasing factors on MaxEnt outputs (Fig. 7).

Discussion

Ecological niche model. The 2018 MaxEnt model successfully identified new locations that were confirmed with subsequent 2019 fieldwork, thus, expanding the known distribution of *Atyphella*. This adds to the growing literature supporting the utility of MaxEnt with limited locality data (e.g.,⁵⁶⁻⁵⁸). The 2018 prediction correctly identified novel locations, the subsequent combined model then refined the overall range. We now have an improved understanding of their range based on the amount of time spent searching and the amount of land that we were able to cover relative to the size of the islands, approaching maximum accuracy potential as discussed by

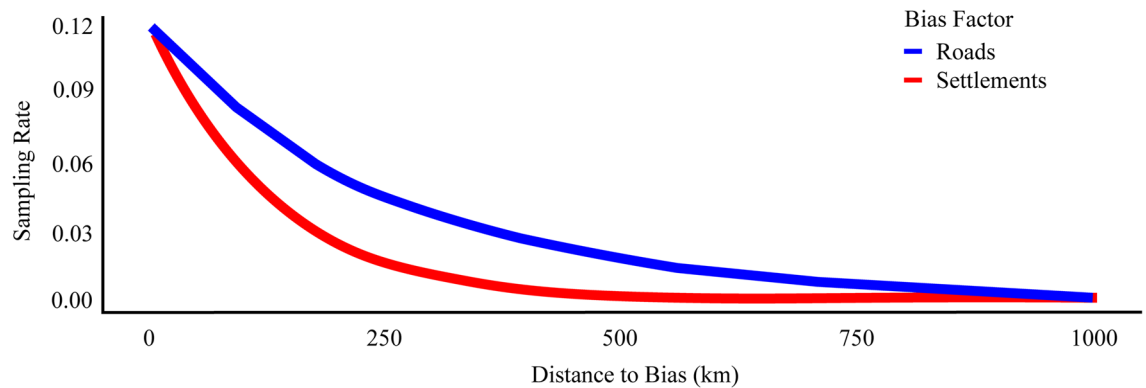


Figure 6. Results of *sampbias* analysis for roads and settlements. The sampling rate is the function of distance to the nearest bias factor. As the distance from a bias factor increases the less likely the location is to be sampled. Negative slopes indicate which factors are in fact biasing collecting locations with steeper slopes having more of an influence.

Hernandez et al.³⁰. The increase in distributional data for *Atyphella* in Vanuatu has begun to minimize the negative effects associated with the Wallacean shortfall which will allow for more confidence in conservation planning. Combining data from the two years led to an increase in predictive power without a corresponding loss of precision of the model (both AUC values over 0.8) which is a result of increasing the number of sample locations and thus accuracy^{31, 58–60}. Accuracy for models based on limited locations has also been shown to increase if the species requires specific habitats^{30,58}. Every location that we provided was critical to the model's training because of the variation between the most contributing factors (Fig. 4). We were able to capture the vast majority of this variation because we sampled such a wide range of coastline across the islands.

There are instances where MaxEnt predictions have been ground tested (e.g.,^{61–63}), however most of these studies were conducted in more easily accessible areas. Additionally, there are studies where MaxEnt has been used to promote conservation in islands settings^{43,64–66}. The models of both Kumar and Stohlgren⁴³ and Thorn et al.,⁶⁶ were limited to Borneo, Raxworthy et al.,⁶⁵ to Madagascar, and Greaves et al.⁶⁴ to the southern island of New Zealand, large single islands. The SDMs from these studies do not deal with challenges that come with predicting and surveying over multiple islands with varying amounts of isolation from each other. Also Kumar and Stohlgren⁴³ and Thorn et al.⁶⁶ considered a small sample to be greater than ten, we were able to successfully validate the utility of Maxent models over many islands with less data, by successive years of field research. This study focused on remote locations with extremely rare endemic species with very specific habitats and was shown to have tremendous value and accuracy.

The 2019 model generated from only four locations is less predictive because the number of locations for consistent “significant predictive ability” is five or above³². Combining data from these models led to a much more robust result, as expected. It is also accepted that outliers and additional records carry significant weight and can influence the prediction when working with small sample sizes if these new locations provide different environmental variables or values^{30,32}. Our 2019 data when analyzed alone illustrated this issue. However, when combining all available data this problem was mitigated by including the full breadth of observed variation.

Collecting bias. There are numerous issues that arise when working with distributional data (i.e., the Wallacean shortfall), but efforts and recommendations are being made to mitigate these effects^{35,67}. Distribution data is essential for studying global patterns of diversity and range shifts for invasive or threatened species, but the extent to which it is collected is inconsistent^{36,68}. Museum collections are often a combination of preferential collecting and intense surveying of certain areas which can bias research results⁶⁸. Steps are being taken to digitize and combine museum records to allow for broader access and likely leading to more accurate analyses^{67–69}. While access to more data is beneficial, it is the quality of the records that need to be assessed⁶⁷. Most studies are very scale dependent. For instance, worldwide patterns can be seen with very coarse data (e.g., country records) while studies focused on smaller geographic areas need finer scale data (e.g., precise coordinates)^{36,67}. It is difficult to get these data in remote locations where access is limited, as is the case with Vanuatu, therefore post hoc analysis can assist identifying bias when interpreting results³⁶.

It is important to identify if there is spatial bias in SDMs because if unknown, the model results can be interpreted incorrectly. If the collection locations are biased and the background locations which are randomly chosen over the entire area are not biased then the results are skewed³³. While there are a few methods to account for spatial bias (e.g., spatial filtering and background manipulation) there is no agreed upon approach^{33,70}. Fourcade et al.⁷¹ found that the most consistent spatial filtering approach was systematic sampling. However, this simply reduces the number of locations and does not account for a lack of data, which is an issue with the dataset for *Atyphella*. As for the background manipulation approach, a buffer is created around the locations from which background locations can be chosen⁷¹. This does not work in Vanuatu because the size of each island is too small to create an appropriate buffer. With the extremely limited locations for *Atyphella*, we cannot use the traditional approaches to correct for spatial bias. These approaches only attempt to correct the prediction and do not identify the original causes of bias. Instead, we conducted a post hoc analysis to identify bias which is shown to work

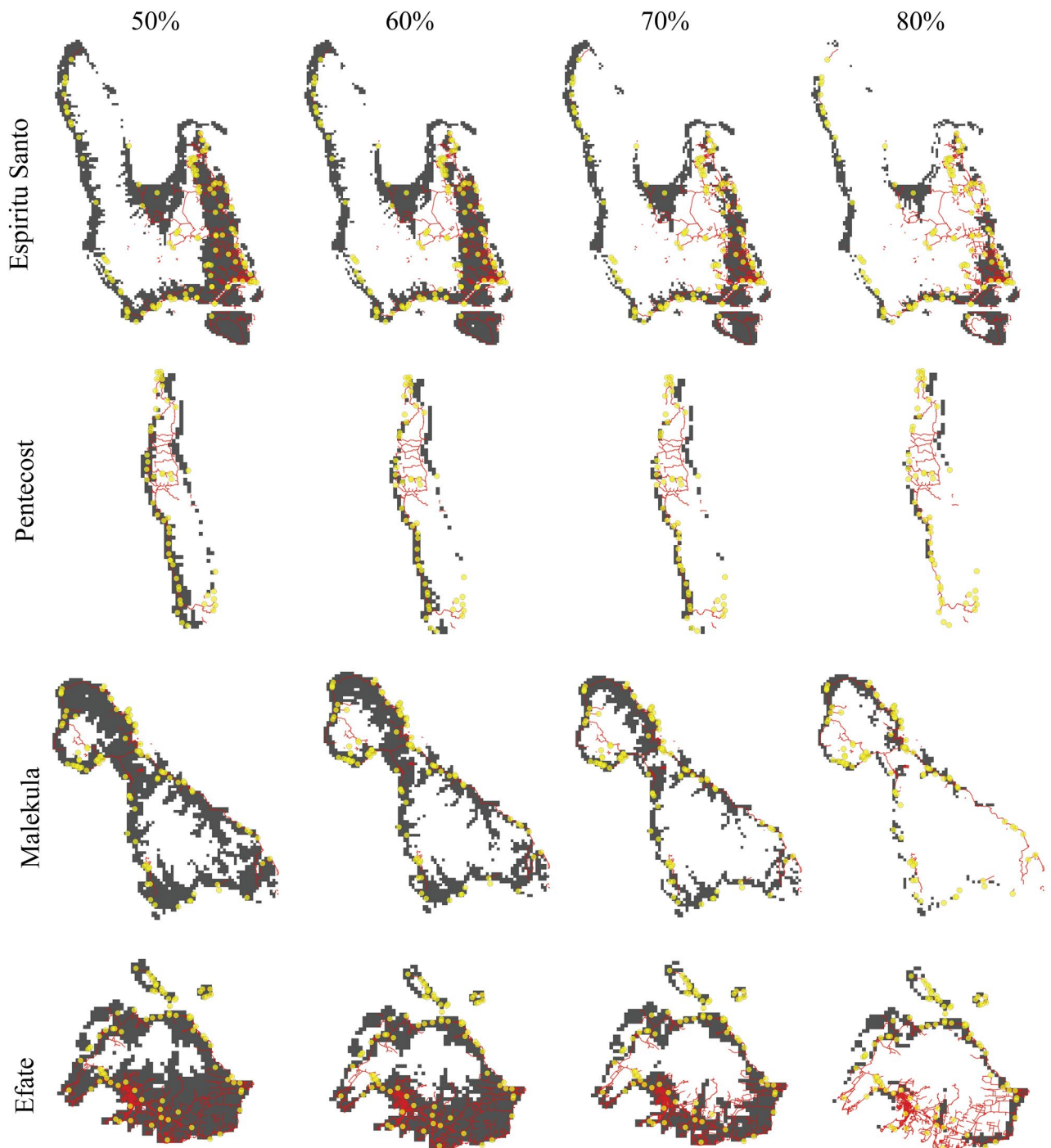


Figure 7. Presence of *Atyphella* predicted by MaxEnt overlaid with potential bias factors: roads (red), settlements (yellow). Rows are four heavily sampled islands, columns are thresholds of predictive results (i.e., 0.6 are those areas predicted to have *Atyphella* with >0.6 probability in the combined model) (MaxEnt version 3.3.3 k, QGIS Zanzibar 3.8, NextGIS: <https://data.nextgis.com/en/region/VU/>, Adobe, San Jose, CA, USA).

well with sparse data and small areas⁵⁵. Identifying the biasing factors allows us to consider potentially affected areas, and make more informed conclusions, better guiding conservation efforts.

The bias in the model stems from the fact that Vanuatu has relatively little infrastructure and it is quickly and constantly changing. The biasing layers available for Vanuatu are roads and settlements, but what is included as a road or settlement varies considerably between the islands. For example, on the more developed islands of Efate, Espiritu Santo, and Tanna the roads were often asphalt or concrete and settlements were often larger. On the other islands, the 'roads' closely resemble dirt trails and structures within the settlements are less common and built from surrounding natural materials (Fig. 8). Even though there are issues with consistency across the layers, this analysis would not have been possible a few years ago because the layers had not yet been developed.



Figure 8. (A) Example of settlements on more remote islands. (C) Example of settlements and roads on more developed islands. (B–D) Variation within roads across the islands (Photo credit: Colin Jensen and Natalie Saxton).

Information that will improve these data layers and expand the possibilities of research similar to ours is currently being compiled for islands across the South Pacific in response to rapid commercialization, globalization, and environmental threats²³.

In addition to infrastructure, cost and time were the other limiting factors. During fieldwork, our research team often reached the end of a road and could continue on foot for short distances but were often limited by social, environmental, or academic factors. The majority of the roads followed the coast because the economy is centered around marine activities (e.g., fisheries) and tourism. The other reason is because the islands are volcanic and road development further inland is minimal and travel is difficult due to the topography. While analyses showed that the collecting was biased, the data that we were able to gather is still extremely valuable and we can more accurately interpret the results in order to correct for this bias in the future. There are large areas of coastline on both Espiritu Santo and Pentecost that do not have roads or settlements and these areas are not predicted to have suitable habitat. The areas on Espiritu Santo and Pentecost that are predicted to have suitable habitat also have roads and settlements nearby. Another example of this bias is seen on Efate. There were three locations on the northern half of the island where *Atyphella* was collected, so it is expected that this half would have high probabilities of suitable habitat associated with it. However, within Efate as we increased the probability of presence, the areas that remained suitable were farther south where the biasing factors were more prominent. Across the entire archipelago, as we increased the minimum threshold for probability of occurrence, the areas considered suitable were largely areas with biasing factors.

The combined model can aid Vanuatu's marine spatial planning committee, by identifying coastal locations that are important for fireflies and firefly protection. The coastlines are under serious threat from rising sea levels, tropical storms, human development, and commerce^{23,72}. Gassner et al.²³ showed that 40.9% of all reef areas are at “high risk” and 13.8% are at an “extremely high risk” due to coastal development (ports, roads, and housing), fishing activities, and human activities. The majority of these high risk areas are the same coastal areas that contain firefly populations. There are currently 14 ports across the archipelago used mainly for commercial fishing, cargo, and tourism. The largest ports are located on Efate and Espiritu Santo which are both home to coastal *Atyphella* species and have additional areas predicted to be appropriate habitat for *Atyphella* to survive. Altering the coastlines to support development lessens the natural protections they provide from tropical storms and destroys habitats for many species²³. Other threats the coastlines are facing come from the ocean itself. Sea levels are rising which is extremely problematic for the low-lying islands of Vanuatu⁷². Gassner et al.²³ also showed how sea levels are predicted to rise a minimum of 0.15 m by 2030. In addition, shallow water along the coastline is quickly warming and becoming more acidic (pH 8.26–8.3) which is a major concern for fireflies as both females and larvae spend a significant amount of their life-cycle in or near the water. If coastal development, ocean levels, water temperatures, and acidity continue to increase there will be significant issues for endemic species dependent on coastal habitats, such as fireflies.

Vanuatu's government is aware of the threats facing their marine and coastal habitats. They were the first Pacific Island nation to have a National Ocean Policy which uses marine spatial planning and climate resistant networks to make policies to conserve the natural diversity within the nation^{23,73}. They are also aware that the natural beauty and biodiversity of Vanuatu is a major draw for tourists. There is already a coastal resort adjacent to a known *Atyphella* habitat which has become a draw for the resort. One of the main tourism slogans is "Discover what matters" with the main attractions being diving and snorkeling and if these areas are not preserved Vanuatu's economy will be negatively impacted²³ (p. 49). Additionally, local leaders have been regulating the fishing industry for years to make it as sustainable as possible, and the national government is following suit in the commercial fishing industry. We can see that the government is committed to protecting what makes Vanuatu unique. One more conservation policy to protect the habitats these extremely unique fireflies need would complement the plans that the country already has in place very well.

Data availability

Presence localities in supplementary material Table S1.

Received: 26 March 2021; Accepted: 10 August 2021

Published online: 30 August 2021

References

- Brooks, T. M. *et al.* Habitat loss and extinction in the hotspots of biodiversity. *Conserv. Biol.* **16**, 909–923 (2002).
- Maschinski, J. *et al.* Sinking ships: Conservation options for endemic taxa threatened by sea level rise. *Clim. Change* **107**, 147–167 (2011).
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).
- Heaney, L. R., Balet, D. S. & Rickart, E. A. Models of oceanic island biogeography: Changing perspectives on biodiversity dynamics in archipelagos. *Front. Biogeogr.* **5**, 249–257 (2013).
- Keppel, G., Lowe, A. J. & Possingham, H. P. Changing perspectives on the biogeography of the tropical South Pacific: Influences of dispersal, vicariance and extinction. *J. Biogeogr.* **36**, 1035–1054 (2009).
- Laurance, W. F. Beyond Island biogeography theory. In *The Theory of Island Biogeography Revisited* (eds Losos, J.B. & Ricklefs, R. E.) 214–237 (Princeton University Press, 2010).
- Cheesman, L. E. Biogeographical significance of Aneityum Island, New Hebrides. *Nature* **180**, 903–904 (1957).
- Cox, B. T. M. & Burns, K. C. Convergent evolution of gigantism in the flora of an isolated archipelago. *Evol. Ecol.* **31**, 741–752 (2017).
- Hamilton, A. M., Klein, E. R. & Austin, C. C. Biogeographic breaks in Vanuatu, a nascent oceanic archipelago. *Pac. Sci.* **64**, 149–159 (2010).
- Coleman, P. J. Geology of the Solomon and New Hebrides islands, as part of the Melanesian re-entrant, Southwest Pacific. *Pac. Sci.* **24**, 289–314 (1970).
- Valente, L. *et al.* A simple dynamic model explains the diversity of island birds worldwide. *Nature* **579**, 92–96 (2020).
- Keppel, G., Buckley, Y. M. & Possingham, H. P. Drivers of lowland rain forest community assembly, species diversity and forest structure on islands in the tropical South Pacific. *Ecology* **98**, 87–95 (2010).
- Cheng, L. Insects in marine environments. *Marine Insects* **1**, 1–4 (1976).
- Ballantyne, L. A. & Buck, E. Taxonomy and behavior of *Luciola* (*Luciola*) *aphrogeneia*, a new surf firefly from Papua New Guinea. *Trans. Am. Entomol. Soc.* **105**, 117–137 (1979).
- Doyen, J. T. Marine beetles (Coleoptera excluding Staphylinidae). In *Marine Insects* (ed. Cheng, L.) 497–519 (American Elsevier, 1976).
- Topp, W. & Ring, R. A. Adaptations of Coleoptera to the marine environment. II. Observations on rove beetles (Staphylinidae) from rocky shores. *Can. J. Zool.* **66**, 2469–2474 (1988).
- Lloyd, J. E. Fireflies (Coleoptera: Lampyridae). In *Encyclopedia of Entomology* 429–1452 (Springer Dordrecht, 2008).
- McDermott, F. A. *Photuris bethaniensis*, a new Lampyrid firefly. *Proc. U. S. Natl. Mus.* **103**, 35–37 (1953).
- Vaz, S. *et al.* On the intertidal firefly genus *Micronaspis* Green, 1948, with a new species and a phylogeny of Cratomorphini based on adult and larval traits (Coleoptera: Lampyridae). *Zool. Anz.* **292**, 64–91 (2021).
- Ballantyne, L. A. & Lambkin, C. Systematics of Indo-Pacific fireflies with a redefinition of Australasian *Atyphella* Olliff, Madagascar *Photurolociola* Pic, and description of seven new genera from the Luciolinae (Coleoptera: Lampyridae). *Zootaxa* **1997**, 1–188 (2009).
- Ballantyne, L. A. *et al.* The Luciolinae of SE Asia and the Australopacific region: A revisionary checklist (Coleoptera: Lampyridae) including description of three new genera and 13 new species. *Zootaxa* **4687**, 1–174 (2019).
- Saxton, N. A., Powell, G. S., Martin, G. J. & Bybee, S. M. Two new species of coastal *Atyphella* Olliff (Lampyridae: Luciolinae). *Zootaxa* **4722**, 270–276 (2020).
- Gassner, P. *et al.* Marine Atlas. Maximizing Benefits for Vanuatu. <https://grid.cld.bz/Marine-Atlas-Maximizing-Benefits-for-Vanuatu1/10/> (2019).
- Saxton, N. A., Powell, G. S., Serrano, S. J., Monson, A. K. & Bybee, S. M. Natural history and ecological niche modelling of coastal *Atyphella* Olliff larvae (Lampyridae: Luciolinae) in Vanuatu. *J. Nat. Hist.* **53**, 2271–2280 (2019).
- Rhoden, C. M., Peterman, W. E. & Taylor, C. A. Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. *PeerJ* **5**, e3632 (2017).
- Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259 (2006).
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E. & Blair, M. E. Opening the black box: An open-source release of Maxent. *Ecography* **40**, 887–893 (2017).
- Warren, D. L. & Seifert, S. N. Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* **21**, 335–342 (2011).
- Stas, M. *et al.* An evaluation of species distribution models to estimate tree diversity at genus level in a heterogeneous urban-rural landscape. *Landsc. Urban Plan.* **198**, 103770 (2020).
- Hernandez, P. A., Graham, C. H., Master, L. L. & Albert, D. L. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* **29**, 773–785 (2006).
- Hernandez, P. A. *et al.* Predicting species distributions in poorly-studied landscapes. *Biodivers. Conserv.* **17**, 1353–1366 (2008).
- Pearson, R. G., Raxworthy, C. J., Nakamura, M. & Townsend Peterson, A. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *J. Biogeogr.* **34**, 102–117 (2007).

33. Phillips, S. J. *et al.* Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol. Appl.* **19**, 181–197 (2009).
34. Silva, D. P., Aguiar, A. G. & Simião-Ferreira, J. Assessing the distribution and conservation status of a long-horned beetle with species distribution models. *J. Insect Conserv.* **20**, 611–620 (2016).
35. Cardoso, P., Erwin, T. L., Borges, P. A. & New, T. R. The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144**, 2647–2655 (2011).
36. Hortal, J. *et al.* Seven shortfalls that beset large-scale knowledge of biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **46**, 523–549 (2015).
37. Lomolino, M. V. Conservation biogeography. In *Frontiers of Biogeography: new directions in the geography of nature* (eds Lomolino, M. V. & Heaney, L. R.) 293–296 (Sinauer Associates, Sunderland, Massachusetts, 2004).
38. Whittaker, R. J. *et al.* Conservation biogeography: Assessment and prospect. *Divers. Distrib.* **11**, 3–23 (2005).
39. Cui, S., Luo, X., Li, C., Hu, H. & Jiang, Z. Predicting the potential distribution of white-lipped deer using the MaxEnt model. *Biodivers. Sci.* **26**, 171 (2018).
40. Moreno, R., Zamora, R., Molina, J. R., Vasquez, A. & Herrera, M. A. Predictive modeling of microhabitats for endemic birds in South Chilean temperate forests using Maximum entropy (Maxent). *Ecol. Inform.* **6**, 364–370 (2011).
41. Raman, S., Shameer, T. T., Sanil, R., Usha, P. & Kumar, S. Protrusive influence of climate change on the ecological niche of endemic brown mongoose (*Herpestes fuscus fuscus*): A MaxEnt approach from Western Ghats, India. *Model. Earth Syst. Environ.* **6**, 1795–1806 (2020).
42. Abdelaal, M., Fois, M., Fenu, G. & Bacchetta, G. Using MaxEnt modeling to predict the potential distribution of the endemic plant *Rosa arabica* Crép., Egypt. *Ecol. Inform.* **50**, 68–75 (2019).
43. Kumar, S. & Stohlgren, T. J. Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. *J. Ecol. Nat.* **1**, 094–098 (2009).
44. Yang, X. Q., Kushwaha, S. P. S., Saran, S., Xu, J. & Roy, P. S. Maxent modeling for predicting the potential distribution of medicinal plant, *Justicia adhatoda* L. Lesser Himalayan foothills. *Ecol. Eng.* **51**, 83–87 (2013).
45. New, T. R. Conserving narrow range endemic insects in the face of climate change: Options for some Australian butterflies. *J. Insect Conserv.* **12**, 585–589 (2008).
46. Booth, T. H., Nix, H. A., Busby, J. R. & Hutchinson, M. F. BIOCLIM: The first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Divers. Distrib.* **20**, 1–9 (2014).
47. Hijmans, R. J., Cameron, S. & Parra, J. *WorldClim, Version 1.4* (University of California, 2005).
48. Hijmans, R. J. *et al.* DIVA-GIS. Version, 7.5. A Geographic Information System for the Analysis of Species Distribution Data. <http://www.diva-gis.org> (2012).
49. Phillips, S. J., Dudík, M. & Schapire, R. E. A maximum entropy approach to species distribution modeling. In *Proceedings of the 21st International Conference on Machine Learning* 655–662 (2004).
50. Merow, C., Smith, M. J. & Silander, J. A. Jr. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069 (2013).
51. Phillips, S. J. & Dudík, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* **31**, 161–175 (2008).
52. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project (2020).
53. Phillips, S. J. A brief tutorial on Maxent. *AT&T Res.* **190**, 231–259 (2005).
54. RStudio Team RStudio: Integrated Development Environment for R. RStudio, PBC. <http://www.rstudio.com/> (2020).
55. Zizka, A., Antonelli, A. & Silvestro, D. Sampbias: Evaluating geographic sampling bias in biological collections. *Ecography* **44**, 25–32 (2020).
56. Almeida, M. C., Cortes, L. G. & De Marco Junior, P. New records and a niche model for the distribution of two Neotropical damselflies: *Schistolobos boliviensis* and *Tuberculobasis inversa* (Odonata: Coenagrionidae). *Insect Conserv. Divers.* **3**, 252–256 (2010).
57. De Siqueira, M. F., Durigan, G., de Marco Júnior, P. & Peterson, A. T. Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. *J. Nat. Conserv.* **17**, 25–32 (2009).
58. Wisz, M. S. *et al.* Effects of sample size on the performance of species distribution models. *Divers. Distrib.* **14**, 763–773 (2008).
59. McCune, J. L. Species distribution models predict rare species occurrences despite significant effects of landscape context. *J. Appl. Ecol.* **53**, 1871–1879 (2016).
60. Rinshofer, L. J. *et al.* Iterative species distribution modelling and ground validation in endemism research: An Alpine jumping bristletail example. *Biodiversity* **21**, 2845–2863 (2012).
61. Peterman, W. E., Crawford, J. A. & Kuhns, A. R. Using species distribution and occupancy modeling to guide survey efforts and assess species status. *J. Nat. Conserv.* **21**, 114–121 (2013).
62. Searcy, C. A. & Shaffer, H. B. Field validation supports novel niche modeling strategies in a cryptic endangered amphibian. *Ecography* **37**, 983–992 (2014).
63. Virzi, T., Lockwood, J. L., Lathrop, R. G., Grodsky, S. M. & Drake, D. Predicting American Oystercatcher (*Haematopus palliatus*) breeding distribution in an urbanized coastal ecosystem using maximum entropy modeling. *Waterbirds* **40**, 104–122 (2017).
64. Greaves, G. J., Mathieu, R. & Seddon, P. J. Predictive modelling and ground validation of the spatial distribution of the New Zealand long-tailed bat (*Chalinolobus tuberculatus*). *Biol. Conserv.* **132**, 211–221 (2006).
65. Raxworthy, C. J. *et al.* Predicting distributions of known and unknown reptile species in Madagascar. *Nature* **426**, 837–841 (2003).
66. Thorn, J. S., Nijman, V., Smith, D. & Nekaris, K. A. I. Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (Primates: Nycticebus). *Divers. Distrib.* **15**, 289–298 (2009).
67. Faith, D. *et al.* Bridging the biodiversity data gaps: Recommendations to meet users' data needs. *Biodivers. Inform.* **8**, 41–58 (2013).
68. Pyke, G. H. & Ehrlich, P. R. Biological collections and ecological/environmental research: A review, some observations and a look to the future. *Biology* **85**, 247–266 (2010).
69. Boakes, E. H. *et al.* Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. *PLoS Biol.* **8**, e1000385 (2010).
70. Kramer-Schadt, S. *et al.* The importance of correcting for sampling bias in MaxEnt species distribution models. *Divers. Distrib.* **19**, 1366–1379 (2013).
71. Fourcade, Y., Engler, J. O., Rödder, D. & Secondi, J. Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS ONE* **9**, e97122 (2014).
72. National Integrated Coastal Management Framework. *National Integrated Coastal Management Framework and Implementation Strategy for Vanuatu*. <https://extwprlegs1.fao.org/docs/pdf/van171039.pdf> (2010).
73. Department of Environmental and Protection and Conservation. *Coastal Development*. https://environment.gov.vu/images/EIA/EIA_G%20Coastal%20development.pdf (2017).

Acknowledgements

The authors would like to thank all the Ni-Vanuatu government officials, especially Dr. Donna Kalfatakvoli, for help in obtaining the necessary permits and visas which made this research possible. We would also like to thank the local guides and landowners for their hospitality and help throughout this process. We appreciate all the BYU undergraduates that assisted in collecting efforts. In addition, we would like to thank Colin Jensen and

Natalie Saxton for providing photos and Makani Fisher for providing preliminary MaxEnt information. Lastly, we thank the reviewers and editor for comments that improved the manuscript. This research was funded through the Roger and Victoria Sant Educational Endowment for a Sustainable Environment, NSF Research Experience for Undergrads Supplemental grants (DEB-1655981), and the BYU Kennedy Center for International Studies.

Author contributions

L.N.S conducted the analyses and wrote the original manuscript draft. G.S.P. and S.M.B conceptualized, designed, and supervised the project. G.S.P. and S.M.B also wrote, edited, and revised the manuscript. S.M.B obtained funding. All authors collected data, designed figures, and reviewed the manuscript.

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/s41598-021-96534-x>.

Correspondence and requests for materials should be addressed to L.N.S.

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) 2021