Characterizing the ecological network of Puerto Rico: The role of landscape connectivity and land use on the distribution and genetic structure of the small Indian mongoose [*Herpestes auropunctatus*]

By

Diana Kamil Guzmán Colón

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Forest and Wildlife Ecology)

at the
UNIVERSITY OF WISCONSIN-MADISON
2020

Date of final oral examination: 12/14/2020

The dissertation is approved by the following members of the Final Oral Committee:

Anna M. Pidgeon, Professor, Forest and Wildlife Ecology
Volker C. Radeloff, Professor, Forest and Wildlife Ecology
Sean Schoville, Professor, Entomology
Erika Marín-Spiotta, Professor, Geography
Jonathan N. Pauli, Professor, Forest and Wildlife Ecology

Acknowledgements

I have grown personally and academically thanks to various communities and individuals at the University, the city of Madison, this state of Wisconsin. I honor these places and these people with my undivided respect and gratitude.

The completion of this dissertation would not have been possible without the advice and support of Volker and Anna, my mentors. I came into the lab like a kid in a candy store with many ideas, lots of plans and many ambitions – so I can only imagine the type of patience they needed to have with my attention-deficit self. Thanks for always being available to share your expertise in Ecology, Conservation and Landscape Ecology. For pushing me to think broadly and outside of my usual comfort zone of insular ecosystems. Lastly, I want to thank them for fostering such a culturally diverse, kind, inclusive, and trusting lab.

The "old" SILVIS guard: David Helmers, Drs: Patricia Alexandre, Brooke Bateman, Catalina Munteanu "*Communist coffee is BETTER than NO coffee*", Chris Hamilton, James Burnham, Naparat Suttidate, Eugenia Bragina, Anita María Venegas, Jess Gorzo, Anu Kramer, Konrad Turlej, you were all such a big part of my formative years and I will never forget our trips to the north. The "new era" folks: Niwaeli Kimambo Kristin Brunk, Maia Persche, Jennifer Schneiderman, Sofia Kozidis, Ashley Olah, Duda Silviera, Laura S. Farwell, David Gudex-Cross, Paul Schilke, Colleen Sutheimer, Martina Hobi, Lena Razenkova, Afag Rizayeva, thank you for all the laughter, sarcasm, beers at the terrace, ornithology-only jokes, blunt questions and the *kindness*. Kasia Lewinska you are one of a kind, I am so grateful for your friendship, mentoring and stimulating conversations about any topic.

I am extremely grateful for the encouragement and advice of my committee members. Dr. Jon Pauli who provided valuable guidance and insight at the conception of this project, I always left our conversations with plenty to ponder. Those conversations helped me shape the way I thought about my dissertation. I am forever indebted to Dr. Sean Schoville for his willingness to offer advice and support my project by providing space in his lab during the first few years of this marathon. I hope that I can be as great as a mentor and teacher. I truly enjoyed the Molecular Ecology class not only as a student but also as someone who loves to observe pedagogical techniques in action. Dr. Erika Marín-Spiotta arrived at the precise moment when I needed lots of encouragement to keep going. Thank you for not hesitating to say "yes" to jump on board half-way through the journey. Your constructive review of my first published chapter, which is near and dear to Puerto Rico, made my manuscript stronger. I also want to thank you for the unparalleled support to Ricky and me, career advice, for all the food and dances we shared and for always supporting our Boricua community.

My two main sources of funding - IGERT and SciMed GERS – were so instrumental in my development as a scientist. The SciMed community, Abbey Thompson and Sara Patterson, provided a space to vent frustrations in a healthy/productive way. Thanks to the community for organizing much needed dialogues about diversity and inclusion, mental health in Academia, and scholarship opportunities. It was within this community that I was able to fulfill my interest in science outreach. The IGERT community (Volker, Erika, Tim Van Deelen, Tony Ives, Adeena Rissman, Shelley Smith, Emily, Sam, Elliot, Andrew, Lori, Chris, Phil, Michael, Sebastián Matinuzzi) - I had the pleasure to belong to a brilliant community of academics asking tough and relevant questions surrounding novel ecosystems. I admire all the people who belonged to this community. My worldview changed by participating in this program, I saw what it was like to be at the frontier of a new field of study and that was illuminating.

To the field support crew (Ashley, Mom, Cheo, Ricky) who accompanied me to set traps for mongooses in the island, thank you! You are part of this! Thanks to the team at Culebra y Vieques islands (David Shaw, Suki Bermúdez, Mike Barandiaran) for providing tissue samples adding valuable information to my dataset. The success of the project, especially the lab portion, has many people behind it: Candice Teschner, Juana Antonio and Paige Kulzer all went through a lot of hair and tissue samples to extract and amplify DNA. The genetics chapter was a direct result of working under the guidance of Dr. Emily Fountain. Emily thank you so much for your patience and willingness to mentor me in a suite of techniques and skills. Thank you for telling it like it is. Emily's support means the world to me, it was instrumental in getting me to the finish line. Thanks to Dr. Zach Peery and Francisco Pelegri for also providing me with lab bench space to do some of my genetics work.

Johanna Buchner, Jennyfer Cruz-Bernal, Carlos Ramirez-Reyes, Ali Paulson, Rose Graves, Evan and Jen Wilson, Caroline Smythe-Kallenborn, you are all my family now and I look forward to celebrate plenty of more milestones in our lives together. We have seen weddings, defenses, easter and Christmas dinners, births, wild birthdays, new jobs, exciting opportunities! Each of you carried me to this point via encouragement, advice, food, drinks, hikes and hobbies. My Boricua community also provided much needed support after our María tragedy and warmth during the harsh cold periods. I am forever grateful to Karla Esquilín-Lebrón, Katiria González, Liz García-Peterson, Dianiris Luciano-Rosario, Jara Rios for making me feel at home plenty of times. Some of us spent countless nights writing at the library with cafecito and galletitas and we saw each other grow into Doctoras en Ciencias.

Life provided me with the honor of having two best friends during this journey: Drs. Isabel Rojas-Viada and Paula L. Perrig. We continue to support each other in so many ways even from a distance. We spent countless hours laughing at ourselves (or at each other), solving problems, discovering new places. I want to thank them for their sincere and constant love and support, for listening to me rambling about (probably nonsense), for being my inspiration to become a better naturalist, for keeping me centered, and for never letting go. Paula you kept me strong, Isabel you kept me energized, and both made me so happy.

Finally, I want to acknowledge my husband, Ricardo Rivera, who has always stood by me, supported me unconditionally and understands my process. Thank you for your honesty, for listening to my ideas and all the productive discussions we have had. Thanks for writing those emails for me when I was too tired to do so. To my family back in Puerto Rico, thanks to their support I am here where I am in this moment in time. They have always encouraged education above all else because they are witness to the opportunities that become available when you possess a degree and do what you love. It was not easy for us this year. I dedicate all this work to my dad, Santos Guzmán-Santiago, and my grandfather Fernando Emilio Colón, both who died during my PhD. I will always love you both, your girls will be alright.

Table of Contents

List of Figures

List of Tables

Dissertation overview

Many of the milestones attained in conservation can be attributed to the translation of theory into conservation practice (Robinson, 2006). One of the cornerstones of conservation practice is protected areas, which currently cover about 14% of the terrestrial surface of earth and have been crucial in reducing biodiversity loss within their boundaries (Butchart et al., 2012, 2015). In addition to protected areas, control of invasive species has facilitated the recovery of threatened species, especially on insular ecosystems (i.e. islands) where a disproportionate amount of the world's biodiversity is found (Jones et al., 2016; Tershy et al., 2015). However, despite decades of conservation efforts, biodiversity continues to decline globally, and again disproportionately on islands (Dirzo et al., 2014; Russell & Kueffer, 2019). On islands, habitat loss and invasive species have historically been the major threats to biodiversity (Millenium Ecosystem Assessment, 2005). Furthermore, other emerging drivers of biodiversity loss – climate change, pollution, and intensification of land use - are affecting both islands and continents as well (Tilman et al., 2017). The future of conservation will require an understanding of Anthropocene landscapes in which humans shape many of the ecological processes that give rise to unique habitat, species interactions, functional and genetic diversity (Geldmann et al., 2014; Margules & Pressey, 2000a; Russell & Kueffer, 2019).

In my dissertation, I focus on an insular ecosystem, Puerto Rico, and two major threats to its biodiversity: human influence and invasive species. I address three key themes: 1.) the potential of spatial graphs (networks) to support conservation management decisions, 2.) determining the best species distribution modeling approach to inform invasive species surveying efforts, and 3.) identifying causal relationships between landscape permeability and invasive species population genetic patterns.

Protected areas exist within an intervening matrix of differing habitat quality (Franklin & Lindenmayer, 2009). Species survival is thus partly regulated by the ability of individuals to move across diverse landscapes (Fahrig et al., 2011). Therefore, landscape connectivity is a critical component for ecosystem conservation, especially when important habitats are fragmented and widely separated. Defined as the degree to which landscape facilitates or impedes movement of organisms among resource patches (Taylor et al., 1993), landscape connectivity is crucial to dispersal and gene flow. Ecological networks are coherent systems composed of core habitat areas linked by ecological corridors that sustain species movement through the landscape in which the network is embedded (G. Bennett, 2004). Landscape connectivity has a net positive effect on biodiversity conservation (Samways & Pryke, 2016). However, the same mechanisms that support connectivity may also facilitate the spread of invasive species, and, therefore, strong ecological networks could increase their spread, density and ecological impacts (Haddad et al., 2014; Wilkerson, 2013). Despite this concern, current conservation planning efforts rarely evaluate invasive species potential movement through ecological networks (Resasco et al., 2014). Because the spread of invasive species has direct effects on biodiversity and ecosystem functioning, it is critical to evaluate how connectivity may facilitate the spread of invasive species through landscapes.

The consequences of invasive species introductions are especially evident in islands (Courchamp et al., 2003; Rohrer et al., 2011). Mammalian species introductions (e.g., of feral cats, pigs, rats and mongooses) have been particularly detrimental for other wildlife species, driving some to extinction and acting as disease vectors (e.g., Leptospirosis and Rabies) (Fagerstone, 2003; Pimentel et al., 2005; Veitch et al., 2011). Collectively, invasive predators are responsible for at least 58% of all bird, mammal, and reptile extinctions (Bellard et al., 2016; Doherty et al., 2016).

Remarkably, 42% of bird extinction on islands are due to mammalian introduced predators (Fagerstone, 2003) and native species continue to be threatened by exotic mammals ((Borroto-Paez, 2009; Engeman et al., 2006). The costs associated with invasive species management are at least $120 billion per year in the United States (Pimentel et al., 2005).

Most mammalian predator introductions on islands are successful because of their high trophic position and broad diet preference (Galiana et al., 2014) thus exerting substantial pressure on the biological communities of islands. Key factors which may also contribute to their persistence include the amount of optimal habitat available, the distribution of habitat, and its structural connectivity, which determines how freely a species can disperse in a heterogeneous landscape (Blackburn et al., 2011; Sakai et al., 2001; Schreiber & Lloyd-Smith, 2009). Because the spread of invasive species has direct effects on biodiversity and ecosystem functioning, it is critical to evaluate how landscape connectivity may facilitate the spread of invasive species through landscapes.

The small Indian mongoose (*Herpestes auropunctatus* [family (*Herpestidae*)]), is one of the top invasive species in the world blamed for the extinction or decrease in numbers of native species, particularly within the island ecosystems (Lowe et al., 2004). The mongoose is a solitary opportunistic carnivore native to parts of the Middle East, India and Asia but was introduced to many Pacific and Caribbean islands at the beginning of the 20th century for biological control primarily for rats and snakes (A Barun et al., 2011). Mongoose population density, feeding behavior, and home range estimates throughout the Caribbean have been found to be variable and can differ among study locations (Berentsen et. al. 2018). This species is the focus of two of my three chapters. Shortly after species establishment, noticeable declines in populations of several native island species were reported (Nadin-Davis et al., 2008; Seaman & Randall, 1962).

The ability of mongooses to recolonize areas from which they have been removed is important considering efforts to mitigate their negative effects (mainly predation) on species of conservation concern.

This dissertation includes three chapters:

***Chapter 1*: Conservation planning for island nations: Using a network analysis model to find novel opportunities for landscape connectivity in Puerto Rico**

The location, size, quality, and connectivity of habitats are all important considerations when deciding which areas to protect. The question is how to incorporate all these land attributes into models that capture both the function and structure of wildlife habitat, and that are scalable across global, regional, and local levels?

I used network analysis, rooted in graph theory, to answer part of that question. First, I developed a map of the human footprint at a 30-m resolution for the island of Puerto Rico to identify areas that have high naturalness. Using the human footprint as a cost-surface for potential species movement, I modelled spatially coherent networks of habitats linked by potential flows of organisms through the landscape matrix. I found that Puerto Rico possesses a compact network of natural areas, with a few patches critical to structural connectivity. More than half of Puerto Rico's current land surface had a low human footprint (56%), but most coastal areas were affected by human use (82%). The ecological network is likely to maintain its resilience in the face of frequent disturbances such as hurricanes.

***Chapter 2:* Field validation of habitat suitability models for the small Indian mongoose: Implications for the use of species distribution models for guiding field surveys, assessing uncertainty and usefulness**

The usefulness of a given modeling framework depends on how it will be applied. For managers who want to control the spread of an invasive species, predicting the potential distribution of species is important. Which environmental features in the island are associated with mongoose presence? Which species distribution best predicts the relative probability of occurrence compared to observed mongoose presence?

I tested different types of species distribution models in their ability to predict out-of-range occurrences with independent validation data in Puerto Rico. I parameterized four types of SDM models (plus their ensemble model) with available mongoose presence data and relevant bioclimatic and land-use covariates. All models slightly differ in their treatment of pseudo absences and correlative approaches. I used the resulting predictive maps to guide my field survey to validate the SDM model predictions. I characterized important habitat variables associated with mongoose presence and assessed the utility of the different SDM approaches in predicting mongoose occurrence outside the model training area. The distribution of habitat suitability and likelihood of species presence differed in geographic space among the four model types. I highlight various observations based on this study: 1.) Maps that maximize specificity through emphasis on correctly identifying potential suitable habitat were highly prone to misclassification, and 2.) multiple evaluation metrics were needed to assess prediction accuracy, and I especially noted differences in metrics that evaluate model calibration. The results emphasize the importance of re-evaluating and iteratively improving SDMs with independent data vs cross-validation, particularly if initial models are trained with data from a limited geographic extent. Predicted mongoose presence in coastal areas in the southwest of Puerto Rico coincided with highest values of predicted terrestrial native vertebrate resident, endemic, and endangered species richness. These areas should be considered of primary conservation concern.

***Chapter 3:* Genetic variability of the small Indian mongoose in Puerto Rico is associated with proximity to modified landscapes.**

Understanding the degree to which invasive species are related among populations is fundamental to managing their populations and ultimately mitigating their impacts. In this chapter I tested whether landscape connectivity facilitated spread of the small Indian mongoose (*Herpestes auropunctatus*) in Puerto Rico. I did this using genetic estimates of gene flow to test several landscape genetic hypotheses about the relationship of structural landscape connectivity with mongoose dispersal. I analyzed the correlation between individual-level genetic distance based on microsatellite loci, and landscape cost distance based on landscape resistance. Under a landscape genetics framework, my objectives were to first, quantify the genetic variation and population structure of the small Indian mongoose in Puerto Rico, and second, assess the influence of landscape connectivity and features on genetic patterns. In addition to null hypotheses (i.e. Panmixia and Isolation by Distance), I considered alternative hypotheses that could explain the genetic variability observed in mongoose individuals, i.e., isolation by barrier, and landscape resistance hypotheses, which include habitat suitability levels, degrees of structural connectivity (patch-level), and the human footprint index.

Although I found no statistical evidence of genetic structuring, there was a spatial pattern in the shared alleles among individuals. Genetically similar individuals were more common near highly human-modified landscapes but there was no evidence of mongooses taking advantage of connected habitats in the landscape.

Conclusion and significance

High demand for natural resources contributes to high vulnerability of wildlife populations in island nations because of their small size and isolation. Understanding the current extent, intensity and spatial pattern of human activities, which typically degrade the conservation value of habitat patches to varying degrees, is key to identifying areas of high conservation value. Furthermore, the extent to which natural areas are connected or fragmented within heterogeneous landscapes may affect species survival and consequently species spatial distributions. Although well-connected landscapes can serve as conduits for vulnerable species, this connectivity can also influence the spread of invasive species, and human-disturbed and fragmented landscapes facilitate biotic invasions. Studying the distribution of invasive species in relation to local environmental variables and functional population connectivity aids in our understanding of our current ecological network. This information on ecological networks is often lacking yet crucial for the development of species and land management plans. It is important to identify areas of high conservation value within islands given their significance for global biodiversity and vulnerability to anthropogenic modification.

The results of my dissertation are directly applicable to conservation planning for islands. I adapted commonly used conservation biology tools (Human footprint, network (graph) theory, species distribution models, conservation genetics) to understand Puerto Rico's ecological network, the distribution of unprotected areas with high conservation value, the degree to which these areas are connected, and how resilient the current network is. I also assessed the utility of species distribution models to design of field surveys and to quantify the accuracy of model predictions. Using a landscape genetics approach, I quantified the functional genetic connectivity of mongooses in Puerto Rico and addressed whether the structural connectivity of the landscape drive patterns of genetic variability in a mongoose population.

There are currently few examples of human footprint maps developed specifically for islands. As a product of this dissertation, I created a human footprint map at a resolution that is relevant to conservation management and characterized the connectivity of patches with low human modification that structurally connects the island's ecological network. These two products will be made publicly available for further modification depending on constant land-use changes, additionally; the network analysis can be adapted to study any species functional connectivity within the context of a modified landscape. I also provided examples of how species with a variety of dispersal distances see the current ecological network and found that although structural connectivity is high in the mountainous areas of Puerto Rico, functional connectivity highly varies for the vulnerable species. My results highlight that urgent conservation action is needed for coastal areas in which valuable conservation areas are fragmented and isolated. At the same time, the network reveals how current forest laws might aid in maintaining connectivity and resilience of natural areas.

The potential distribution of an exotic species is a key indicator of its ability to spread and is therefore frequently considered in prioritization processes focused on eradication. Mongoose are among the most successful invasive mammals in Caribbean and Pacific islands and also the most detrimental for local species. The challenge is hence to develop methods for detecting and controlling mongoose. My dissertation identified how features of the Puerto Rican landscape influence the distribution of mongoose and their neutral genetic variability across multiple environments. I showed that the invasive mongoose maintains a widespread and relatively unstructured panmictic population in Puerto Rico, and that these patterns are not related to the structural connectivity in the island. Proximity to wetlands strongly predicted mongoose presence, although its relative importance varied among my four species distribution models. Strong predictors of mongoose occurrence also indicated that higher temperatures and areas with low precipitation are also important. These observations are consistent with the fact that mongoose density is higher in coastal areas. Although I was not able to discern the complex genetic processes that characterize mongoose demography (i.e. evidence of propagule pressure, genetic drift, founder effects), I did find that mongooses might be exploiting aspects of modified environments. This suggests that although mongoose can travel far from modified habitats, they might do so infrequently, and that management of the species should focusing efforts within vulnerable areas and at the edges of modified habitat. In terms of where to find mongoose populations, I showed that species suitability models are useful, but model type should match the main objective of the project. For example, MaxEnt can be useful for preliminary assessments of species distributions and field survey design with limited presence and spatial data since it was able to identify potential species occurrence. Indeed, MaxEnt correctly identified observed presences. However, boosted regression trees might be the best choice after recalibration with new presence data, when the objective is to find individuals in low suitability areas.

My study is of particular interest to researchers, managers, and policy makers dealing with dynamic landscapes such as islands, and particularly island nations. Human influence and invasive species effects on the local biodiversity depend on the local context of the landscape. My research provides novel information about the current state of the ecological network in Puerto Rico and how an important invasive species is distributed and functionally connected through the island. Islands share unique challenges in integrating conservation, human resource use, and altered patterns of disturbance. Thus, extending conservation theory to application is relevant when decisions need to be urgently made in the face of uncertainty.

References

Barun A, Hanson C, Campbell K, Simberloff D (2011) A review of small Indian mongoose management and eradications on islands. In: Veitch CR, Clout MN, Towns DR (eds) Island invasives: eradication and management. Proceedings of the International Conference on Island Invasives. International Union for Conservation of Nature, Gland, Switzerland, pp 17–25

Bellard C, Cassey P, Blackburn TM (2016) Alien species as a driver of recent extinctions. Biol

Lett 12:20150623. https://doi.org/10.1098/rsbl.2015.0623

Bennett G (2004) Integrating biodiversity conservation and sustainable use : lessons learned from ecological networks. IUCN Commission on Ecosystem Management (CEM), Syzygy, NL

Blackburn TM, Pyšek P, Bacher S, et al (2011) A proposed unified framework for biological invasions. Trends Ecol. Evol. 26:333–339

Borroto-Paez R (2009) Invasive mammals in Cuba: an overview. Biol Invasions 11:2279–2290

Butchart SHM, Clarke M, Smith RJ, et al (2015) Shortfalls and Solutions for Meeting National and Global Conservation Area Targets. Conserv Lett 8:329–337. https://doi.org/10.1111/conl.12158

Butchart SHM, Scharlemann JPW, Evans MI, et al (2012) Protecting Important Sites for Biodiversity Contributes to Meeting Global Conservation Targets. PLoS One 7:e32529. https://doi.org/10.1371/journal.pone.0032529

Courchamp F, Woodroffe R, Roemer G (2003) Removing protected populations to save endangered species. Science (80-) 302:1532–1532

Dirzo R, Young HS, Galetti M, et al (2014) Defaunation in the Anthropocene. Science (80-.). 345:401–406

Doherty TS, Glen AS, Nimmo DG, et al (2016) Invasive predators and global biodiversity loss. Proc Natl Acad Sci U S A 113:11261–11265. https://doi.org/10.1073/pnas.1602480113

Engeman R, Whisson D, Quinn J (2006) Monitoring invasive mammalian predator populations sharing habitat with the critically endangered Puerto Rican parrot Amazona vittata. Oryx 40:95. https://doi.org/10.1017/S0030605305001286

Fagerstone K (2007) Mitigating Impacts of Terrestrial Invasive Species. Encycl. Pest Manag. Vol. II 347–349

Fahrig L, Baudry J, Brotons L, et al (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. Ecol Lett 14:101–12. https://doi.org/10.1111/j.1461-0248.2010.01559.x

Franklin JF, Lindenmayer DB (2009) Importance of matrix habitats in maintaining biological diversity. Proc Natl Acad Sci U S A 106:349–50. https://doi.org/10.1073/pnas.0812016105

Galiana N, Lurgi M, Montoya JM, López BC (2014) Invasions cause biodiversity loss and community simplification in vertebrate food webs. Oikos 123:721–728. https://doi.org/10.1111/j.1600-0706.2013.00859.x

Geldmann J, Joppa LN, Burgess ND (2014) Mapping Change in Human Pressure Globally on Land and within Protected Areas. Conserv Biol 28:1604–1616. https://doi.org/10.1111/cobi.12332

Haddad NM, Brudvig LA, Damschen EI, et al (2014) Potential Negative Ecological Effects of Corridors. Conserv. Biol. 28:1178–1187

Jones HP, Holmes ND, Butchart SHM, et al (2016) Invasive mammal eradication on islands results in substantial conservation gains. Proc Natl Acad Sci U S A 113:4033–4038. https://doi.org/10.1073/pnas.1521179113

Lowe S, Browne M, Boudjelas S, De Poorter M. (2004) 100 of the world’s worst invasive alien species . Auckland, New Zealand

Margules CR, Pressey RL (2000) Systematic conservation planning. Nature 405:243–253. https://doi.org/10.1038/35012251

Millenium Ecosystem Assessment (2005) Ecosystems and Human Well-being: Biodiversity Synthesis. Washington, DC

Nadin-Davis S, Velez J, Malaga C, Wandeler A (2008) A molecular epidemiological study of rabies in Puerto Rico. Virus Res 131:8–15

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002

Resasco J, Haddad N, Orrock J (2014) Landscape corridors can increase invasion by an exotic species and reduce diversity of native species. Ecology 95:2033–2039

Robinson JG (2006) Conservation biology and real-world conservation. Conserv. Biol. 20:658–669

Rohrer J, Hanley J, Nadig A, et al (2011) Small Indian mongoose – management and eradication using DOC 250 kill traps , first lessons from Hawaii. In: Veitch CR, Clout MN, Towns DR (eds) Island invasives: eradication and management. IUCN, Gland, Switzerland, pp 225–227

Russell JC, Kueffer C (2019) Island Biodiversity in the Anthropocene. Annu Rev Environ Resour 44:31–60. https://doi.org/10.1146/annurev-environ-101718-033245

Sakai A, Allendorf F, Holt J (2001) The population biology of invasive species. Annu Rev … 32:305–332

Samways MJ, Pryke JS (2016) Large-scale ecological networks do work in an ecologically complex biodiversity hotspot. Ambio 45:161–172. https://doi.org/10.1007/s13280-015-0697-x

Schreiber SJ, Lloyd-Smith JO (2009) Invasion dynamics in spatially heterogeneous environments. Am Nat 174:490–505. https://doi.org/10.1086/605405

Seaman GA, Randall JE (1962) The mongoose as a predator in the Virgin Islands. J Mammal 43:544–546

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity Is a Vital Element of Landscape Structure. Oikos 68:571. https://doi.org/10.2307/3544927

Tershy BR, Shen KW, Newton KM, et al (2015) The importance of islands for the protection of biological and linguistic diversity. Bioscience 65:592–597

Tilman D, Clark M, Williams DR, et al (2017) Future threats to biodiversity and pathways to their prevention. Nature 546:73–81

Veitch CR, Clout MN, Towns DR (2011) Island Invasives : Eradication and Management

Wilkerson ML (2013) Invasive plants in conservation linkages: a conceptual model that addresses an underappreciated conservation issue. Ecography (Cop) 36:1319–1330. https://doi.org/10.1111/j.1600-0587.2013.00182.x

Citation: Guzmán-Colón, D. K., Pidgeon, A. M., Martinuzzi, S., & Radeloff, V. C. (2020). Conservation planning for island nations: Using a network analysis model to find novel opportunities for landscape connectivity in Puerto Rico. Global Ecology and Conservation, 23, e01075. https://doi.org/10.1016/j.gecco.2020.e01075

Chapter 1. Conservation planning for island nations: Using a network analysis model to find novel opportunities for landscape connectivity in Puerto Rico

Abstract

Oceanic islands are important habitats for many endemic species. Global conservation assessments, however, are too coarse to characterize areas of high human influence or landscape connectivity at a resolution that is useful for conservation planning on most islands. Our goal was to identify landscape elements that are essential for the maintenance of structural connectivity among natural habitat patches on islands. Using the Caribbean island of Puerto Rico as a case study, our specific objectives were to: (1) develop a map of the human footprint, and (2) characterize the connectivity of patches exhibiting low human modification that structurally connect the island's ecological network. We used the human footprint as a measure of impediments to connectivity among Puerto Rico's natural areas using network analysis. We found that more than half of Puerto Rico's current land surface had a low human footprint (56%), but that coastal areas were highly affected by human use (82%). Puerto Rico possesses a compact network of natural areas, with a few patches in the interior mountains critical to structural connectivity. The number of isolated patches is very high; more than 60% of the patches were 2000 m or more apart. Identifying sites that are key hubs to connectivity on islands and ensuring they remain undeveloped is one strategy to balance land use and conservation, and to facilitate the persistence of endemic species. We show here how to improve general conservation assessment methods to be more relevant for islands. There is potential to support an interconnected network of natural areas that promotes landscape connectivity in Puerto Rico among non-coastal habitats, because the human activities are concentrated along the coast whereas the interior mountain range has a relatively low human footprint.

Keywords: Island Conservation, Conservation planning, Networks, Anthropogenic influence

Introduction

Mitigating the effect of human influence on ecological landscapes has become a global conservation issue (Ellis & Ramankutty, 2008). Less than a quarter of the Earth's terrestrial surface remains "wild" and 20% has been classified as semi-natural (Ellis et al. 2010). For island nations, limited size, isolation, and high demand for natural resources all contribute to high vulnerability to human modification (Mimura et al., 2007; Vitousek et al., 1997; Wong et al., 2005). Although occupying only 5% of the global land area, islands are priorities for conservation because of their high levels of endemism, the small population sizes of many island species, and diverse functional traits found among many island species (Borges et al., 2018; Kier et al., 2009). Thus, it is important to identify areas of high conservation value within islands given their significance for global biodiversity and vulnerability to anthropogenic modification. Understanding the current extent, intensity and spatial pattern of human activities, which typically degrade the conservation value of patches to varying degrees, is key to this process. Human footprint maps, *i.e.*, standardized efforts that synthesize multiple anthropogenic threats to biodiversity, have been developed at regional (González-Abraham et al., 2015; Leu et al., 2008;

Tapia-Armijos et al., 2017; Woolmer et al., 2008) and global scales (Geldmann et al., 2014; Sanderson et al., 2002). Unfortunately, however, the resolution of these maps is in most cases too coarse to capture patterns of the human footprint within islands, and there are few examples of human footprint maps developed specifically for islands. The question is how to derive human footprint maps at a scale suitable for conservation planning and management on islands.

In addition to mapping the human footprint, characterizing landscape connectivity is important for conservation planning to identify barriers to the movement of organisms or processes among habitat patches (Taylor et al. 1993; Crooks and Sanjayan 2006; Dobson et al. 1999; Margules and Pressey 2000; Mitchell et al. 2013) . In particular, measures of connectivity based on graph-theory, such as network models, provide a strong framework for evaluating multiple aspects of habitat connectivity (Pascual et al., 2007; Urban & Keitt, 2001), ranging from simple patch and landscape structural indices to more complex spatially explicit metapopulations (Calabrese & Fagan, 2004; Rojas et al., 2016; Rozenfeld et al., 2008). There are few connectivity assessments of islands (but see Zhang & Wang, 2006), and it is unclear if islands have landscape patterns that either uniquely foster or hinder connectivity.

Ecological networks, defined here as a set of spatially linked patches that are relatively uninfluenced by humans, have been the focus of conservation actions around the world aimed at increasing landscape connectivity to conserve biodiversity and other ecosystem functions (Biondi et al., 2012; Damschen et al., 2019; Hermoso et al., 2018). Network topology, the physical configuration of patches connected by links in a network, is an emergent property which stems from features such as the total network area, patch quality, patch density, and permeability of the matrix (Opdam et al., 2006). Network topology can help to predict the spread of information and disease, vulnerability to disturbance, and stability of a system (Albert &

Barabási, 2002; Gastner & Newman, 2006; Melian & Bascompte, 2002). Network analysis has revealed overlooked patterns of resource partitioning for certain species (M. S. Araújo et al., 2008), helped disentangle the effects of habitat loss and fragmentation across multiple scales (Dilts et al., 2016), and helped to identify the spatial decision-making patterns of loggers in the Amazon basin (Walker et al., 2013). Conservation plans for islands could benefit from network analysis for efficient characterization of landscape connectivity. Further, network analysis can easily incorporate more information as it become available for a given species or a process of interest (Dale & Fortin, 2010; Rayfield et al., 2011; Urban et al., 2009).

The goal of this study was to identify those landscape elements that are essential for the maintenance of structural connectivity within a matrix of anthropogenic threats, to serve as resource for landscape conservation planning. Specifically, we aimed to 1) map the human footprint at a scale relevant to management (*i.e.* 30-m resolution), and 2) characterize the connectivity of patches with low human modification that structurally connect the island's ecological network. We used Puerto Rico, an island territory of the United States located in the Greater Antilles archipelago, as a case study.

Methods

Study area

Puerto Rico is an 8,900 km^2 island territory of the USA, one of more than thirty island nations or territories located in the Caribbean (Figure 1). With a population of 3,725,789, Puerto Rico is densely settled (~438 persons/km^2) (United States Census Bureau, 2010), although it has seen a rapid population decline since 2005 (around 20% of the total population, Makoff & Setser, 2017; Meléndez & Hinojosa, 2017).

Altitude ranges from sea level to 1300 m, annual temperatures range from 19.4 °C to 29.7 °C, and precipitation ranges from 701 mm to 4,598 m (Daly et al., 2003). The island features diverse geology and topography, which is expressed as three major physiographic regions: coastal and interior plains, hills and central range, and the Luquillo mountain range (Figure 1). Land cover in Puerto Rico is dominated by forests (39 %) and grasslands (32 %) with 11 % of the area classified as urban with a high level of urban sprawl, the rest is composed of a mix of cultivated land, scrub/shrub, wetlands and shoreline (Gould et al., 2008; Wang et al., 2016). Protected areas occupy 16 % of the total land surface (Caribbean Landscape Conservation Cooperative, 2016). The island supports about 3,100 plant species (more than 250 endemics and up to 300 naturalized exotics) and 378 terrestrial vertebrate species (14 endemic bird species, 15 endemic amphibians, 70 endemic reptiles) (R. Joglar, 2005; R. L. Joglar et al., 2007; Miller & Lugo, 2009).

Human footprint model

Datasets

We selected ten spatial datasets to create a map the human footprint, within the broad categories of land use, human access, and electrical infrastructure (Table 1).

Variable scoring for human footprint

We assigned a score from 0-1 to each variable, representing irreversibility of habitat modification (Table 2). A score of 0 indicates no documented modification and a score of 1 indicates severe human modification and irreversibility. For example, the conversion of a forest patch to pasturelands alters species composition and soil attributes, however, some ecosystem functionality is retained or is restorable (Shimamoto et al., 2018). Therefore, pasturelands were assigned a lower human modification score than urbanized areas, but a higher score than forests (González-Abraham et al., 2015). Our scores were based on published studies relevant to Puerto

Rico and followed the scoring system used in other studies (Etter et al., 2011; González-Abraham et al., 2015; Sanderson et al., 2002; Theobald, 2013; Woolmer et al., 2008)

Land cover

We simplified the 13-class land cover classification scheme to five classes (Table 2). Mixed primary and secondary forests, wetlands, scrub, shrubs and shoreline, grasslands, and cultivated lands (pastures and agriculture). We eliminated: "Open Water" and "Natural Barrens" (represented <0.4% of the island's area or 3,576 ha). The latter included Rocky cliffs and shelves, Gravel beaches and stony shoreline, Fine to coarse sandy beaches, mixed sand and gravel beaches, Salt and mudflats, and they were included in the "shore" category.

Puerto Rico's landscape has experienced rapid land cover change during the 20th century (Helmer et al. 2002). The landscape underwent a 64% reduction in agricultural lands between 1977 and 1992 and these areas mostly reverted to mixed secondary forests. Some of the former agricultural land has converted to pasture (48,000 ha) or urban/developed lands (~ 7,300 ha) (Helmer, 2004). While some pasturelands have also reverted back to secondary forest in the past decades (a total of 64,000ha)(Helmer, 2004). We did not assign specific values to protected areas because we also wanted to capture human modification within their boundaries.

Grassland cover represents a mix of active or abandoned pastures that are is actively maintained by intentional fire and grazing. Other land covers influenced by humans include shrub and scrublands, which have been recently abandoned or are described as marginally active and semi-active pastures. Wetlands are a natural land cover on the island, but when they are not flooded, they are used for grazing or hay cultivation, and some wetlands develop on disturbed saline swamplands. Most shore land is subject to artificial maintenance and periodically flooded (Gould et al., 2008).

Dams

Dams can prevent the migration of riverine species (e.g. shrimp), between freshwater and saltwater (Holmquist et al., 2008; Ligon et al., 1995; March et al., 2003). In Puerto Rico's freshwater ecosystems, even low-head structures on inland streams were found to decrease connectivity for aquatic species (Cooney & Kwak, 2013). To characterize both downstream and upstream effects, we assigned scores at two levels: at the dam site and at the watershed level. We defined large dams as either a) ≥15 m high from the lowest foundation to the crest, spillway discharge of >2000m^3, or b) having a reservoir volume of >1 000 000m^3; Clarke 2000). We defined the zone of influence of each dam to be a circle with a diameter equal to the dam's crest length plus 100 m to account for any spatial inaccuracies of the data, following Woolmer et al. (2008). All dam locations were assigned the highest modification score of 1. To assess the influence of large dams at the watershed level, we calculated the reservoir volume (m^3) per area (km^2) for each watershed influenced by a dam, and using the distribution of values for volume/area, we assigned a human modification score for entire watersheds designating the top 75% of volume/km^2 as 0.5, the median volume/km^2 as 0.3, and the lowest 25% received a score of 1, following the approach developed by WWF-Canada (2003) and implemented by Woolmer et. al. (2008). Watersheds without dams were given a score of 0 (Table 2).

Impervious surfaces

Urban development has especially long-lasting ecological consequences (Blair, 2004; Hansen et al., 2005). We used land cover data to identify impervious surfaces and open developed spaces (Table 2). Open-developed spaces have a mixture of constructed materials, and managed vegetation such as parks or gardens.

Roads and railways

Compared to Australia (11 km/km^2) or Netherlands (331 km/ km^2), Puerto Rico's road network is very dense (301 km/ km^2). Roads fragment landscapes affecting wildlife demography and water quality, promote spread of invasive species, and limit accessibility to natural areas (Fahrig & Rytwinski, 2009; Forman & Alexander, 1998). In tropical dry forests, roads alter sediment production and run-off. Even when a relatively small percentage of the land in dry forests are disturbed, there's an increase in run-off and sediment delivery to into coastal waters (Ramos-Scharrón & Thomaz, 2017).

We divided the roads into primary (highways, expressways) and secondary (local) roads. We set the maximum human modification score for primary roads at 1, and a maximum score for secondary roads at 0.8, and assigned incrementally lower modification scores with increasing distance from roads, out to 1000 m (Table 3). Similarly, we assigned a modification score that declined with distance from railways, out to 500 m, following Woolmer et al (2008). The basis for the buffer values follow a "road-effect zone" based on common ecological effects extending different distances from a road (Forman & Alexander, 1998). Given that we could not find specific studies on effect zones for roads in PR, we assumed that there are road effect even at >500m as previous studies have demonstrated that roads lead to contagious development (Ibisch et al., 2016) . While the linear decay rate followed the methodology established in the re-scaled Global Human Footprint analysis by Woolmer et al. 2008. We believe that we were able to capture road density influence when two road buffer zones overlap within a pixel, these scores are not additive, using Theobald's 2013 fuzzy algebraic sum we aimed to reduce errors due to partial dependence between layers (i.e., road buffers) that can coincide in the same pixel.

Mines

Sand extraction is common in Puerto Rico (Orris & Carbonaro, 1992; Rodriguez, 2017). We distinguished active from inactive mines assigned scores following the same decay function used for roads, with a score of 0.5 for active mines and a maximum score of 0.3 for inactive mines (Table 3).

Power Plants and Electrical substations

We divided Puerto Rico's electrical power infrastructure into two categories: main power plants (score of 1) and electrical substations (0.3) (Table 3), with a distance decay function with distance. Although transmission line data was available, the majority of the electric lines ran parallel to the roads, thus we decided to exclude these from the analysis.

Human footprint calculation

We used a "fuzzy algebraic sum" to calculate the human footprint in Puerto Rico (Bonham-Carter, 1992; Theobald, 2013), and to address the issue of non-independence among variables. The overall value of human modification H_i at each cell i is calculated as:

$$H_i = 1 - \prod_{j=1}^{k} (1 - h_j)$$

where h_j is the human modification score of each layer (j=1…k), with values ranging from zero to one.

Level of agreement between the Global HF and our final HF model

Lastly, we compared the 1 km^2 resolution global human footprint map developed by Sanderson et al. 2002 with our HF map by calculating a Kappa statistic as a degree of agreement of scores between the two maps. This analysis was based on a sample of 1,669 random points with a minimum distance of one km (a total of 8.7% of the Global HF map of the island).

Network datasets

Landscape representation

Patches

We characterized structural connectivity using attributes from both patches (habitat area and human modification score) and links (see section 3.1.2 below for description). We considered patches with >25 ha area and with a human footprint ranging from 0 to 0.3 as high quality. We selected the patch size based on the average home range sizes of IUCN red-listed species and species of concern from the Puerto Rico GAP analysis (Gould et. al. 2008) (Table 4). We used Core Mapper within Gnarly Landscape Utilities in ArcGIS 10.3 (http://www.circuitscape.org/gnarly-landscape-utilities; accessed 30 March 2018) to calculate the average habitat value within a 25-ha circular moving window. The resulting habitat patch layer contained 352 patches with a total area of 3,121km^2 (35% of the total land surface; protected areas represented 16% of the area of these patches).

Links by resistance distance

We used the HF map layer as a cost surface to model pair-wise connectivity between all patches. This layer represents landscape resistance to movement from 0 (no resistance) to 1 (maximum resistance). We used Circuitscape to create the cost surface layer using an eight neighbor rule (McRae et al., 2008).

Network construction

We converted the set of connected patches and their pair-wise resistance values into an adjacency matrix, in which connections present between two patches represented the total resistance distance between zero and one. The network was constructed and analyzed using the “igraph” package in R (Csardi & Nepusz, 2006). Finally, we removed links between patches with >0.6 resistance distance.

Links by Euclidian distance

We built five additional adjacency matrices to construct networks with various dispersal distance thresholds for hypothetical species (from 300 to 2600 m, in 600m steps). We selected the specific distances based on home range sizes of five species prioritized on the IUCN red list and Puerto Rico GAP analysis (Gould et. al. 2008) (Table 4). Centroid-to-centroid Euclidian distances between habitat patches were calculated using SDM toolbox GIS extension for ArcGIS 10.3 (www.sdmtoolbox.org/data/sdmtoolbox/; accessed 15 April 2018).

Network analysis

Network parameters

We quantified several route-specific properties of network connectivity, including flux, redundancy, and vulnerability (Rayfield et. al. 2011) as well as network resilience (Table 5). Within the island’s network, “components” refer to subsets of patches that are structurally connected to each other but are disconnected from other subsets in the network. For this study, the network metrics dealing with “distance” refer to the number of links or paths needed to reach any two patches in the network. Redundancy refers to the presence of multiple or alternate potential movement routes among habitat patches. Route vulnerability captures the stability of the network, and the degree to which the landscape structure funnels movement.

Network topology

We evaluated network topology by comparing the actual network in Puerto Rico to a null model derived from 10,000 networks with the same number of patches and links that were randomly arranged (see Table 5metrics). We fitted the connectivity distribution and the clustering coefficient distribution of each network to a power law (M. Newman, 2005). This processing was done using package “igraph” in R.

Assessing network resilience

We wanted to investigate how the connectivity of the network changed as patches were removed to reveal the network’s robustness to patch loss. Thus, we quantified the effects of patch removal following the principles of percolation theory (Stauffer & Aharony, 1994). In particular, we assessed the fraction of patches that could be removed before the largest component disintegrated into smaller components (*i.e*., network contained 50% or less of its original patches, or reached zero components (Franceschet, 2012; M. E. J. Newman, 2010; Stauffer, 1987). We progressively removed patches using three strategies: 1) in decreasing order from most connected to least connected, 2) random, and 3) in decreasing order of patch area size.

Results

Human footprint in Puerto Rico

Our human footprint map for Puerto Rico provides a spatially detailed view of current human influences on the island (Figure 2. The human footprint in Puerto Rico. The highest human footprint scores (blue colors) are near main roads, coastal areas, and densely populated municipalities. The areas in beige represent the lowest values of human footprint. High and very high categories covered 32% of the island, while the two lowest categories covered 55%. The totals do not add to 100% due to rounding.). More than half of the island’s area (56%) has very low or low human footprint (HF class < 3) and highest HF areas (>5) are concentrated around main roads, cities, and coastal areas.

While most of Puerto Rico has a very low and low human footprint (55% of the area), there were far more patches in the highest HF class than the lowest class (16,787 versus 4,344 for in the lowest HF class), indicating fragmentation in areas where human activities are concentrated (Table 6).

Local HF index comparison with Global HF

Our HF map and the 1 km resolution Global HF developed by Sanderson et. al. (2002) were spatially distinct (unweighted Kappa Index of 0.21 (CI = 0.19-0.23; SE= 0.007) on a scale of -1 to 1, with 1 indicating perfect agreement, 0 indicating what would be expected by chance and a negative value indicating disagreement. In our visual assessment, the influence of roads and dense urban areas were well represented in the Global HF, but it overlooked the heterogeneity of human influence that was apparent at a finer scale. Overall, the Global HF underestimates the amount of the most natural areas and overestimates the amount of human modification (Figure 3).

Spatial patterns of anthropogenic impacts

The three main physiographic regions of the island differ in the degree of human impact experienced (Figure 4). Puerto Rico has higher human footprint scores in the coastal and plain areas than in the interior mountains. In total, 82% of the coastal region fell within medium to high HF classes.

Puerto Rico ecological network

Network topology

We used Puerto Rico's human footprint model to define a landscape ecological network. A total of 152 patches and 481 links were present, organized in a theoretical scale-free network (R^2 = 0.95). Such networks are distinguished by having a few highly connected patches ('hubs') and many patches with few connections (Barabási and Albert 1999). We found that PR's network contained 33 densely connected groups (Q = 0.47; Newman 2006) ranging 1 to 42 patches ($\bar{x}$ =

4). The network had one major hub (Area = 73 km^2) in the central northwest of the island between the municipalities of Arecibo and Utuado (18.343511, -66.749858) which had the most links (31 or 6%).

Network parameters

We compared network parameters between PR ecological network and a simulated random network (Table 7) and found several differences. The clustering coefficient for the PR network was threefold that of the random network, and the connectivity correlation was positive in PR's network but negative in the random network. These characteristics imply that there are tight clusters and a few patches with a disproportionate number of connections, specifically in the central northwestern Karst region of the island.

The values for number of components, size of the largest component, and diameter, along with clustering coefficient and connectivity correlation suggest that the PR network is more compact than random, however, the average shortest path length was larger in the PR network. Thus, the random network had a higher number of short paths between patches. Since size of the largest component is related to the diameter, a better way to compare size of the network is to calculate the ratio of the size to diameter for the largest component. Both networks had the same ratio of 0.5. The random network's average shortest path length was shorter (<2.5) even though it had the largest component, compared to PR network's average shortest path length, which was relatively short (3.4). Both numbers are characteristic of a highly compact network in which hypothetical individuals would potentially interact with others through a path of <3 links even though patches were spread over a large area.

We further explored how the number of connections for each patch were distributed compared to a simulated random network. We found that PR's network displayed a skewed distribution (skewness = 1.07; Figure 6), indicating a complex network, and supporting our finding for a scale-free network in which there are relatively few, but very well connected patches (M. E. J. Newman, 2003).

Finally, we examined connectivity correlation of the PR network, by calculating the mean patch degree of neighbors connected to a patch of k connections (Figure 7). For our simulated random network, there was no relationship between number of patch connections and the mean number of connections of a patch's neighbors, and all patches in the random network were accessible. In contrast, PR's network had a positive relationship ($r = 0.29$) in which patches with the greatest number of connections tended to have neighbors with more connections.

Distance thresholds and patch area

Most patches within the PR network were connected. Unconstrained by distance, only 15% of the patches were completely isolated ($k = 0$), all them in the coastal region. However, when accounting for dispersal threshold distances, 98% of patches were isolated from other patches at a maximum dispersal distance of 1000 m. Even at 2000 m, 60% were still isolated from each other (Figure 8a). The number of components decreased as the threshold distance increased (Figure 8b). Both the percent of isolated patches and the number of components decreased as dispersal threshold distance was increased.

Network robustness

We calculated the robustness of the PR network using three patch removal strategies. First, we quantified the effects of progressively removing the most highly connected patches, second, we removed patches randomly, and third, we progressively removed the largest patches. In the first case, when 32 (of 152, or 21%) highly connected patches were removed, 50% of all connections were lost. Under random removal of patches, 49% of patches needed to be removed to reach that threshold. When removing patches by size (area), we found that 42% of the patches in the network needed to be removed to reach the threshold.

Discussion

Our goals were to develop a human footprint map and characterize the ecological network of Puerto Rico. We found that global analysis did not capture Puerto Rico's heterogeneity, and underestimated natural areas. Our findings highlight the potential of using network analysis to reveal the spatial context and connectivity among natural habitat patches, and the need for island- specific characterization of the human footprint, rather than relying on products developed at the global scale. Further, our network model showed that Puerto Rico possesses a robust network characterized by a concentration of potential habitat patches with high connectivity values in the western mountainous region of the island. However, coastal areas are fragmented and disconnected from the main network, a cause for concern regarding conservation of littoral ecosystem processes.

Patterns of the human footprint

The spatial analysis of Puerto Rico's HF showed clear gradients in human-caused pressures among Puerto Rico's three physiographic regions. The most obvious pattern was the strong difference between the coastal and central mountain regions in terms of number and isolation of patches. Developed areas sprawled from main roads, and impervious surfaces are increasingly encroaching on the protected areas (Castro et al., 2016). Our observations agree with other studies that have shown a positive relationship between roads, human density and conversion to urban cover (e.g. Estes, et. al. 2012; Freitas, Hawbaker, & Metzger, 2010; Hawbaker, et. al. 2005).

In coastal areas, fragmentation was high and the few areas with low human footprint score were isolated. Unfortunately, in addition to direct anthropogenic pressures, the coastal areas are also the most vulnerable to sea level rise as a consequence of climate change (Jury, 2018; Strauss & Kulp, 2018). On the other hand, most of the connected low-HF patches were concentrated in the central mountain range region. The resulting PR network has the largest component and most important hubs in a region where a multi-sectorial effort to protect 3,900 acres of diverse ecosystems has been carried out since 1999 (Ley Núm. 182 Del Año 2014, 2014). The PR network identified priority areas in the western karst and mountainous region that are important for conservation and landscape connectivity, and supports the current law which requires harmonizing the protection of the natural environment, while promoting sustainable development(Ley de Bosques de Puerto Rico, 1975; Ley Orgánica Del Departamento de Recursos Naturales y Ambientales, 1972).

It has been argued that limiting human influence into areas of conservation value may be the most cost-effective and direct way of achieving global sustainability goals (Ibisch et al., 2016b). Our findings suggest that some unprotected land may be suitable for conservation or sustainable mixed uses in Puerto Rico. Currently, protected areas cover around 16% of the island (Caribbean Landscape Conservation Cooperative, 2016) however, our study found that very low and low HF areas cover 53% (Figure 5). Identifying unprotected areas of conservation value is very important since unprotected secondary forests are the dominant forest type in the island (Chazdon et al., 2009). Secondary forests >20 years old in the island exhibit novel plant species assemblages and share some characteristics with mature forests in terms of structure, heterogeneity and complexity (Herrera-Montes & Brokaw, 2010) meaning that ecosystem function could recover in a relatively short period of time if human modification is limited.

Further, second-growth forests will be critical spots for conserving evolutionary diversity for multiple species and assume a pivotal role in terms of carbon sequestration and carbon stocks (Chazdon, 2008; Edwards et al., 2017).

Global land-use trends in indicate that agricultural expansion takes place on fertile soils, while abandoned farmlands are most common on marginal areas with poor soils (Cramer et al., 2008). Abandoned farmlands present an opportunity for ecological restoration efforts on islands, but management of degraded landscapes is highly varied and the impacts of second-growth forest on biodiversity also varies substantially (Queiroz et al., 2014). Islands share similar histories of land use dynamics yet trajectories of forest regeneration over time are varied (Aide et al., 1995; Blondel & Médail, 2009; Chazdon, 2003). In the Caribbean island of St Croix for example, 40-year-old post-agricultural tropical dry forests that regenerated from former plantations shares similar structural characteristics but differs in species composition(Atkinson & Marín-Spiotta, 2015), a pattern similar to post-agricultural forests in Puerto Rico. In contrast, there is a wide range of responses to pastoral or agricultural abandonment in the Mediterranean islands (Médail, 2016). While agricultural abandonment has led to a general increase in matorral and second-growth forests on some Mediterranean islands, the relationship between successional processes, biodiversity, and land-management is still uncertain (Rühl & Pasta, 2007; Schaich et al., 2015). In some cases, the end of traditional agricultural land use in Mediterranean islands has led to severe soil erosion or extinction of endemic species (Médail, 2016; Petanidou et al., 2008). Although global analyses are good at capturing general trends, specific conservation problems can differ by island. In this study we have identified unprotected land areas of high conservation value.

For islands, balancing the trade-offs between urban development and local and global conservation goals is especially challenging. However, our characterization of human modification patterns and connectivity via network analysis provide the quantitative knowledge needed to find that balance. For example, protected area effectiveness is highly dependent on the possibility of movement through unprotected landscapes (Saura et al., 2018). That is why we included landscape context in the form of human modification scores to quantify the important areas for structural connectivity, and our approach is adaptable and hence suitable for diverse conservation goals.

Human footprint at the local level

Global models of the human footprint (Sanderson et al., 2002; Venter et al., 2016) are essential for understanding broad trends of human impacts in terrestrial ecosystems. However, global models are likely to be of limited use for islands as they can be incomplete. For example, most of the datasets used to build the Global HF do not consider oceanic islands which are thus excluded from global analyses. Therefore, to effectively interpret how the human footprint may influence ecological processes on oceanic islands, fine-scale island-specific data are needed. For example, our study revealed that relatively undisturbed natural areas have been underestimated by the global human footprint model, with the result that opportunities for conservation could be overlooked. Other studies have also found greater heterogeneity in fine scale models (Perkl, 2017) relative to global efforts, for example landscapes in the western United States were considered relatively undisturbed in global models but 13% of the region was covered by anthropogenic features in an analysis of fine-scale data (Leu et al., 2008). Similar to our findings for Puerto Rico, in both southern Patagonia and the Northern Appalachian Region of the USA, the global HF model underestimates the wildest areas (Inostroza et al., 2016; Woolmer et al., 2008). Overall, efforts to reassess the human footprint at greater resolution than global models are important and useful for conservation planning at the regional level.

Network topology and parameters

We found that the PR network follows the principle of preferential attachment, in which patches tend to connect to other patches that have an existing high level of connectivity (Barabasi and Albert 1999). Additionally, the PR system is highly compact in that within clusters, the majority of patches are connected to other patches via a path of < 3 connections. Nevertheless, the high skewed patch-degree distribution and positive connectivity correlation highlights the fact that there are many patches with few or no links to other patches. In other words, there are key patches that hold the network together and those hubs tend to be clustered (Minor & Urban, 2008). The clustering coefficient was higher for the real PR network than for the simulated random network, indicating that there are redundant or alternative paths in the real network. The PR network structure is attributable to the large extent of second-growth forest cover present in the western central mountain range in Puerto Rico. Prioritizing patches according to their degree of connectivity within this network can support mechanisms that promote species coexistence by facilitating colonization and promoting species' ability to survive in the aftermath of diverse disturbances (*e.g.* hurricanes, conversion to agriculture or pasture) (Uriarte et al., 2012). It has been shown that distance between patches influences forest recovery in Puerto Rico (Hogan et al., 2016), however in our network model, we found that only a few patches within El Yunque National Forest in the east of the island served as key hubs that provided connectivity for otherwise isolated habitat patches of the eastern coastal areas. Nevertheless, these protected areas in the east of the island are crucial for the maintenance of endemic biodiversity and for various ecosystem processes and services (Lugo, 2005), but they are in constant threat of human modification.

There are positive and negative conservation implications that emerge from a scale-free network such as the one that we found. In terms of stability or resilience, if patches were to disappear randomly, there is a high likelihood that hubs would not be affected and that overall, components would remain connected. From our analysis, it appears that the PR network could sustain a fair amount of random patch loss before overall landscape connectivity is compromised, but loss of hubs would be highly detrimental (Barabási & Bonabeau, 2003; Urban & Keitt, 2001).

Patch isolation and distance thresholds

The combination of habitat isolation, dispersal limitation and low reproduction rates of organisms exacerbates the likelihood of extinctions (Kadmon & Allouche, 2007). Our assessment of isolated patches revealed that habitat connections for species with limited dispersal ability were few and many patches were isolated. For example, there are three release sites for the captive-bred and endangered Puerto Rican parrot that differed in the degree of structural connectivity. Two of these sites, the Rio Abajo and Maricao State Forests belong to the main component of the network in the west mountainous region of the island, while the other site, El Yunque National Forest appeared isolated. Based on species dispersal distance only the parrot could reach the patches within the main the network in the western region. On the other hand, parrot dispersal is restricted on habitat patches in El Yunque that have a strong human influence in its buffer zone with only a few links connecting them to coastal habitat. These distinct spatial patterns and the fact that El Yunque is considered sub-optimal habitat for this species, are a serious consideration for the Amazon population recovery. The level of human influence in eastern Puerto Rico contribute to patch isolation and potential species movement.

Limitations

Connectivity and resistance surface models derived from naturalness indices may not contain enough information for identifying the needs of habitat specialist species. However, models of the HF are repeatable, relatively simple and further, connectivity models based on naturalness have been shown to provide a good proxy for focal species (Krosby et al., 2015). There's wide acceptance that biodiversity conservation must adopt a dynamic approach when tackling climate change, biological invasions and habitat fragmentation (Harrison et al., 2006; Willis et al., 2009). Using HF as a resistance surface for landscape connectivity models can be the first step to analyze island natural areas networks within modified landscapes and can later be complemented with analyses that focus on individual species. Further, the graph-theoric approach for connectivity presented in this study is dynamic in the sense that patch attributes are modifiable depending on the conservation objective being tested.

Conclusion

The main driver of biodiversity decline is human pressure on Earth's ecosystems. Island ecosystems remain disproportionally threatened by a number of anthropogenic stressors, and as a result, it is imperative to not only study the spatial arrangement of those remaining natural and semi-natural areas, but to understand the existing structural connectivity among them. The smaller geographic size of islands often means that there are many vulnerable species due to small areas of habitat, small population sizes and high endemism. Conservation and economic development goals often clash, placing unsustainable demands on resources thus leading to exploitation of natural resources and habitat modification (Graham et al., 2017). In Puerto Rico, forests in the interior mountains are subject to less human activity based on our human footprint model than coastal areas. Furthermore, patches within these forests were of utmost importance for connecting isolated coastal patches to the main network. Our analysis suggests that Puerto Rico's ecological network is likely to maintain its resilience in the face of frequent disturbances such as hurricanes, if other disturbances do not substantially alter the network. The relatively robust ecological network is in large measure attributable to agricultural abandonment and subsequent second growth forest. Other islands might exhibit a different distribution of human modification patterns given the different histories of extraction and diverse natural forces that alter their habitats (Velmurugan, 2018), however, it is plausible that an analysis of coastal habitat patches in other island nations reveal a similar pattern of patch isolation and fragmentation. We found that existing legal protections in the western mountain region of Puerto Rico were crucial for the tightly connected ecological network found and that protection of the rainforest is critical for maintaining natural areas connectivity in the eastern region. Our study provides a guide to planning and prioritizing conservation efforts based on both naturalness and importance to maintaining overall landscape connectivity, and having such analyses completed can be useful when windows of opportunity for conservation actions occur.

References

Aide TM, Zimmerman JK, Herrera L, et al (1995) Forest recovery in abandoned tropical pastures in Puerto Rico. For Ecol Manage 77:77–86. https://doi.org/10.1016/0378-1127(95)03576-V

Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97. https://doi.org/10.1103/RevModPhys.74.47

Araújo MS, Guimarães PR, Svanbäck R, et al (2008) Network analysis reveals contrasting effects of intraspecific competition on individual vs. population diets. Ecology 89:1981–1993. https://doi.org/10.1890/07-0630.1

Atkinson EE, Marín-Spiotta E (2015) Land use legacy effects on structure and composition of subtropical dry forests in St. Croix, U.S. Virgin Islands. For Ecol Manage 335:270–280. https://doi.org/10.1016/J.FORECO.2014.09.033

Barabasi A-L, Albert R (1999a) Emergence of scaling in random networks. Science 286:509–12. https://doi.org/10.1126/science.286.5439.509

Barabasi A-L, Albert R (1999b) Emergence of scaling in random networks. Science 286:509–12. https://doi.org/10.1126/science.286.5439.509

Barabási A-L, Bonabeau E (2003) Scale-Free Networks. Sci Am 288:60–69. https://doi.org/10.1038/scientificamerican0503-60

Biondi E, Casavecchia S, Pesaresi S (2012) Natura 2000 and the Pan-European Ecological Network : a new methodology for data integration. 1741–1754. https://doi.org/10.1007/s10531-012-0273-7

Blair R (2004) The effects of urban sprawl on birds at multiple levels of biological organization. Ecol Soc 9:

Blondel J, Médail F (2009) Biodiversity and conservation. In: Woodward JC (ed) The Physical Geography of the Mediterranean. Oxford University Press, Oxford, pp 615–650

Bonham-Carter GF (1992) Fuzzy logic method. In: Geographic Information Systems for Geoscientists, Modelling with GIS

Borges PA V., Cardoso P, Kreft H, et al (2018) Global Island Monitoring Scheme (GIMS): a proposal for the long-term coordinated survey and monitoring of native island forest biota. Biodivers Conserv 27:2567–2586. https://doi.org/10.1007/s10531-018-1553-7

Calabrese J, Fagan W (2004) A comparison-shopper's guide to connectivity metrics. Front Ecol … 2:529–536

Caribbean Landscape Conservation Cooperative (2016) Puerto Rico Protected Areas Database. San Juan, Puerto Rico

Castro J, Quiñones M, Gould WA (2016) Characterization of the Network of Protected Areas in Puerto Rico. Caribb Nat 1–16

Chazdon RL (2003) Tropical forest recovery: legacies of human impact and natural disturbances. Perspect Plant Ecol Evol Syst 6:51–71. https://doi.org/10.1078/1433-8319-00042

Chazdon RL (2008) Beyond deforestation: restoring forests and ecosystem services on degraded lands. Sci 320:1458–60. https://doi.org/10.1126/science.1155365

Chazdon RL, Peres CA, Dent D, et al (2009) The Potential for Species Conservation in Tropical Secondary Forests. Conserv Biol 23:1406–1417. https://doi.org/10.1111/j.1523-1739.2009.01338.x

Cooney PB, Kwak TJ (2013) Spatial Extent and Dynamics of Dam Impacts on Tropical Island Freshwater Fish Assemblages. Bioscience 63:176–190. https://doi.org/10.1525/bio.2013.63.3.6

Cramer V, Hobbs R, Standish R (2008) What's new about old fields? Land abandonment and ecosystem assembly. Trends Ecol Evol 23:104–112. https://doi.org/10.1016/j.tree.2007.10.005

Crooks KR, Sanjayan MA (2006) Connectivity conservation. Cambridge University Press, New York

Csardi G, Nepusz T (2006) The igraph software package for complex network research. InterJournal, Complex Syst

Dale MRT, Fortin M-J (2010) From Graphs to Spatial Graphs. Annu Rev Ecol Evol Syst 41:21–38. https://doi.org/10.1146/annurev-ecolsys-102209-144718

Daly C, Helmer EH, Quiñones M (2003) Mapping the climate of Puerto Rico, Vieques and Culebra. Int J Climatol 23:1359–1381. https://doi.org/10.1002/joc.937

Damschen EI, Brudvig LA, Burt MA, et al (2019) Ongoing accumulation of plant diversity through habitat connectivity in an 18-year experiment. Sci 365:1478–1480. https://doi.org/10.1126/science.aax8992

Dilts TE, Weisberg PJ, Leitner P, et al (2016) Multiscale connectivity and graph theory highlight critical areas for conservation under climate change. Ecol Appl 26:1223–1237. https://doi.org/10.1890/15-0925

Dobson A, Ralls K, Foster M, et al (1999) Corridors: Reconnecting Fragmented Landscapes. In: Soulé ME, Terborgh J (eds) Continental Conservation: Scientific Foundations of Regional Reserve Networks. Island Press, Washington, pp 129–160

Edwards DP, Massam MR, Haugaasen T, Gilroy JJ (2017) Tropical secondary forest regeneration conserves high levels of avian phylogenetic diversity. Biol Conserv 209:432–439. https://doi.org/10.1016/J.BIOCON.2017.03.006

Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the world. Front Ecol Environ 6:439–447. https://doi.org/10.1890/070062

Estes AB, Kuemmerle T, Kushnir H, et al (2012) Land-cover change and human population trends in the greater Serengeti ecosystem from 1984–2003. Biol Conserv 147:255–263. https://doi.org/10.1016/J.BIOCON.2012.01.010

Etter A, McAlpine CA, Seabrook L, Wilson KA (2011) Incorporating temporality and biophysical vulnerability to quantify the human spatial footprint on ecosystems. Biol Conserv 144:1585–1594. https://doi.org/10.1016/j.biocon.2011.02.004

Fahrig L, Rytwinski T (2009) Effects of Roads on Animal Abundance: an Empirical Review and

Synthesis. Ecol Soc 14:art21. https://doi.org/10.5751/ES-02815-140121

Forman RTT, Alexander LE (1998) Roads and their major ecological effects. Annu Rev Ecol Syst 29:207–231. https://doi.org/10.1146/annurev.ecolsys.29.1.207

Franceschet M (2012) The large-scale structure of journal citation networks. J Am Soc Inf Sci Technol 63:837–842. https://doi.org/10.1002/asi.22608

Freitas SR, Hawbaker TJ, Metzger JP (2010) Effects of roads, topography, and land use on forest cover dynamics in the Brazilian Atlantic Forest. For Ecol Manage 259:410–417. https://doi.org/10.1016/J.FORECO.2009.10.036

Gastner MT, Newman MEJ (2006) Optimal design of spatial distribution networks. Phys Rev E 74:016117. https://doi.org/10.1103/PhysRevE.74.016117

Geldmann J, Joppa LN, Burgess ND (2014) Mapping Change in Human Pressure Globally on Land and within Protected Areas. Conserv Biol 28:1604–1616. https://doi.org/10.1111/cobi.12332

González-Abraham C, Ezcurra E, Garcillán PP, et al (2015) The human footprint in mexico: Physical geography and historical legacies. PLoS One 10:1–17. https://doi.org/10.1371/journal.pone.0121203

Gould W, Alarcón C, Fevold B, et al (2008) The Puerto Rico Gap Analysis Project. Volume 1: Land cover, vertebrate species distributions, and land stewardship. Gen Tech Rep IITF-GTR-39 165 p.

Graham NR, Gruner DS, Lim JY, Gillespie RG (2017) Island ecology and evolution: challenges in the Anthropocene. Environ Conserv 44:323–335. https://doi.org/10.1017/S0376892917000315

Hansen AJ, Knight RL, Marzluff JM, et al (2005) Effects of exurban development on

Biodiverity: Patterns, mechanisms, and research needs. Ecol Appl 15:1893–1905. https://doi.org/10.1890/05-5221

Harrison PA, Berry PM, Butt N, New M (2006) Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. Environ Sci Policy 9:116–128. https://doi.org/10.1016/j.envsci.2005.11.003

Hawbaker TJ, Radeloff VC, Hammer RB, Clayton MK (2005) Road Density and Landscape Pattern in Relation to Housing Density, and Ownership, Land Cover, and Soils. Landsc Ecol 20:609–625. https://doi.org/10.1007/s10980-004-5647-0

Helmer EH (2004) Forest conservation and land development in Puerto Rico

Hermoso V, Morán-Ordóñez A, Brotons L (2018) Assessing the role of Natura 2000 at maintaining dynamic landscapes in Europe over the last two decades: implications for conservation. Landsc Ecol 33:1447–1460. https://doi.org/10.1007/s10980-018-0683-3

Herrera-Montes A, Brokaw N (2010) Conservation value of tropical secondary forest: A herpetofaunal perspective. Biol Conserv 143:1414–1422

Hogan JA, Zimmerman JK, Uriarte M, et al (2016) Land-use history augments environment-plant community relationship strength in a Puerto Rican wet forest. J Ecol 104:1466–1477. https://doi.org/10.1111/1365-2745.12608

Holmquist JG, Schmidt-Gengenbach JM, Yoshioka BB (2008) High Dams and Marine-Freshwater Linkages: Effects on Native and Introduced Fauna in the Caribbean. Conserv Biol 12:621–630. https://doi.org/10.1111/j.1523-1739.1998.96427.x

Ibisch PL, Hoffmann MT, Kreft S, et al (2016a) A global map of roadless areas and their conservation status. Science 354:1423–1427. https://doi.org/10.1126/science.aaf7166

Ibisch PL, Hoffmann MT, Kreft S, et al (2016b) A global map of roadless areas and their conservation status. Science (80-) 354:1423–1427. https://doi.org/10.1126/science.aaf7166

Inostroza L, Zasada I, König HJ (2016) Last of the wild revisited: assessing spatial patterns of human impact on landscapes in Southern Patagonia, Chile. Reg Environ Chang 1–15. https://doi.org/10.1007/s10113-016-0935-1

Joglar R (2005) Biodiversidad de Puerto Rico. Editorial del Instituto de Cultura Puertorriqueña

Joglar RL, Álvarez AO, Aide TM, et al (2007) Conserving the Puerto Rican herpetofauna. Appl Herpetol 4:327–345. https://doi.org/10.1163/ej.9789004194083.i-439

Jury MR (2018) Puerto Rico sea level trend in regional context. Ocean Coast Manag 163:478–484. https://doi.org/10.1016/J.OCECOAMAN.2018.08.006

Kadmon R, Allouche O (2007) Integrating the Effects of Area, Isolation, and Habitat Heterogeneity on Species Diversity: A Unification of Island Biogeography and Niche Theory. Am Nat 170:443–454. https://doi.org/10.1086/519853

Kier G, Kreft H, Lee TM, et al (2009) A global assessment of endemism and species richness across island and mainland regions. Proc Natl Acad Sci U S A 106:9322–7. https://doi.org/10.1073/pnas.0810306106

Krosby M, Breckheimer I, John Pierce D, et al (2015) Focal species and landscape “naturalness” corridor models offer complementary approaches for connectivity conservation planning. Landsc Ecol 30:2121–2132. https://doi.org/10.1007/s10980-015-0235-z

Leu M, Hanser SE, Knick ST, et al (2008) The Human Footprint in the West : A Large-Scale Analysis of Anthropogenic Impacts. Ecol Appl 18:1119–1139

Ley del Bosque Modelo de Puerto Rico (2014) Ley Núm. 182 del año 2014

Ligon FK, Dietrich WE, Trush WJ (1995) Downstream Ecological Effects of Dams. Bioscience 45:183–192. https://doi.org/10.2307/1312557

Lugo AE (2005) The Ecosystems of the Luquillo Mountains. Acta Científica 19:3–6

Makoff G, Setser BW (2017) Puerto Rico Update: PROMESA, Population Trends, Risks to the Fiscal and Economic Plan - and Now Maria. CIGI Pap. No. 146 1–28

March JG, Benstead JP, Pringle CM, Scatena FN (2003) Damming Tropical Island Streams: Problems, Solutions, and Alternatives. Bioscience 53:1069–1078. https://doi.org/10.1641/0006-3568(2003)053[1069:dtisps]2.0.co;2

Margules CR, Pressey RL (2000) Systematic conservation planning. Nature 405:243–253. https://doi.org/10.1038/35012251

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724. https://doi.org/10.1890/07-1861.1

Médail F (2016) Plant biodiversity and vegetation on Mediterranean islands in the face of global change. In: Moatti JP (ed) The Mediterranean region under climate change. IRD Éditions, Marseille, pp 363–376

Meléndez E, Hinojosa J (2017) Estimates of Post-Hurricane Maria Exodus from Puerto Rico. New York, NY

Melian CJ, Bascompte J (2002) Food web structure and habitat loss. Ecol Lett 5:37–46. https://doi.org/10.1046/j.1461-0248.2002.00280.x

Miller GL, Lugo AE (2009) Guide to the ecological systems of Puerto Rico. San Juan, PR

Mimura N, Nurse L, McLean RF, et al (2007) Small islands. In: Parry ML, Canziani OF, Palutikof PJ, et al. (eds) Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change . Cambridge University Press, Cambridge, UK, pp 687–716

Minor ES, Urban DL (2008) A graph-theory framework for evaluating landscape connectivity and conservation planning. Conserv Biol 22:297–307. https://doi.org/10.1111/j.1523-1739.2007.00871.x

Mitchell MGE, Bennett EM, Gonzalez A (2013) Linking Landscape Connectivity and Ecosystem Service Provision: Current Knowledge and Research Gaps. Ecosystems 16:894–908. https://doi.org/10.1007/s10021-013-9647-2

Newman M (2005) Power laws, Pareto distributions and Zipf's law. Contemp Phys 46:323–351. https://doi.org/10.1080/00107510500052444

Newman MEJ (2003) The Structure and Function of Complex Networks. SIAM Rev 45:167–256. https://doi.org/10.1137/S003614450342480

Newman MEJ (2010) Networks: An Introduction. Oxford University Press

Opdam P, Steingröver E, Rooij S Van (2006) Ecological networks: A spatial concept for multi-actor planning of sustainable landscapes. Landsc Urban Plan 75:322–332. https://doi.org/10.1016/j.landurbplan.2005.02.015

Orris GJ, Carbonaro M (1992) Puerto Rico: Industrial mineral mines, prospects, and occurrences in the Mineral Resources Data System (MRDS)

Pascual M, Chapuis MP, Mestres F, et al (2007) Introduction history of Drosophila subobscura in the New World: A microsatellite-based survey using ABC methods. Mol Ecol 16:3069–3083. https://doi.org/10.1111/j.1365-294X.2007.03336.x

Perkl RM (2017) Measuring landscape integrity (LI): development of a hybrid methodology for planning applications. J Environ Plan Manag 60:92–114. https://doi.org/10.1080/09640568.2016.1142863

Petanidou T, Kizos T, Soulakellis N (2008) Socioeconomic Dimensions of Changes in the Agricultural Landscape of the Mediterranean Basin: A Case Study of the Abandonment of Cultivation Terraces on Nisyros Island, Greece. Environ Manage 41:250–266. https://doi.org/10.1007/s00267-007-9054-6

Queiroz C, Beilin R, Folke C, Lindborg R (2014) Farmland abandonment: threat or opportunity for biodiversity conservation? A global review. Front Ecol Environ 12:288–296. https://doi.org/10.1890/120348

Ramos-Scharrón CE, Thomaz EL (2017) Runoff Development and Soil Erosion in a Wet Tropical Montane Setting under Coffee Cultivation. L Degrad Dev 28:936–945. https://doi.org/10.1002/ldr.2567

Rayfield B, Fortin MJ, Fall A (2011) Connectivity for conservation: A framework to classify network measures. Ecology 92:847–858. https://doi.org/10.1890/09-2190.1

Rodriguez R (2017) Sand and Gravel Resources of Puerto Rico. Guaynabo, Puerto Rico

Rojas ER, Sueur C, Henry P-Y, et al (2016) Network Analysis Shows Asymmetrical Flows within a Bird Metapopulation. PLoS One 11:e0166701. https://doi.org/10.1371/journal.pone.0166701

Rozenfeld AF, Arnaud-Haond S, Herná Ndez-García E, et al (2008) Network analysis identifies weak and strong links in a metapopulation system

Rühl J, Pasta S (2007) Plant succession on Sicilian terraces. Ann di Bot 7:. https://doi.org/10.4462/ANNBOTRM-9090

Sanderson EW, Jaiteh M, Levy M a., et al (2002) The Human Footprint and the Last of the Wild. Bioscience 52:891. https://doi.org/10.1641/0006-3568(2002)052[0891:THFATL]2.0.CO;2

Saura S, Bertzky B, Bastin L, et al (2018) Protected area connectivity: Shortfalls in global targets and country-level priorities. Biol Conserv 219:53–67. https://doi.org/10.1016/J.BIOCON.2017.12.020

Schaich H, Kizos T, Schneider S, Plieninger T (2015) Land Change in Eastern Mediterranean Wood-Pasture Landscapes: The Case of Deciduous Oak Woodlands in Lesvos (Greece). Environ Manage 56:110–126. https://doi.org/10.1007/s00267-015-0496-y

Shimamoto CY, Padial AA, da Rosa CM, Marques MCM (2018) Restoration of ecosystem services in tropical forests: A global meta-analysis. PLoS One 13:e0208523. https://doi.org/10.1371/journal.pone.0208523

Stauffer D (1987) Introduction to Percolation Theory. Phys. Today 40:122–123

Stauffer D, Aharony A (1994) Cluster numbers. In: Introduction to Percolation Theory. Taylor & Francis, London, p 45

Strauss B, Kulp S (2018) Sea-Level Rise Threats in the Caribbean: Data, tools, and analysis for a more resilient future. Princeton

Tapia-Armijos MF, Homeier J, Draper Munt D (2017) Spatio-temporal analysis of the human footprint in South Ecuador: Influence of human pressure on ecosystems and effectiveness of protected areas. Appl Geogr 78:22–32. https://doi.org/10.1016/j.apgeog.2016.10.007

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity Is a Vital Element of Landscape Structure. Oikos 68:571. https://doi.org/10.2307/3544927

Theobald DM (2013) A general model to quantify ecological integrity for landscape assessments and US application. Landsc Ecol 28:1859–1874. https://doi.org/10.1007/s10980-013-9941-6

United States Census Bureau (2010) United States Census Bureau. https://www.census.gov/topics/population/data.html

Urban DL, Keitt T (2001) Landscape Connectivity : A Graph-Theoretic Perspective. Ecology

82:1205–1218

Urban DL, Minor ES, Treml EA, Schick RS (2009) Graph models of habitat mosaics. Ecol Lett 12:260–273. https://doi.org/10.1111/j.1461-0248.2008.01271.x

Uriarte M, Clark JS, Zimmerman JK, et al (2012) Multidimensional trade-offs in species responses to disturbance: implications for diversity in a subtropical forest. Ecology 93:191–205. https://doi.org/10.1890/10-2422.1

Velmurugan A (2018) The Nature and Characters of Tropical Islands. In: Sivaperuman C, Velmurugan A, Singh AK, Jaisankar I (eds) Biodiversity and Climate Change Adaptation in Tropical Islands. Academic Press, San Diego, CA, pp 3–30

Venter O, Sanderson EW, Magrach A, et al (2016) Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nat Commun 7:1–11. https://doi.org/10.1038/ncomms12558

Vitousek PM, Mooney H a, Lubchenco J, Melillo JM (1997) Human Domination of Earth' s Ecosystems. Science (80-) 277:494–499. https://doi.org/10.1126/science.277.5325.494

Walker R, Arima E, Messina J, et al (2013) Modeling spatial decisions with graph theory: logging roads and forest fragmentation in the Brazilian Amazon. Ecol Appl 23:239–254. https://doi.org/10.1890/11-1800.1

Wang C, Yu M, Gao Q, et al (2016) Fusion of optical and SAR remote sensing images for tropical forests monitoring. In: American Geophysical Union

Willis SG, Hole DG, Collingham YC, et al (2009) Assessing the Impacts of Future Climate Change on Protected Area Networks: A Method to Simulate Individual Species' Responses. Environ Manage 43:836–845. https://doi.org/10.1007/s00267-008-9125-3

Wong PP, Marone E, Lana P, et al (2005) Island Systems. In: Rashid M. Hassan, Robert Scholes

NA (ed) Ecosystems and Human Well-being: Current State and Trends, Illustrated. Island Press, pp 663–678

Woolmer G, Trombulak SC, Ray JC, et al (2008) Rescaling the Human Footprint: A tool for conservation planning at an ecoregional scale. Landsc Urban Plan 87:42–53. https://doi.org/10.1016/j.landurbplan.2008.04.005

World Wildlife Federation Canada (2003) The Nature Audit: Setting Canada's Conservation Agenda for the 21st Century. Report No. 1. Toronto

Zhang L, Wang H (2006) Planning an ecological network of Xiamen Island (China) using landscape metrics and network analysis. Landsc Urban Plan 78:449–456. https://doi.org/10.1016/J.LANDURBPLAN.2005.12.004

(1972) Ley Orgánica del Departamento de Recursos Naturales y Ambientales. Puerto Rico

(1975) Ley de Bosques de Puerto Rico

Figures and tables

Table 1. Variables, data sources and spatial resolution or scale of data used in the creation of the human footprint index for Puerto Rico.

Category	Variable	Source	Spatial resolution or scale of data
Land Use			
	Land cover:	NOAA's Coastal Change Analysis program. High-Resolution Land Cover Classification scheme	30-m grid
	Dams	US Army Corps of Engineers National Inventory of Dams. 2013.	Point vector layer
	Watersheds	United States Geological Survey Watershed Boundary Dataset. 2014	Generally developed at 1:24,000/1:12,000 scale
	Mines (Open)	Mineral Resources Data System (MRDS): U.S. Geological Survey Digital Data. 2005	Point vector layer
Access			
	Roads	USA: Census – Tiger/Line Files. 2010	1:100000
	Railways	Puerto Rico Transportation and Roads Authority	Line vector layer
Electrical Power Infrastructure		Puerto Rico Planning Board	Point vector layer

Table 2. Impervious surfaces and reservoir volume density scores of human modification for each variable used in the analysis.

Variable	Class	Score	
	Mixed Forests	0	0
Land cover	Wetlands	0.1	1
	Scrub/Shrub/Shore	0.3	3
	Grasslands and Herbaceous	0.5	5
	Cultivated/Pastures, Hay	0.6	6
Developed land	Impervious	1	
	Open developed spaces	0.7	
Dams (Volume of water /land area)			
	≥ 310409 m^3/km^2	0.3	
	14239 – 310409 m3/km2	0.2	
	< 14239 m3/km2		
		0.1	

Table 3. Human modification scores for each infrastructure variable used in the analysis

Variable	0 – 10 m	11 – 90 m	91 – 500m	501 – 1000m
Roads				
Primary	1	0.8	0.6	0.4
Secondary	0.8	0.6	0.4	0.2
Railways	0.6	0.4	0	0
Mines				
Active	0.5	0.3	0.1	0
Inactive	0.3	0.1	0	0
Electrical Power				
Power Stations	1	0.6	0.2	0.1
Substations	0.3	0.2	0.1	0

Table 4. Several native species average home range (ha) and maximum daily movement recorded in literature.

Species	Average home range (ha)	Maximum daily movement (m)	Reference
Yellow shouldered blackbird (*Angelaius xanthomus*)	256	10,000	Post 1981; Post 2011
Puerto Rican Amazon (*Amazona vittata*)	22	2,058	Lindsay et. al. 1991; Snyder 1987
Puerto Rican Nightjar (*Anthrostomus noctitherus*)	5.2	360	Vilella 1995; Vilella 2010
Puerto Rican Boa (*Epicrates inornatus*)	11	26	Puente-Rolón, et. al. 2004; Wunderle and Mercado 2004
Various frog species Eleuctherodactylus spp.	0.0005	5	Woolbright 1985; Ovaska 1992;

Table 5. Topological network-level measures (Adapted from Rayfield et. al. 2011 and references within; Minor & Urban, 2008).

.

Connectivity Property	Metrics	Definition	Ecological Relevance
Route-specific flux	Degree Distribution	Probability distribution of patch degrees (connections) over the entire network. P(k)	Distribution of potential source and sink habitat patches.
	Network order (Number of Patches *N*)	Total number of patches within network	Number of habitat patches in the habitat network
	Network size (Number of Connections *k*)	Total number of links within network	Number of pairs of directly connected habitat patches
	Number of components	Total number of groups of structurally connected patches and links	Number of distinct, unconnected groups of habitat patches. Spatially isolated components

	Average shortest path length (characteristic path length)	Average length of the shortest path connecting patch pairs. Implies efficiency of movement within a network. This metric requires discovering all possible paths between patches *i* and *j*, and then finding the shortest path length (l_{ij}). Note that length refers to the number of links or paths between any two patches.	When short (<6 steps; Travers & Milgram, 1969), all patches tend to be easily reachable. Could imply a patchy landscape rather than a hierarchical organization.
	Diameter (component level) and size of largest component (network level)	Measures the greatest distance of a path between any pair of patches in the network. Can be	Indicates the compactness of a component, short diameter implies fast

		interpreted as the easiness of an organism to functionally reach all other patches in the network.	movement through the network.
Route Redundancy	Clustering Coefficient	Measures the average fraction of the patch's neighbors that are also neighbors with each other.	A high clustering coefficient can facilitate dispersal among patches and more resilient to patch removal
Route Vulnerability	Connectivity Correlation	Measures the relationship between average connections of a focal patch relative to the average number of connections of its neighbors	High values can indicate a presence of sub-networks. Sub-networks (compartmentalization) tend to reduce the spread of a surging disturbance
Resilience	Iterative removal of patches based on	Number of patches whose removal	Simulates the overall connectivity of the

	three percolation strategies: random, area, degree	disconnects the largest component of the network	network and relative importance of habitat patches when they are destroyed

Table 6. Total number of patches and their proportion in the island of Puerto Rico per human footprint category.

Value	**Numbers of Patches**	**Proportion in landscape**
Low (0-0.3)	4,344	0.55
Med (>0.3-0.7)	15,271	0.13
High (>0.7-1)	16,787	0.32

Table 7. Graph diagnostic metrics for assessing connectivity of networks.

	Puerto Rico Network	Random Network
Number of components	33	24
Size of largest component	112	129
Diameter	5.55	6
Average shortest path length	3.27	2.73
Clustering coefficient	0.25	0.09
Connectivity correlation	0.29	-0.07

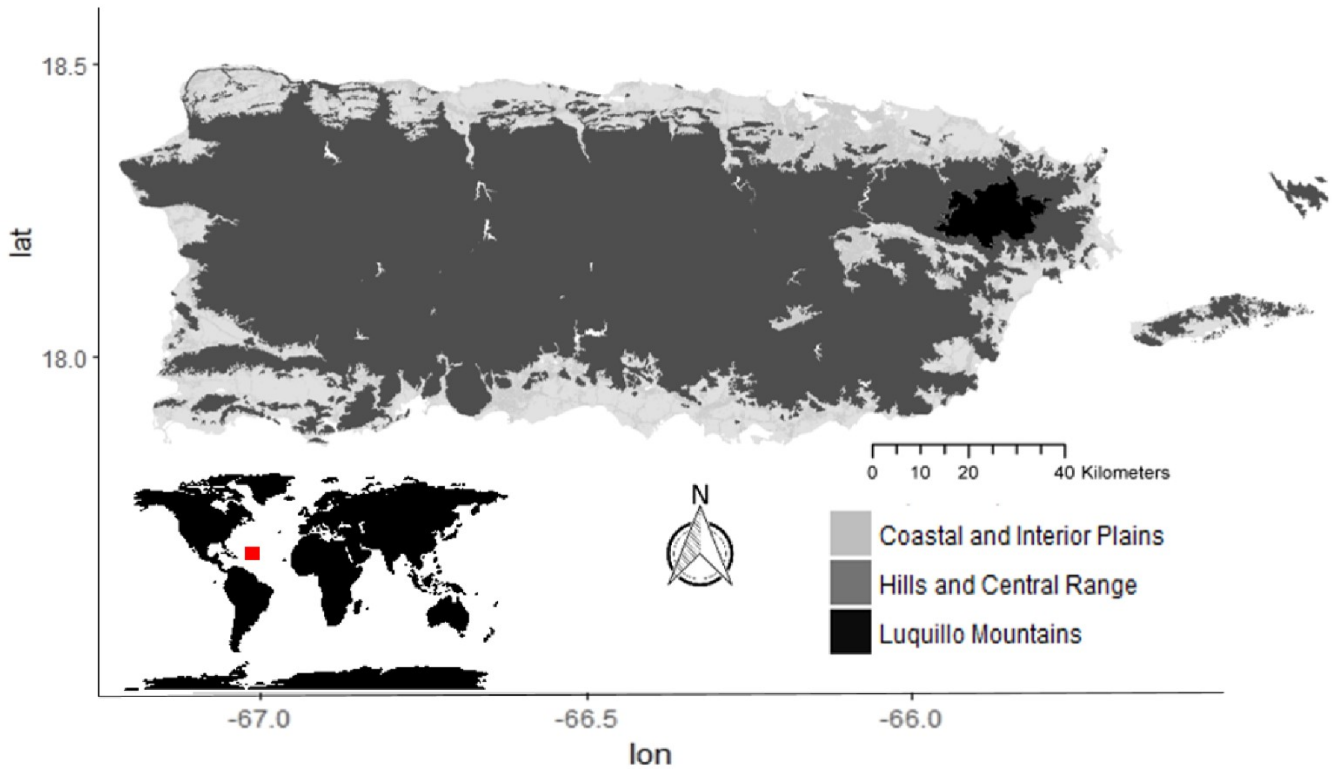


Figure 1. Puerto Rico, one of more than thirty island nations or territories located in the Caribbean (red square), has three major physiographic regions: Coastal and Interior plains, Hills and central mouFigure ntain range and the Luquillo Mountains.

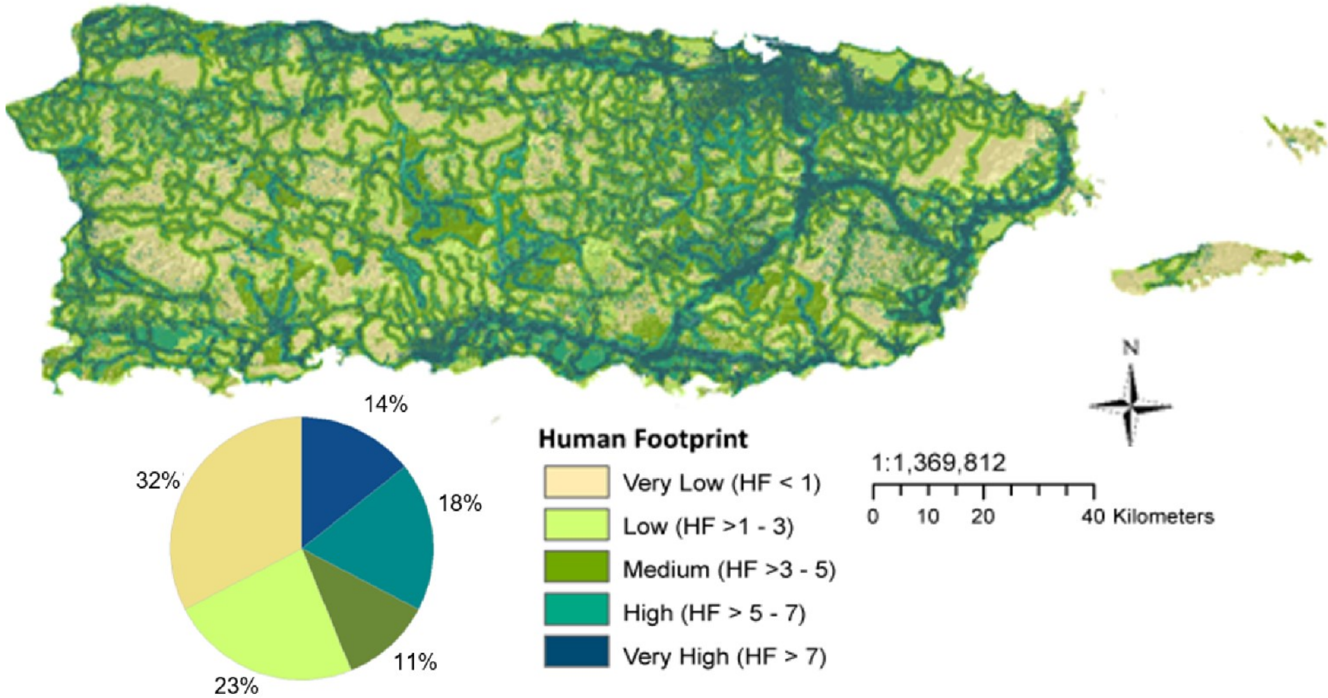


Figure 2. The human footprint in Puerto Rico. The highest human footprint scores (blue colors) are near main roads, coastal areas, and densely populated municipalities. The areas in beige represent the lowest values of human footprint. High and very high categories covered 32% of the island, while the two lowest categories covered 55%. The totals do not add to 100% due to rounding.

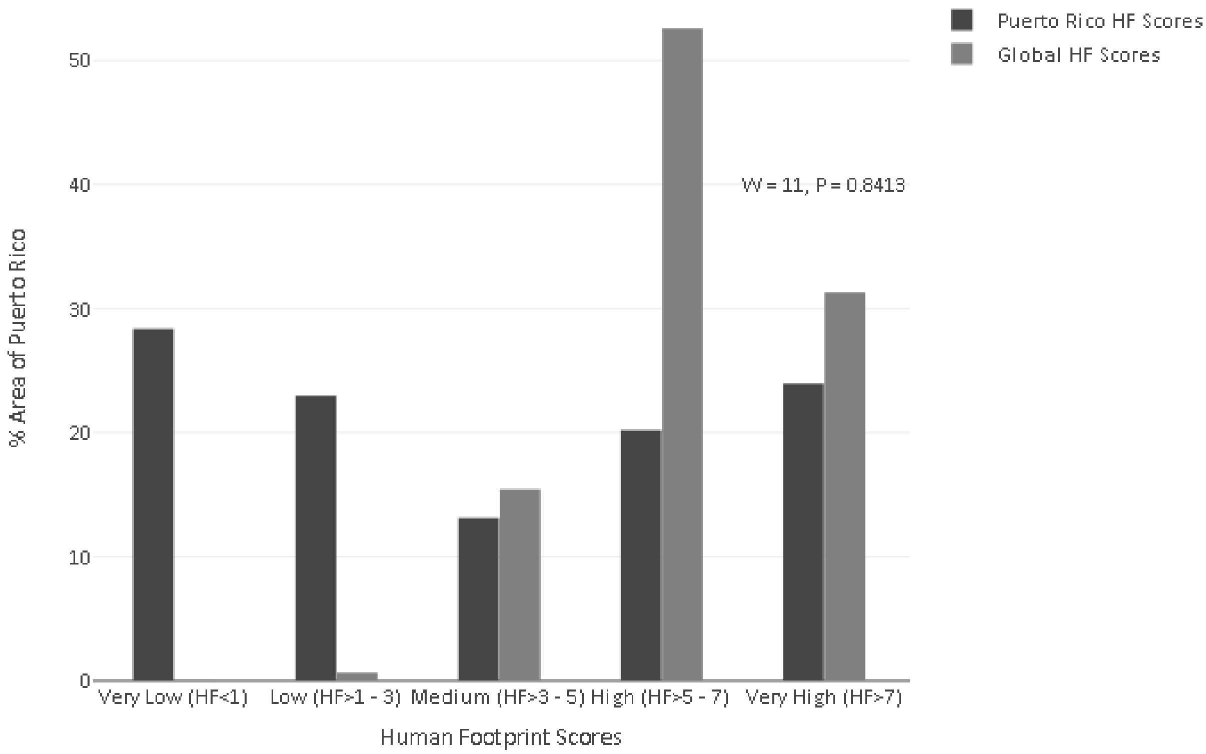


Figure 3. Human footprint scores based on our model (black columns) compared to the global HF model (gray columns).

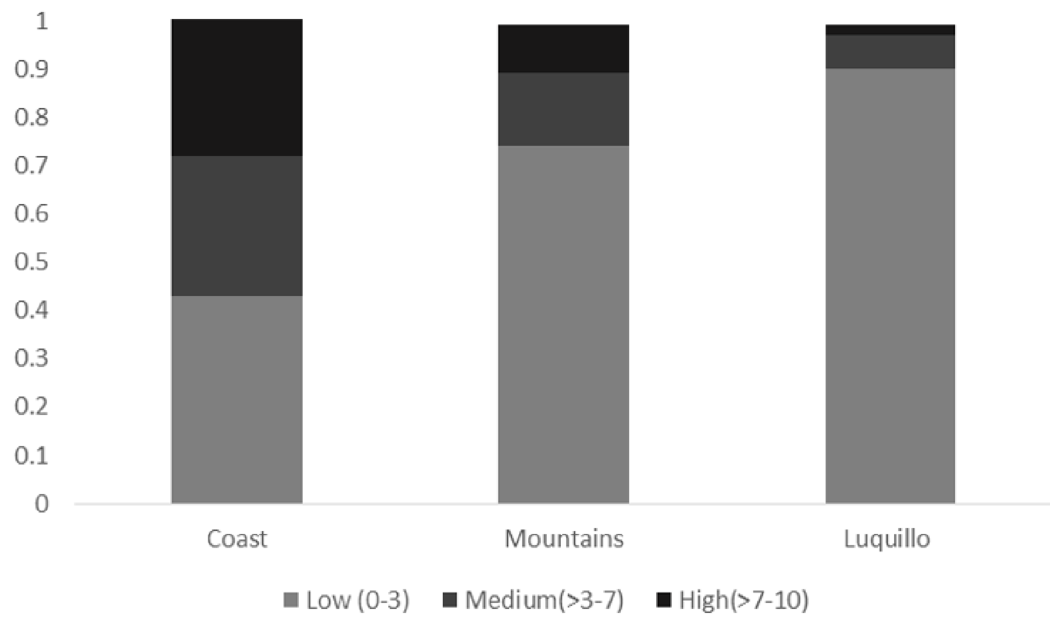


Figure 4. Proportion of land with low, medium, and high human impact according to our HF model, within the three broad physiographic regions in Puerto Rico: Coast, Central Mountains, and Luquillo mountains.

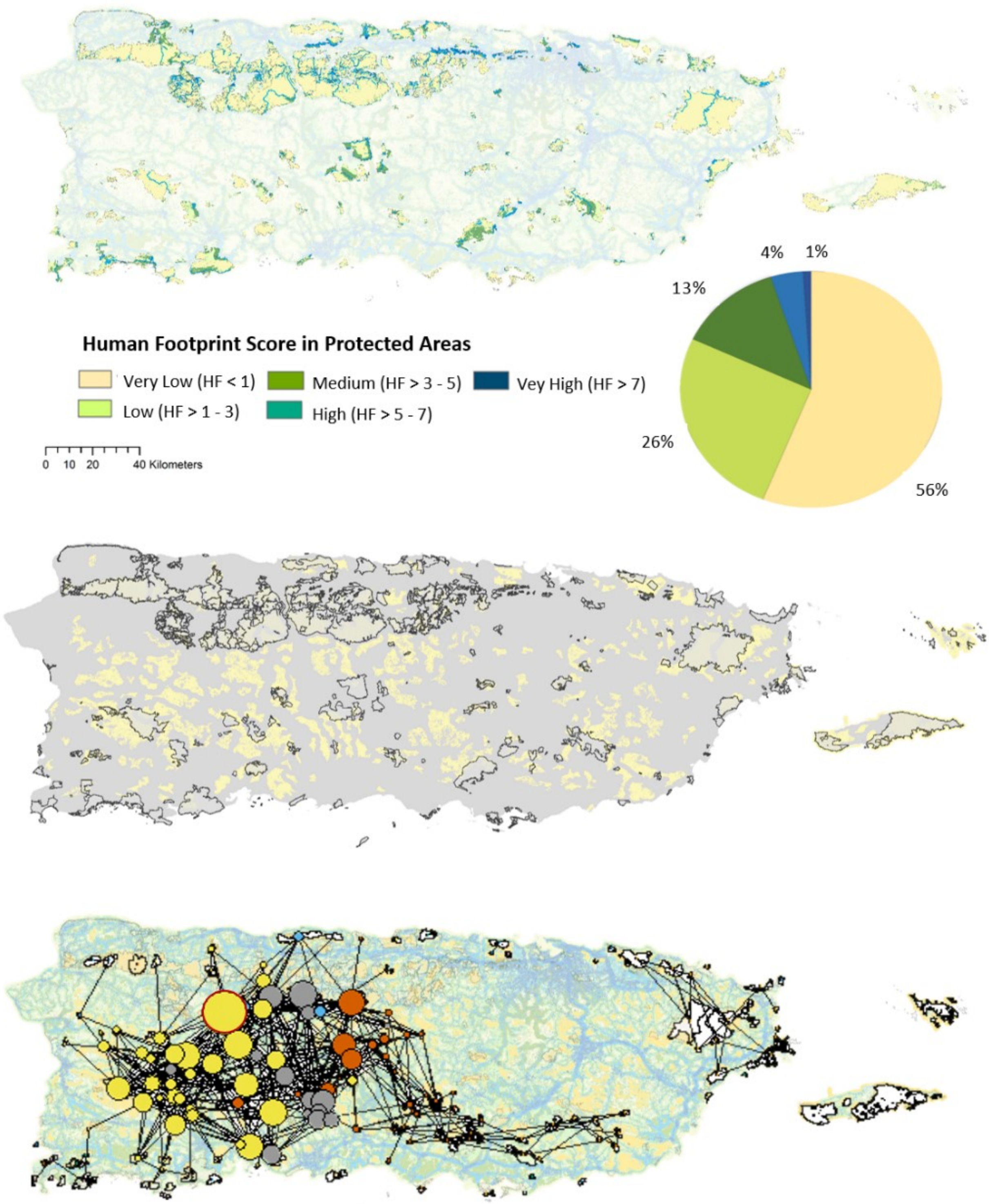
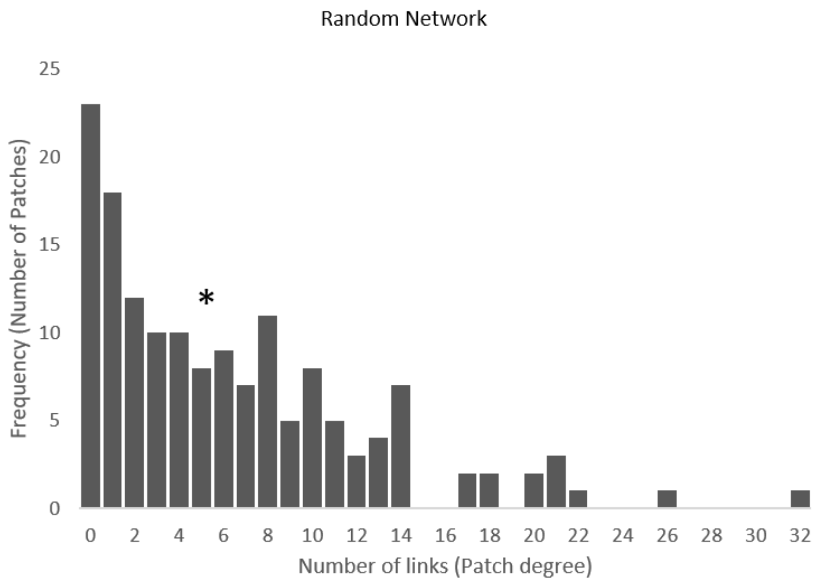


Figure 5. Protected areas in Puerto Rico and their Human Footprint Scores (Top). Low and Very Low human modified areas accounted for 82% (1, 129.96 km^2 out of a total of 1, 378km^2). Medium to Very High modified areas covered 18% of protected areas (248 km^2). Patches with a very low score and > 25 ha (0.25 km^2) outside of protected areas (center) are scattered though the central mountainous regions and almost non-existent in coastal areas. The Puerto Rico

ecological areas network (bottom). Patch size represents number of connections and patch color represents the components to which each patch belongs. The patch with the most connections, the hub, was identified as part of the Karst region in the island (delineated in red).



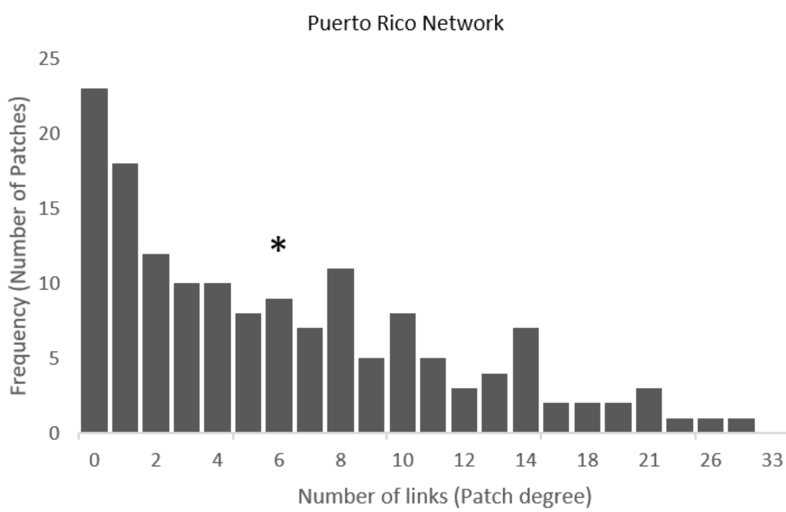
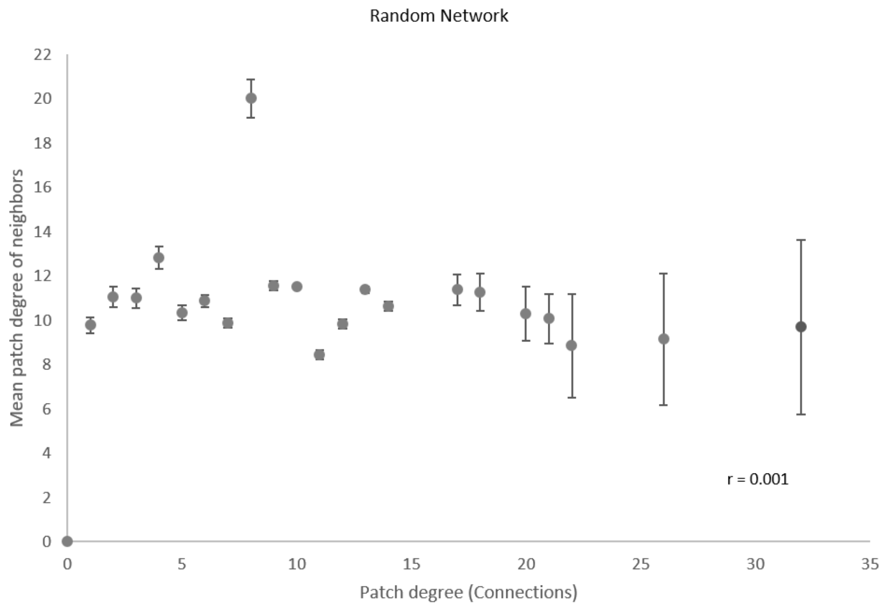


Figure 6. Patch-degree distribution for Puerto Rico's ecological network. In this graph, there are far more patches with few a links. The distribution of patch degrees is heavily skewed to the right for both the random network (1.49) (Top), and for Puerto Rico's network (1.07) (bottom). The asterisk indicates mean patch degree.



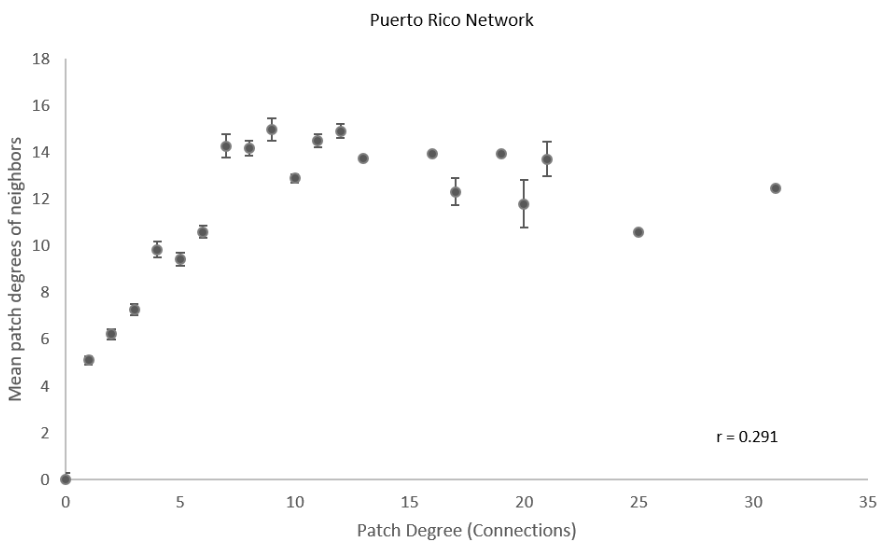


Figure 7. Connectivity correlation for random network (top), and the Puerto Rico network (bottom). PR network had a positive relationship between mean patch degrees of neighbors and number of connections of a patch. In PR network, the higher the patch degree, the more neighbors connected to each patch. The pattern suggests that patches with a minimum of 7 or 8 connections to other patches (given our parameters) are enough to reach neighbors with the mean maximum possible of connections in the system

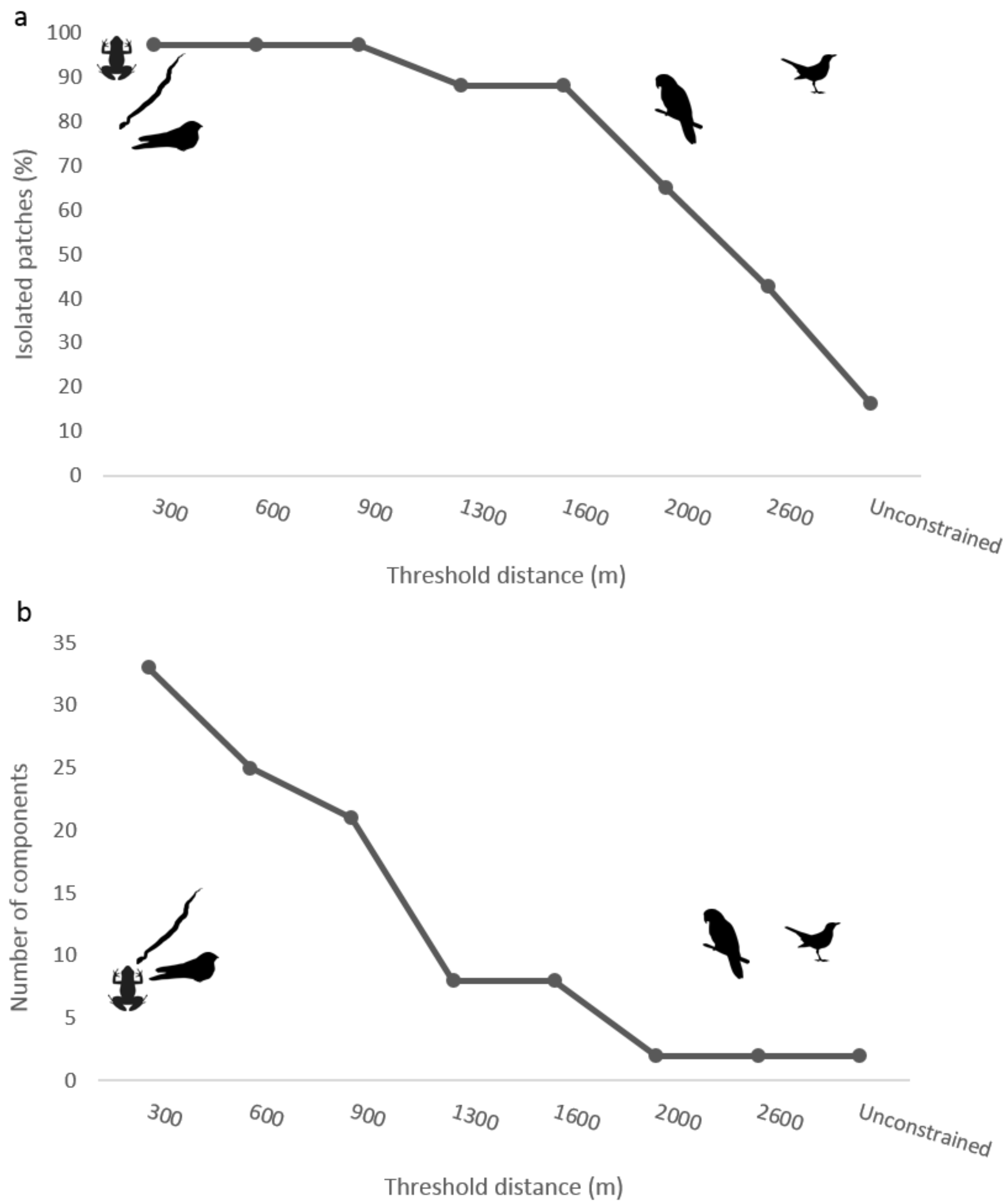


Figure 8. The (a) percentage of isolated patches and (b) number of components found in the PR network according to several threshold distances. Number of components included isolated patches. Five species listed in the PR-GAP analysis were used to classify threshold distances (see table 4).

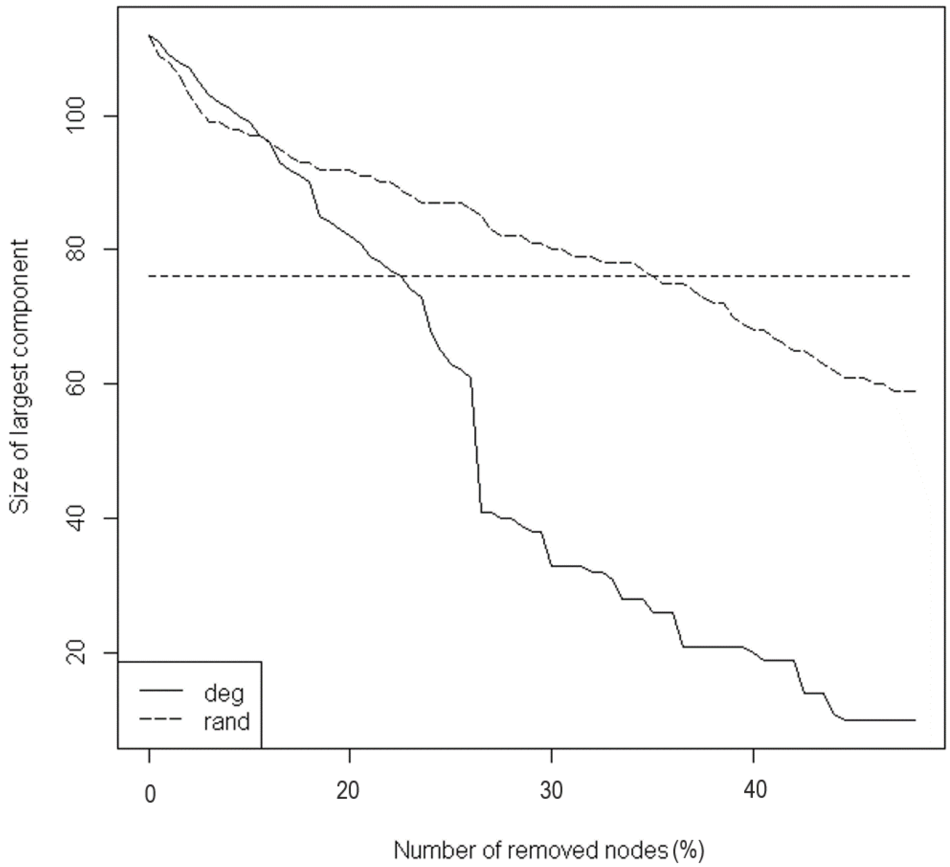


Figure 9. Effect of patch removal on the connectivity of the Puerto Rico ecological network. The relative size of the largest component decreases as an increasing number of patches are removed. Patches were removed according to three different percolation strategies: removal of patches with the highest number of links first (deg), random removal (rand), and by area (largest to smallest). The horizontal dashed line represents half of the total of patches in the network (n = 76).

Chapter 2. Field validation of habitat suitability models for the small Indian mongoose: Implications for the use of species distribution models for guiding field surveys, assessing uncertainty and usefulness.

Target Journal: Biological Invasions

Abstract

Accurate knowledge of invasive species occurrence and distribution is crucial for decision making and assessments of risk. Species distribution models (SDMs) are commonly used to predict environmental suitability; however, few studies can provide independent validation for SDM outputs. I had two goals. First, we aimed to develop a habitat suitability model to predict the presence of an invasive mammal [small Indian mongoose: (*Herpestes auropunctatus*)] in the island of Puerto Rico. Second, I aimed to use field surveys to validate predicted distributions of mongoose in Puerto Rico from four alternative modelling approaches: Ecological Niche Factor Analysis, MaxEnt, Generalized Additive Model, and Boosted Regression Trees. The likelihood of habitat suitable for mongoose was predicted to be very high in coastal regions, mongoose presence had a positive association with proximity to wetlands in all models. Evaluation metrics suggested that models MaxEnt, GAM and BRT performed well (AUC = 0.84-0.89), but not necessarily agreed between predicted probabilities of occurrence and observed presence in the field. A Kruskal-Wallis rank test indicated that the estimated probability of occurrence differed significantly among the models, resulting in different maps of potential species distribution. Overall, we found that MaxEnt can be useful for preliminary assessments of species distributions and field survey design, however, boosted regression trees might be the best choice after further iterations with independently gathered presence data. The results emphasize the importance of re-evaluating and iteratively improving SDMs with independent data vs cross-validation, particularly if initial models are trained with data from a limited geographic extent. This study

highlights both the importance of field validation of SDM models and to reflect on the purpose of the model for end users.

Keywords: Species Distribution Model, invasive species, model validation, model performance, Potential distribution models

Introduction

Introduced species have the potential to become harmful to local biodiversity, human health, and the economy (Pimentel et al., 2005), making it important to predict the potential distribution of such species (Elith et al., 2006). Fundamentally, the invasive potential of an introduced species depends on its survival in each of four main stages of the invasion pathway: transport, introduction, establishment, and spread (Blackburn et al., 2011), and management efforts can target any of these steps. However, the first two, transport and establishment, are the most elusive stages for population management efforts due to the diversity of transport vectors and factors that influence the survival of introduced species in a new environment (Lockwood et al., 2013). Therefore, the most common management approach is slowing the spread via on-the-ground control or eradication programs (Le Maitre et al., 2008). Species distribution models can assist such efforts greatly if they predict where a species is likely to occur based on a suite of environmental conditions. However, validating model predictions in the field and deciding whether the model is useful for its intended application (e.g. establishing containment boundaries, early detection, or rapid response measures) are less common (M. B. Araújo & Guisan, 2006; Antoine Guisan & Thuiller, 2005b).

Environmental suitability and species dispersal abilities are two major factors that determine the colonization of new geographical areas (Lockwood et al., 2013). Environmental suitability defines the fundamental niche of a species, i.e., the set of environmental conditions under which populations can survive and reproduce (Soberón, 2007), which is often much larger than the realized niche, i.e., where a species actually occurs because of species interactions, and areas that are yet-to-be colonized. Species distribution models predict those areas that are hypothetically available to the species across a study area ("potential niche") by establishing a statistical relationship between occurrence locations and a set of climatic and landscape features. The estimated relationships can then be used to predict a species' distribution in areas where an introduced species has not yet colonized which is valuable for monitoring programs, risk mapping and decision-making (El-Gabbas & Dormann, 2018; Guillera-Arroita et al., 2015; Manzoor et al., 2018). However, the challenge of predicting invasions outside their actual distribution is that invasive species can spread into parts of the fundamental niche that are not available in their native range (Early & Sax, 2014; Le Maitre et al., 2008). Therefore, species distribution models for introduced species require frequent updating with new occurrence records (Cook et al., 2019; Uden et al., 2015). Predictions for initial models can guide targeted surveys though, that then increase model reliability (Crall et al., 2013; Lauzeral et al., 2015; West et al., 2016a). Unfortunately, while many thousands of species distribution models have been developed, few have been validated with an independently collected dataset (West et al., 2016b; Westwood et al., 2019).

The challenge for predicting distributions into areas that are yet-to-be colonized is that species distribution models are usually validated with data from the same region in which they were fitted, e.g., via cross-validation, but that does not quantify the usefulness or applicability in the areas that will be colonized (Randin et al., 2006). Only a geographically independent data set can truly validate the accuracy of predictions in space (A Guisan & Zimmermann, 2000; Yates et al., 2018).

The usefulness in predicting future colonization depends on both the accuracy of the prediction and on the specific conservation goals (M. B. Araújo & Townsend Peterson, 2012). For example, when predicting non-invasive species scenarios with stable distributions, a model that predicts presence where the species is absent is over-prediction (i.e. commission error). However, when predicting invasive species with the aim to discover new populations, a model that predicts presence in a location where a species had not been recorded so far is desirable (M. B. Araújo & Townsend Peterson, 2012) because those ‘over-predictions’ may represent the areas of potential spread that are suitable but not yet established (Antoine Guisan & Thuiller, 2005a). Similarly, at the later stages of the invasion, when the management goal is to eradicate new populations before they can fully establish, the modeling goal is also to predict all potential presences, and minimize errors of omission (Ward, 2007). A false absence may result in a failure to establish appropriate control measures (Baxter & Possingham, 2011). Thus, mismatches between prediction and observation do not necessarily mean that the model is wrong and may be the result of specific management and modeling goals, but a model that vastly over-predicts presences is rarely useful either. Deciding how to balance errors of commission and omission will depend on the objective of the study.

The goal of our study was to model the distribution of the small Indian mongoose (*Herpestes auropunctatus*) in Puerto Rico where it is an established invasive species, and to test different types of species distribution models in their ability to predict out-of-range occurrences with independent validation data. The specific objectives of this study were to: 1.) model landscape-level habitat associations of the mongoose in Puerto Rico; 2.) compare the usefulness of 4 modeling frameworks to predict out-of-range mongoose occurrence; 3.) use the species distribution model that maximized the true positive rate of probability values (i.e. maximized sensitivity) to stratify a random sampling scheme to collect an independent mongoose presence dataset; 4.) re-calibrate the species distribution model with the new dataset and compare expected versus observed presence points to assess each model's usefulness in guiding survey efforts in new locations for invasive mongoose monitoring. I chose this species because of its threat to island biodiversity and to human health, because mongooses are a disease vector of the rabies virus and the bacterial disease leptospirosis (Benavidez et al., 2019a; Berentsen et al., 2018; D. B. Hoagland et al., 1989)

Methods

Study area

Puerto Rico is situated in the northeastern Caribbean (18.2208° N, 66.5901° W). The island encompasses 8,900 km^2 of land and has a tropical maritime climate. Temperatures are warm year-round (average 25 ºC) (Daly et al., 2003). There is one distinct wet (June through November) and one dry (December through May) period (Waide et al., 2013), and rainfall averages 1, 687 mm/year. The highest elevation point is 1,338 m (Garcia-Martino et al., 1996). There are six life zones in Puerto Rico (Miller & Lugo, 2009), ranging from subtropical-dry to low-montane rain forest. Climate shifts with elevation, and lower temperatures, greater precipitation, cloud cover, and humidity occur at higher elevations (Waide et al., 2013). In terms of landcover, forests cover 53 % of the main island, and developed areas 11 %, the latter mostly occurring as linear developments along highways and roads with urban centers near coastal plains and valleys distributed throughout the island. The remaining 35 % is composed of pastureland, grasslands and agriculture, and natural barrens (Gould et al., 2008).

Study species

The small Indian mongoose is a solitary opportunistic predator native to the Middle East, South and Southeast Asia (Nellis & Everard, 1983). It was originally introduced to the Caribbean region between 1872 and 1900 (Coblentz & Coblentz, 1985). Mongoose are on the top one-hundred most harmful invasive species because of their impacts on faunal communities (Lowe et al., 2000). The mongoose is a mostly diurnal predator with a high reproductive output (i.e. litters of 2-4 pups, two to three times per year) (Nellis & Everard, 1983). In Puerto Rico, home ranges vary widely from 3.1 to 52.2 ha, (D. Hoagland & Kilpatrick, 1999; J. H. Quinn & Whisson, 2005). Mongoose density varies among Caribbean islands from 0.19 to 9.0 mongooses/ha with overlapping home ranges (Johnson et al., 2015b; Pitt et al., 2015; J. Quinn et al., 2006; Roy Horst et al., 2001). The average daily travel distance of mongoose ranges from 300 to 500 meters (Pimentel, 1955). Mongoose feed on a wide variety of food items, but mostly small vertebrates and invertebrates (71-89%), and plant material (15-29%) (Pearson & Baldwin, 1953; Pimentel, 1955; Vilella & Zwank, 1993), with items differing in importance depending on habitat (Gorman, 1975; Powers et al., 2020). The density and diet of mongooses varies according to habitat characteristics, and they are most numerous in dry, hotter, and lower elevation areas (Guzman-Colon & Roloff, 2014; Johnson et al., 2015b). Population management strategies include localized trapping (Berentsen et al., 2018).

Environmental and land cover data

I obtained 19 WORLDCLIM climate variables at ~1 km resolution (0.0083 degrees; 30-arc seconds) that represent 30-year climate averages for Puerto Rico. I tested variables for collinearity, and retained candidate temperature and precipitation variables that were not strongly correlated (Pearson correlation coefficient $|r| < 0.7$ or $|r| > -0.7$) and most likely to impose resource-based limits (e.g. seasonality) (Merow et al., 2013), i.e., annual mean temperature, annual mean precipitation, temperature seasonality, and precipitation seasonality (Table 8). Our land cover data were obtained from the United States Geological Services – GAP analysis project (2011) as a raster layer of seventy-one classes at 15 m resolution. I reclassified the raster into eight land cover classes: evergreen forests, freshwater, grasslands, plantations, semi-deciduous forests, shrublands, urban, and wetlands. I excluded shoreline and ocean water. For each land cover type, I calculated the Euclidian distance from each mongoose presence location to the nearest pixel of, e.g., forest (Table 8). The variables 'annual mean precipitation' and 'evergreen forest' were highly correlated with 'elevation' but not with each other. I did not include elevation.

Field methods

Preliminary presence data

To develop our initial presence-background model in Maxent, I used mongoose occurrence data from a capture-recapture study completed at El Yunque National Forest and Northeast Ecological Corridor in 2011 (number of unique captures = 36) (Figure 10) (Guzman-Colon et al., 2019). These presence points were distributed over an elevational range of 0 – 955 m.

Guided survey

I designed a stratified survey to test both the ability of the models to predict presence of mongoose and to compare the performances of the four model types. The stratified survey was based on the relative probability output obtained from the initial MaxEnt model based on 2011 trapping. Our aim was to capture the full range of habitat suitability across the island, and so I stratified our sampling into three classes of habitat suitability index values as determined by the 2011 data (HSI: 0 - 0.3; > 0.3 - 0.7; > 0.7 - 1.0). In each of the three HSI categories, I located nine random sites in the mountainous region (n = 27) and nine random sites within the coastal areas (n = 27) for a total of 54 sites. Within each site I created one 1-km^2 grid. Within each grid at each site, I randomly placed eight traps within each grid. I generated the eight random trap points within each grid using the Sample Design Toolbox for Arc GIS (Figure 11).

I set and monitored the traps from May to September of 2017 because mongoose disperse during the summer months (Hayes, 1989). Three grids were monitored each day and each grid was monitored for a period of two consecutive days.

Species distribution modeling frameworks

I compared 4 species distribution modeling types. For each, I first trained and tested the model within El Yunque National Forest watershed with data from the 2011 capture-recapture study described in 2.4.1, and second, extrapolated our predictions to the rest of the island. Third, I assessed how well each model predicted mongoose presence according to our empirical mongoose trap data.

I evaluated our initial models via 10-fold cross-validation by partitioning our original dataset into a training set to train the model and testing set to evaluate it. I used the Receiver Operating Characteristic-Area Under the Curve (AUC) metric as a measure of model discrimination (i.e. ability to distinguish between occupied and unoccupied sites). I also evaluated sensitivity, specificity and True Skill Statistic (TSS).

Maxent

I used package "maxnet" (Phillips et al., 2017) in R (Rstudio Team, 2020) to build our initial species distribution model with presence data from 2011. MaxEnt uses the environmental (covariate) data from the presence records and the background (pseudo absence points) sample to estimate the ratio of the conditional density of the covariates at the presence sites over the density of covariates across the study area. I chose 10,000 background points within a bias grid that delineated the extent of the eastern watershed using the function sampleRast() in the enmSdm R package (Smith 2019).

Environmental Niche Factor Analysis

Environmental niche factor analysis is based on marginality and tolerance. Marginality assesses to what extent a species' mean environmental space differs from the background mean environmental space across the study area. Tolerance assesses to what extent a species' variance in environmental space differs from the background environmental variance.

I used package adehabitatHS in R (Calenge, 2006) for our environmental niche factor analysis. The predictor variables were transformed into a set of three uncorrelated factors. The first factor maximized marginality by identifying which environmental condition at a presence site differed most from the conditions in the background area. The second factor maximized tolerance; it maximized the ratio of the variance of the background points to that of mongoose distribution. The process resulted in a set of two eigenvectors that can be interpreted as the environmental suitability of mongooses. These suitability values, however, are represented as distance from the cluster centroid in the environmental variable space. Thus, I transformed ENFA values into presence-background data using the function roc() from package PRoc.

Generalized Additive Model (GAM)

The third SDM was a general additive regression model (GAM). This model allow non-linear relationships between predictors and presence/absence data. I fitted the model using gam() function in R by specifying a binomial distribution and a logistic link. The best model explained the greatest amount of deviance reduction with terms falling below $p < 0.05$. I selected significant predictors using step-wise backward selection of variables and used the Akaike information criterion (AIC) to select the best and final model given the lowest value for AIC.

Boosted Regression Trees

I implemented a boosted regression tree analysis with the function gbm.step from Elith et.al. (2008) and followed their cross-validation implementation to identify an optimal number of trees that maximize the ability of the model to make accurate predictions while avoiding model complexity. I applied a learning rate (i.e. determines the contribution of each tree to the growing model) of 0.01, tree complexity of 1 (fits an additive model) and bag fraction of 0.5 per model. The optimal number of trees identified by the cross-validation implementation was 2000.

Ensemble and uncertainty

I used all final model predictions to create an ensemble and uncertainty model using package 'esdm' (Woodman et al., 2019). The ensemble model was an unweighted average of the predictions of the 4 individual models, while uncertainty was calculated using among-model variance.

Model comparison of relative probabilities generated by various models

I validate the accuracy of the species distribution models with new field data from the entire island. To do so, I classified each model's prediction accuracy using an error matrix by comparing binary categorical output maps of prediction (< 0.5 = absence and > 0.51 = presence with occurrence data collected in the field. I calculated several threshold-independent (i.e. AUC, continuous Boyce index, Presence-Only Calibration plot) and threshold-dependent (i.e., sensitivity, specificity, and true skill statistic) measures of model accuracy to evaluate model performances. The continuous Boyce index is the Spearman rank correlation coefficient between the proportion of sites in each prediction class and the expected proportion of predictions in each prediction class based on the proportion of the landscape that is in that class (Phillips & Elith, 2010a). AUC measures model discrimination ability, and the continuous Boyce index and presence-only calibration plots measured how well predicted values match probabilities at presence and background locations (Fieberg et al., 2018; Phillips & Elith, 2010b). I tested the differences among models in their ability to distinguish between suitable and non-suitable areas with a Kruskal-Wallis test in R (R Development Core Team, 2008; Randin et al., 2006).

Results

MaxEnt model and comparison of predicted habitat suitability to survey data

The MaxEnt model, first trained, tested, and validated with data from a 2011 field survey within the eastern watershed environment, had good evaluation metric scores when applied to the island-wide empirical data acquired at 54 trapping sites. The model predicted the conditional probability of presence across all validation runs well (AUC = 0.89), distinguished occupied vs non-occupied sites (CBI = 0.79), with a true-positive rate (sensitivity) of 0.88, true-negative rate (specificity) of 0.63, and a True Skill Statistic of 0.45. Of the 54 sites surveyed, 18 had mongoose present (33%). The frequency of mongoose presence was higher in coastal areas (n = 21) than in the mountain regions (n = 10).

Relative habitat suitability values extracted at each observed mongoose presence sites were not consistent among the four models (Figure 12). Almost 80% of observed presence points were classified in the lowest suitability class by environmental niche factor analysis. On the other hand, MaxEnt classified 73% of observed presences as highly suitable. Most of the observed presences were classified in mid- to high-suitability classes by BRT and GAM, but BRT assigned 42% of observed presence points to the lowest two suitability classes, while GAM assigned only 10 % to the lowest two classes.

Comparison of model performance based on field validation

After 10-fold cross-validation I assessed several metrics evaluating the performance of the species distribution models and found inconsistent results. For example, MaxEnt outperformed all models based on mean AUC (0.89) but was ranked as lowest according to Continuous Boyce Index (0.47) (Table 2). The boosted regression tree model, although also performing well based on AUC (0.84), was not ranked well by the Continuous Boyce Index, which indicated wide variation in mongoose presence around the mean value (0.56). According to all threshold independent metrics, GAM had relatively good performance (AUC = 0.86, CBI = 0.73), while the environmental niche factor analysis model had poor performance overall (AUC = 0.75, CBI = 0.61). The environmental niche factor analysis habitat suitability output showed the lowest prediction accuracy according to all metrics. As a result, many sites that were identified by the model as non-optimal had mongoose present in the field survey. Sensitivity scores (% of correctly classified presences at a threshold of > 0.55) were high for BRT (0.97) and MaxEnt (0.88).

The Kruskal Wallis rank test for comparison of AUC and Continuous Boyce Index scores for all models confirmed they were significantly different ($H = 66.7$ $p < 0.001$). The ad-hoc Tukey test indicated that mean environmental niche factor analysis AUC was statistically different that the rest of the models, but the GAM, MaxEnt and boosted regression tree models did not differ ($p > 0.5$).

According to the presence-only calibration plots, the relative probability of mongoose occurrence tended to increase with the probability of presence predicted by the initial models. However, all models seem to be poorly calibrated at some prediction values. In general, the models (except BRT) did not show a close match between the relative probability of presence generated after projecting all four SDM models outside the training area and the predicted probability of presence calculated after model calibration with survey data (Figure 13). MaxEnt and ENFA tended to overestimate prevalence in many of the presence categories. ENFA did not predict presences beyond $p = 0.4$, and underestimated species presence in this same category, whereas MaxEnt underestimated prevalence when $p > 0.8$. The recalibrated GAM model underestimated prevalence when $p > 0.2$. Although BRT closely matched the diagonal line in the POC plots, it severely overestimated values when $p > 0.8$.

Final mongoose suitability models and important predictors

Distance from wetlands consistently had a high relative contribution in all four models. According to the response plots of MaxEnt and BRT, marginality values on ENFA and coefficient on GAM, habitat suitability for mongoose decreased with distance from wetlands. The variable with the highest mean percent contribution in the MaxEnt model was annual mean temperature (BIO1) (67%), followed by proximity to wetlands (37%) and precipitation seasonality (31%). The latter variable was also important in the BRT model (25%).

Predicted suitability outputs from all models varied in their spatial patterns. However, habitat suitability predictions were similar in coastal areas and for transition zones into the central mountains (Figure 14e). The average predictions across all four models matched our expectations based on observations of mongoose affinity to coastal and drier environments in the Caribbean. Spatial differences, however, were apparent among models: the environmental niche factor analysis predicted only low suitability areas across the entire island (0.20 – 0.37) whereas GAM predicted medium to high suitability across the extent of the island (0.20 – 1) (Figure

14c;d). The MaxEnt model predicted a wide range of suitability values with highest values (0.60 – 1) identified in coastal areas (Figure 14b). The boosted regression tree projection of habitat suitability was much more restricted, predicting a few areas of suitable habitat in the coastal area (Figure 14a). Although most models assigned high suitability values in the coastal area, uncertainty of among model predictions was also highest on coastal areas of the north, west, a few areas on the southwest, and within the eastern watersheds training area, suggesting that the predictions in those areas must be viewed with caution (Figure 14). Evaluation metrics for the ensemble models were not substantially better than those of the individual models (AUC = 0.74, CBI = 0.36, Specificity = 0.74).

Discussion

I characterized important habitat variables associated with mongoose presence and assessed the utility of various species distribution model approaches in predicting mongoose presences. The distribution of habitat suitability and likelihood of mongoose presence differed vastly in space. Although the models predicted different patterns of habitat distributions, their selection of the most important environmental variables to predict mongoose presence were quite similar. Proximity to wetlands strongly predicted mongoose presence, however, its relative importance varied among the four species distribution models. Strong predictors of mongoose occurrence also indicated that higher temperatures and areas with low precipitation are also important. The close association of mongoose and wetlands is an interesting relationship since it has been observed that mongoose avoid water and rarely swim (Nellis & Everard, 1983), the relationship of mongoose presence near wetlands in this study might likely be due to wetlands being commonly found in the coastal areas where mongoose are abundant. However, tropical wetlands are also known to be productive areas with a diversity of food sources (e.g. terrestrial insects and fruits) (Wantzen & Junk, 2000), but the relationship between prey availability in wetlands and mongoose presence is yet to be elucidated.

I discovered that when these models were validated with an independent dataset, MaxEnt, GAM and environmental niche factor analysis had underestimated the predicted probability of presence in the high-suitability classes, but overestimated presence at lower suitability classes. Discrimination metrics, such as AUC, have several drawbacks including that AUC weights commission and omission errors equally, and that AUC scores are positively related to spatial extent (Jiménez-Valverde, 2012; Lobo et al., 2010). This latter characteristic is highly problematic when predicting invasive species since the geographic range of interest for the predictions is usually much broader than the range for which training data is available. Thus, it is important to also include calibration metrics (i.e. Continuous Boyce Index and Presence-Only plots) for model evaluation, especially when new data can highlight model shortcomings such as a missing covariates or model overfitting. In this study, I found that models with good discrimination metrics can have a low calibration capacity.

Overall, GAM performed well according to AUC and CBI, but its measure of sensitivity vs specificity and output map indicate that it has the potential for fitted responses that extrapolate unrealistically. MaxEnt can overpredict when modelling in environmental space outside of the range included during model training (Pokharel et al., 2016), and this can pose a problem when the goal is to detect potential colonization areas. I found that when MaxEnt was validated with external data, it tended to overestimate the likelihood of mongoose presence, but for higher predicted suitability values (> 0.8) it underestimated observed presence. This suggests that MaxEnt can be useful for preliminary assessments of habitat suitability and for guiding field surveys since it was able to identify potential species occurrence. Nevertheless, MaxEnt may be limited in its ability to discern the likelihood of actual mongoose presence. The environmental niche factor analysis model did not perform well in our study, and tends to perform poorly when extrapolating outside the training region because the model per se characterizes the niche according to the geographic limit of the training area (Hirzel et al., 2001). I found that boosted regression trees, although not the top performer, displayed reasonable refinement after re-calibration with external data as demonstrated by the presence-only calibration plot and continuous Boyce index values. Boosted regression tree models may be the best choice if the management goal is to reduce well-established populations because the model assigned high suitability values to areas where mongoose were actually found. Finally, the assumption for ensemble models is that if each method performs well on its own, they should do even better together (Ali et al., 2015). However, our ensemble model had lower discrimination and calibration performance as shown by AUC and CBI scores, than any of the four models individually. Still, the ensemble model provided valuable information about the uncertainty in predictions among models. In our results, coastal areas had the most agreement, but within those coastal areas, some patches in the north, west south west and around the El Yunque National Forest in the east had disagreements in predictions. Ultimately, biological insight and expert knowledge is valuable when choosing a SDM model type for invasive species, and that choice also depends on the management goal.

Predicted mongoose presence on coastal areas in the southwest (near Cabo Rojo National Wildlife Refuge and other natural reserves; 17.9752° N, 67.1686° W), north (Tortuguero; 18.461098° N, -66.422660° W) and south central region (Jobos Estuary; 17.9462° N, 66.1921° W) coincide with highest values of predicted terrestrial vertebrate resident, endemic, and endangered species richness (Gould et. al. 2008), and therefore are areas of primary conservation concern. Efforts to understand mongoose ecology, manage populations, or further examine their role in the ecosystem can be carried out with confidence in areas of consensus that have low variance in predictions.

Conclusion

Many regions of the world experience threat to biodiversity by species invasions. Predictions of invasive species distributions are paramount for decision making but require models of their relationship to the environment. Species distribution models are used frequently for making predictions of occurrence, but predictions are compromised by scant information about species' habitat requirements, limited availability of presence records, poor transferability of models in space and time, and by models that do not serve the purpose of management. I highlight various observations based on this study: 1.) Maps that maximize specificity designed to correctly identify potential suitable habitat in guided surveys are still highly prone to misclassification by certain model frameworks (e.g. MaxEnt), 2.) Multiple evaluation metrics are needed to assess prediction accuracy, especially those metrics that evaluate model calibration (e.g., Presence-only calibration plots), 3.) These results emphasize the importance of an adaptive approach to re-evaluate and improve SDMs at subsequent modeling iterations, particularly after training initial models with limited data and geographic extent (Conroy et al., 2011; Howell et al., 2009). The results also highlight that while a model might not be optimal, the utility of the SDM is contingent upon the specific management goal for invasive species management and monitoring.

References

Ali S, Tirumala SS, Sarrafzadeh A (2015) Ensemble learning methods for decision making: Status and future prospects. In: Proceedings - International Conference on Machine Learning and Cybernetics. IEEE Computer Society, pp 211–216

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. J Biogeogr 33:1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Araújo MB, Townsend Peterson A (2012) Uses and misuses of bioclimatic envelope modeling. Ecology 93:1527–1539. https://doi.org/10.1890/11-1930.1

Baxter PWJ, Possingham HP (2011) Optimizing search strategies for invasive pests: learn before you leap. J Appl Ecol 48:86–95. https://doi.org/10.1111/j.1365-2664.2010.01893.x

Benavidez KM, Guerra T, Torres M, et al (2019) The prevalence of Leptospira among invasive small mammals on Puerto Rican cattle farms. PLoS Negl Trop Dis 13:e0007236. https://doi.org/10.1371/journal.pntd.0007236

Berentsen ARR, Pitt WC, Sugihara RT (2018) Ecology of the Small Indian Mongoose (Herpestes auropunctatus) in North America

Blackburn TM, Pyšek P, Bacher S, et al (2011) A proposed unified framework for biological invasions. Trends Ecol. Evol. 26:333–339

Calenge C (2006) The package “adehabitat” for the R software: A tool for the analysis of space and habitat use by animals. Ecol Modell 197:516–519. https://doi.org/10.1016/j.ecolmodel.2006.03.017

Coblentz BE, Coblentz BA (1985) Control of the Indian Mongoose Herpestes auropunctatus on St John, US Virgin Islands. Biol Conserv 33:281–288. https://doi.org/10.1016/0006-3207(85)90018-7

Conroy MJ, Runge MC, Nichols JD, et al (2011) Conservation in the face of climate change: The roles of alternative models, monitoring, and adaptation in confronting and reducing uncertainty. Biol Conserv 144:1204–1213. https://doi.org/10.1016/j.biocon.2010.10.019

Cook G, Jarnevich C, Warden M, et al (2019) Iterative Models for Early Detection of Invasive

Species across Spread Pathways. Forests 10:108. https://doi.org/10.3390/f10020108

Crall AW, Jarnevich CS, Panke B, et al (2013) Using habitat suitability models to target invasive plant species surveys. Ecol Appl 23:60–72. https://doi.org/10.1890/12-0465.1

Daly C, Helmer EH, Quiñones M (2003) Mapping the climate of Puerto Rico, Vieques and Culebra. Int J Climatol 23:1359–1381. https://doi.org/10.1002/joc.937

Early R, Sax DF (2014) Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. Glob Ecol Biogeogr 23:1356–1365. https://doi.org/10.1111/geb.12208

El-Gabbas A, Dormann CF (2018) Wrong, but useful: regional species distribution models may not be improved by range-wide data under biased sampling. 2196 | Ecol Evol 8:2196–2206. https://doi.org/10.1002/ece3.3834

Elith J, H Graham C, P Anderson R, et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography (Cop) 29:129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Fieberg JR, Forester JD, Street GM, et al (2018) Used-habitat calibration plots: a new procedure for validating species distribution, resource selection, and step-selection models. Ecography (Cop) 41:737–752. https://doi.org/10.1111/ecog.03123

Garcia-Martino AR, Warner GS, Scatena FN, Civco DL (1996) Rainfall, Runoff and Elevation Relationships in the Luquillo Mountains of Puerto Rico

Gorman M (1975) The diet of feral Herpestes auropunctatus (Carnivora: Viverridae) in the Fijian Islands. J Zool 175:273–278

Gould W, Alarcón C, Fevold B, et al (2008) The Puerto Rico Gap Analysis Project. Volume 1: Land cover, vertebrate species distributions, and land stewardship. Gen Tech Rep IITF-

GTR-39 165 p.

Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, et al (2015) Is my species distribution model fit for purpose? Matching data and models to applications. Glob Ecol Biogeogr 24:276–292. https://doi.org/10.1111/geb.12268

Guisan A, Thuiller W (2005a) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guisan A, Thuiller W (2005b) Predicting species distribution: Offering more than simple habitat models. Ecol Lett 8:993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology

Guzman-Colon DK, Roloff GJ (2014) Small Indian mongoose (Herpestes auropunctatus) population abundance and effects of habitat features on trapping success in protected areas of eastern Puerto Rico. Caribb Nat 19:1–12

Guzman-Colon DK, Roloff GJ, Montgomery RA (2019) Environmental features associated with trapping success of mongoose (Herpestes auropunctatus) in eastern Puerto Rico. Caribb J Sci 49:

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

Hoagland D, Kilpatrick C (1999) Genetic variation and differentiation among insular populations of the small Indian mongoose (Herpestes javanicus). J Mammal 80:169–179

Hoagland DB, Horst GR, Kilpatrick CW (1989) Biogeography and population biology of the mongoose in the West Indies. In: Woods CA (ed) Biogeography of the West Indies. Sand Hill Crane Press, Gainesville, Florida, USA, pp 611–634

Howell JE, Moore CT, Conroy MJ, et al (2009) Conservation of northern bobwhite on private lands in Georgia, USA under uncertainty about landscape-level habitat effects. Landsc Ecol 24:405–418. https://doi.org/10.1007/s10980-008-9320-x

Jiménez-Valverde A (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. Glob Ecol Biogeogr 21:498–507. https://doi.org/10.1111/j.1466-8238.2011.00683.x

Johnson SR, Berentsen AR, Ellis C, et al (2015) Estimates of small Indian mongoose densities: Implications for rabies management. J Wildl Manage n/a-n/a. https://doi.org/10.1002/jwmg.998

Lauzeral C, Grenouillet G, Brosse S (2015) The iterative ensemble modelling approach increases the accuracy of fish distribution models. Ecography (Cop) 38:213–220. https://doi.org/10.1111/ecog.00554

Le Maitre DC, Thuiller W, Schonegevel L (2008) Developing an approach to defining the potential distributions of invasive plant species: a case study of Hakea species in South Africa. Glob Ecol Biogeogr 17:569–584

Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. Ecography (Cop) 33:103–114. https://doi.org/10.1111/J

Lockwood JL, Hoopes MF, Marchetti MP (2013) Establishment Success: The Influence of Biotic Interactions. In: Invasion Ecology , 2nd edn. Wiley-Blackwell, West Sussex

Lowe S, Browne M, Boudjelas S, De Poorter M (2000) 100 of the World's Worst Invasive Species: A selection from the Global Invasive Species Database

Manzoor SA, Griffiths G, Lukac M (2018) Species distribution model transferability and model grain size-finer may not always be better. Sci Rep 8:. https://doi.org/10.1038/s41598-018-

25437-1

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

Miller GL, Lugo AE (2009) Guide to the ecological systems of Puerto Rico. San Juan, PR

Nellis DW, Everard COR (1983) Biology of the mongoose in the Caribbean

Pearson O, Baldwin P (1953) Reproduction and age structure of a mongoose population in Hawaii. J Mammal 34:436–447

Phillips SJ, Anderson RP, Dudík M, et al (2017) Opening the black box: an open-source release of Maxent. Ecography (Cop) 40:887–893. https://doi.org/10.1111/ecog.03049

Phillips SJ, Elith J (2010a) POC plots: Calibrating species distribution models with presence-only data. Ecology 91:2476–2484. https://doi.org/10.1890/09-0760.1

Phillips SJ, Elith J (2010b) POC plots: calibrating species distribution models with presence-only data. Ecology 91:2476–2484. https://doi.org/10.1890/09-0760.1

Pimentel D (1955) Biology of the Indian Mongoose in Puerto Rico. J Mammal 36:62–68

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002

Pitt WC, Sugihara RT, Berentsen AR (2015) Effect of travel distance , home range , and bait on the management of small Indian mongooses , Herpestes auropunctatus. Biol Invasions 1743–1759. https://doi.org/10.1007/s10530-014-0831-x

Pokharel KP, Ludwig T, Storch I (2016) Predicting potential distribution of poorly known species with small database: the case of four-horned antelope *Tetracerus quadricornis* on the Indian subcontinent. Ecol Evol 6:2297–2307. https://doi.org/10.1002/ece3.2037

Powers KE, Raymond JE, Titus KL, et al (2020) Assessment of Small Indian Mongoose Diet on St. John, USVI, Using Stable Isotope Analyses. Caribb Nat 1–14

Quinn J, Whisson D, Cano F (2006) Managing the Small Indian Mongoose (Herpestes javanicus) in the Midst of Human Recreation: What is the Optimal Approach? In: Timm RM, O'Brien JM (eds) Proceedings of the 22nd Vertebrate Pest Conference. University of California, Davis, pp 393–398

Quinn JH, Whisson D a (2005) The effects of anthropogenic food on the spatial behaviour of small Indian mongooses (Herpestes javanicus) in a subtropical rainforest. J Zool 267:339–350. https://doi.org/10.1017/S0952836905007491

R Development Core Team (2008) R: A language and environment for statistical computing

Randin CF, Dirnböck T, Dullinger S, et al (2006) Are niche-based species distribution models transferable in space? J Biogeogr 33:1689–1703. https://doi.org/10.1111/j.1365-2699.2006.01466.x

Roy Horst G, Hoagland D, William Kilpatrick C (2001) The Mongoose in the West Indies. In: Biogeography of the West Indies. CRC Press, pp 409–424

Rstudio Team (2020) RStudio: Integrated Development for R

Soberón J (2007) Grinnellian and Eltonian noches and geographic distributions of species. Ecol Lett 10:1–9

Uden DR, Allen CR, Angeler DG, et al (2015) Adaptive invasive species distribution models: a framework for modeling incipient invasions. Biol Invasions 17:2831–2850. https://doi.org/10.1007/s10530-015-0914-3

Vilella FJ, Zwank PJ (1993) Ecology of the small Indian mongoose in a coastal dry forest of

Puerto Rico where sympatric with the Puerto Rican nightjar. Caribb J Sci 29:24–29

Waide RB, Comarazamy DE, González JE, et al (2013) Climate variability at multiple spatial and temporal scales in the Luquillo Mountains, Puerto Rico

Wantzen KM, Junk WJ (2000) The importance of stream-wetland-systems for biodiversity: a tropical perspective. In: Gopal B, Junk W, Davies J (eds) Wetlands: assessment, function and conservation. Backhuys, Leiden, The Netherlands, pp 11–34

Ward DF (2007) Modelling the potential geographic distribution of invasive ant species in New Zealand. Biol Invasions 9:723–735. https://doi.org/10.1007/s10530-006-9072-y

West AM, Kumar S, Brown CS, et al (2016a) Field validation of an invasive species Maxent model. Ecol Inform 36:126–134. https://doi.org/10.1016/j.ecoinf.2016.11.001

West AM, Kumar S, Brown CS, et al (2016b) Field validation of an invasive species Maxent model. Ecol Inform 36:126–134. https://doi.org/10.1016/j.ecoinf.2016.11.001

Westwood R, Westwood AR, Hooshmandi M, et al (2019) A field-validated species distribution model to support management of the critically endangered Poweshiek skipperling (Oarisma poweshiek) butterfly in Canada. Rev Conserv Sci Pract ID19-200. https://doi.org/10.1111/csp2.163

Woodman SM, Forney KA, Becker EA, et al (2019) esdm: A tool for creating and exploring ensembles of predictions from species distribution and abundance models. Methods Ecol Evol 10:1923–1933. https://doi.org/10.1111/2041-210X.13283

Yates KL, Bouchet PJ, Caley MJ, et al (2018) Outstanding Challenges in the Transferability of Ecological Models. Trends Ecol Evol 33:790–802. https://doi.org/10.1016/J.TREE.2018.08.001

Figures and Tables

Table 8. Environmental variables included in the species distribution models. Temperature and precipitation seasonality are represented by coefficients of variation across months within a year. Maximum temperature of the warmest month is the maximum temperature value across all months within a year. Worldclim acronyms are shown in parenthesis (e.g. Bio1).

Predictor	Spatial Resolution	Data Source
Annual Mean Temperature (Bio1) Temperature Seasonality (Bio4) Max Temperature of Warmest Month (Bio4) Precipitation Seasonality (Bio15)	Resampled to 30 m	BIOCLIM
Distance from observed Mongoose presence to: Evergreen forests Freshwater (e.g. rivers, lakes) Pastures Plantations Semi-deciduous forest Shrublands Urban Areas Wetlands	Resampled to 30 m	Puerto Rico Gap Analysis Project

Table 9. Comparison of performance metrics of mongoose distribution models when applied to field collected presence data. Bold text indicates acceptable discrimination between presence and background points.

	Threshold independent (± SD)		Threshold dependent		
Model	AUC	CBI	Sensitivity	Specificity	TSS
MaxEnt	0.89 (± 0.04)	0.44 (± 0.12)	0.88	0.64	0.52
GAM	0.86 (± 0.01)	0.73 (± 0.16)	0.57	1	0.57
ENFA	0.75 (± 0.41)	0.61 (± 0.41)	0.51	0.66	0.17
BRT	0.84 (± 0.07)	0.56 (± 0.49)	0.97	0.05	0.02
Ensemble	0.74	0.36	0.68	0.74	0.42

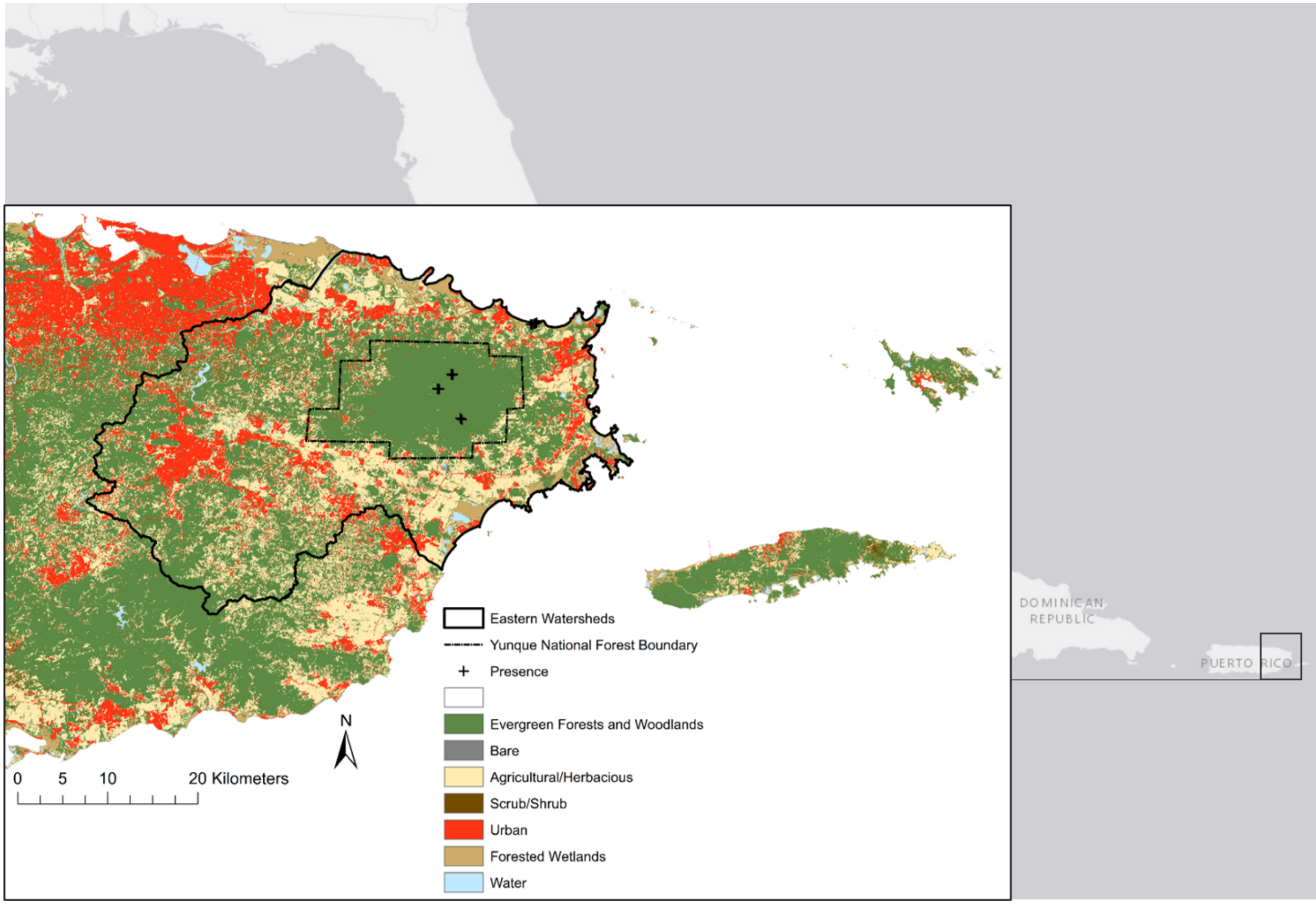


Figure 10. Study area for an initial species distribution modelling of small Indian mongoose (*Herpestes auropunctatus*) in Puerto Rico. The left bottom inset shows El Yunque National Forest proclamation zone (black dotted line), eastern watershed zone (solid black boundary) and traps (black crosses) within the Yunque Protected area and the North East Ecological Corridor. I trained the initial SDM model used to guide a subsequent field survey at the watershed level. The eastern watershed zone includes a mix of mostly agricultural/herbaceous and evergreen forests with fragmented patches of urban areas.

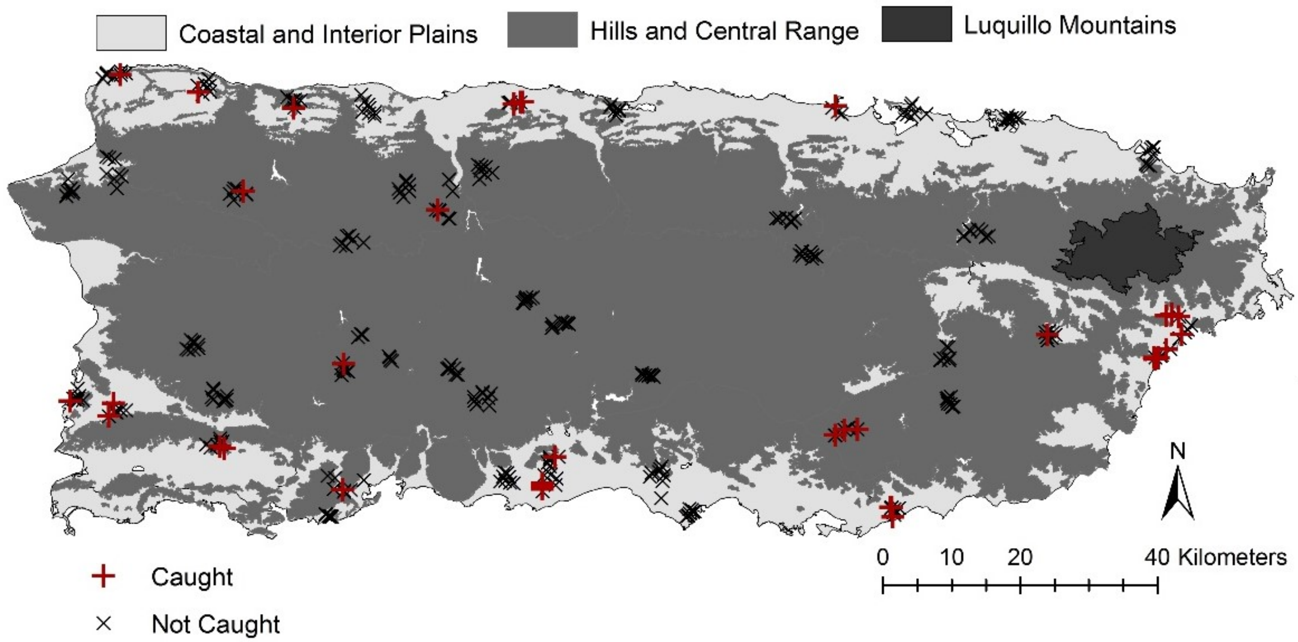


Figure 11. We placed fifty-four 1-km^2 grids with eight traps randomly located within each grid. The grids were distributed equally among three classes of habitat suitability index values (HSI: 0 - 0.3; > 0.3 - 0.7; > 0.7 - 1.0) with nine replicates for each of the three HSI categories on the mountainous region (n = 27) and on the coastal areas (n = 27) within the island.

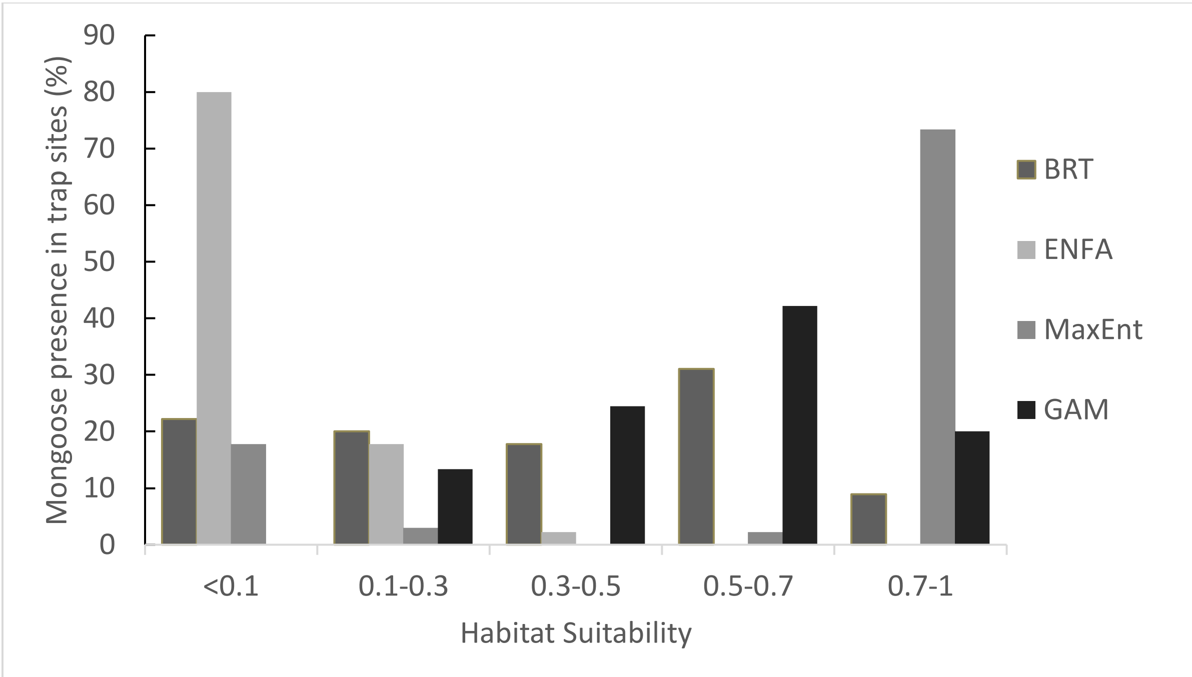


Figure 12. Percent of mongoose observed in sampled sites for each relative habitat suitability class generated by each model.

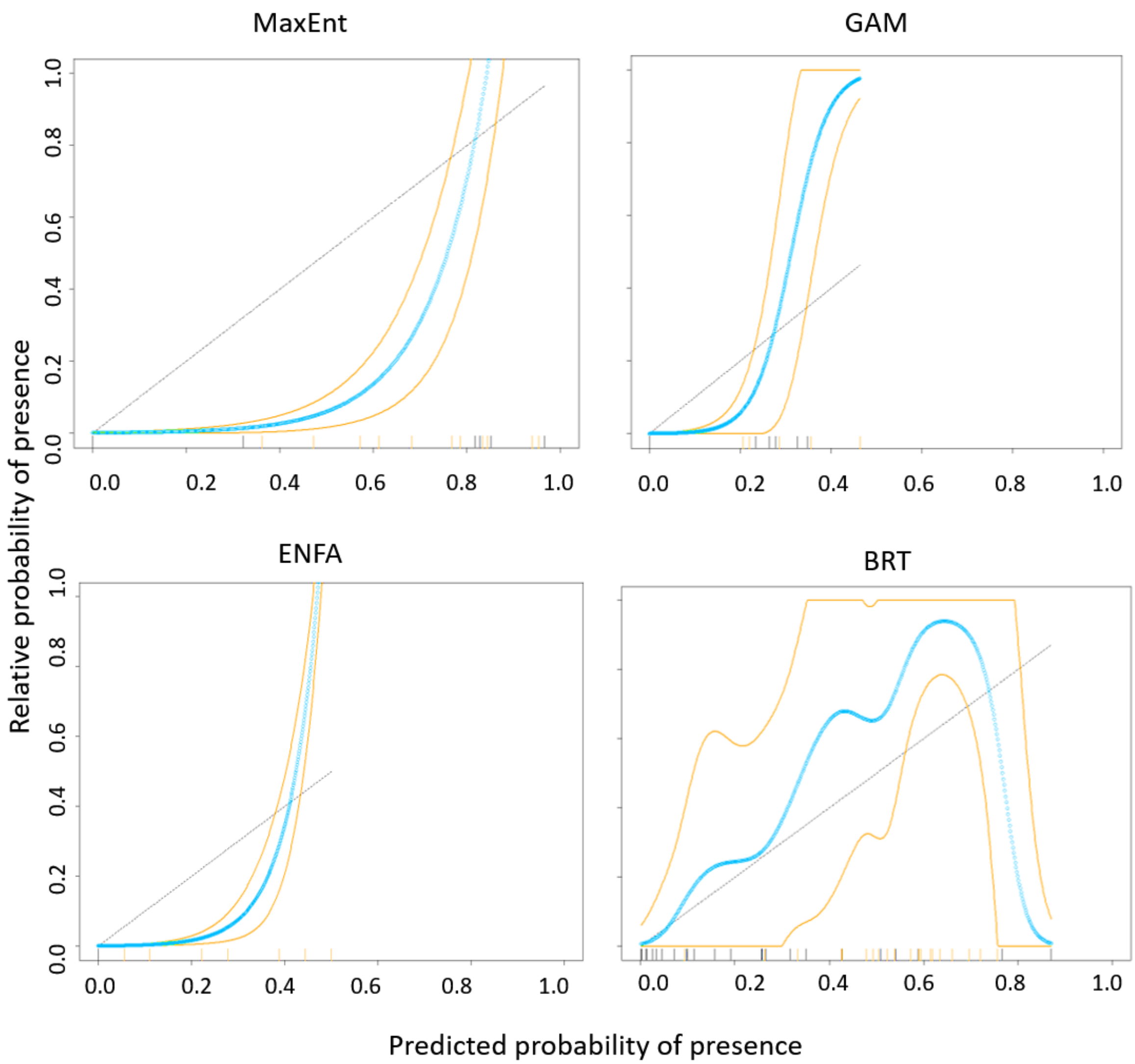


Figure 13. Comparison of presence-only smoothed calibration plot (POC) for our four SDM types when validated with external data (presence data gathered in a guided survey). The y-axis represents the relative probability of presence projected by a model, the x-axis is the predicted probability of presence calibrated externally with the 2017 field data. Values (blue line) above the diagonal indicate model underestimation of species prevalence (i.e. observed proportion of presences). Values below the diagonal indicate overestimation of species prevalence. Orange lines represent ±2 standard deviations.

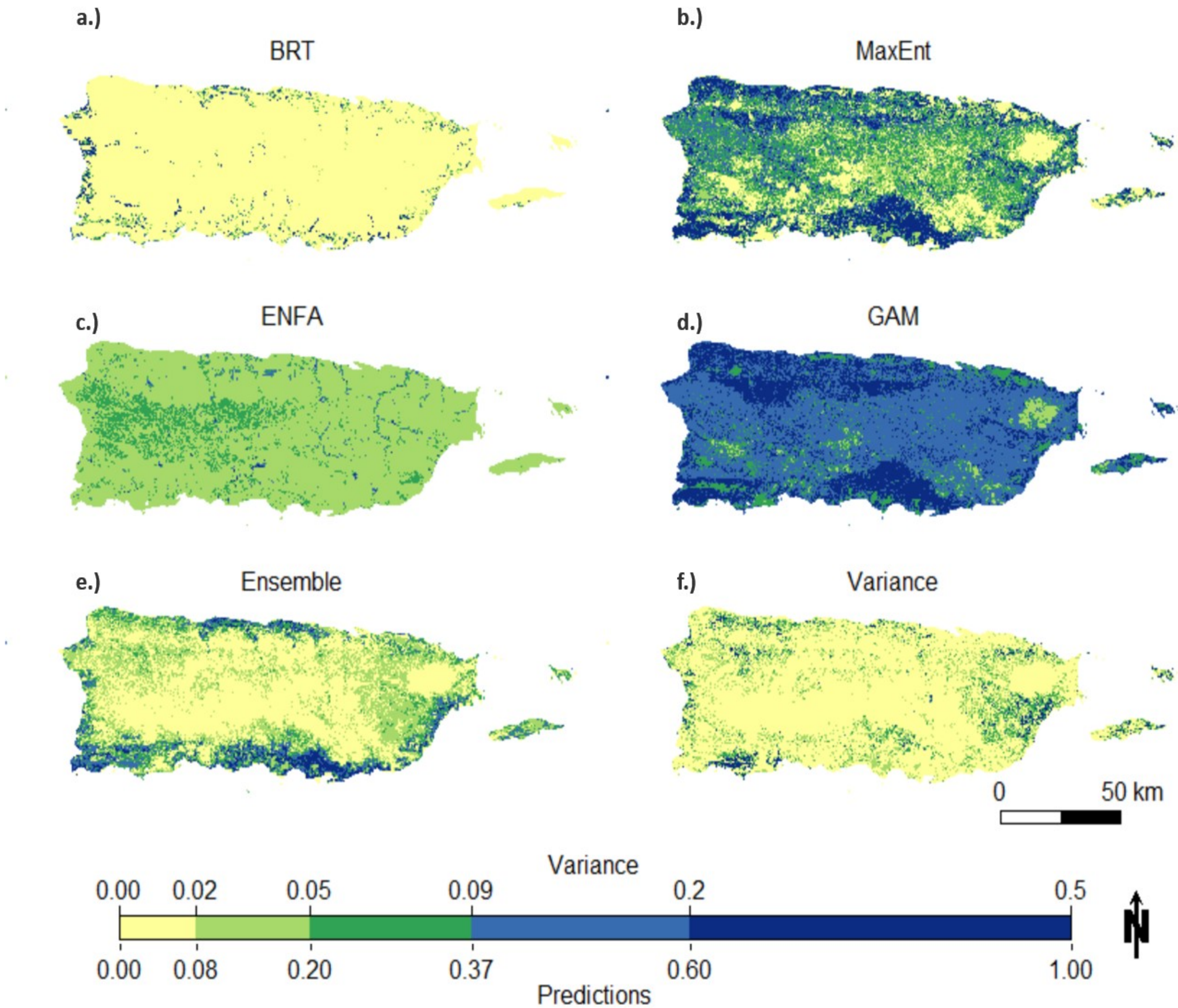


Figure 14. Relative habitat suitability for mongoose (Herpestes auropunctatus) in Puerto Rico according to four species distribution models (a-d), and ensemble of all predictions and the variance in those predictions; (a) Boosted regression tree, (b) MaxEnt, (c) Environmental niche factor analysis, and (d) General additive model. Ensemble model maps created with the unweighted average of the predictions of the 4 individual models (f) and associated variance (e). Variance in predictions reflects among-model uncertainty.

Chapter 3. Genetic variability of the small Indian mongoose in Puerto Rico is associated with proximity to modified landscapes.

Target Journal: Biological Invasions

Abstract

Mammalian predators contribute disproportionately to declines in global biodiversity. Effective management of invasive mammalian predators requires understanding connectivity and dispersal among their populations. The goal of this study is to describe the small Indian mongoose (*Herpestes auropunctatus*) genetic patterns and determine whether landscape characteristics influence genetic connectivity (the proportion shared alleles by individuals). The mongoose is one of the top invasive predators in the world and it has been introduced to several islands of the Caribbean and Pacific. We used a landscape genetic approach to determine how landscape features in Puerto Rico influence gene flow for mongooses by analyzing correlations between individual genetic distance and landscape resistance models. We evaluated several alternative hypotheses about mechanisms shaping genetic structure, including the Puerto Rico human footprint, mongoose habitat distribution, structural landscape connectivity and elevation. Additionally, we assessed evidence of isolation-by-distance (IBD) and isolation-by-barrier (IBB) in the mongoose population. We found a trend supporting weak genetic structure ($Rst = 0.035; p = 0.062$), and observed that genetically similar individuals clustered near highly human-modified landscapes ($r = -0.21$, $p < 0.05$). This work directly addresses the concerns about the ability of connected natural areas to allow for the spread of invasive species. Except for human-modified landscapes, we found no other relationship between mongoose genetic distance and hypothesized barriers quantified as resistance.

Keywords: landscape connectivity, landscape genetics, invasive species, Caribbean, invasive predators

Introduction

Invasive species can disrupt ecosystem processes by altering evolutionary or competitive interactions of species in invaded areas, sometimes promoting hybridization, and leading to biodiversity loss or loss of community complexity in invaded habitat (Mesgarana et al., 2016; Mollot et al., 2017; Pyšek et al., 2012). Studies at a global scale have highlighted how invasive species in combination with human-driven land-use practices have had negative impacts on biodiversity (Saunders et al., 1991). Understanding and managing invasive populations are thus a key focus for conservation strategies (Leung et al., 2002). Realistic control strategies rely on good prediction of potential pathways of invasion and degree of connectivity with other populations. A good understanding of genetic exchange in the landscape can lead to the identification of isolated populations which might be easier to manage than populations functionally connected by dispersal (Robertson & Gemmell, 2004). While there are many approaches to estimate population connectivity (O'Brien et al., 2006; Rayfield et al., 2011), robust validation of connectivity estimates is based on inferring gene flow (Schwartz et al., 2009).

The basis for maintaining population functional connectivity is the continued dispersal of individuals through the landscape to maintain genetic diversity within populations (Frankham, 2008). Dispersal itself is considerably influenced by the structural characteristics of the landscape (Gauffre et al., 2008). If landscape attributes and environmental factors regulate genetic processes in invasive species, identifying potential dispersal routes of invasive species can help guide actions to reduce the occurrence of new invasions, demonstrate how landscapes shape population genetic structure, and identify landscape barriers against gene flow. An individual-based, landscape genetic approach for the analysis of population genetic connectivity is a powerful way to quantify how environmental features effect population structure because it directly correlates patterns of genetic relatedness between individuals with resistance in movement between these individuals (Barros et al., 2019; Montgelard et al., 2014).

Caribbean islands, considered one of the world's biodiversity hotspots (Myers et al., 2000), are in peril; their biodiversity is at high risk from invasive species introductions via trade-related pathways and the islands are vulnerable to species extinction through habitat degradation and loss (Waugh, 2009). Because landscapes can have a direct influence on gene flow, in this study, we investigate how landscape characteristics might shape the genetic structuring of invasive mongoose populations in Puerto Rico.

The mongoose is listed as one of the top invasive species in the world (Lowe et al., 2004). It was introduced to many Pacific and Caribbean islands at the beginning of the 20th century (Arijana Barun et al., 2011), and was identified as a major threat in the Caribbean region (Kairo et al., 2003). Shortly after mongoose establishment, noticeable declines in several native island species were reported (Nadin-Davis et al., 2008; Seaman & Randall, 1962). An additional topic of concern is rabies spread (Felipe Cano and Are Berentsen, USDA Forest Service and APHIS, *pers. comm.*) and, recently, the mongoose has been found to carry the bacterial disease leptospirosis (Benavidez et al., 2019b). The prevalence of mongoose throughout Puerto Rico may be related to their adaptability to new environments and their overall ecological generalism (Powers et al., 2020), but their population genetic structure and how it is related to the landscape is still not understood.

Introduced mongoose populations are genetically isolated from those of their native range (i.e. Middle East to South Asia) and show phenotypic variation consistent with ecological release from competition (Simberloff et al., 2000). For example, mongooses on islands are slightly larger both overall and in the size of their upper canine tooth and exhibit greater sexual dimorphism than mongooses in their native range. However, there is no evidence on whether this morphological variation is genetically controlled. Some island populations have retained much genetic variation while others have not, suggesting that observed patterns of genetic diversity may be due to multiple independent mongoose introductions along with rapid population growth on some islands but not others (D. Hoagland & Kilpatrick, 1999; C.-G. Thulin et al., 2006) . Nevertheless, those patterns are viewed as tentative given that only single sites (<2 per island) with low sample sizes (n=<50) were represented in genetic analyses (C. E. Bennett, 2007; C. Thulin & Gyllenstrand, 2002).

The overarching goal of this study was to determine the degree to which genetic structure of the small Indian mongoose and landscape structure are related on the island of Puerto Rico. Under a landscape genetics framework, our objectives were to first quantify genetic variation and population structure, and second, assess the influence of landscape connectivity and features on the genetic patterns. We considered alternative hypotheses that could explain the genetic variability observed, including isolation by distance (IBD), isolation by barrier (IBB), and landscape resistance, where resistance included habitat suitability levels, degrees of structural connectivity (patch-level), and human footprint index.

Methods:

Study Area and Species

Mongoose presence in the Caribbean is more apparent in coastal habitats including dry forests, scrub areas, and pasturelands ((Pimentel, 1955; Vilella, 1998; Vilella & Zwank, 1993)than in interior habitats. Individuals are mostly solitary with overlapping home ranges (3.1-52.2ha) and their mean dispersal distance is 550m (Johnson et al., 2015a). Different numbers of introductions of mongoose on different islands are most likely responsible for the genetic diversity patterns observed in the Caribbean (D. Hoagland & Kilpatrick, 1999). Nuclear and mitochondrial homogeneity across islands may have been influenced by human-mediated transfers of mongooses (C. E. Bennett, 2007; D. Hoagland & Kilpatrick, 1999).

Spatial sampling design

Mongoose tissue and hair samples were collected in two ways: from opportunistic encounters from US Forest Service and Fish and Wildlife Service lands, and from a randomized trapping scheme. Forty samples were acquired opportunistically, including from the island of Vieques, 21 km east of the main island, between 2011-2017. When designing the sampling scheme, I assumed that mongooses exhibit a continuous distribution through the landscape (Oyler-McCance et al., 2012). I used ArcGIS v.10.1 to select 54 sites island-wide, allocating grids in a stratified-random design to encompass two major physiogeographical regions that generally differ in elevation, precipitation, and vegetation patterns (Torres-Valcárcel et al., 2014). I stratified each region with nine replicates according to a combination of High and Low Habitat

Suitability scores with High and Low connectivity (Guzmán-Colón et. al., *in prep*). I generated a set of 8 trap points separated by a minimum of 200 – m) within each grid using the Sampling Design Toolbox for ArcGIS v. 10.3.

DNA Extraction and Molecular Markers

I extracted genomic DNA with DNeasy Blood and Tissue Kit (QIAgen GmbH) using 25 mg of tissue material and a range of 5 - 10 hair follicles following manufacturer protocol. All extracted hair follicles were further purified via ethanol precipitation. I also included a negative control during each extraction to monitor contamination. I quantified extracted DNA from each individual via spectrophotometry using a NanoDrop CD-1000 spectrophotometer.

I selected ten microsatellites primer pairs developed by Thulin et. al. 2002 for the small Asian mongoose (*Herpestes javanicus*). I amplified DNA target fragments using primers tagged with an Illumina adapter in separate PCRs. Each PCR was conducted using Phusion Taq polymerase and the reactions consisted of: 1X PCR buffer, 0.25 mM of dNTPs, 0.25 unit/uL of polymerase, 0.2 µM of each primer, 10 µL DNA extract and ultra-pure water to bring the total reaction volume to 25 µL. My PCR parameters consisted of one step at 98°C for 30 sec followed by 40 cycles of 98°C for 15 sec, 55°C-65°C for 30 sec, and 72°C for 30 sec, and then a final elongation step at 72°C for 10 min. I then checked for positive amplification via gel electrophoresis. All positively amplified products were cleaned using AMPure XP beads to remove primer dimer. I quantified the cleaned PCR products using Qubit and combined equimolar all PCR products by individual. After I pooled PCR products by individual, I attached unique identifying barcodes via PCR. The PCR reactions were carried out using Kapa HiFi HotStart Master Mix. I purified all PCR products using AMPure XP beads and then quantified the cleaned products using Qubit. I then pooled all amplicons equimolar into a single library for pair-end sequencing on an Illumina

MiSeq platform. Library QC, dilution and sequencing were conducted at University of Wisconsin Biotechnology Center.

Microsatellite genotyping and errors

I characterized individual genotypes using program MEGASAT (Zhan et al., 2017) which reads sequence files from MiSeq and automatically scores microsatellite genotypes.

I calculated genotyping error rates using a repeat-genotyping method which quantifies the mismatches between genotypes called from duplicated individuals. Out of the 192 samples genotyped, 84 were resampled to estimate genotyping error. I tested a series of parameters to ultimately choose the highest number of successful reads and minimize genotype discrepancies between samples and duplicates. The combination of parameters consisted of varying the mismatch tolerance from 2-4 base pairs and sequence depth threshold from 30-50 in MEGASAT. Error rate per locus was calculated as the total number of mismatches per locus over the total number of alleles genotyped per locus (Zhan et. al. 2017). This error calculation was repeated for all parameter combinations. Finally, I removed samples with incomplete genotypes from further analysis.

Genetic Structure

I assessed the distribution of gene diversity and estimated the hierarchical variance components attributed to populations, sample sites and individuals via an analysis of molecular variance (AMOVA). To investigate population differentiation, I used pairwise R_{st} comparison among all pairs of sampling sites using the stepwise mutation model (9, 9999 permutations for significance). Both AMOVA and R_{st} calculations were done in Genalex (Peakall & Smouse, 2012). In addition, I used a Principal Coordinate Analysis (PCoA) with package *adegenet* in R (version) to show the pattern of genetic differentiation of mongoose individuals.

Because of a lack of significant genetic structure in samples (see Results section), I used an individual-based genetic distance measure (1-Dps; or 1- the proportion of shared alleles) to create a matrix of pairwise genetic distance that would be correlated with different matrices of geographical and cost (resistance) distances.

I used resistance maps which represented six different hypotheses about the resistance of landscape attributes to potential gene flow among individuals. Four landscape variables were used to represent resistance models: (1) the Puerto Rico human footprint, (2) the Island-wide structural Patch connectivity, (3) Mongoose habitat suitability, and (4) Elevation. The other two resistance models represented isolation by distance and isolation by barrier, where barriers were defined as oceanic regions. Each map had a raster cell size set to 30 x 30 m, which is smaller than the average home-range size of the study species (i.e. > 0.02 km^2). Having a sampling grain size that is smaller than the home-range size has been found adequate to infer landscape effects on gene flow because it allows representation of smaller fragmented elements in the landscape and can be adjusted to a coarser resolution as needed (Dale & Fortin, 2010; Wiens, 1989). I didn't simplify continuous landscape values into discrete classes since this practice decreases power to detect landscape effects on gene flow (S. a Cushman & Landguth, 2010). I used Circuitscape (v. 4) to calculate cumulative resistance distance among mongoose locations throughout each of the resistance maps.

1) Isolation by distance: IBD was treated as a null model. On an unvarying landscape, I assumed movement could occur without resistance in any direction, but would decline linearly with geographical distance. I built IBD values by calculating the Euclidean distance using UTM coordinates between all individuals sampled using function "mat_geo_dist" in the package *graph4lg* in R.

2.) Isolation by Barrier: I acquired samples from two islands in the PR archipelago thus I treated the sea as a possible barrier to dispersal. Individuals within an island had a pair-wise resistance value of 0 (none), otherwise, resistance was set to 1 (high).

3.) Isolation by resistance: My third set of hypotheses proposed that landscape and environmental features of the island can either hinder or promote genetic connectivity of mongooses.

a.) Human Footprint: The human footprint (a quantification of the impact human driven activities have in the land surface) differentially influences genetic connectivity for mammal species (Thatte et al., 2020). For example, road density reduces gene flow, thus affecting population genetic connectivity regardless of the dispersal ability of species (Coulon et al., 2006; Epps et al., 2007; Thatte et al., 2020). My resistance scores were assigned according to the human footprint value calculated in Guzmán-Colón et. al. 2019 and were scaled from 0 (low) to 1 (high).

b.) Island-wide Patch connectivity: Multiple studies suggest that landscape connectivity increases species persistence by allowing dispersal, promoting genetic diversity, and aiding in species conservation goals (i.e. mitigating inbreeding and sustaining evolutionary processes (Andrello et al., 2015; Gómez-Fernández et al., 2016; Schoville et al., 2018). I previously constructed an ecological network consisting of a set of high-quality habitat patches and links (Guzmán-Colón et al., 2020). I assigned resistance values according to the degree of connectivity for each patch and link.

c.) Elevation: Elevation gradients can directly result in reproductive isolation of populations due to local topographic and climatic factors that restrict gene flow (Osborne et. al. 2013; Funk and Murphy 2010). Mongoose abundance in Puerto Rico is significantly greater at lower elevation zones (Johnson et. al. 2015; Guzmán-Colón et. al. 2016) where there is a relatively flat landscape. Indeed, other species of mongoose (*Herperstes ichneumon*) experience high elevation as a barrier in western Iberia (Barros et al., 2015).

d.) Mongoose Habitat Suitability: I used the inverse of habitat suitability values as a measure of environmental resistance. I developed a habitat suitability distribution map representing preferred mongoose habitat based on land cover and environmental variables (Guzman-Colon et. al., *Chapter 2*). Habitat suitability values range from 0 to 1, with higher values indicating higher suitability. Those values were reversed to assign resistance scores.

Causal modeling

I used a causal modeling framework to infer which landscape features best explained the genetic structure of mongooses (S. A. Cushman et al., 2006; Reding et al., 2013). I used three explanatory models of causality by implementing simple Mantel and partial Mantel tests containing seven hypotheses against the IBD null model: 1.) Simple Mantel tests between genetic distance and landscape resistances; 2.) Partial Mantel tests between genetic distances and landscape resistances, while partialling out the effects of Euclidean distance; 3.) Partial mantel test between Euclidian distances, partialling out the effects of landscape resistance measures (Figure 15).

In the causal model, evidence of landscape resistance to gene flow (i.e. a lower proportion of shared alleles among individuals) would be indicated if model 1 and 2 Mantel tests are significant ($p < 0.05$; Mantel $r <> 0$) and in model 3, if the Mantel test is non-significant. We implemented models with Pearson correlation with 10,000 permutations for a 95% confidence interval using the *vegan* package for R.

Results

Microsatellite analyses

From the combined opportunistic and stratified sampling, I sampled genotypes for 192 individual samples of which 82 were duplicates (42.7%). I retained 55 unique samples after screening for errors and selecting a mismatch tolerance of two and sequence depth of 50 (50% from the total samples analyzed excluding duplicates). The genotyping errors ranged from 0.01to 0.06, with a mean error for all loci and samples combined of 0.043. Unfortunately, three of the ten microsatellites that were amplified during PCR were not sequenced reliably and thus I excluded them from subsequent analyses.

All remaining loci were variable with total numbers of alleles ranging between 3 and 8 per locus. All microsatellites displayed a range of polymorphism from low (2) to moderately high (8). Mean expected (He) and observed (Ho) heterozygosity for the sample was He = 0.196 +/- 0.05 and Ho = 0.218 +/- 0.07. The overall mongoose dataset showed higher than expected heterozygosity than predicted under Hardy-Weinberg expectations. Three of the seven loci were out of Hardy-Weinberg proportions. One pair of loci of the 21 showed linkage disequilibrium after Bonferroni correction.

The genetic diversity indices across the three sampling sites (which we treat as populations) are summarized in Figure 16. The number of different alleles (Na), private alleles (NP) and effective alleles (NE) averaged across all loci ranged from 0.3 to 2.1, 0.2 to 1, and 0.3 to 1.1, respectively for the three sampling sites (Coastal, Mountain and Vieques). I found that the mountain population was inbred and had the lowest allelic diversity (Shannon's Index; I = 0.069) over all pairs of loci.

Structure

The analysis of molecular variance (AMOVA) indicated that 70% of the total variation was due to differences among individuals across all populations whereas 11% of the variation was due to heterozygosity within individuals. At the site level, there was a slight but non-significant indication of population structure ($R_{st} = 0.035$; $p = 0.062$). My calculated inbreeding coefficient showed that there is a moderate degree of inbreeding ($R_{is} = 0.202$, $p < 0.05$), and this pattern was also evident as a deficiency of observed heterozygotes among individuals with respect to that expected for the total population. Pairwise comparisons among populations were not significant in all cases ($p <0.005$), although there was slight ($R_{st} = 0.396$) indication of genetic differentiation between coastal and mountain area populations.

The principal coordinate analysis agreed with the population structure R_{st} statistics and showed no indication of significant structure in the data set (Figure 17).

Isolation by distance and isolation by resistance through mantel correlations

Simple Mantel tests showed that out of the six hypotheses, only the Puerto Rico human footprint had a significant positive correlation with genetic distance across the study area ($r = -0.21$, $p < 0.05$) (Model 1). Mongooses near anthropogenic activities in the landscape tend to be more similar genetically. The isolation by distance plot confirms that the genetic patterns in our samples are not explained by geographic distance; however, we identified high density of genetic relatedness at short distances (0-2500 m) (Figure 18).

Partial mantel test analyses for Model 2 also supported the human footprint resistance ($r = -0.207$; $p < 0.05$) after partialling out the effects of Euclidian distance. None of the other hypotheses showed a significant effect on genetic distance with simple mantel tests or after partialling out Eucledian distance. We did not find support for Model 3. We expected that the hypothesis supported by the data in Models 1 and 2, but not for Model 3, and indeed the human footprint showed this relationship.

Discussion

My results suggest that closely related mongooses occur near highly modified land-cover areas (i.e. areas with a high Human Footprint) but they are not structured by degree of geographic isolation or features such as habitat structural connectivity, habitat distribution or elevation. I did not find a strong pattern of IBD, and in fact, the causal modelling framework supported the human footprint hypothesis with no independent support for the IBD hypothesis. Instead, the IBD plot revealed a continuous cloud of genetic differentiation with no sign of discontinuity, and local densities appeared to be concentrated between 1, 500 to 2, 500 m. I found that mongooses in Puerto Rico present a slight degree of genetic variation among and within individuals across the sampled population, but this variation was not enough to confidently delineate populations. Despite the good dispersal ability of the mongoose and seemingly panmictic population, there were signatures that populations violated equilibrium assumptions. Specifically, mean observed heterozygosity was higher than expected. This pattern could be due to recent stochastic demographic processes, or it could also indicate an isolate-breaking effect when mongooses from different populations were introduced in the island. For example, the practice of introducing propagules from Jamaica to other Caribbean islands was fairly common in the 1900's (C.-G. Thulin et al., 2006). But, although it is plausible that multiple introduction events occurred, the only instance recorded in Puerto Rico was one shipping event containing twenty mongooses in which the location of introductions remain unknown (Espeut 1889; Hoagland 1989). As in

Jamaica and other Caribbean islands, it is believed that rapid population growth in Puerto Rico after the introduction event precluded genetic differentiation despite the small founding group (Bennet 20016; Hoagland 1989). Founding effects or bottlenecks can cause rare alleles to be lost in the population, and indeed, we found a low number of private alleles despite higher observed heterozygosity. We also found, consistent with other allozyme studies at the broader Caribbean extent and stepping-stone models, that average allelic richness is lower in Puerto Rico than in Jamaica and their native range.

Understanding the causes of observed patterns of genetic structure in mongoose populations ultimately depends on associating the behavior and movement of individuals to population-level patterns of gene flow (Cushman et. al. 2006). In the end, what is reflected at the population level are the individual decisions that led to dispersal and reproduction. One of the fundamental factors governing behavior is the distribution of resources in the landscape, but also proximity to abundant food sources that can support a high population size (Fedriani et al., 2001). Patchy but abundant resources such as those available in the proximity of humans can lead to high density and tolerance among conspecifics, without imposing significant fitness costs (Elmhagen et al., 2014). Anthropogenic food sources can lead to constrained space use and overlapping home-ranges, as has been noted in banded mongoose (*Mungos mungo*; (Laver & Alexander, 2018), raccoons (*Procyon lotor*; Prange et al. 2004), and black bears (*Ursus americanus*; (Rogers, 1987). Thus, modified landscapes and not landscape connectivity per se could act as a facilitator of gene flow for certain omnivorous species. The molecular variation patterns within individuals and pair-wise genetic relatedness in mongoose populations explained by proximity to modified habitats make sense in that mongooses tend to restrict their normal area of use around known food reserves, but occasional individuals can also travel and forage over large areas (Johnson et. al. 2017; Quinn 2015).

It is worth noting that this analysis is based on a limited genetic dataset and some loci that were discarded after genotyping. Individual assignments to populations are more accurate when using genetic clusters previously identified instead of sampled populations. Additionally, low genetic diversity, which is characteristic of mongoose populations in the Caribbean, makes it difficult to achieve high accuracy of relatedness estimates because there is less genetic variation available to differentiate between individuals, effectively reducing the variance in relatedness. However, our field sampling design allowed us to sample individuals across a broad geographical range in the island which is representative of different environmental conditions. We also selected the best combination of depth and base pair mismatch tolerance when scoring genotypes to reduce genotyping errors.

In conclusion, we found that genetic similarities among mongooses in our sampled population is facilitated by the areas of high anthropogenic influences. Invasive species are a major threat to island biodiversity world-wide and potentially on the island of Puerto Rico. While we can model population connectivity based on habitat suitability or expert opinion, genetic analyses like this one may assist in identifying potential invasion routes as well as revealing population connectivity. This study offers some clues about the role of landscape connectivity in facilitating mongoose gene flow, suggesting that the degree of structural connectivity within the island was not associated with gene flow. However, a clear management application of this study is that the association between mongoose genetic connectivity and anthropogenic influence highlights the importance of maintaining protected areas and corridor buffer zones geographically separate from development and other intense anthropogenic land uses to prevent mongoose establishment in these sensitive areas.

References

Andrello M, Mouillot D, Somot S, et al (2015) Additive effects of climate change on connectivity between marine protected areas and larval supply to fished areas. Divers Distrib 21:139–150. https://doi.org/10.1111/ddi.12250

Barros T, Carvalho J, Fonseca C, Cushman SA (2019) Assessing the complex relationship between landscape, gene flow, and range expansion of a Mediterranean carnivore. Eur J Wildl Res 65:. https://doi.org/10.1007/s10344-019-1274-6

Barun A, Simberloff D, Tvrtkovic N, Pascal M (2011) Impact of the introduced small Indian mongoose (Herpestes auropunctatus) on abundance and activity time of the introduced ship rat (Rattus rattus) and the small mammal community on Adriatic islands, Croatia. NeoBiota 11:51. https://doi.org/10.3897/neobiota.11.1819

Benavidez KM, Guerra T, Torres M, et al (2019) The prevalence of Leptospira among invasive small mammals on Puerto Rican cattle farms. PLoS Negl Trop Dis 13:e0007236. https://doi.org/10.1371/journal.pntd.0007236

Bennett CE (2007) Reconstructing the Evolutionary and Population History of an Introduced Carnivore, The Small Indian Mongoose. Columbia University

Coulon a, Guillot G, Cosson J-F, et al (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. Mol Ecol 15:1669–79. https://doi.org/10.1111/j.1365-294X.2006.02861.x

Cushman S a, Landguth EL (2010) Spurious correlations and inference in landscape genetics. Mol Ecol 19:3592–602. https://doi.org/10.1111/j.1365-294X.2010.04656.x

Cushman SA, Mckelvey KS, Hayden J, Schwartz MK (2006) Gene Flow in Complex Landscapes: Testing Multiple Hypotheses with Causal Modeling. Source Am Nat 168:486–499. https://doi.org/10.1086/506976

Dale MRT, Fortin M-J (2010) From Graphs to Spatial Graphs. Annu Rev Ecol Evol Syst 41:21–38. https://doi.org/10.1146/annurev-ecolsys-102209-144718

Elmhagen B, Hersteinsson P, Norén K, et al (2014) From breeding pairs to fox towns: The social organisation of arctic fox populations with stable and fluctuating availability of food. Polar Biol 37:111–122. https://doi.org/10.1007/s00300-013-1416-3

Epps CW, Wehausen JD, Bleich VC, et al (2007) Optimizing dispersal and corridor models using landscape genetics. J Appl Ecol 44:714–724. https://doi.org/10.1111/j.1365-2664.2007.01325.x

Fedriani JM, Fuller TK, Sauvajot RM (2001) Does availability of anthropogenic food enhance densities of omnivorous mammals? An example with coyotes in southern California. Ecography (Cop) 24:325–331. https://doi.org/10.1111/j.1600-0587.2001.tb00205.x

Frankham R (2008) Inbreeding and Extinction: Island Populations. Conserv Biol 12:665–675. https://doi.org/10.1111/j.1523-1739.1998.96456.x

Gauffre B, Estoup a, Bretagnolle V, Cosson JF (2008) Spatial genetic structure of a small rodent in a heterogeneous landscape. Mol Ecol 17:4619–29. https://doi.org/10.1111/j.1365-294X.2008.03950.x

Gómez-Fernández A, Alcocer I, Matesanz S (2016) Does higher connectivity lead to higher genetic diversity? Effects of habitat fragmentation on genetic variation and population structure in a gypsophile. Conserv Genet 17:631–641. https://doi.org/10.1007/s10592-016-0811-z

Guzmán-Colón DK, Pidgeon AM, Martinuzzi S, Radeloff VC (2020) Conservation planning for island nations: Using a network analysis model to find novel opportunities for landscape connectivity in Puerto Rico. Glob Ecol Conserv 23:e01075. https://doi.org/10.1016/j.gecco.2020.e01075

Hoagland D, Kilpatrick C (1999) Genetic variation and differentiation among insular populations of the small Indian mongoose (Herpestes javanicus). J Mammal 80:169–179

Johnson SR, Berentsen AR, Ellis C, et al (2015) Estimates of small Indian mongoose densities: Implications for rabies management. J Wildl Manage n/a-n/a. https://doi.org/10.1002/jwmg.998

Kairo M, Ali B, Cheesman O, et al (2003) Invasive species threats in the Caribbean region: report to the Nature Conservancy

Laver PN, Alexander KA (2018) Association with humans and seasonality interact to reverse predictions for animal space use. 6:5. https://doi.org/10.1186/s40462-018-0123-7

Leung B, Lodge DM, Finnoff D, et al (2002) An ounce of prevention or a pound of cure: Bioeconomic risk analysis of invasive species. Proc R Soc B Biol Sci 269:2407–2413. https://doi.org/10.1098/rspb.2002.2179

Lowe S, Browne M, Boudjelas S, De Poorter M. (2004) 100 of the world's worst invasive alien species . Auckland, New Zealand

Mesgarana MB, Lewisb MA, Adesc PK, et al (2016) Hybridization can facilitate species invasions, even without enhancing local adaptation. Proc Natl Acad Sci U S A 113:10210–10214. https://doi.org/10.1073/pnas.1605626113

Mollot G, Pantel JH, Romanuk TN (2017) The Effects of Invasive Species on the Decline in Species Richness: A Global Meta-Analysis. In: Advances in Ecological Research.

Academic Press Inc., pp 61–83

Montgelard C, Zenboudji S, Ferchaud AL, et al (2014) Landscape genetics in mammals. Mammalia 78:139–157. https://doi.org/10.1515/mammalia-2012-0142

Myers N, Mittermeier RA, Mittermeier CG, et al (2000) Biodiversity hotspots for conservation priorities. Nature 403:853–858

Nadin-Davis S, Velez J, Malaga C, Wandeler A (2008) A molecular epidemiological study of rabies in Puerto Rico. Virus Res 131:8–15

O'Brien D, Manseau M, Fall A, Fortin MJ (2006) Testing the importance of spatial configuration of winter habitat for woodland caribou: An application of graph theory. Biol Conserv 130:70–83. https://doi.org/10.1016/j.biocon.2005.12.014

Oyler-McCance SJ, Fedy BC, Landguth EL (2012) Sample design effects in landscape genetics. Conserv Genet 14:275–285. https://doi.org/10.1007/s10592-012-0415-1

Pimentel D (1955) Biology of the Indian Mongoose in Puerto Rico. J Mammal 36:62–68

Powers KE, Raymond JE, Titus KL, et al (2020) Assessment of Small Indian Mongoose Diet on St. John, USVI, Using Stable Isotope Analyses. Caribb Nat 1–14

Prange S, Gehrt SD, Wiggers EP (2004) Influences of Anthropogenic Resources on Raccoon (Procyon lotor) Movements and Spatial Distribution. J Mammal 85:483–490. https://doi.org/10.1644/1383946

Pyšek P, Jarošík V, Hulme PE, et al (2012) A global assessment of invasive plant impacts on resident species, communities and ecosystems: The interaction of impact measures, invading species' traits and environment. Glob Chang Biol 18:1725–1737. https://doi.org/10.1111/j.1365-2486.2011.02636.x

Rayfield B, Fortin MJ, Fall A (2011) Connectivity for conservation: A framework to classify network measures. Ecology 92:847–858. https://doi.org/10.1890/09-2190.1

Reding DM, Cushman S a., Gosselink TE, Clark WR (2013) Linking movement behavior and fine-scale genetic structure to model landscape connectivity for bobcats (Lynx rufus). Landsc Ecol 28:471–486. https://doi.org/10.1007/s10980-012-9844-y

Robertson BC, Gemmell NJ (2004) Defining eradication units to control invasive pests. J Appl Ecol 41:1042–1048. https://doi.org/10.1111/j.0021-8901.2004.00984.x

Rogers LL (1987) Effects of Food Supply and Kinship on Social Behavior, Movements, and Population Growth of Black Bears in Northeastern Minnesota. Wildl Monogr 97:3–72

Saunders DA, Hobbs RJ, Margules CR (1991) Biological Consequences of Ecosystem Fragmentation: A Review. Conserv. Biol. 5:18–32

Schoville SD, Dalongeville A, Viennois G, et al (2018) Preserving genetic connectivity in the European Alps protected area network. Biol Conserv 218:99–109. https://doi.org/10.1016/j.biocon.2017.12.017

Schwartz MK, Luikart G, McKelvey KS, Cushman SA (2009) Landscape genomics: A brief perspective [Chapter 9]. In: Cushman SA, Huettmann F (eds) Spatial Complexity, Informatics, and Wildlife Conservation. Springer, pp 165–174

Seaman GA, Randall JE (1962) The mongoose as a predator in the Virgin Islands. J Mammal 43:544–546

Simberloff D, Dayan T, Jones C, Ogura G (2000) Character displacement and release in the small Indian mongoose, Herpestes javanicus. Ecology 81:2086–2099

Thatte P, Chandramouli A, Tyagi A, et al (2020) Human footprint differentially impacts genetic connectivity of four wide-ranging mammals in a fragmented landscape. Divers Distrib 26:299–314. https://doi.org/10.1111/ddi.13022

Thulin C-G, Simberloff D, Barun A, et al (2006) Genetic divergence in the small Indian mongoose (Herpestes auropunctatus), a widely distributed invasive species. Mol Ecol 15:3947–56. https://doi.org/10.1111/j.1365-294X.2006.03084.x

Thulin C, Gyllenstrand N (2002) Highly variable microsatellite loci for studies of introduced populations of the small Indian mongoose (Herpestes javanicus). Mol Ecol … 453–455. https://doi.org/10.1046/j.1471-8278

Torres-Valcárcel Á, Harbor J, González-Avilés C, Torres-Valcárcel A (2014) Impacts of Urban Development on Precipitation in the Tropical Maritime Climate of Puerto Rico. Climate 2:47–77. https://doi.org/10.3390/cli2020047

Vilella FJ (1998) Biology of the Mongoose (Herpestes javanicus) in a Rain Forest of Puerto Rico. Biotropica 30:120–125. https://doi.org/10.1111/j.1744-7429.1998.tb00374.x

Vilella FJ, Zwank PJ (1993) Ecology of the small Indian mongoose in a coastal dry forest of Puerto Rico where sympatric with the Puerto Rican nightjar. Caribb J Sci 29:24–29

Waugh JD (2009) Trade and Invasive Species in the Caribbean: A Universe of Risk. Gland, Switzerland

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3:385–397

Zhan L, Paterson IG, Fraser BA, et al (2017) Megasat: Automated Inference of Microsatellite Genotypes From Sequence Data. Mol Ecol Resour 17:247–256. https://doi.org/10.1111/1755-0998.12561

Figures and Tables

Table 10. Pairwise R_{ST} comparisons between small Indian mongooses (*Herpestes auropunctatus*) sampled in three sites in Puerto Rico.

Population	Coast	Mountain	Vieques
Coast	-	0.396	0.011
Mountain	0	-	0.083
Vieques	0.096	0.081	-

Table 11. Results from the causal modelling framework. The human footprint (bold) hypothesis is supported by the modelling approach. R is the Mantel correlation coefficient based on Spearman's rank correlation.

Variables	Simple Mantel		Partial Mantel			
	MODEL 1 - Genetic x Landscape		MODEL 2 - Genetic ~ Landscape | Euclidean distance		MODEL 3 - Genetic ~ Euclidean distance | Landscape	
	R	P < 0.05	R	P < 0.05	R	P < 0.05
Human Footprint	**-0.2076**	**0.028**	**-0.207**	**0.022**	**0.065**	**0.150**
Habitat Distribution	0.0125	0.406	0.012	0.4334	0.066	0.158
Connectivity	0.108	0.113	0.101	0.137	0.054	0.195
Barrier	-0.042	0.669	-0.043	0.618	0.067	0.140
Elevation	-0.101	0.861	-0.087	0.800	0.042	0.243
Euclidean Distance	-0.060	0.174				

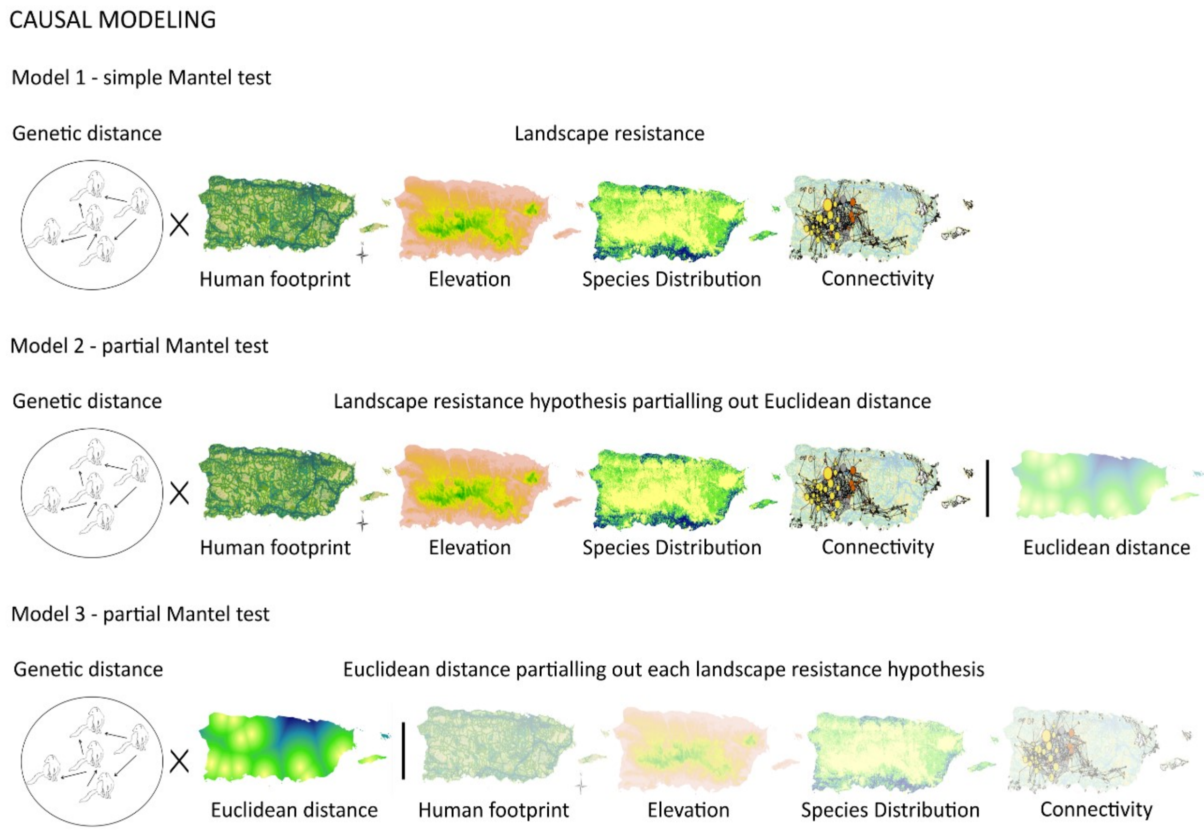


Figure 15. The causal modelling approach tested in our study. We used simple mantel tests to test pair-wise genetic distance and various isolation by landscape resistance hypotheses (Model 1). Partial mantel tests tested Isolation by landscape resistance partialling out the Euclidean distance between the sampled individuals (Model 2). Model 3 are partial mantel tests examining isolation by distance partialling out landscape-resistance.

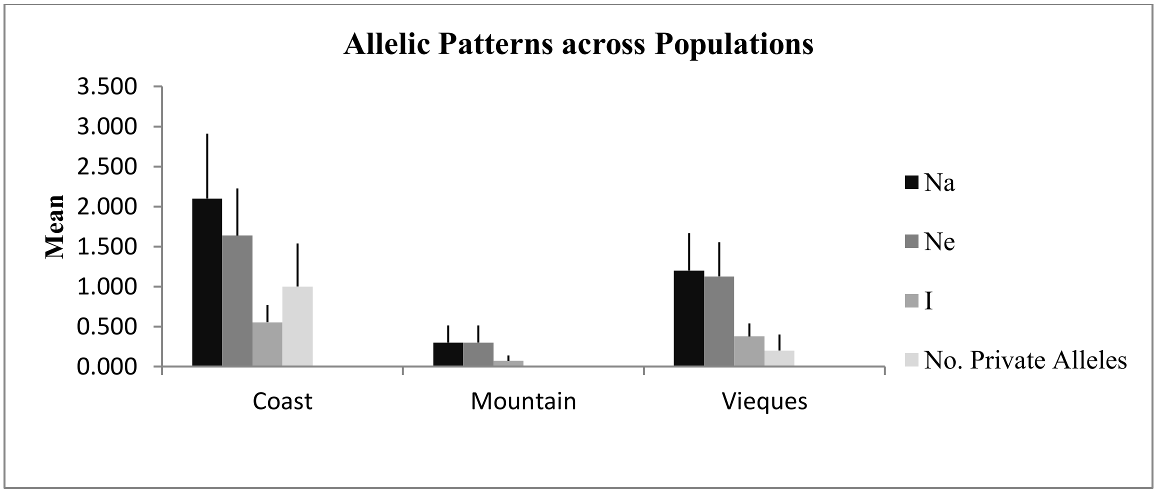


Figure 16. Allelic patterns across **sites**. Na = Number of different alleles. Ne = effective number of alleles. I = Shannon's Information Index. No. of private alleles **(PA)** = Number of alleles unique to a single population.

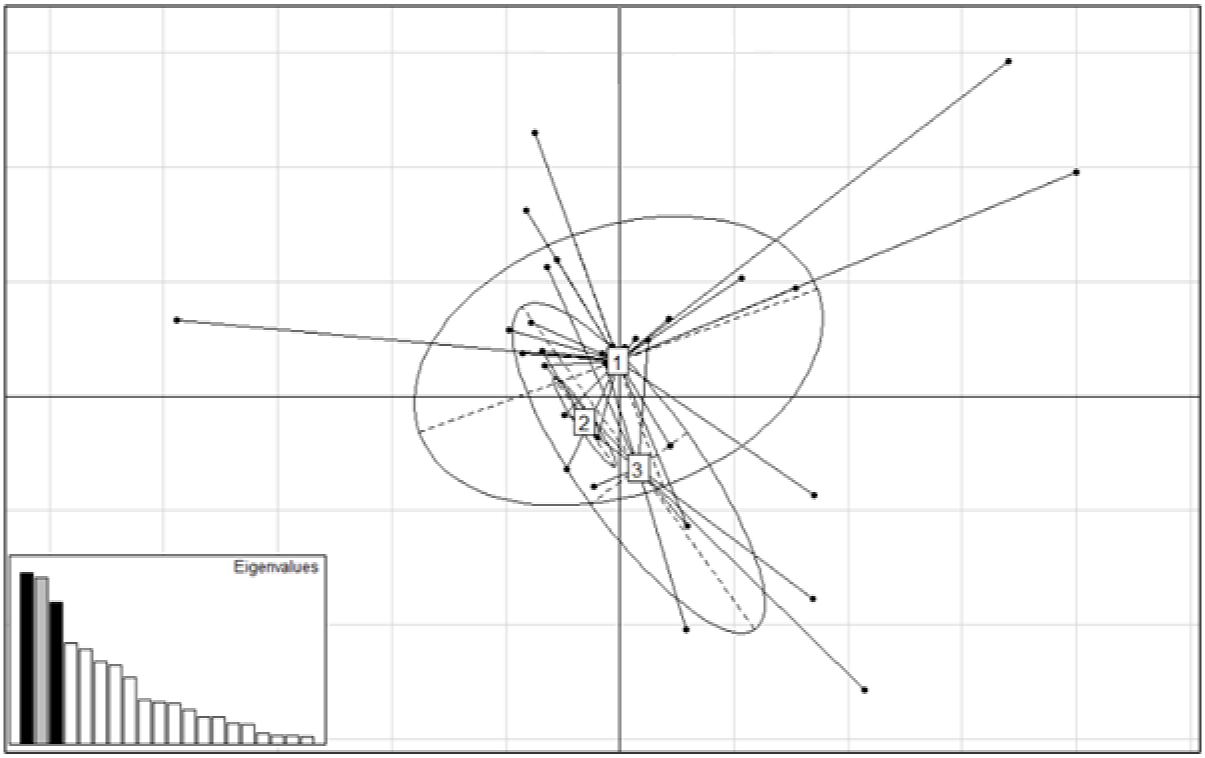


Figure 17. Principal Coordinate Analysis (PCoA) showing the similarity relations among of the small Indian mongoose (*Herpestes auropunctatus*) in Puerto Rico. 1 = Coast, 2 = Mountain, 3 = Vieques. Individuals closer to one another are more similar than those ordinated further away. The eigenvalues are the amount of variance explained in each dimension.

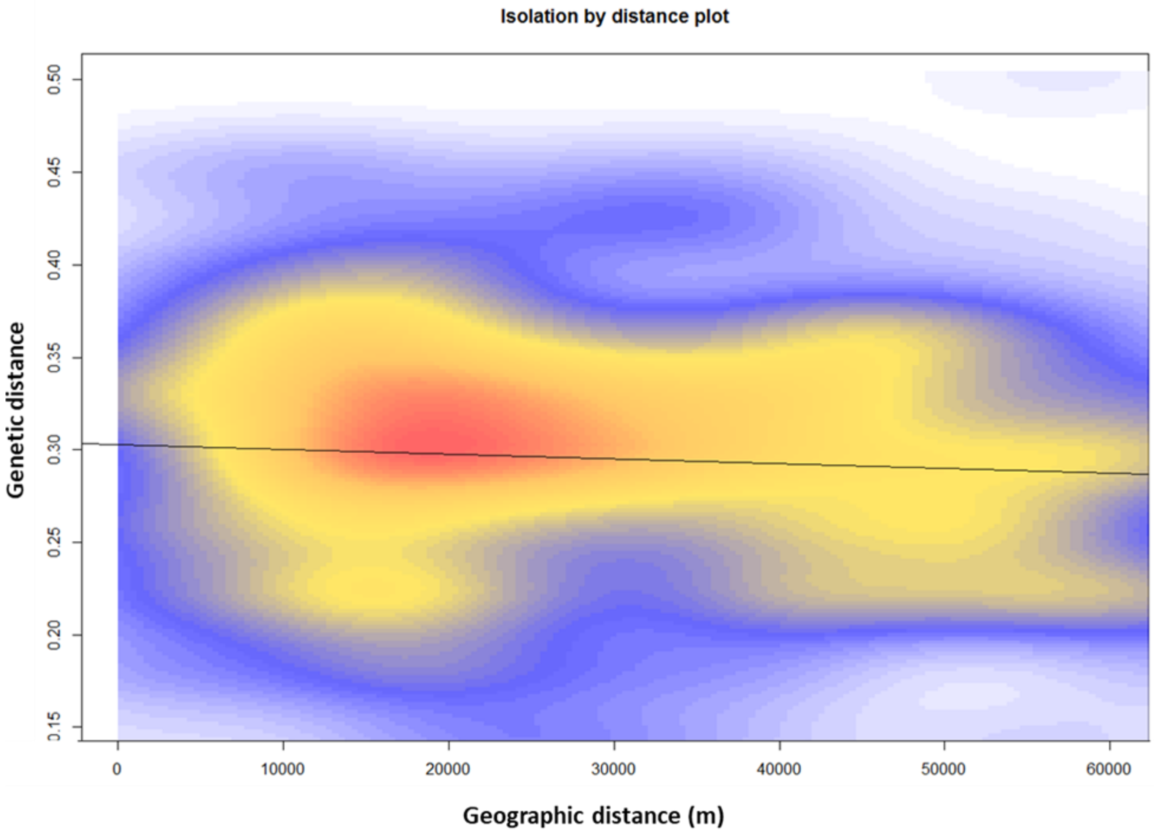


Figure 18. Scatterplot of geographic vs. genetic distance and 2-dimensional kernel density estimation to assess local density of points. A gradient from low to high density is represented by a blue-to-red color palette. Included are the estimated correlation coefficient with a simulated p-value based on 1000 permutations to test for IBD. all correlations were negative, which indicates more differentiation within sites than between pairs of sites