## Monito Gecko - supplemental occupancy and abundance analysis

Sphaerodactylus micropithecus (Schwartz 1977) (Sauria: Gekkonidae) is a small gecko (adult snout-vent length: 17-36 mm) that is grey with dark spotting. The geographic distribution of the species is limited to 14.5 ha Monito Island, a small limestone island 5 km NW of Mona Island in the Mona Passage which is within a protected area administered by the Commonwealth of Puerto Rico Department of Natural and Environmental Resources. Based on its single-island distribution, apparent decline from rat predation, past habitat destruction from military activities, and the risk of stochastic events on small island populations, the species was listed as Endangered by the U.S. Fish and Wildlife Service in 1982, by the Commonwealth of Puerto Rico in 1985, and by the International Union for the Conservation of Nature in 1996 (Tolson 1996).

A Recovery Plan was established in 1986 by the U.S. Fish and Wildlife Service. One of the main threats identified was depredation of geckos by invasive rats (Rattus rattus), and in 1999 the Puerto Rico DNER carried out a successful rat eradication from Monito Island (Garcia 2002). However, until this study there had been no standardized monitoring of the gecko population, and it was not known if the gecko had benefited from the eradication of rats.

Surveys implemented in 1982 and 1984, intended to establish baseline data, were conducted during the day and indicated that the species’ distribution was mostly limited to two small regions of the island (Dodd and Ortiz 1983 in Ortiz 1983; Ortiz 1983, Hammerson 1984). However, during a visit to Monito Island in May 2014 by biologists from the U.S. Fish and Wildlife Service and Puerto Rico DNER, geckos were detected at several locations around the island and included a count of 23 geckos in one night (USFWS 2016). The Service estimates that geckos are detectable 30% of the time based on congeneric species around the Caribbean.

A multi-partner effort to establish baseline data for the species, its abundance and distribution largely unknown, was undertaken in May 2016 using nocturnal surveys.

First, we needed to choose how many sites to visit. We kept detection stable at 0.3 for each scenario, and we assumed that the sites were 50% occupied.

Scenario 1: Simulate data using the multinomial-Poisson model with a repeated constant-interval removal design for 40 sites for 4 primary periods - 2 days, 2 nights

# Simulate data using the multinomial-Poisson model with a  
# repeated constant-interval removal design. for FOUR primary periods - 2 days, 2 nights  
n <- 40 # number of sites  
T <- 4 # number of primary periods  
J <- 2# number of secondary periods  
lam <- 3  
phi <- 0.5  
p <- 0.3

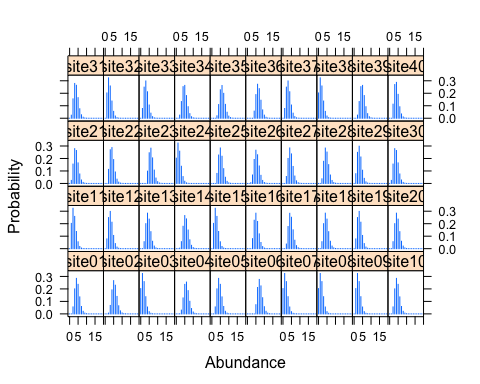
Create a dataset for simulation

head(umf1)

## Data frame representation of unmarkedFrame object.  
## y.1 y.2 y.3 y.4 y.5 y.6 y.7 y.8  
## 1 2 0 0 1 0 1 0 0  
## 2 0 2 1 0 1 0 2 0  
## 3 0 0 0 0 0 0 0 0  
## 4 1 1 0 1 2 0 2 0  
## 5 1 1 1 0 0 0 0 1  
## 6 2 3 0 0 2 0 0 1  
## 7 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0  
## 9 0 0 0 0 0 0 1 0  
## 10 0 0 1 1 1 0 0 1

Run null model

##   
## Call:  
## gmultmix(lambdaformula = ~1, phiformula = ~1, pformula = ~1,   
## data = umf1, K = 50)  
##   
## Abundance:  
## Estimate SE z P(>|z|)  
## 1.41 0.384 3.68 0.000232  
##   
## Availability:  
## Estimate SE z P(>|z|)  
## 0.0587 1.29 0.0453 0.964  
##   
## Detection:  
## Estimate SE z P(>|z|)  
## -1.2 0.744 -1.62 0.106  
##   
## AIC: 505.3817

Figure 1. Simulation for results of 8 surveys across 40 sites using empirical Bayes estimates 

Scenario 2 - simulate data using the multinomial-Poisson model with a repeated constant-interval removal design for 40 sites for 4 surveys over 2 nights

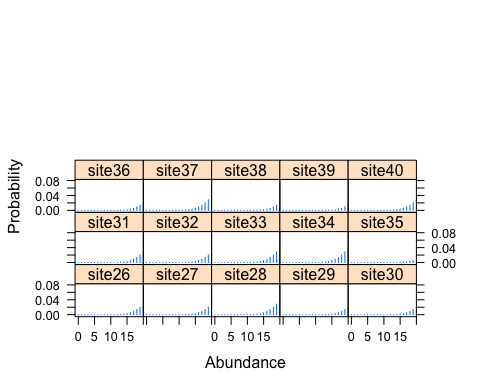
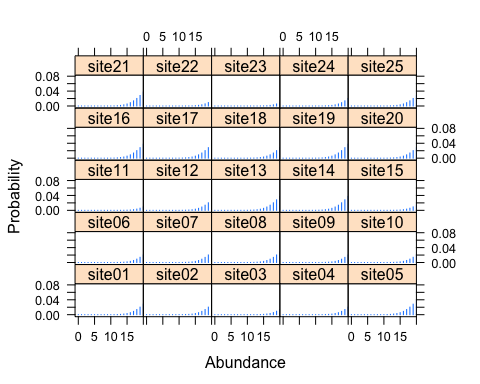
Run second sampling model scenario

head(y.ijt2) #simulated dataset

## [,1] [,2] [,3] [,4]  
## [1,] 0 1 0 0  
## [2,] 0 0 0 1  
## [3,] 1 1 1 0  
## [4,] 0 0 1 1  
## [5,] 0 0 0 0  
## [6,] 1 0 1 0

### Run model with null covariates  
(m2 <- gmultmix(~1, ~1, ~1, data=umf2, K=50))

##   
## Call:  
## gmultmix(lambdaformula = ~1, phiformula = ~1, pformula = ~1,   
## data = umf2, K = 50)  
##   
## Abundance:  
## Estimate SE z P(>|z|)  
## 3.33 0.897 3.71 0.00021  
##   
## Availability:  
## Estimate SE z P(>|z|)  
## 3.58 32.2 0.111 0.911  
##   
## Detection:  
## Estimate SE z P(>|z|)  
## -4.35 1.26 -3.45 0.000566  
##   
## AIC: 233.5455

Figure 2. Simulation for results of 4 surveys across 40 sites using empirical Bayes estimates 

## Fit statistics

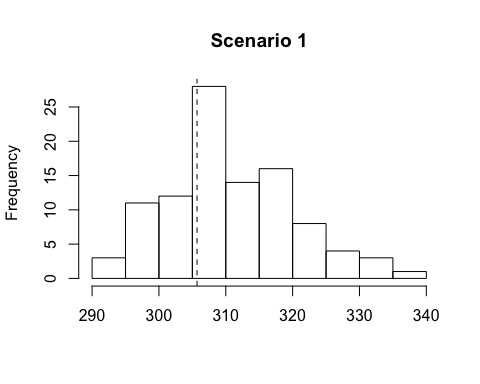
## t0 = 305.6991   
## Running in parallel. Bootstrapped statistics not reported.

##   
## Call:   
## parboot(object = topModel1, statistic = chisq, nsim = 100, report = 1)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 306 -5.01 9.34 0.713  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 292 295 305 309 318 330 337  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

## t0 = 150.3426   
## Running in parallel. Bootstrapped statistics not reported.

##   
## Call:   
## parboot(object = topModel2, statistic = chisq, nsim = 100, report = 1)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 150 0.43 3.72 0.515  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 140 143 147 150 152 158 162  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

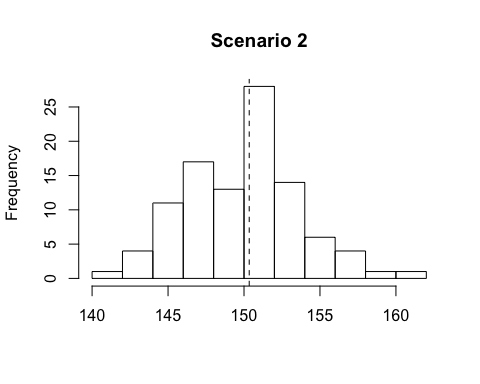
##########################################  
#### Check model GOF  
##########################################  
  
plot(pb.plot1, main="Scenario 1")

 Figure 3. Model fit results of 8 surveys across 40 sites

pb.plot1

##   
## Call:   
## parboot(object = topModel1, statistic = chisq, nsim = 100, report = 1)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 306 -5.01 9.34 0.713  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 292 295 305 309 318 330 337  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

plot(pb.plot2, main= "Scenario 2")

 Figure 4. Model fit results of 4 surveys across 40 sites

pb.plot2

##   
## Call:   
## parboot(object = topModel2, statistic = chisq, nsim = 100, report = 1)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 150 0.43 3.72 0.515  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 140 143 147 150 152 158 162  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

## Monito Gecko field surveys

Because the simulated models using both 4 and 8 surveys suggest that enough data will result to fit occupancy models using a removal design, we choose the scenario with the least amount of effort due to the difficult logistics of working on Monito Island.

Survey site selection: A total of 50 survey points were computed using a geographically random method in ArcGIS. Of these, a total of 40 survey points (10 per quadrant) were located on the island (on average ± 3 m) using a handheld GPS unit pre-programmed with the geographic coordinates of each survey point. Each 2 person team was provided with a map and GPS coordinates to all of the survey points.

Survey protocols: We organized four 2-person teams (8 persons) to complete survyes across sites. Each team was assigned 10 survey sites and was responsible for marking the survey plots, collecting habitat data, and conducting gecko surveys. Monito Island was divided into four quadrants (southwest, northwest, northeast, southeast) each of which was randomly assigned to a team.

See text for more details.

Once plots were marked, each team recorded habitat covariates in each plot: % shrub cover (woody plants under 2 m tall), % herbaceous cover (grasses and sedges), % ground cover (exposed rock or soil cover), % leaf litter, and % canopy cover (trees taller than approximately 1.8 m (6 ft)).

Because we were using the habitat variables for predictive models, we tested for collinearity among each covariate across the sites. Ground cover was collinear with shrub cover (r = 0.98), and so only shrub cover was used. Add ? here

The survey plots comprised 35.73 % leaf litter (Range: 5 – 90 %), 1% herbaceous material (Range: 0 – 15 %), 8.01% canopy (Range: 0 – 60 %), 35.75% ground cover (Range: 5 – 90 %), and 62.55 % shrub cover (Range: 9 – 95 %). Because ground cover and shrub cover were highly correlated, we only used shrub cover in the analysis to reduce redundancy.

head(covMonito)

## shrub herb ground leaf canopy  
## 1 90 0.0 10 0 0  
## 2 70 0.0 30 0 0  
## 3 85 0.0 15 1 0  
## 4 90 0.0 10 50 5  
## 5 50 0.0 50 25 0  
## 6 40 0.5 60 0 0

cor(covMonito)

## shrub herb ground leaf canopy  
## shrub 1.0000000000 -0.12573753 -0.98799059 0.06151453 0.0009461528  
## herb -0.1257375257 1.00000000 0.00731348 -0.13147996 -0.0700742136  
## ground -0.9879905860 0.00731348 1.00000000 -0.02929358 0.0191948601  
## leaf 0.0615145255 -0.13147996 -0.02929358 1.00000000 0.3126030292  
## canopy 0.0009461528 -0.07007421 0.01919486 0.31260303 1.0000000000

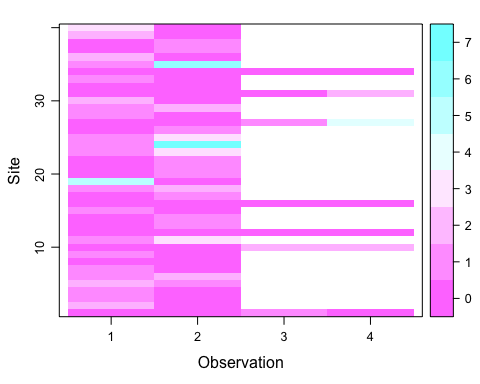
summary(covMonito)

## shrub herb ground leaf   
## Min. : 9.00 Min. : 0.00 Min. : 5.00 Min. : 0.000   
## 1st Qu.:50.00 1st Qu.: 0.00 1st Qu.:19.00 1st Qu.: 0.000   
## Median :62.50 Median : 0.00 Median :30.00 Median : 0.000   
## Mean :62.55 Mean : 0.95 Mean :35.73 Mean : 8.075   
## 3rd Qu.:80.00 3rd Qu.: 1.00 3rd Qu.:49.25 3rd Qu.: 4.000   
## Max. :95.00 Max. :15.00 Max. :90.00 Max. :60.000   
## canopy   
## Min. : 0.000   
## 1st Qu.: 0.000   
## Median : 0.000   
## Mean : 2.675   
## 3rd Qu.: 1.250   
## Max. :50.000

The data was formatted for analysis using each plot sample as a separate survey history.

summary(umf) #summary of observed dataset

## unmarkedFrame Object  
##   
## 40 sites  
## Maximum number of observations per site: 4   
## Mean number of observations per site: 2.35   
## Sites with at least one detection: 35   
##   
## Tabulation of y observations:  
## 0 1 2 3 4 5 6 7 <NA>   
## 48 27 11 4 1 1 1 1 66   
##   
## Site-level covariates:  
## shrub herb ground leaf   
## Min. : 9.00 Min. : 0.00 Min. : 5.00 Min. : 0.000   
## 1st Qu.:50.00 1st Qu.: 0.00 1st Qu.:19.00 1st Qu.: 0.000   
## Median :62.50 Median : 0.00 Median :30.00 Median : 0.000   
## Mean :62.55 Mean : 0.95 Mean :35.73 Mean : 8.075   
## 3rd Qu.:80.00 3rd Qu.: 1.00 3rd Qu.:49.25 3rd Qu.: 4.000   
## Max. :95.00 Max. :15.00 Max. :90.00 Max. :60.000   
## canopy   
## Min. : 0.000   
## 1st Qu.: 0.000   
## Median : 0.000   
## Mean : 2.675   
## 3rd Qu.: 1.250   
## Max. :50.000

Figure 5. Results of the observed data during 2016 surveys on Monito Island. The scale marks surveys where observations ranged from zeo geckos to seven gecko individuals. A total of four surveys occurred at each site. The first two surveys occured on night 1 by independent observers. The second two surveys occured on night 2 by two different observers. If geckos were observed during either survey on night 1, the site was removed from sampling on night 2. 

The R package ‘unmarked’ version 0.10.2 provides methods to estimate site occupancy, abundance, and density of animals that cannot be detected with certainty (Fiske and Chandler, 2011). The occupancy of the island by Monito Geckos was fit with a model described by Mackenzie et al. (2002) and implemented with the function ‘occu’ where the occupancy and detection of the observed geckos is described by a Bernoulli process. We assumed all sites were closed based on the life history and activity of the Monito Gecko.

The abundance (N\_ of the gecko on Monito Island was estimated using a N-mixture model with a negative binomial distribution for overdispersed data (α). Because the species is cryptic, surveys were likely to result in many absences, and the negative binomial distribution represents a zero-inflated dataset (Johnson et al., 1992). Multiple models were fit with each combination of covariates was fit using the function ‘pcount’ (Royle 2004; Kéry et al. 2005) according to the equation) where the mean of the negative binomial distribution is and and the covariates describing the abundance and detection were transformed with a logit link.

Occupancy of the geckos on Monito Island was 27.8 % (CI: 11.3 – 68.6 %). The mean number of geckos per 400 m2 plot from the abundance model was 73.3 geckos (Range: 1 – 101; Table 5). The ‘true’ abundance from the model per plot indicates a total of 1,112 geckos present within the 40 surveyed plots in Monito Island (95% CI: 362 – 2,281). Interpolated across the entire island (14.5 ha), Monito Island hosts approximately 7,661 geckos (50% CI: 5,344 – 10,590). The interpolated zones represent the approximate number of geckos at any randomly selected 30 m2 area. See text for more details.

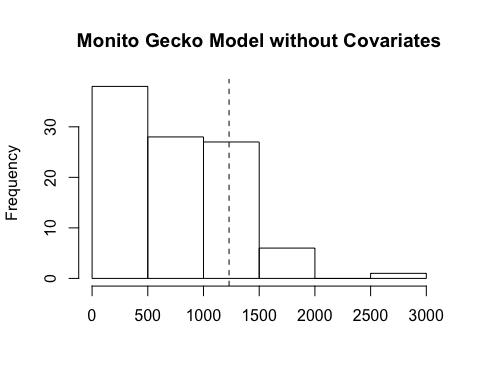
## Data frame representation of unmarkedFrame object.  
## y.1 y.2 y.3 y.4 shrub herb ground leaf canopy  
## 1 0 0 1 0 90 0.0 10 0 0  
## 2 2 0 NA NA 70 0.0 30 0 0  
## 3 1 0 NA NA 85 0.0 15 1 0  
## 4 1 0 NA NA 90 0.0 10 50 5  
## 5 2 1 NA NA 50 0.0 50 25 0  
## 6 1 2 NA NA 40 0.5 60 0 0  
## 7 1 0 NA NA 80 1.0 19 40 0  
## 8 0 0 NA NA 25 0.5 65 0 0  
## 9 1 0 NA NA 60 15.0 25 0 1  
## 10 0 0 2 2 55 0.5 45 0 0  
## 11 1 3 NA NA 80 0.0 15 1 0  
## 12 0 0 0 0 70 0.0 30 0 2  
## 13 0 1 NA NA 90 0.0 10 1 0  
## 14 0 1 NA NA 50 1.0 49 20 0  
## 15 1 0 NA NA 80 0.0 20 2 10  
## 16 0 0 0 0 85 0.0 15 0 0  
## 17 0 1 NA NA 60 0.0 40 60 0  
## 18 1 2 NA NA 60 0.0 40 0 6  
## 19 5 0 NA NA 80 1.0 19 0 0  
## 20 0 1 NA NA 94 1.0 5 0 0  
## 21 0 1 NA NA 75 1.0 24 0 0  
## 22 0 1 NA NA 15 1.0 80 0 0  
## 23 1 3 NA NA 50 0.0 50 0 0  
## 24 1 7 NA NA 55 0.0 45 0 0  
## 25 1 3 NA NA 40 0.0 60 35 20  
## 26 0 1 NA NA 80 0.0 20 0 0  
## 27 0 0 1 4 9 1.0 90 0 0  
## 28 1 0 NA NA 70 0.0 30 30 50  
## 29 1 2 NA NA 50 1.0 49 0 2  
## 30 2 0 NA NA 50 10.0 40 0 2  
## 31 0 0 0 2 60 0.5 40 0 0  
## 32 0 0 NA NA 20 1.0 75 0 0  
## 33 1 0 NA NA 80 0.0 20 10 0  
## 34 0 0 0 0 95 0.0 5 0 0  
## 35 1 6 NA NA 30 0.0 70 0 5  
## 36 2 0 NA NA 60 0.0 40 1 4  
## 37 0 1 NA NA 85 0.0 14 1 0  
## 38 0 1 NA NA 65 1.0 25 1 0  
## 39 2 0 NA NA 39 1.0 60 25 0  
## 40 3 0 NA NA 80 0.0 20 20 0

##   
## Call:  
## pcount(formula = ~1 ~ 1, data = umf, K = 120, mixture = "NB")  
##   
## Abundance:  
## Estimate SE z P(>|z|)  
## 3.43 0.48 7.13 9.88e-13  
##   
## Detection:  
## Estimate SE z P(>|z|)  
## -3.47 0.503 -6.9 5.04e-12  
##   
## Dispersion:  
## Estimate SE z P(>|z|)  
## 1.03 0.514 2.01 0.0448  
##   
## AIC: 259.8191

##   
## Call:  
## pcount(formula = ~leaf + herb ~ 1 + canopy + leaf + herb, data = umf,   
## mixture = "NB")  
##   
## Abundance:  
## Estimate SE z P(>|z|)  
## (Intercept) 3.29319 0.5391 6.1089 1.00e-09  
## canopy 0.00564 0.0196 0.2879 7.73e-01  
## leaf 0.00314 0.0325 0.0967 9.23e-01  
## herb 0.02319 0.1265 0.1833 8.55e-01  
##   
## Detection:  
## Estimate SE z P(>|z|)  
## (Intercept) -3.28385 0.5679 -5.783 7.35e-09  
## leaf -0.00837 0.0333 -0.251 8.02e-01  
## herb -0.05625 0.1340 -0.420 6.75e-01  
##   
## Dispersion:  
## Estimate SE z P(>|z|)  
## 1.05 0.515 2.03 0.042  
##   
## AIC: 269.4217

## Estimate lowerCI upperCI  
## psi 0.96849081 11.99053383 78.79127930  
## p 0.03016975 0.01148223 0.07690533

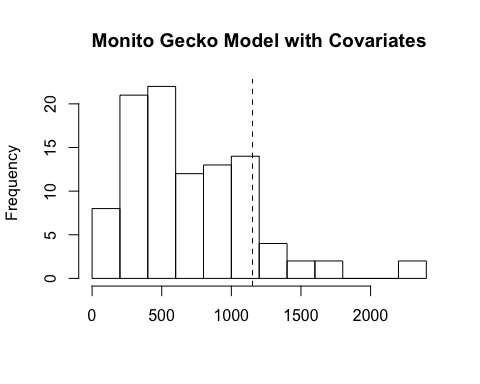
To test the fit of the best ranked models the data well, a parametric bootstrap sampling from the function ‘parboot’ was used. The function ‘parboot’ uses the model object previously fitted in ‘pcount’ to refit the model iteratively and test for goodness-of-fit (GOF). We specified a Bayesian posterior test of the latent abundance which allows us to compare the modeled abundance to distributions simulated from the fit model. The bootstrap P value for the models was run 1,000 times; a non-significant value (P > 0.05) suggests the model estimate has an appropriate GOF. The final model was examined to find the abundance (N) and detection for each site. Using AIC, we assumed that the most simple model will have a mathematically convergent solution.

Figure 6. Model fit results without covariates 

##   
## Call:   
## parboot(object = MonitoNull, statistic = Nhat, nsim = 100, report = 5)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 1229 459 532 0.198  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 80 106 258 765 1167 1660 2957  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

Figure 7. Model fit results with covariates

plot(P.Cov.Mon, main= "Monito Gecko Model with Covariates")

 Figure 7. Model fit results with covariates

##   
## Call:   
## parboot(object = MonitoCov, statistic = Nhat, nsim = 100, report = 5)  
##   
## Parametric Bootstrap Statistics:  
## t0 mean(t0 - t\_B) StdDev(t0 - t\_B) Pr(t\_B > t0)  
## 1 1152 465 442 0.099  
##   
## t\_B quantiles:  
## 0% 2.5% 25% 50% 75% 97.5% 100%  
## [1,] 110 158 364 581 943 1687 2320  
##   
## t0 = Original statistic compuated from data  
## t\_B = Vector of bootstrap samples

The final model estimates for each site were used as a dataset to fit a geographic linear interpolation across the 14.5 ha island using a 30 m2 grid based on the pixel size of the GIS data. The overall method was developed for use with terrestrial low-detection species and previously peer-reviewed (Angeli et al. 2018, Fitzgerald et al. 2015). We assumed that the distribution and abundance between sites was representated accurately by our surveys. See text for associated figures.

The Monito gecko is distributed across at least one quarter of Monito Island based on the final occupancy results. The modeled population estimates reveal 7,661 geckos (50% CI: 5,344 – 10,590) across the entire island (Table 6). The density of 19 geckos per sq. m of the species is in fact lower than the 134 – 528 geckos per sq. m observed in removal plots of the congeneric species Sphaerodactylus macrolepis in the Virgin Islands (Rodda et al., 2001). Habitat covariates recorded could not explain the differences in gecko detection or abundance across plots.

Future efforts could incorporate new and different covariates to better measure the gecko’s habitat preference, and in particular the micro-habitats influenced by substrate temperature, moisture, pH, soil type, and topography, invertebrate prey diversity and abundance, and seabird presence. To enable direct comparison to the baseline established here and to maximize detectability, we recommend that this sampling protocol is repeated for future surveys, and conducted at the same time of year and during the night when geckos were easy to detect.