SUPPLEMENTAL INFORMATION

DUPED BY FALSE-NEGATIVE ERRORS: A CASE STUDY WITH THE MEDITERRANEAN GECKO

CALEB A. ALDRIDGE, THOMAS P. ROGERS, W. TAYLOR BANKS, AND SCOTT A. RUSH

The following material is provided by the authors and was not subjected to peer review or editing by *Herpetological Conservation and Biology*.

R code.—R code to estimate the number of sampling occasions needed in occupancy model. Estimates of ψ and *p* to be estimated from detection histories in Gomez-Zlatar *et al.* (2006). On page 429 in Gomez-Zlatar *et al.* (2006): "Of the 160 walls that we surveyed, 26 walls had at least one gecko on the first visit only, 20 walls had geckos on just the second visit, and 39 walls had geckos during both sampling visits. Seventy-five walls had no geckos on either visit."

```
mat <- matrix(c(rep(c(1, 0), 26),  # 26 detected on 1st visit only
                rep(c(0, 1), 20),  # 20 detected on 2nd visit only
                rep(c(1, 1), 39),  # 39 detected on both visits
                rep(c(0, 0), 75)), # 75 with no detection
              byrow = T, ncol = 2)
# Check number of sites is equal to that reported in Gomez-Zlatar et al.
(2006).
nrow(mat)

# Load unmarked package.
library(unmarked)

# Place detection histories into specially formatted unmarked data frame.
dat <- unmarkedFrameOccu(y = mat)
# Have a look at the data summary.
summary(dat)
```

```
# Have a closer look at the detection history.

detHist(dat)

# Estimate constant Psi & p using methods of MacKenzie 2002, 2006.

fm <- occu(~ 1 ~ 1, data = dat)

# Back transform log-odds to probability.

(p <- backTransform(fm, type = 'det'))  # p

(psi <- backTransform(fm, type = 'state'))  # psi

# Formula used to estimate number of sampling occasions MacKenzie and Royle
(2005).

se <- 0.1  # standard error of estimate (corresponds with confidence level)

ceiling(log(se) / log(1 - p@estimate))

# Reference standard design tables in MacKenzie and Royle (2005) too!
```

Gomez-Zlatar, P., M.P. Moulton, and R. Franz. 2006. Microhabitat use by *Hemidactylus turcicus* (Mediterranean Geckos) in North Central Florida. Southeastern Naturalist 5:425–434.

MacKenzie, D.I. and J.A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Other model comparisons.—It is common for studies to account for imperfect detection by adjusting estimates of ψ_a by multiplying with an estimate of apparent detection (p_a; Chen et al 2009; Gorosito et al. 2016). We used a typical modeling approach for estimating p_a and compared the results to p estimated in our occupancy model. To estimate p_a it is assumed that the target organism is present at sites but may not be encountered due to covariates and unknown factors. To meet this assumption, we subset our data from all three sampling occasions to sites only where the Mediterranean Gecko was detected on at least one occasion. We then fit detection–non-detection data using a mixed-effects logistic regression with a random covariate for site to account for among-site variability and non-independence of detections at the same site (Chen et al. 2009; McIntyre et al. 2020). Model assumptions for the mixed-effects logistic regression were tested using goodness-of-fit statistics and permutation tests (n = 4999) in the *DHARMa* package (Hartig 2019). We used the *rsq* package (Zhang, 2017) to calculate the r^2 of the fixed effects. We compared p_a and p and the effects of covariates on p_a and p in the same way as we did for ψ_a and ψ above.

Tests of uniformity (Kolmogorov–Smirnov D = 0.12, P-value = 0.44) and dispersion (observed dispersion = 1.16, P-value = 0.36) did not indicate violation of assumptions in the mixed-effects logistic regression, so we proceeded with model interpretation. No covariates were statistically significant at the α = 0.05 level in the mixed-effects logistic regression (Table S1). Apparent detection (p_a) was estimated to be 0.69 (CIs = [0.20, 0.95]) when all variables were held at their mean value (Figure S1). The fixed effects r^2 for the mixed-effects logistic regression model was 0.44.

Supplemental Table 1. Mixed-effects logistic regression results for apparent detection probability (p_a). Coefficient estimates given in log-odds scale and probability scale (in parentheses). P-values <0.05 indicate statistically significant results.

Covariate (p_a)	**Coefficient[a]**	**S.E.**	***t*-value**	***P*-value**
Intercept	0.30 (0.57)	1.48	0.20	0.84
Well-defined eaves	2.88 (0.95)	2.48	1.16	0.25
Presence of artificial lighting	0.86 (0.70)	1.57	0.55	0.58
Minutes after sunset	1.74 (0.85)	1.11	1.57	0.12
Pedestrian traffic (ped. minutes^{-1})	-2.89 (0.05)	1.96	-1.47	0.14

[a]Numbers in parentheses are in probability scale

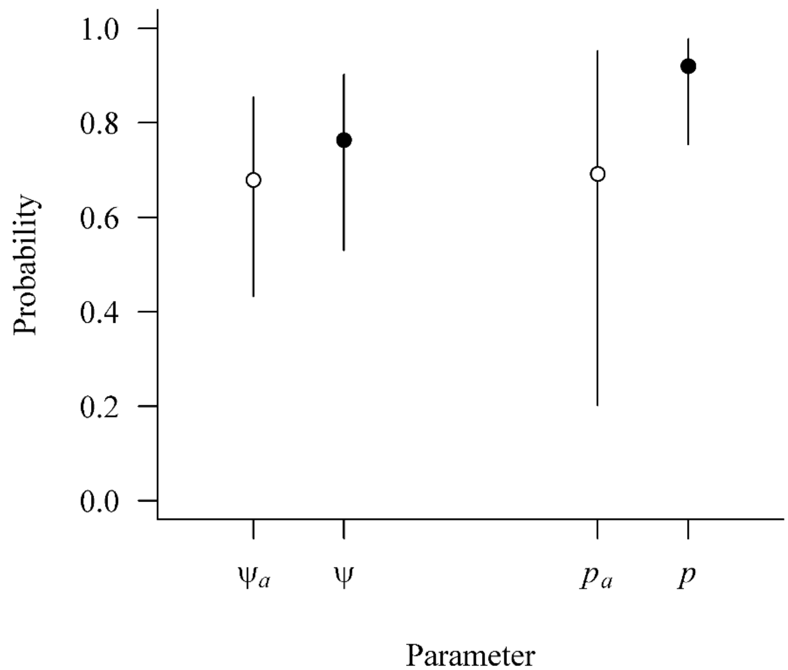
One approach to modeling ψ_a might be to aggregate data from multiple occasions to meet model assumptions (e.g., pseudo-replication). We modeled aggregated data and present results in (Table S2; Figure S1). One obvious drawback to this approach is loss of information and wasted effort—there are models to accommodate nested model structures (i.e., hierarchical models).

Supplemental Table 2. Logistic regression results for apparent occupancy (ψ_a) using aggregated data from all three sampling occasions. Coefficient estimates given in log-odds scale and probability scale (in parentheses). P-values <0.05, indicated in bold, are considered statistically significant.

Covariate	**Coefficient[a]**	**S.E.**	***t*-value**	***P*-value**
Intercept	-0.21 (0.45)	1.42	-0.15	0.88
Well-defined eaves	2.82 (0.94)	1.61	1.75	**0.08**
Presence of artificial lighting	0.17 (0.54)	1.44	0.12	0.91
Minutes after sunset	-2.67 (0.07)	1.43	-1.86	**0.06**
Pedestrian traffic (ped. minutes^{-1})	0.38 (0.59)	1.42	0.27	0.79

Note: r^2 = 0.53; ψ_a = 0.84 (0.57, 0.96) when covariates held at their means; Kolmogorov-Smirnov test *P*-value = 0.79, Uniformity test *P*-value = 0.81

[a]Number in parentheses are in probability scale



Supplemental Figure 1. Estimates of occupancy (ψ_a; ψ) and detection probabilities (p_a; p). Open circles represent estimates of apparent occupancy (ψ_a) from logistic regression and apparent detection probability (p_a) mixed-effects logistic regression. Occupancy model estimates of unbiased occupancy (ψ) and detection probability (p) indicated by filled circles. Error bars represent 95% CIs.

Chen, G., M. Kéry, J. Zhang, and K. Ma. 2009. Factors affecting detection probability in plant distribution studies. Journal of Ecology 97:1383–1389.

Gorosito, I.L., M.M. Bermúdez, R.J. Douglass. And M. Busch. 2016. Evaluation of statistical methods and sampling designs for the assessment of microhabitat selection based on point data. Methods in Ecology and Evolution 7:1316–1324.

Hartig, F. 2019. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.2.4. https://CRAN.R-project.org/package=DHARMa

McIntyre, T., T.L. Majelantle, D.J. Slip, and R.G. Harcourt. 2020. Quantifying imperfect camera-trap detection probabilities: implications for density modelling. Wildlife Research 47:177–185.

Zhang, D. 2017. A coefficient of determination for generalized linear models. The American Statistician 71: 310–316.

Modified sampling design.—We used the results of our occupancy model to modify our sampling design and test it at urban centers in two unsampled towns nearby, Columbus and West Point in Lowndes and Clay Counties Mississippi, respectively. First, we substituted temporal replication (sampling occasions in time) with spatial replication (sampling occasions in space), defining the town as the site and building walls as the sampling occasions (Kéry and Royle 2008, 2016; Guillera-Arroita et al. 2010). This approach is advantageous when the goal is to rapidly assess occupancy of the target species over a large spatial area, as temporal replication can be logistically and financially prohibitive. In addition, the spatial resolution of the Mediterranean Gecko's distribution has traditionally been reported at the county level (see Nelson and Carey 1993; Meshaka et al. 2006; Lee 2008; White and Husak 2015; White et al. 2017; see also Geographic Distribution notes in *Herpetological Review*). Next, we calculated the minimal number of occasions needed to be ≥99% confident that a site was unoccupied when sampling under good conditions (Kéry 2002). This coincides with the number of replicates need to detect the target organism in at least one replicate when a site is truly occupied. We defined good conditions as buildings with well-defined eaves in areas with little to no pedestrian traffic and waiting to begin surveys until approx. an hour after sunset. As a final point of comparison, we calculated the travel necessary to reliably sample all these urban areas under temporal (one site surveyed *N* times) and spatial replication (*N* sites surveyed on one time), assuming a route from Starkville to Columbus (41.8 km; www.maps.google.com), from Columbus to West Point (29.8 km), and then returning to Starkville from West point (33.6 km).

Our modified survey protocol targeted buildings with well-defined eaves in areas with little to no pedestrian traffic and did not begin until approx. an hour after sunset. The number of spatial replicates needed to detect Mediterranean Gecko at new sites, those not yet surveyed, was estimated to be four (Table S3). We detected Mediterranean Gecko in both Columbus (75% of subsites; Rogers *et al.* 2020) and West Point (50% of subsites; Banks *et al.* 2020). The difference in travel distance when using temporal and spatial replication (four replicates) was 315.6 km (approx. 4.5 hr of travel time).

Supplemental Table 3. Estimates of detection probability and the number of replicate samples needed to be 90% and 99% confident that Mediterranean Gecko would be detected on at least one occasion and, or absent from a site.

Scenario	**Covariates**	**Values**	***p* [95% CIs]**	***n* 90%**	***n* 99%**
Worst-case	Well-defined eaves	Absent	<0.01 [<0.01, 0.10]	4408	8815
	Minutes after sunset	30			
	Pedestrian traffic	4			
Typical-case	Well-defined eaves	Either	0.78 [0.58, 0.90]	2	4
	Minutes after sunset	60			
	Pedestrian traffic	0.5			
Best-case	Well-defined eaves	Present	0.97 [0.87, 0.99]	1	2
	Minutes after sunset	90			
	Pedestrian traffic	0			

Note: *p* = probability of detection and *n* = integer estimate.

Banks, T., T.P. Rogers, and C.A. Aldridge. 2020. Geographic distribution: *Hemidactylus turcicus* (Mediterranean Gecko). USA: MISSISSIPPI: CLAY CO. Herpetological Review 51:541.

Guillera-Arroita, G., M.S. Ridout, and B.J.T. Morgan. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1:131–139.

Kéry, M. 2002. Inferring the absence of a species – a case study of snakes. Journal of Wildlife Management 66:330–338.

Kéry, M. and J.A. Royle. 2008. Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. Journal of Applied Ecology 45:589–598.

Kéry, M. and J.A. Royle. 2016. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS, Volume 1 Prelude and Static Models. Academic Press, London, UK.

Lee, J.R. 2008. Geographic distribution of *Hemidactylus turcicus* (Reptilia: Squamata: Gekkonidae) in Mississippi. Journal of the Mississippi Academy of Sciences 53:184188.

Meshaka, W.E., Jr., S.D. Marshall, J. Boundy, and A.A. Williams. 2006. Status and geographic expansion of the Mediterranean Gecko, *Hemidactylus turcicus*, in Louisiana: implications for the Southeastern United States. Herpetological Conservation and Biology 1:45–50.

Nelson, D.H. and S.D. Carey. 1993. Range extension of the Mediterranean Gecko (*Hemidactylus turcicus*) along the Northeastern Gulf Coast of the United States. Northeast Gulf Science 13:53–58.

Rogers, T., T. Banks, and C.A. Aldridge. 2020. Geographic distribution: *Hemidactylus turcicus* (Mediterranean Gecko). USA: MISSISSIPPI: LOWNDES CO. Herpetological Review 51:77.

White, J.W. and M.S. Husak. 2015. New county records and range expansion of the Mediterranean Gecko (*Hemidactylus turcicus*) in Southwestern Oklahoma. The Southwestern Naturalist 60:99–101.

White, J.W., M.S. Husak, and R.E. Willis. 2016. Geographic distribution of the nonnative Mediterranean Gecko (*Hemidactylus turcicus*) in Oklahoma. The Southwestern Naturalist 61:338–341.