

Power of GRAMP Surveys to Monitor Amphibian Abundance

Submitted to:

M. Segura and B. Mitchell

I&M Network:

Gulf Coast

Submitted by:

Ellen Cheng, University of Georgia

Date:

February 5, 2017

Contents

INTRODUCTION	3
Background	3
Factors influencing statistical power	4
Simulation-based power analyses	5
Amphibian dynamics	6
METHODS	6
Field data	6
Model estimation	7
Model selection and evaluating model support	10
Power analyses	11
Intrinsic population growth (r)	12
Change in probability of occupancy	13
RESULTS	15
JELA	17
JELA: <i>Hyla</i> mix	19
JELA: <i>Lithobates clamitans</i>	26
JELA: <i>Eleutherodactylus planirostris</i>	34
GUIS	42
GUIS: <i>Eleutherodactylus planirostris</i>	44
GUIS: <i>Hyla</i> mix	52
SAAN	59
SAAN: <i>Eleutherodactylus cystignathoides</i>	61
SAAN: <i>Incilius nebulifer</i>	69
VICK	76
VICK: <i>Eurycea guttolineata</i>	78
VICK: <i>Scincella lateralis</i>	85
DISCUSSION	92
Statistical power and coherent interannual variation	92
Statistical power and survey variables	93
Model selection	93
Covariate relationships	94
Survey considerations	94
LITERATURE CITED	96

APPENDIX A1: Additional power analysis results	99
JELA Hyla mix	99
JELA Lithobates clamitans	101
JELA Eleutherodactylus planirostris	102
GUIS Eleutherodactylus planirostris	104
GUIS Hyla mix	105
SAAN Eleutherodactylus cystignathoides	106
SAAN Incilius nebulifer	107
VICK Eurycea guttolineata	108
VICK Scincella lateralis	109

INTRODUCTION

Background

Since 2011, the Gulf Coast Inventory and Monitoring Network (GULN) of the National Park Service (NPS) has been conducting monthly herpetological surveys as part of the nationwide NPS Vital Signs Monitoring program. Under this program, NPS Inventory and Monitoring (I&M) networks track the status and trends of selected natural resources and processes that represent the overall condition of park ecosystems. Amphibians and reptiles were identified by GULN as a critical vital sign for the region’s parks because herps are a widely distributed taxa that occupy many habitat types, are sensitive indicators of environmental health, and include species that are protected at the state level.

To ensure monitoring data can effectively support management decision-making for park resources, an important component of the Vital Signs Monitoring program is to evaluate the ability of survey protocols to detect changes in population status. In this report we assess the statistical power of the GULN reptile and amphibian monitoring protocol (GRAMP) to detect long-term trends in herp abundance. Statistical power is the probability of detecting a change in a population, given that a change has occurred. Our aim was to estimate power to detect population trends under different scenario combinations of monitoring duration, within-year survey frequency, number of traps, trend magnitude, and statistical significance level (alpha).

We used data from the first five years of GRAMP surveys at four network parks (one site per park; Table 1) to inform simulation-based power analyses for generalized linear mixed models (GLMMs). GLMM is an extension of simple linear regression appropriate for the repeated measures data collected in GRAMP surveys. A separate analysis was conducted for the most commonly observed species at each GRAMP site, for a total of nine site-species analyses (Table 2). For most site-species, count data were summarized as the number of individuals captured at each trap in each survey (i.e., a trap-count), then analyzed using a negative binomial or Poisson GLMM. Data for VICK species were analyzed differently due to insufficient captures, so results are not directly comparable with the other GRAMP sites. For VICK species, count data were converted into a binary response (captured or not captured) for each transect in each survey, then analyzed using a logistic regression (GLMM with logit link). GRAMP sites are widely separated geographically, so were not expected to exhibit similarities in amphibian composition or dynamics.

Table 1. Study sites and number of each trap type. In this report, sites are referenced by their park acronym (one site per park). At least two trap types are deployed at each site: PVC = number of PVC sets (one large and one small pipe on a tree constitute a set); CB = number of coverboard traps; MT = number of minnow traps.

Park name	Park acronym	PVC	CB	MT	Site information
Jean Lafitte National Historical Park and Preserve	JELA	26	30	25	Primarily canopy forest of diverse hardwoods, tupelo and cypress. The forest is temporarily inundated with sheet-flow flooding during seasonal storms. The adjacent Bayou des Familles provides permanent waters for amphibian breeding.
Gulf Islands National Seashore	GUIS	24	30	25	Forest (oak, pine, and tupelo) and bluff-scrub habitat adjacent to Beaver Pond, a 4-hectare coastal freshwater pond. Minnow traps are placed in a permanently inundated tupelo stand.
San Antonio Missions National Historical Park	SAAN	10	60	0	Successional wooded habitat along the banks of the San Antonio River, and adjacent “old field” habitats. The site is surrounded by dense urban developments. The region faced severe drought for several years prior to monitoring year 2012. Changes in water management since 2012 increased the availability of water at the site.
Vicksburg National Military Park	VICK	15	60	0	Second-growth hardwood forest. The site has no permanent water except for small seeps. Almost all captures of <i>E. guttolineata</i> at this site occurred at two traps near seeps.

Table 2. Site-species analyzed in this report. We analyzed data for the most commonly captured species at each site. Numbers are the total captures of each species from each trap type (PVC = PVC sets, CB = coverboard traps, MT = minnow traps), summed across all traps and survey dates. ‘Hyla mix’ represents the combined data for *Hyla squirella* and *Hyla cinerea*, because these sister species share the same ecological niche and some morphs can be difficult to visually distinguish. ‘% of captures’ is the proportion of total captures at the site accounted for by that species. For every species, the trap type used in power analyses is in bold. Minnow traps are not deployed at SAAN and VICK. ‘Habitat and breeding’ provides general biological information about the species, but actual breeding dates may differ by site and year.

Site	Scientific name	Common name	% of captures	PVC	CB	MT	Habitat and breeding
JELA	<i>Hyla mix</i>	Squirrel & Green Treefrogs	37.8	1452	3	11	Found in moist habitats. Lays eggs in shallow pools, often following rainfall, between March and October.
JELA	<i>Lithobates clamitans</i>	Green Frog	24.0	0	31	899	Adults are highly aquatic, but juveniles may be found on land when conditions are wet. Breeds from April to August, laying eggs in semipermanent or permanent fresh water. In many parts of their range, tadpoles overwinter in permanent waters.
JELA	<i>Eleutherodactylus planirostris</i>	Greenhouse Frog	7.5	0	292	0	Native to Cuba and other islands in the West Indies. Lives in moist leaf litter, emerging on warm rainy days. Breeds from May to September. Eggs are laid terrestrially in moist soil and leaf litter.
GUIS	<i>Eleutherodactylus planirostris</i>	Greenhouse Frog	28.8	0	441	0	(See description above)
GUIS	<i>Hyla mix</i>	Squirrel & Green Treefrogs	25.9	387	3	6	(See description above)
SAAN	<i>Eleutherodactylus cystignathoides</i>	Rio Grande Chirping Frog	55.2	0	1882	NA	Lives in dense vegetation of moist habitats. Breeds from April to May. Eggs are laid terrestrially in moist soil and leaf litter.
SAAN	<i>Incilius nebulifer</i>	Coastal Plains Toad	38.0	0	1295	NA	Found in many habitat types. Breeds from March to August, often following spring or summer rains exceeding 3 cm. Eggs are laid in many types of still water.
VICK	<i>Eurycea guttolineata</i>	Three-lined Salamander	22.2	0	86	NA	Often found under wood, rocks, or other cover near water. Breeds in fall and winter. Eggs may be attached to cover in subterranean sections of streams and seeps.
VICK	<i>Scincella lateralis</i>	Little Brown Skink	16.0	0	62	NA	A fossorial species often found buried in deep substrate, especially leaf litter. Breeds in March through August, laying eggs in moist soil or under various types of cover.

Factors influencing statistical power

Statistical theory provides us with preliminary insights on factors that influence power to detect long-term population trends. These insights can be used to inform survey design. Statistical power typically increases with sample size, effect size (the magnitude of population change), and significance level, and decreases with variability in the response variable. Improving

statistical power in surveys by increasing the first three factors has trade-offs. A larger sample size (number of traps, number of surveys per year, or number of years of monitoring) may improve power by increasing precision of the population estimate, but may require more money and time for surveys. Specifying a larger effect size means the population may undergo greater change before management interventions are activated. And increasing significance level can increase power but will simultaneously increase Type I error, i.e., the probability of inferring a significant trend when there is none.

The fourth factor—variability in the response—affects statistical power by masking the signal of a population change. The impacts of this “noise” on power to detect a trend can be reduced by including relevant covariates in analyses. For example, if amphibian capture rates are correlated with temperature at the time of a survey, then including temperature as a model covariate can “remove” this temperature-generated variability from the regression model, thereby increasing the model’s signal-to-noise ratio (and statistical power to detect a population change). Variation among years may be reduced by incorporating a year-specific covariate, such as hydroperiod, in a GRAMP model.

Response variation that cannot be explained by covariates can have varying impacts on statistical power, depending on the source of the variation. For example, unexplained variation in GRAMP capture numbers can be partitioned among several sources: 1) variation among traps, e.g., some traps may consistently have higher numbers of captures; 2) survey variation, e.g., environmental conditions at the time of a survey may influence site-wide capture success; 3) variation among years, e.g., captures may be generally higher in some years than others due to unmeasured climatic covariates (aka ‘coherent interannual variation’); and 4) any remaining unexplained variation.

When the objective is to detect long-term trends with hierarchical sampling—such as GRAMP sampling over multiple years with multiple surveys within each year and multiple traps per survey—statistical power is often most sensitive to sample sizes and unexplained variation at the highest level of sampling (Maas and Hox 2005), in this case the year-level variation. That is, the power to detect long-term trends should be especially sensitive to coherent interannual variation (throughout this report, this term will refer only to *unexplained* year-to-year variation in captures). Statistical power can be improved by reducing coherent interannual variation and by increasing the number of years of surveys (Urquhart 1998; Urquhart 1999). With GRAMP surveys, the influence of unexplained variation at the lowest sampling level (variation among traps) may be further diminished by the fact that the same traps at a site are visited each survey. If, instead of repeated measures sampling, a random sample of traps were visited each survey, then variation in capture numbers among traps could potentially have greater impact on power to detect population trends (Urquhart and Kincaid 1999; Urquhart et al. 1998).

Simulation-based power analyses

Statistical power can be estimated analytically for studies with fixed effects and normally distributed response variables. But many ecological studies incorporate hierarchical sampling (requiring random effects) and non-normal response data. In the case of GRAMP surveys, the same traps at a site are surveyed repeatedly over time, so ‘trap ID’ is included as a random effect in models. Furthermore, the response data for most GRAMP sites are counts of amphibians. Count data tend to be non-normally distributed, so should not be analyzed using statistical tests that assume normality. Analytical power calculations are not readily available for these data types, which require complex modeling frameworks such as GLMMs. Simulation-based power analyses have the flexibility to accommodate GLMMs, and may produce more accurate results even in cases when analytical power analyses are possible (Johnson et al 2015).

The idea behind simulation-based power analyses for GLMMs with count data is relatively straightforward. If pilot data have already been collected, they can be modeled to estimate the regression intercept (i.e., overall mean count for the dataset), coefficients (these quantify the relationship between counts and categorical fixed effects or continuous covariates), and the variances and covariances of random effects. Count data often follow a negative binomial distribution, which is more flexible than a Poisson distribution because a dispersion parameter (K) allows the variance of counts to vary from the mean (Rice 1995). Similar to the residual error term in a linear regression model, unexplained variation in a negative binomial mixed model is absorbed in the dispersion parameter, which should be included in power analyses (Johnson et al 2015).

A GLMM estimated from pilot data identifies the relationship between a response variable (e.g., trap-counts) and predictor variables and quantifies the remaining unexplained variation. Using the model relationships and natural intrinsic variability of the system, estimated from the pilot data, we can repeatedly simulate trap-counts for a particular scenario of survey design and effect size. Statistical power is then calculated as the proportion of simulated data sets for which the population change is statistically significant in the expected direction (increase or decline). For example, using the regression coefficients and variance estimates from a GLMM fit to GRAMP data, we could simulate trap-count data for a particular scenario, such as a population decline of 4% per year for 10 years, six GRAMP surveys per year, and a fixed plot design of 30 traps. The annual population decline of 4% could be simulated by reducing the mean trap-count by 4% each year, for 10 years. Thus, if the estimated mean trap-count from actual GRAMP data were 5 animals/trap, we could simulate mean captures of 5 animals/trap in the first year, then 4.8 animals/trap ($5 * 0.96 = 4.8$) in the second year and 4.61 animals per trap ($4.8 * 0.96 = 4.61$) in the third year, etc.

Due to intrinsic variation in the system that is being simulated, we would generate a large number of hypothetical 10-year data sets, typically 500 or more. Each simulated data set would then be analyzed to determine if a significant population decline is detected at the specified significance level, e.g., $\alpha = 0.20$ for a two-tailed test. If a statistically significant population decline is detected in 350 out of 500 simulated data sets, the power of the simulated survey design to detect the declining trend would be $350 / 500 = 70\%$. It would also be useful to estimate false trends, which in this case would be the proportion of simulated data sets in which a significant increase is detected although a deterministic decline is simulated. False trends tend to occur if the intrinsic variation of a system is large relative to the rate of decline that is simulated. Simulations can be repeated for different realistic combinations of survey frequency, monitoring duration, and other survey factors, to compare alternatives for increasing the statistical power of a survey.

Amphibian dynamics

Herps, and in particular amphibians, are a challenging taxa in which to detect long-term trends because many species undergo large annual population fluctuations, sometimes over orders of magnitude (Green 2003; Marsh 2001; Pechmann et al. 1999; Semlitsch et al. 1996). However, some amphibian species fluctuate less than others. Plethodontid salamander populations tend to be more stable than salamandrid or ambystomatid salamander populations. *Eleutherodactylus* and other frog species with direct terrestrial development tend to exhibit greater year-to-year stability than do species with aquatic larval phases (Green 2003).

Amphibian power analyses frequently find that long durations of monitoring and large effect sizes are required to achieve 80% power (a common standard for power analyses) to detect a decline. For example, Hayes and Steidl (1997) estimated that for *Pseudacris ornata* (Ornate Chorus Frog), a population decline of 5% per year would be detected with 80% power at $\alpha = 0.10$, only after 32 years of monitoring. By this time, the population would be 20% of its initial population size. Using counts of *Lithobates sylvaticus* (Wood Frog) egg masses, Scherer and Tracey (2011) found that a decline in egg mass numbers of 10% per year would be detected with 80% power at $\alpha = 0.15$, after 9 years of monitoring.

A factor that may further complicate trend detection for amphibians is the episodic nature of annual population dynamics, especially for species with aquatic larvae. A common dynamic pattern for amphibians is occasional years of very high recruitment punctuating indeterminate spans of low-recruitment years in which the population declines because mortality exceeds recruitment (Alford and Richards 1999; Meyer et al. 1998; Pechmann & Wilbur 1994; Semlitsch et al. 1996). One consequence of episodic dynamics is that false trends (described in the previous section) are common with short time series (Meyer et al. 1998). With short time series for episodic populations it is also difficult to reliably identify covariates linked to year-to-year variation, determine the nature of density dependence, and estimate variability for power analyses. For example, Marsh (2001) found that estimates of CV varied up to 10-fold when calculated from 5-year time series that were subsamples of longer amphibian time series.

METHODS

Field data

We estimated power to detect trends in the most commonly observed amphibian species at four GRAMP sites (Table 2). GRAMP surveys use arboreal PVC pipe (PVC), terrestrial cover-board (CB), and/or aquatic minnow-trap (MT) fixed-point sampling to assess temporal trends in herp relative abundance and assemblage composition. Each site has between 70 and 81 traps distributed in up to six linear clusters, with at least two meters between traps and five meters between clusters. At each site, herps captured during monthly surveys by GULN biologists are counted and recorded to species by trap ID. Any herps opportunistically observed outside of traps during field sampling are recorded as “casual observation”.

At all sites, up to two HOBO data-loggers placed among traps record air temperature (F) and relative humidity (%) at 10-minute intervals throughout the year (but HOBO data for SAAN are only available since 2014). These environmental covariates are monitored because they affect herp activity levels (Bellis 1962; Joppa et al. 2009; Reynolds et al. 1994; Roloff et al. 2011), which may influence the animals’ use of artificial cover such as PVC and coverboard traps. Specifically, amphibians may use artificial cover when temperature and humidity are low, and below certain thresholds the animals may move to deeper, more protective cover than artificial cover can provide (Dr. John Maerz, University of Georgia, pers. comm.). Detailed site and survey descriptions can be found in GRAMP annual reports.

Precipitation levels may influence amphibian activity and recruitment success (e.g., Baldwin et al. 2006; Reynolds et al. 1994; Semlitsch & Pechmann 1985). The Global Historical Climatology Network (GHCN; www.ncdc.noaa.gov) is an integrated database of climate data compiled from more than 20 sources and subjected to standardized quality assurance reviews. For each GRAMP site we obtained daily and monthly precipitation totals from the nearest GHCN weather station with the most complete data for monitoring years 2011–2016 (Table 3), and used these data as covariates in model estimation.

Table 3. Sources of precipitation data for GRAMP sites. We obtained daily and monthly precipitation totals from these Global Historical Climatology Network weather stations for monitoring years 2011 through 2016. All weather stations are within 30 km of the GRAMP site (last column). Data and metadata for the weather stations are available from www.ncdc.noaa.gov.

Site	Station ID	Station name	Distance to station (km)
JELA	USW00012916	New Orleans international airport LA US	26
GUIS	USW00013899	Pensacola regional airport FL US	12
SAAN	USW00012970	San Antonio Stinson municipal airport TX US	2
VICK	US1MSWR0007	Vicksburg 12.3 S MS US	23.5

Model estimation

We used GRAMP data collected from monitoring years 2011 to 2015 (a monitoring year begins on October 1 of the previous calendar year) to identify the best-supported GLMM separately for each site-species. When available, data for the first few months of monitoring year 2016 were also included. Our approach to estimating GLMM parameters from the field data was similar for most site-species and is described below, followed by explanations of any modifications for particular site-species. All models and analyses were developed using R version 3.2.3 (R Core Team 2015).

We wanted to compare simulation scenarios using current trap levels versus increasing the number of traps. Therefore, for all sites except VICK, we used trap-count as the GLMM response variable. We defined trap-count as the number of individuals recorded at each trap in each survey, for the site-species of interest. We used maximum likelihood estimation to estimate parameters for a global GLMM and for subsets of the global model. Combined, these models were the set of candidate models for each site-species. The fixed effect variable ‘transect’ was included in all candidate models. All candidate models also had the same random effects—trap ID and monitoring year—to account for variation among traps and for coherent interannual variation. Subsets of the global model for a site-species were created by combining these three parameters with all possible combinations of other fixed effect variables, including a null model.

VICK data were analyzed differently from other sites due to very low capture numbers at VICK. For the two VICK species, approximately 98% of trap-counts were zero. We were not able to obtain reasonable model fit to these count data. Therefore, we summarized trap-counts by transect and analyzed the data with mixed effects logistic regression (i.e., GLMM with a logit link instead of a log link). The binary response variable in VICK models was ‘1’ if the species was captured on a transect during a survey, and ‘0’ if the species was not captured on a transect during a survey. Because VICK data were modeled with logistic regression, power analysis results for VICK are not directly comparable with the other GRAMP sites.

Every species analyzed for this report was primarily captured in a single trap type (Table 2). To maximize the degrees of freedom in the GLMM for each site-species, we only used capture data for the primary trap type. For example, 97% of *Lithobates clamitans* (Green Frog) captures at JELA were in minnow traps. Therefore, only captures at minnow traps were used in JELA analyses for this species. Any species detections logged as “casual observation” during field sampling were excluded from analyses.

Capture rates for many species exhibited seasonal patterns. We expected this pattern largely reflected correlation in environmental factors, such as temperature, that were included in candidate models. However, we also included a day-of-year covariate and binary breeding season variable to model any seasonal variation not accounted for by these other covariates. If the coefficient for a second-order term of a quadratic variable was highly non-significant in most candidate models, we additionally evaluated models without the quadratic term. For JELA, GUIS and SAAN, if the 95% confidence interval for the negative binomial dispersion parameter overlapped zero in most candidate models, we additionally evaluated models with a Poisson distribution.

We did not remove correlated predictors from candidate models because a combination of partially correlated variables sometimes yields a better predictive model than if variables had been omitted to eliminate multicollinearity. Multicollinearity should not affect estimates of model fit or a model’s predictive abilities (Shmueli 2010). We did not model interactions because data were sparse and we did not have reason to expect strong interactions among predictor variables. We also did not incorporate autoregressive structure in candidate models, due to the low likelihood of resolving autoregressive structure in a 5-year data set for a taxa with high coherent interannual variation.

Parameter estimation was implemented in the R package glmmADMB (Fournier et al. 2012; Skaug et al. 2016). In R, the global model for JELA, GUIs and SAAN species was coded as:

```
global.model <- glmmadmb(count ~ (transect - 1) + nb + my.day + I(my.day^2) +
  temp.f + I(temp.f^2) + rh + I(rh^2) + precip.7 + I(precip.7^2) + mean.prior.precip +
  (1 | monitor.yr) + (1 | transect/trap.id), data = data, family = "nbinom",
  link = "log")
```

The global model for VICK species was similar, but used a binary response variable (captured versus not captured) instead of a count, excluded the trap random effect because captures were analyzed at the transect level, and used the logit link function instead of a log link function.

Global model variables are defined below, followed by an explanation of changes to the global model for particular site-species.

Fixed effects:

- **Transect**—Traps were clustered in transects, and we expected some transects would have consistently higher mean trap-counts than others, possibly due to proximity to water or other unmeasured covariates. In addition, GULN was interested in evaluating if power could be improved by increasing the number of traps in existing versus new transects. Methods for each case are described in the ‘Power analyses’ section, below. We included ‘transect’ as a categorical trap-specific parameter in all candidate models for JELA, GUIs and SAAN.
- **Breeding season (‘nb’)**—A binary (0 / 1) variable was included to distinguish surveys conducted in breeding versus non-breeding months because activity levels may differ between these seasons and may therefore influence capture rates. We broadly classified October through March as the non-breeding season, and April through September as the breeding season. The actual dates of breeding differ among species and, even for the same species, may differ among sites and years. In the absence of site- and year-specific breeding dates for each species we applied our general classification of breeding season for all species. Candidate models either included or excluded this binary survey-specific covariate.
- **Day-of-year (‘my.day’)**—Modeling trap-count as a function of the number of days since the start of the monitoring year (October 1) is one way to account for seasonal effects that cannot be explained by seasonal patterns in measured covariates. Candidate models included ‘day-of-year’ as a quadratic or linear (in some cases of non-breeding season monitoring) covariate term, or excluded this survey-specific covariate.
- **Temperature, Fahrenheit (‘temp.f’)**—Air temperature may influence the probability of detecting amphibians and reptiles at PVC and coverboard traps by modifying activity levels of these cold-blooded animals (Joppa et al. 2009; Reynolds et al. 1994; Roloff et al. 2011). Air temperature recorded at 10-minute intervals was available from HOBO data logger(s) placed at each site. Typically, surveys of PVC and coverboard traps commenced at approximately 7:30 AM and were completed between 10:30 AM to noon. For each survey at a site, we estimated temperature as the mean temperature between 8:00 AM and 10:00 AM for the 5-day period centered on the survey date. We averaged temperature across five days because sometimes temperature data were not available for a survey date, but were available for a day or two before or after. In that case, we averaged temperature across the days for which data were available. If temperature data were not available for any of the five days centered on a survey date, we classified the covariate data as ‘missing data’ and excluded the corresponding capture data from analyses. Candidate models included ‘temperature’ as a quadratic or linear covariate term, or excluded this survey-specific covariate.
- **Relative humidity, % (‘rh’)**—Relative humidity may influence the probability of detecting amphibians at PVC and coverboard traps by modifying the animals’ use of cover to reduce evaporative water loss through their skin (Bellis 1962; Reynolds et al. 1994). Relative humidity recorded at 10-minute intervals was available from HOBO data logger(s) placed at each site. We estimated relative humidity in the same way as we estimated survey temperature—it was the mean between 8:00 AM and 10:00 AM for the 5-day period centered on the survey date. As with temperature, if relative humidity data were not available for any of the five days centered on a survey date, we excluded the corresponding capture data from analyses. In most cases, temperature and relative humidity data were missing for the same survey dates. Candidate models included ‘relative humidity’ as a quadratic or linear covariate term, or excluded this survey-specific covariate.
- **Mean daily precipitation in the seven days prior to a survey, inches (‘precip.7’)**—Recent precipitation may influence substrate moisture and may increase the availability of temporary standing pools of water. The availability of moist microhabitat at a site may influence amphibians’ activity patterns and ranges (Reynolds et al. 1994; Semlitsch & Pechmann 1985), which may affect their use of PVC and coverboard traps. For each GRAMP site we obtained daily precipitation data from a nearby GHCN weather station (Table 3). ‘Precip.7’ was calculated as the daily precipitation averaged across the seven days prior to a survey. Candidate models included ‘precip.7’ as a quadratic or linear covariate term, or excluded this survey-specific covariate.

- Mean monthly precipitation during the prior monitoring year, inches ('mean.prior.precip')—Precipitation can increase amphibian breeding activity and can impact larval survival through changes in local hydroperiod (Baldwin et al. 2006; Paton & Crouch 2002; Pechmann et al. 1989; Reynolds et al. 1994; Semlitsch 2002). These factors may influence post-breeding population sizes. For each GRAMP site we obtained precipitation data from a nearby GHCN weather station (Table 3). 'Mean.prior.precip' was calculated as monthly precipitation averaged over the prior monitoring year. For example, for monitoring year 2014 (October 2014 through September 2015), 'mean.prior.precip' was estimated from the precipitation for October 2013 through September 2014. Thus, 'mean.prior.precip' accounted for precipitation during the prior breeding season as well as during months preceding the prior breeding season—factors that may influence the local hydroperiod. Candidate models either included or excluded this monitoring year-specific covariate.

Random effects:

- Monitoring year ('monitor.yr')—Captures may be overall higher or lower in certain years due to climatic or other variables not accounted for in regression models. 'Monitor.yr' was included as a random effect in all candidate models to account for coherent interannual variation in capture numbers. By including 'monitor.yr' as a random effect only, we made explicit two assumptions: 1) there was no linear trend in the pilot data, and 2) the coherent interannual variation observed in the pilot data was a random sample of coherent interannual variation for the site-species. Regarding the first assumption, pilot data for some site-species suggested an ongoing population trend, but short time series of amphibian counts are prone to false trends (Meyer et al. 1998). With our assumption of no linear trend in the pilot data, our estimate of power was potentially conservative. Regarding the second assumption, estimates of coherent interannual variation based on only five years of amphibian data should be considered highly tentative (Marsh 2001).
- Trap ID ('trap.id')—Data were collected from the same traps each time a site was surveyed. Counts from the same trap are expected to be more correlated than counts from different traps. Traps in the same transect may also be more correlated than traps from different transects. Therefore, we included 'trap.id' as a random effect nested within 'transect', in all candidate models for JELA, GUIS and SAAN (data for VICK were analyzed at the transect level). In this way we controlled for these potential sources of non-independence in trap counts, and reflected in analyses that we considered the surveyed traps to represent a random sample of potential traps for a transect.

Based on examination of the pilot data, we made the following site-specific modifications to the global model:

JELA

- Most captures occurred during the non-breeding season, which was defined in our analyses as October through March. Analyzing year-round data led to poor model fit, even when the binary predictor variable for breeding season ('nb') was included. Therefore, we restricted analyses to data from the non-breeding season.
- Precipitation data for the week prior to a survey was not available for many survey dates, from the nearby GHCN weather station. Instead of using 'precip.7' as a covariate for JELA models, we used 'precip.30', calculated as the mean daily precipitation for the 30 days prior to a survey.
- For *E. planirostris*, the global model could not be estimated with all parameters included. We initially removed the temperature covariate from the global model, because temperature was not an important covariate for this species at GUIS. However, with temperature removed the global model still could not be estimated. We ultimately removed 'mean.prior.precip', to allow convergence of the global model.
- Transect was not included as a parameter for analyses of *L. clamitans* data because all minnow traps (the primary trap type for this species) were in a single transect. We also did not include relative humidity in *L. clamitans* models because we felt it did not make biological sense that relative humidity could influence aquatic captures.

GUIS

- Most captures occurred during the non-breeding season, which was defined in our analyses as October through March. Analyzing year-round data led to poor model fit, even when the binary predictor variable for breeding season ('nb') was included. Therefore, we restricted analyses to data from the non-breeding season.

SAAN

- Water management practices at SAAN changed in monitoring year 2012, increasing availability of water at the site. Amphibian trap-counts at SAAN also increased considerably around this time. The higher amphibian capture numbers at SAAN since 2012 may have been due, in part, to the change in water management practices. To eliminate potential biases due to this factor, we omitted monitoring year 2011 data from SAAN analyses. During monitoring year 2011 (monitoring began in June 2011), there were no captures of *E. cystignathoides* and only one capture of *I. nebulifer*, so exclusion of monitoring year 2011 data primarily reduced the proportion of zero-counts in the data.
- Temperature and relative humidity data were not available for monitoring years 2011 through 2013 at SAAN. Therefore, SAAN models did not include these covariates.

VICK

- Temperature and relative humidity data were missing for approximately one-quarter of VICK surveys. For both VICK species, when temperature and relative humidity were included in logistic regression analyses, neither variable was retained in the best-supported model. Therefore, we re-ran the analyses (global model and all subsets) without temperature and relative humidity data, which allowed us to use all survey data (i.e., no response values were omitted due to missing covariates).

Model selection and evaluating model support

When mixed models differ only in their fixed effects, as was the case in our analyses, Akaike information criteria (AIC; Burnham & Anderson 2002) are commonly used for model selection. However, uncertainties about how to count the parameter equivalence of random effects are yet unresolved (Spiegelhalter et al. 2002; Symonds & Moussalli 2011). Also still debated are issues of averaging across models with polynomial functions, natural versus zero method averaging, and whether variances of random effects can be averaged (Burnham & Anderson 2002; Stephens et al. 2007).

In the absence of broadly accepted guidelines for model selection with mixed models, we used AIC scores, estimated with the R package *bbmle* (Bolker & R Development Core Team 2014), to select the best-supported (lowest AIC) among candidate models for each site-species. We then examined the range of coefficient and random effect variance estimates in the “confidence set” of top models. We reasoned that if many models in the confidence set for a site-species had similar coefficient estimates and (more importantly) similar variance estimates for the year random effect, then uncertainties regarding selection of a best-supported model from the candidate set should have minimal influence on power analysis results. The confidence set for each site-species comprised models with an AIC within six units of the best-supported model, excluding any that were more complex forms of models with lower AIC. For example, if one model with AIC = 100 included the predictors A and B, and another model with AIC = 101.2 had predictors A, B and C, then we excluded the latter from the confidence set because the additional parameters were likely uninformative (Arnold 2010; Richards et al. 2011).

Prior to conducting power analyses, for each site-species we evaluated the fit of the best-supported model. Temporal autocorrelation was not apparent for any site-species, based on plots of residuals against survey dates. We examined histograms of observed versus predicted data, to identify discrepancies that might suggest poor model fit. We also simulated data sets from the best-supported model for each site-species, to compare test statistics calculated from the observed versus simulated data. To simulate count data, we followed the approach of Johnson et al. (2015), implemented in the R package *glm.simm* and outlined below. For VICK species, we used a modified approach to generate binary response data.

Simulated count data were generated using regression parameter estimates and sample sizes (e.g., number of traps, survey frequency) obtained from the pilot data. “Noise” in expected trap-counts among years and traps was simulated by drawing from normal distributions with mean = 0 and variances estimated from the best-supported model (the random effects variances for ‘monitor.yr’ and ‘trap.ID’, respectively). The remaining variation was incorporated through the dispersion parameter (K) for the negative binomial distribution from which simulated trap-counts were drawn.

For each site-species, each iteration of a simulation generated a time series of trap counts equal in length to the pilot data. We generated 100 iterations of simulated data per site-species. We calculated two test statistics from the observed data and each iteration of simulated data: 1) the proportion of zero-counts, i.e., trap-counts in which no animals were captured; and 2) Freeman-Tukey goodness-of-fit (Freeman & Tukey 1950). Ideally, observed values of the statistics for each site-species would be close to the median of the statistics calculated from the simulated data. If observed test statistics fall in the extreme lower or upper distributions of statistics calculated from the simulated data, the best-supported model does not describe the observed data well. Due to the small number of simulation iterations for each site-species (limited by the slow speed of fitting GLMMs to sparse data), histograms of the test statistics from the simulated data were considered to be rough approximations of the test statistics’ theoretical distributions.

Power analyses

After confirming reasonable fit of the best-supported models to the pilot data, we used those models to inform simulation-based power analyses. In simulations we assumed constant detection probability, so changes in population size translated to changes in the number of individuals captured. For all sites, we simulated effect sizes representing annual change (decline or increase) in population numbers. However, the specific interpretation of “annual change” is different for VICK compared to other sites, because VICK data were analyzed with a logistic regression model and with the transect as the unit of analysis. A detailed explanation of simulated effect sizes is described in the sections ‘Intrinsic population growth (r)’ and ‘Change in probability of occupancy’.

Separately for each of the nine site-species, we assessed how power to detect a population trend was influenced by monitoring duration, survey frequency, number of traps, trend magnitude, and statistical significance level (Table 4). Likelihood ratio tests were used to determine if simulated trends were statistically significant. Bootstrapping can more accurately determine significance in mixed models, but computation time would have prohibitively long. The scenarios evaluated for survey frequency and trap effort differed among site-species, based on GULN requests and sometimes by data limitations.

Table 4. Power analysis scenarios completed for this report. Site-species abbreviations in column headings 4–12 are as follows: J.Hm = JELA *Hyla mix*, J.Lc = JELA *Lithobates clamitans*, J.Ep = JELA *Eleutherodactylus planirostris*, G.Ep = GUIS *Eleutherodactylus planirostris*, G.Hm = GUIS *Hyla mix*, S.Ec = SAAN *Eleutherodactylus cystignathoides*, S.In = SAAN *Incilius nebulifer*, V.Eg = VICK *Eurycea guttolineata*, V.Sl = VICK *Scincella lateralis*. In table cells, ‘Y’ indicates analyzed scenarios. (Column 1) ‘Number of traps’ is the number of simulated traps where ‘Baseline’ = current number of traps; ‘50%, existing’ = 50% more traps distributed equally among existing transects; and ‘50%, new’ = 50% more traps in simulated new transects. (Column 2) ‘Duration’ is the number of years of population trend simulated. (Column 3) ‘Survey frequency’ is whether surveys are conducted monthly or in alternate months (‘alt’). Because analyses for JELA and GUIS considered only non-breeding season months (October through March), ‘monthly’ means 6 surveys per year and ‘alt’ means 3 surveys per year. For SAAN and VICK, ‘monthly’ means 12 surveys per year and ‘alt’ means 6 surveys per year. For each of the 48 scenario combinations in this table, we assessed statistical power at 7 effect sizes (-0.06, -0.04, -0.02, 0.0, 0.02, 0.04, and 0.06). For JELA, GUIS and SAAN, the effect sizes represent intrinsic population growth (r). For VICK, the effect sizes represent change in the probability of occupancy, on the logit scale. For each of the 336 power analysis scenarios completed (48 $\times$ 7), we estimated statistical power at two levels of alpha (0.10 and 0.20, two-tailed tests) based on 500 simulation iterations per alpha level.

Number of traps	Duration	Survey frequency	J.Hm	J.Lc	J.Ep	G.Ep	G.Hm	S.Ec	S.In	V.Eg	V.Sl
Baseline	5	alt	Y	Y	Y	Y	Y	Y	Y		
Baseline	5	monthly	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline	10	alt	Y	Y	Y	Y	Y	Y	Y		
Baseline	10	monthly	Y	Y	Y	Y	Y	Y	Y	Y	Y
50%, existing	5	alt	Y		Y						
50%, existing	5	monthly	Y		Y						
50%, existing	10	alt	Y		Y						
50%, existing	10	monthly	Y		Y						
50%, new	5	alt	Y		Y						
50%, new	5	monthly	Y		Y						
50%, new	10	alt	Y		Y						
50%, new	10	monthly	Y		Y						

The completed power analysis scenarios in Table 4 differed from GULN requests in the following ways:

- GULN had requested analyses of statistical power to detect an overall 10% or 25% population decline or increase for each scenario combination. The range of intrinsic population growth (r) we simulated encompassed GULN’s request, but included additional trends of lesser and greater magnitude to provide a clearer understanding how statistical power changed over a wider range of effect sizes. That is, we simulated 7 instead of the requested 4 levels of population change for each of the 48 scenario combinations.
- For JELA and GUIS, GULN had requested power analyses at current and seasonally reduced sampling intensities.

Current sampling intensity at both sites is 12 months per year. However, the number of captures for the JELA and GUIIS species we analyzed was very low during breeding season months (approximately April through September) compared to non-breeding season months. When we tried to fit models to year-round capture data, model fit was poor. Therefore, we conducted power analyses for 6 and 3 months per year (all during the non-breeding season) rather than for 12 and 6 months per year.

- For VICK, GULN had requested power analyses with the current number of traps and with 50% more traps on existing versus new transects. The pilot data for VICK were too sparse to analyze at the trap level. Therefore, we analyzed data at the transect level and conducted power analyses only for the current number of traps.

We conducted power analyses following the same approach as our simulations for testing model fit (Johnson et al. 2015). Deterministic population trends were incorporated into simulations by including a continuous fixed effect ('t') for the number of years since start of monitoring. For JELA, GUIIS and SAAN, the regression coefficient for this linear time variable represented percent change in the number of individuals captured at a trap during a survey. We assumed changes in capture numbers accurately portrayed changes in population size, so hereafter use 'intrinsic population growth' (r) to describe simulated effect sizes for these three sites. For VICK, the regression coefficient for the linear time variable represented change (on the logit scale) in the probability of capturing a species on a transect during a survey. We assumed changes in probability of capture accurately portrayed changes in probability of occupancy of survey-transects, so hereafter use 'change in probability of occupancy (logit scale)' to describe simulated effect sizes for VICK.

In all simulations, we incorporated stochastic variation around the deterministic population trend to reflect natural year-to-year variability in the response variable. The magnitude of this coherent interannual variation was determined by the estimated variance for the 'monitor.yr' random effect in the best-supported model for each site-species. We did not correlate interannual variation with population size in our simulations. In reality, year-to-year variation may decrease as a population gets smaller and increase as a population gets larger, but we were unable to determine this relationship from the pilot data.

In simulations, we specified survey frequency through the day-of-year ('my.day') variable. The GRAMP monitoring year begins on October 1, which was considered Day 0 in simulations. So, for example, in a simulation of surveys in alternate months during the non-breeding season, 'my.day' was set at the values 31, 92, and 151 to simulate surveys conducted on November 1, January 1, and March 1 of each year.

We simulated three levels of trap effort: current number of traps (baseline scenario), 50% more traps distributed equally among existing transects (for JELA only), and 50% more traps in new transects (for JELA only). Traps were nested within transects and were considered random samples from a normal distribution centered on each transect coefficient, with among-trap variance estimated from the best-fit regression model for the site-species. For each iteration of the scenario, a new set of traps was randomly drawn. Because transects were treated as fixed effects in model estimation, simulating new transects for JELA involved generating a mean value for each new transect, based on the mean of the estimated transect coefficients, and then also allowing for trap-level variability in the new transects.

If environmental covariates (temperature, relative humidity, recent precipitation, or prior year's precipitation) were in the best-supported model for a site-species, we fixed their values at their mean levels calculated from the pilot data. In total we simulated 336 power analysis scenarios (48 scenario combinations x 7 intrinsic population growth rates). Each power analysis scenario was iterated 500 times. Due to high stochastic variation in amphibian populations and the short duration of monitoring that we simulated (5 and 10 years), it would not be unusual to detect a statistically significant increase in a population that is undergoing a deterministic decline, or vice versa (i.e., a 'false trend'). Therefore, regardless of the trend direction simulated, we report both the proportion of iterations in which a significant decline was detected and the proportion in which a significant increase was detected. Statistical significance was based on a two-tailed test with $\alpha = 0.10$ and with $\alpha = 0.20$.

In the next two sections we explain how intrinsic population growth and change in probability of occupancy (logit scale) relate to overall population change.

Intrinsic population growth (r)

For JELA, GUIIS and SAAN we simulated seven levels of intrinsic population growth (-0.06, -0.04, -0.02, 0.0, 0.02, 0.04 and 0.06), with negative values representing exponential population decline and positive values representing exponential population increase. At $r = 0$, a simulated population exhibits no deterministic population change.

The expected value from the linear predictor of a negative binomial or Poisson regression model is on the logarithmic scale. To back-transform to the scale of the response variable (trap-counts), we must exponentiate the expected value. When we back-transform ' r ' to the response scale, we get an estimate of the finite population growth rate, λ . The equation that relates r and λ is:

$$\lambda = e^{rt}$$

(where λ = finite population growth rate, e = natural log, r = intrinsic population growth, and t = number of years since start of monitoring)

For example, if $r = -0.06$ (6% decline per year, on a logarithmic scale), then λ after one year will be 0.94. That is, the population size after one year will be 94% of the initial population size:

$$\lambda = e^{(-0.06)(1)}$$

$$\lambda = 0.942$$

And after three years of an annual 6% decline:

$$\lambda = e^{(-0.06)(3)}$$

$$\lambda = 0.835$$

Thus, if a population starts with 100 individuals and declines at an intrinsic rate of 6% per year, then after three years the population size will be 84 individuals. We can reach this same conclusion with the logic that if the population starts with 100 individuals and each year the population is 0.942 of the prior year's population (this is λ), then after three years the population size will be $100 \times 0.942 \times 0.942 \times 0.942 = 83.6$ (~84) individuals. Table 5 shows the long-term population change equivalent for the monitoring durations and intrinsic population growth rates we simulated.

Table 5. Conversion of intrinsic population growth rates (r) to proportion of initial population size. ‘Duration’ is number of years for which the population trend is sustained; ‘ r ’ is intrinsic population growth rate; ‘Percent of initial population’ is the percentage change in population size after 5 or 10 years (duration) of the specified trend. For example, if an annual population decline of $r = -0.06$ is sustained for 5 years, then at the end of the 5 years the population will be 74% of its initial size. If an annual population increase of $r = 0.06$ is sustained for 5 years, then at the end of 5 years the population will be 135% of its initial size.

Duration	r	Percent of initial population (%)
5	-0.06	74.1
5	-0.04	81.9
5	-0.02	90.5
5	0	100
5	0.02	110.5
5	0.04	122.1
5	0.06	135.0
10	-0.06	54.9
10	-0.04	67.0
10	-0.02	81.9
10	0	100
10	0.02	122.1
10	0.04	149.2
10	0.06	182.2

Change in probability of occupancy

For VICK, the same effect sizes (-0.06, -0.04, -0.02, 0.0, 0.02, 0.04, and 0.06) were incorporated in regression simulations, but a logit link function was used instead of a log link because we were modeling a binary outcome variable. Logistic regression is often used when observations can be recorded as a binary outcome variable (e.g., occupied versus not occupied). A simple logistic regression equation, a model without any predictor variables, can be written as:

$$\text{logit}(p) = \beta_0$$

(where β_0 is the probability of success for a binary variable, on the logit scale)

If the probability (p) that a transect is occupied by a species during a survey is 0.20, then:

$$\beta_0 = \text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \log\left(\frac{0.20}{0.80}\right) = -1.39$$

(where the ratio $\frac{p}{1-p}$ gives us the ‘odds of success’ for a binary variable. The logit transform is also known as the log odds of success)

The logit transform used in logistic regression converts a probability ranging from 0 to 1 to a log odds ranging from negative infinity to positive infinity. If probability of occupancy is 0.50, the log odds is 0. If probability of occupancy is less than 0.50, the log odds is negative; greater than 0.50, the log odds is positive. In other words, by modeling $\text{logit}(p)$ as a linear function of predictors, the estimates of p are properly constrained to be between the values of 0 and 1.

In a power analysis simulation with probability of occupancy changing over time from a baseline level, a logistic regression model with no additional predictors can be written as:

$$\text{logit}(p_t) = \beta_0 + (x_1)(t)$$

For example:

$$\text{logit}(p_t) = 0.40 + (-0.06)(5)$$

$$\text{logit}(p_t) = 0.1$$

$$p_t = 0.52$$

Here, initial probability of occupancy (p_0) is 0.60 ($\text{logit}(0.60) = 0.40$, which is β_0 in the above equation) and probability of occupancy declines with time (x_1 is a negative number). After 5 years, probability of occupancy is 0.52.

The change in probability of occupancy from the first year of monitoring depends on the initial probability of occupancy and on how much time has passed (monitoring duration). Table 6 shows the probability of occupancy for a sample of initial probabilities, monitoring durations and effect sizes.

Table 6. Probability of occupancy for a sample of initial probabilities, monitoring durations and effect sizes. ‘Duration (t)’ is number of years for which the occupancy trend is sustained; ‘Effect size’ represents an annual change in the probability of occupancy, on the logit scale; ‘ p_t ’ is the probability of occupancy after 5 or 10 years (duration) of the specified trend; ‘ $p_t - p_0$ ’ is the change in probability of occupancy from the initial to the final year of monitoring. Results are shown for three initial values for probability of occupancy: 0.10, 0.50 and 0.70. For example, if the initial probability of occupancy is 0.10 and an annual increase in occupancy of 0.060 (on the logit scale) is sustained for 10 years, then at the end of 10 years the probability of occupancy will be 0.168, an increase of 0.068 from initial probability of occupancy.

If initial probability of occupancy = 0.10 ($\beta_0 = -2.20$)...

Duration (t)	Effect size	p_t	$p_t - p_0$
5	-0.060	0.076	-0.024
5	-0.040	0.083	-0.017
5	-0.020	0.091	-0.009
5	0.000	0.100	0.000
5	0.020	0.109	0.009
5	0.040	0.119	0.019
5	0.060	0.130	0.030
10	-0.060	0.057	-0.043
10	-0.040	0.069	-0.031
10	-0.020	0.083	0.017
10	0.000	0.100	0.000
10	0.020	0.119	0.019
10	0.040	0.142	0.042
10	0.060	0.168	0.068

If initial probability of occupancy = 0.50 ($\beta_0 = 0$)...

Duration	Effect size	p_t	$p_t - p_0$
5	-0.060	0.426	-0.074
5	-0.040	0.450	-0.050
5	-0.020	0.475	-0.025
5	0.000	0.500	0.000
5	0.020	0.525	0.025
5	0.040	0.550	0.050
5	0.060	0.574	0.074
10	-0.060	0.354	-0.146
10	-0.040	0.401	-0.099
10	-0.020	0.450	-0.050
10	0.000	0.500	0.000
10	0.020	0.550	0.050
10	0.040	0.599	0.099
10	0.060	0.646	0.146

If initial probability of occupancy = 0.70 ($\beta_0 = 0.85$)...

Duration	Effect size	p_t	$p_t - p_0$
5	-0.060	0.634	-0.066
5	-0.040	0.656	-0.044
5	-0.020	0.679	-0.021
5	0.000	0.700	0.000
5	0.020	0.721	0.021
5	0.040	0.740	0.040
5	0.060	0.759	0.059
10	-0.060	0.562	-0.138
10	-0.040	0.610	-0.090
10	-0.020	0.656	-0.044
10	0.000	0.700	0.000
10	0.020	0.740	0.040
10	0.040	0.777	0.077
10	0.060	0.810	0.110

RESULTS

In this report, regression coefficients and random effect variances are presented on a logarithmic scale for JELA, GUIS, and SAAN (in accordance with negative binomial and Poisson regressions), and on a logit scale for VICK (in accordance with logistic regressions). In Results, we provide a brief summary followed by figures and tables for each site-species. In Discussion, we summarize overall findings across site-species.

The best-supported regression models for the nine site-species are presented in Table 7. Day-of-year ('my.day') was an important predictor of trap-counts for most site-species. The importance of other covariates varied among site-species. For example, temperature ('temp.f') was an important predictor of trap-counts for *Hyla mix* and *L. clamitans*, but not for *E. planirostris*, *E. guttolineata*, or *S. lateralis*.

Table 7. Summary of predictor variables in the best-supported regression model for each site-species. Models were fit based on year-round (YR) or non-breeding (NB) season data, as indicated in the ‘Season’ column. Parameters that were included in all candidate models (transect, monitoring year, and trap ID) are not shown in this table. Predictor variables are abbreviated as follows in the column headings: ‘nb’ = breeding season, a binary variable indicating if the survey was conducted during breeding or non-breeding months, only relevant for models based on year-round data; ‘my.day’ = number of days since October 1, the start of each monitoring year; ‘temp.f’ = air temperature, ‘rh’ = relative humidity, ‘precip.7’ = mean daily precipitation in the 7 days prior to a survey (or for JELA, 30 days prior); and ‘mean.prior.precip’ = mean monthly precipitation during the prior monitoring year. For JELA, GUIS and SAAN species, the last column shows the negative binomial dispersion parameter, with standard error (SE) in parentheses. If the best-supported model had a Poisson distribution, ‘Poisson’ is written instead. VICK species were analyzed using logistic regression, which does not include a dispersion parameter. In the table cells, ‘Y’ means the variable was in the best-supported model for that site-species, a dash ‘-’ means it was not in the best-supported model, and a blank cell means the variable was not analyzed for that site-species.

Site	Season	Species	nb	my.day	temp.f	rh	precip.7	mean. prior. precip	Dispersion (SE)
JELA	NB	<i>Hyla mix</i>		Y	Y	Y	-	-	4.16 (0.95)
JELA	NB	<i>Lithobates clamitans</i>		Y	Y		-	Y	0.90 (0.14)
JELA	NB	<i>Eleutherodactylus planirostris</i>		Y	-	Y	Y		Poisson
GUIS	NB	<i>Eleutherodactylus planirostris</i>		-	-	Y	Y	Y	1.56 (0.55)
GUIS	NB	<i>Hyla mix</i>		Y	Y	-	-	-	5.01 (3.36)
SAAN	YR	<i>Eleutherodactylus cystignathoides</i>	Y	Y			Y	-	1.06 (0.1)
SAAN	YR	<i>Incilius nebulifer</i>	Y	-			-	-	0.37 (0.03)
VICK	YR	<i>Eurycea guttolineata</i>	-	Y	-	-	-	-	
VICK	YR	<i>Scincella lateralis</i>	Y	Y	-	-	-	-	

JELA

We analyzed data for three JELA species that accounted for almost 70% of all captures at the site. Data for *Hyla squirella* and *Hyla cinerea* were combined as *Hyla* mix, because these sister species share the same ecological niche and some morphs can be difficult to visually distinguish. The three JELA species exhibited strong seasonality in trap-counts, with almost all captures occurring during non-breeding season months (October through March). We therefore focused our JELA analyses only on non-breeding season data.

Trap-counts for *Hyla* mix were highest in monitoring year 2012 and slowly declined in subsequent years. *L. clamitans* and *E. planirostris* populations peaked in monitoring year 2013. *E. planirostris* was not captured at JELA in monitoring years 2010 or 2011, but it's not known if this species was not present at JELA in those years or was present but not captured. In monitoring year 2014, trap-counts of *E. planirostris* were again extremely low. Approximately 13% of survey data were omitted from regression analyses for JELA species due to missing relative humidity or temperature information. JELA received low precipitation most years. Prior year's precipitation was highest for monitoring year 2011 and lowest for 2014, i.e., monitoring year 2010 was a relatively wet year and 2013 a very dry year.

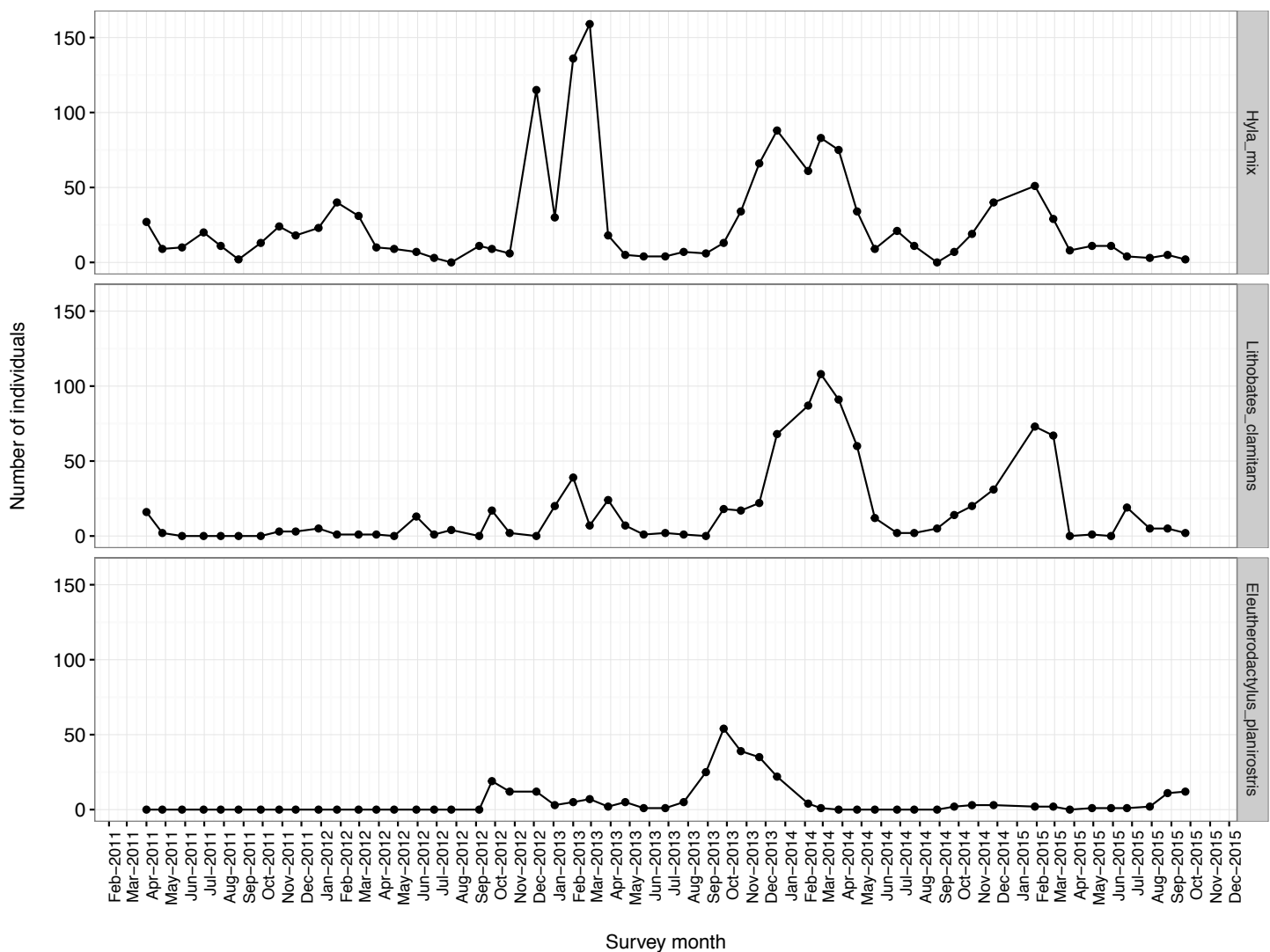


Figure 1. (JELA) Number of individuals captured each survey event, summed across traps, for three species (horizontal panels).

Table 8. Summary of JELA covariate data corresponding with survey dates, non-breeding season only. Minimum, median, and maximum values, and % of non-breeding season data missing for covariates: ‘temp.f’ = temperature (Fahrenheit), ‘rh’ = relative humidity (%), ‘precip.30’ = mean daily precipitation in 30 days prior to survey (inches). For JELA, we used precip.30 instead of precip.7 as a survey covariate because precip.7 data were not available for 17% of survey dates.

Statistic	temp.f	rh	precip.30
Minimum	42.38	58.60	0.00
Median	59.24	82.75	0.01
Maximum	70.99	98.51	0.18
% missing data	8.7	13.0	0

Table 9. (JELA) For each monitoring year, the mean monthly precipitation (in inches) for the prior monitoring year.

monitor.yr	mean.prior.precip
2011	3.63
2012	1.07
2013	1.63
2014	0.12
2015	0.69

JELA: *Hyla* mix

Trap-counts of JELA *Hyla* mix peaked between December 2012 and February 2013, reaching a maximum of 16 individuals in a single PVC set in a survey. Captures were highest in Transects F and G. When only non-breeding season months were considered, 35% of trap-counts were zero-counts (no animals captured). The best-supported model for JELA *Hyla* mix included day-of-year, temperature, and relative humidity. This was the only model in the AIC confidence set, i.e., no candidate model with likely informative parameters had AIC values within six units of the best-supported model.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 44% for 3 or 6 surveys per year. The proportion of false trends for these scenarios did not exceed 2%. Increasing the number of traps by 50% in existing or new transects improved power by only a few percentage points.

	2011-04-01	2011-04-26	2011-05-27	2011-06-30	2011-07-27	2011-08-24	2011-09-28	2011-10-27	2011-11-22	2011-12-28	2012-01-26	2012-02-29	2012-03-28	2012-04-25	2012-05-30	2012-06-27	2012-07-24	2012-08-26	2012-09-26	2012-10-24	2012-12-05	2013-01-03	2013-02-01	2013-02-27	2013-03-28	2013-04-24	2013-05-23	2013-06-26	2013-07-25	2013-08-29	2013-09-26	2013-10-23	2013-11-21	2013-12-19	2014-02-06	2014-02-26	2014-03-26	2014-04-24	2014-05-22	2014-06-26	2014-07-23	2014-08-28	2014-09-24	2014-10-22	2014-11-25	2015-01-29	2015-02-27	2015-03-25	2015-04-29	2015-05-29	2015-06-23	2015-07-29	2015-08-26	2015-09-23	
F_pvc1		1		2				1	1										1			10	2	9	15		1					1	2	1	3	3	1			1				2	2	3									
F_pvc2	1		1	1			1			1	3	1					1		1	1		9	7	12	14	1							4	8	7	5	4	3	1	2				3		4	3		1	2	1				
F_pvc3						2		1												1	1		6	1	11	10	3		1		1			1	4	2	1	3	1	2					1					1	1	1			
F_pvc4	1	1					1		1	1	1	1							1	1	1	2			8	4			1		1	1	1	6	3	1	5	5	1	1	3	1			4	2					1				
G_pvc1	4			2	1	1					1	2	1	1	1								5		6	7				1	1		1	3	2	6	6	2	4	1		3				1	1	1							
G_pvc2	1	2	1								2	1					1					4	1	8	8	1	1				1			2	3	4	3		3	1	1				1	2	2	1							
G_pvc3	3	1	2	2	1			1	1	2	1	5	5	1	1					1			11	3	12	12				1			1	3	3	3	3	2	1	1	1			2	2	1	2					1			
G_pvc4	1	1	2	1	1			1	1	1		2	1			1	1		2	1		10	5	16	15						1	2	2	5	9	6	6	5	3	1	3	3			2	1		2	1		1				
H_pvc1								1												1	2		1	1						1	1	1	3	3	2	6	3	2		1	1			2	2										
H_pvc2	1				1						2	1	1									4	2	1	5								1	3	1	3	4						2	1											
H_pvc3					1	1				2	3	2		1	1	1						4		7	7								2		1	4	3			2				1							1				
H_pvc4	2		1	1	2			1	2	1								1				3		6	5					1			3	5		3	3	1		1				3	3	1									
I_pvc1	2				1	2				2	1	3	4					1				2		5	5								3	3		3	5	2	1	1				5	4			2	2						
I_pvc2		1						1			1	1	1			1				1		5		4								1	1	3	5	2	6	2			1	1			2	5	4			2	3				
I_pvc3		1						1			2	2	1	1	1							7		4	3		2						1	3	5	3	1	4	4	3			1	3	6	5	4	1	2	1			1		
I_pvc4					1	2		2	1	1			1													1													1				1												
J_pvc1					1		1	2	1	1		2	1	1	1	2						9		9	9							2	2	1	4	4	2	1						2				1							
J_pvc2										1	1									1		3	3	3	4									1	1	1	1		2				1	1											
J_pvc3			1																1		1	3		5	6						1	1		1	4		1	1	1					2	1		1		1		1	1			
J_pvc4					1			1		1	1									1		2	2	5	7	1		1	2		1	1	2	3	3	4	2	4	2	2	2	1			1	1	2	3	1	3	3	1		1	1
K_pvc1	3				2		1	1	1	1	2	2	2	3	1							1		2	1								5	4	3	7	5	6	2			1					1								
K_pvc2	2		1	1				1	2	3	4	3										1			2	1								1	2	2	4	4	2						1	1	1								
K_pvc3	1				1			2		2	2	2	1							1		4	2	2	7	3								4	1	4	1	3	3						2	1									
K_pvc4	4				1	1			1	1	3	2	1		2				1	2	1	4		5	4	2		1		3			1	3	4	4	2	4	4	2			2		2	1	3	1	2	1				1	
K_pvc5				1				4			3	5							1		1	2		2	3	2	1	1					1	2	1	2	2	4	3	2	1	1	2		2	1	2	3	3						
K_pvc6	1					1			1													2	2	1																			1	2	1	1	1								

Figure 2. (JELA *Hyla* mix) Number of individuals captured per PVC trap set each survey event, year-round data. Rows are PVC trap sets, columns are survey events. The initial capital letter (F–K) in each row name indicates the transect. White cells indicate no animals were detected. Otherwise, number of detected individuals ranges from a low of 1 (blue) to a high of 16 (red). Note that this matrix shows data from all surveys, but regression analysis for JELA *Hyla* mix focused on non-breeding season data (October to March of each monitoring year) only.

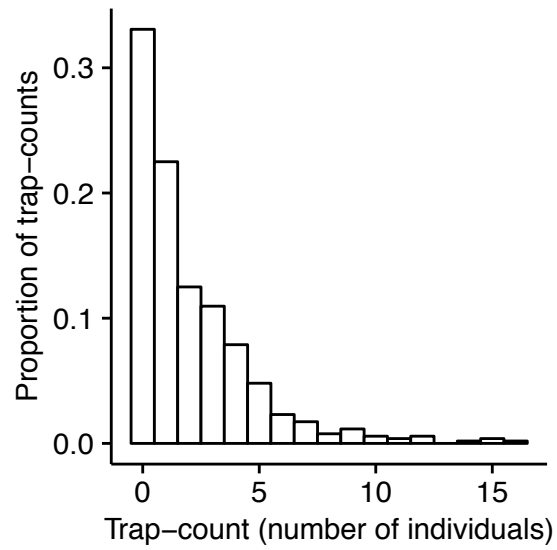


Figure 3. (*JELA Hyla mix*) Density histogram of trap-counts, non-breeding season only. For this site-species, 35% of trap-counts analyzed were zero.

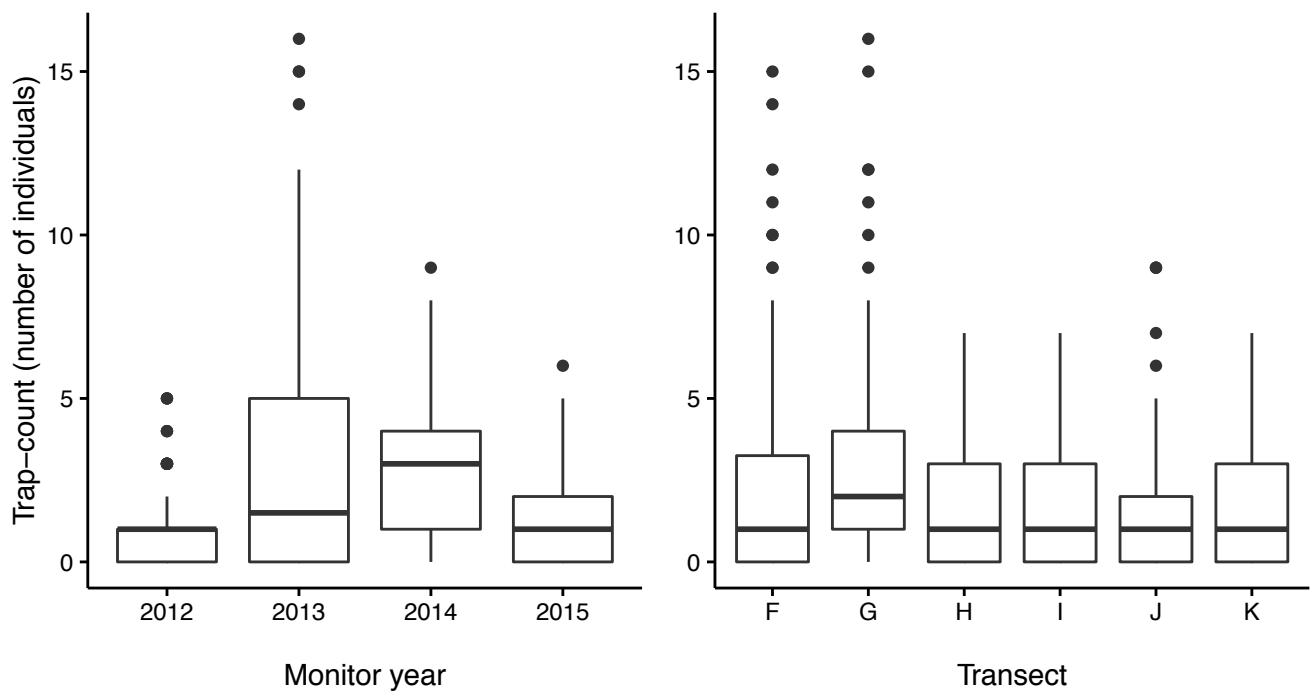


Figure 4. (*JELA Hyla mix*) Boxplots of detections by monitoring year (left) and transect (right), non-breeding season only.

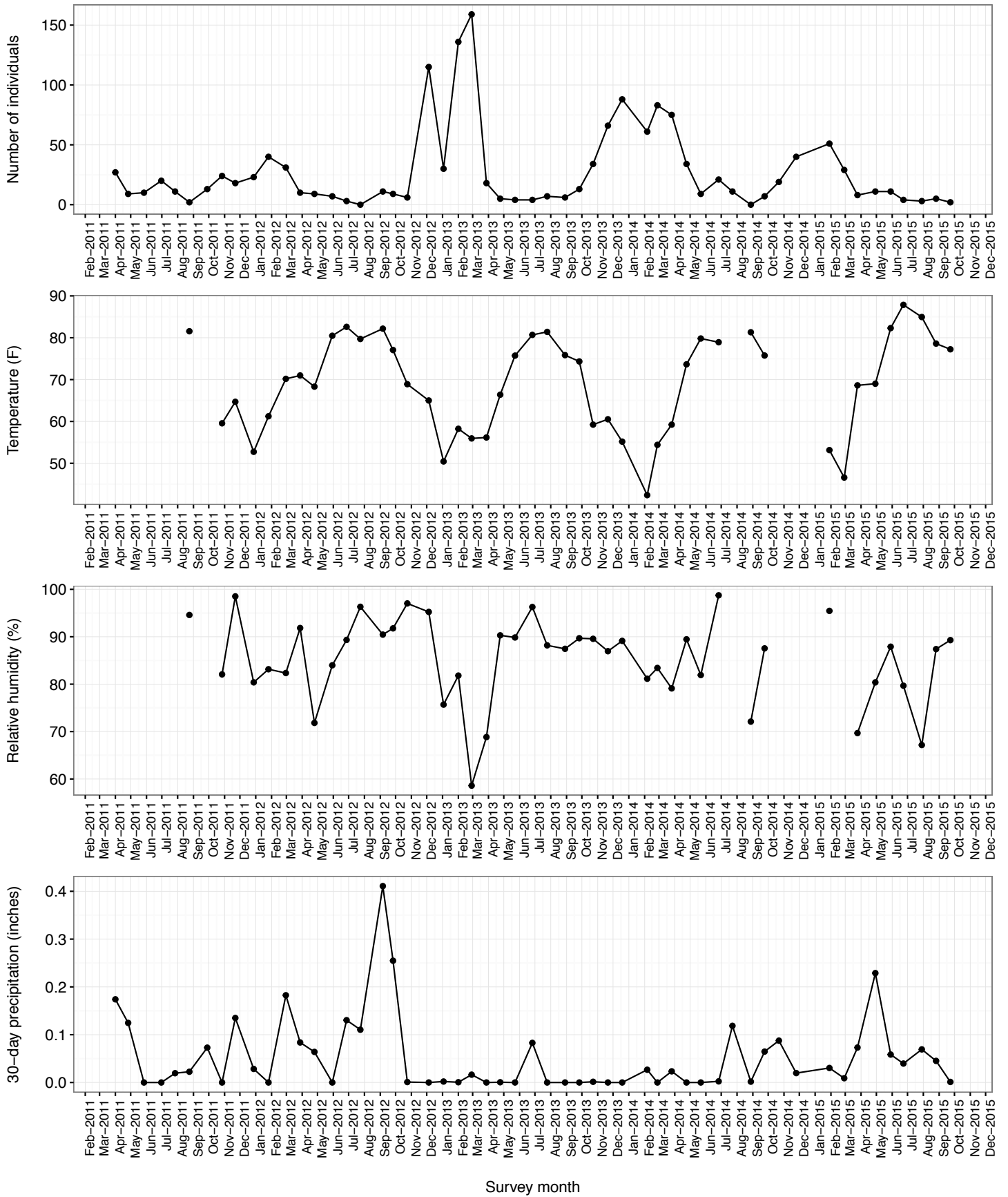


Figure 5. (*JELA Hyla mix*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 30 days prior to each survey event.

Table 10. (JELA Hyla mix) Best-supported regression model for JELA Hyla mix, non-breeding season only.
The fixed-effect covariates ‘my.day’ (day-of-year), ‘temp.f’ (temperature), and ‘rh’ (relative humidity) were centered and scaled so regression coefficients could be presented in standardized units.

```
##
## Call:
## glmnmb(formula = count ~ transect + my.day + I(my.day^2) +
##      temp.f + I(temp.f^2) + rh + I(rh^2) + (1 | monitor.yr) +
##      (1 | trap.id) - 1, data = jela.hmix.scale.nb.select.rh, family = "nbinom",
##      link = "log")
##
## AIC: 1810.2
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## transectJELA_F  1.2756    0.3567   3.58 0.00035 ***
## transectJELA_G  1.3773    0.3559   3.87 0.00011 ***
## transectJELA_H  0.7741    0.3613   2.14 0.03215 *
## transectJELA_I  0.6910    0.3648   1.89 0.05818 .
## transectJELA_J  0.8002    0.3611   2.22 0.02670 *
## transectJELA_K  0.9718    0.3414   2.85 0.00442 **
## my.day          0.0943    0.0678   1.39 0.16424
## I(my.day^2)     -0.6757    0.0730  -9.25 < 2e-16 ***
## temp.f          0.1024    0.0768   1.33 0.18227
## I(temp.f^2)     -0.1752    0.0349  -5.03 5.0e-07 ***
## rh              0.0769    0.0754   1.02 0.30769
## I(rh^2)         0.1408    0.0337   4.18 2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=520, monitor.yr=4, trap.id=26
## Random effect variance(s):
## Group=monitor.yr
##      Variance StdDev
## (Intercept)  0.3264 0.5713
## Group=trap.id
##      Variance StdDev
## (Intercept)  0.1215 0.3486
##
## Negative binomial dispersion parameter: 4.159 (std. err.: 0.94875)
##
## Log-likelihood: -890.087
```

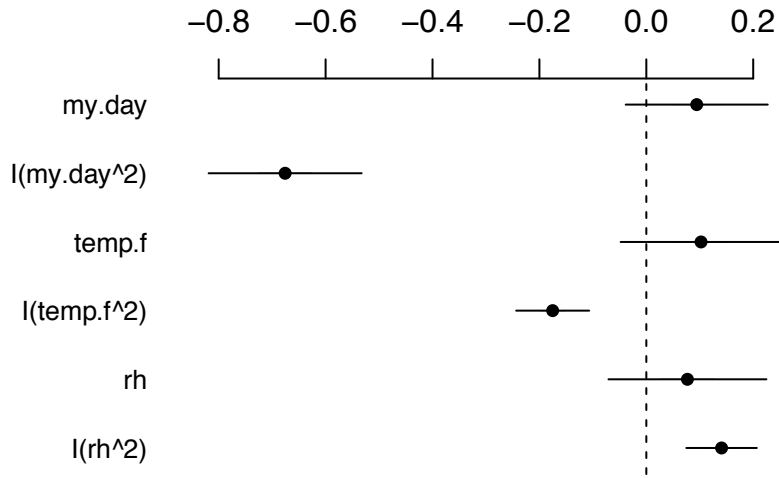


Figure 6. (*JELA Hyla mix*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes. For *JELA Hyla mix*, the confidence set only includes the best-supported model.

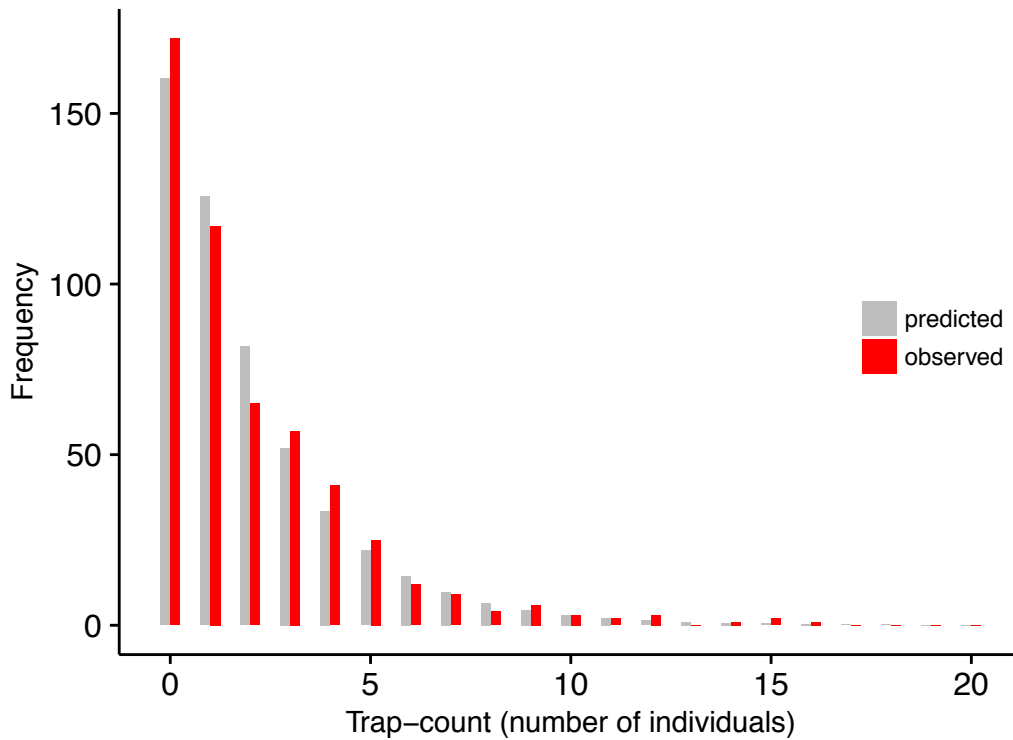


Figure 7. (*JELA Hyla mix*) Frequency histogram of trap-counts for observed data vs. model predictions, non-breeding season only.

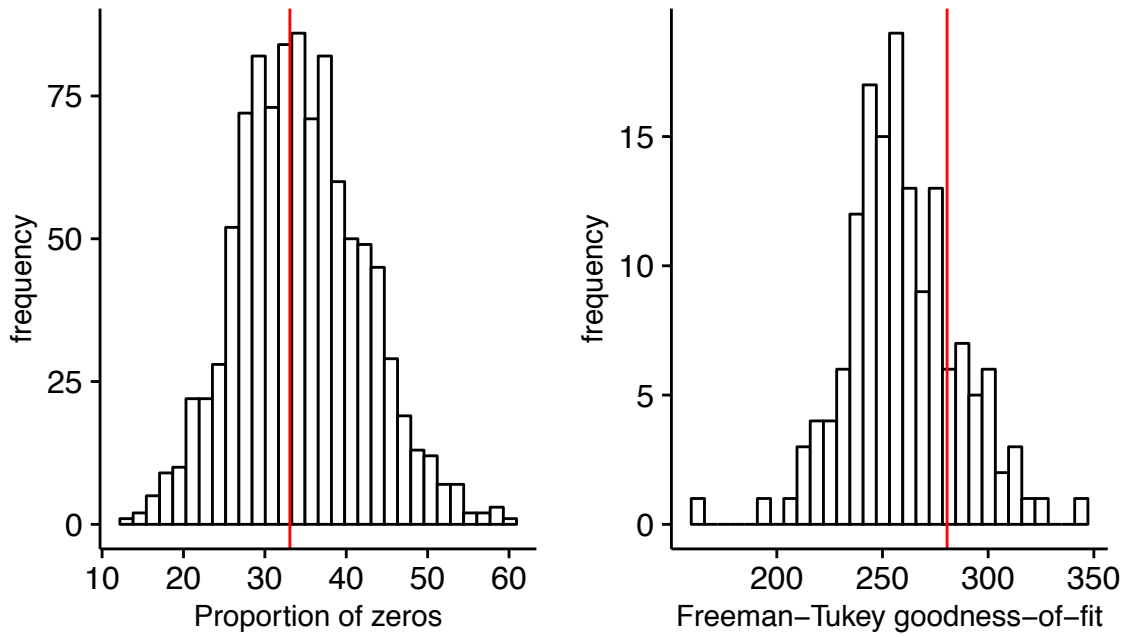


Figure 8. (*JELA Hyla mix*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is adequate. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

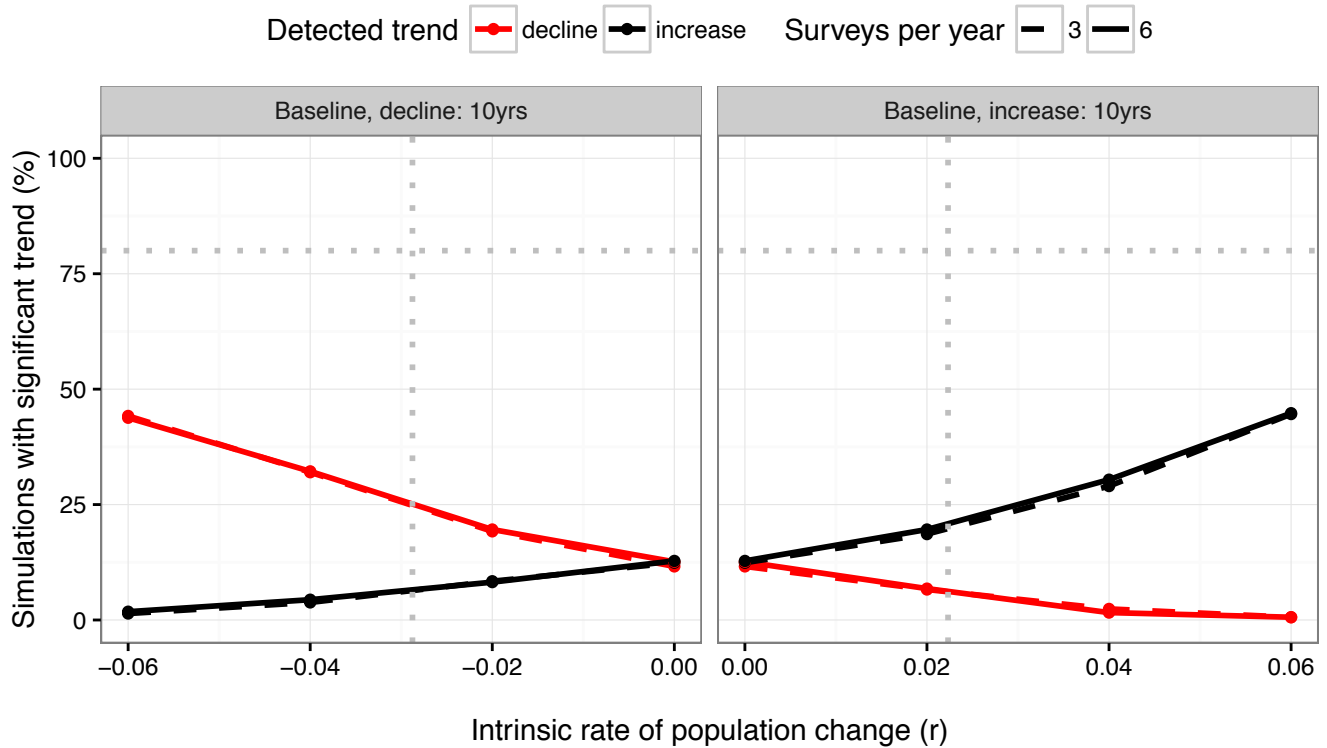


Figure 9. (*JELA Hyla mix*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), non-breeding season only. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 3 surveys per monitoring year (once in each month of November, January, and March). Solid lines represent 6 surveys per monitoring year (monthly from October through March). The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 11. (JELA Hyla mix) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate alpha. Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p -values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for JELA Hyla mix 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	32 / 0	44 / 1
-0.06	Baseline	6	33 / 1	44 / 2
-0.06	50% more (old transects)	3	33 / 1	47 / 1
-0.06	50% more (old transects)	6	34 / 1	49 / 1
-0.06	50% more (new transects)	3	34 / 1	47 / 1
-0.06	50% more (new transects)	6	35 / 1	48 / 1
-0.04	Baseline	3	20 / 2	32 / 4
-0.04	Baseline	6	20 / 2	32 / 4
-0.04	50% more (old transects)	3	23 / 2	31 / 3
-0.04	50% more (old transects)	6	23 / 2	34 / 3
-0.04	50% more (new transects)	3	22 / 1	33 / 3
-0.04	50% more (new transects)	6	28 / 5	34 / 7
-0.02	Baseline	3	11 / 4	19 / 8
-0.02	Baseline	6	13 / 5	20 / 8
-0.02	50% more (old transects)	3	15 / 3	22 / 7
-0.02	50% more (old transects)	6	15 / 3	22 / 6
-0.02	50% more (new transects)	3	15 / 3	22 / 7
-0.02	50% more (new transects)	6	15 / 3	22 / 6
0.00	Baseline	3	6 / 8	12 / 12
0.00	Baseline	6	7 / 8	13 / 13
0.00	50% more (old transects)	3	8 / 7	14 / 12
0.00	50% more (old transects)	6	8 / 7	14 / 11
0.00	50% more (new transects)	3	8 / 7	14 / 12
0.00	50% more (new transects)	6	8 / 7	14 / 12

JELA: *Lithobates clamitans*

Most captures of JELA *L. clamitans* occurred between December 2013 and April 2014, with a smaller surge in trap-counts between November 2014 and February 2015. The largest number of captures in a single PVC set in a survey was 15 animals. Few animals were captured in the first 1.5 years of monitoring. When only non-breeding season months were considered, 63% of trap-counts were zero-counts (no animals captured).

The best-supported model for JELA *L. clamitans* included day-of-year, temperature, and prior year's precipitation. The AIC confidence set included three other models with AIC values within six units of the best-supported model. These models differed from the best-supported model in that they did not include prior year's precipitation and/or day-of-year was a linear rather than a quadratic function of trap-counts. The confidence set models that did not include prior year's precipitation had much larger coherent interannual variation (0.75 and 0.81, compared to 0.33 for the best-supported model). In all confidence set models, day-of-year and temperature were still important predictors of trap-counts. Regression coefficients, trap variance, and the negative binomial dispersion parameter were similar across all confidence set models.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 36% for 3 surveys per year and 41% for 6 surveys per year. The proportion of false trends for these scenarios was 1%.

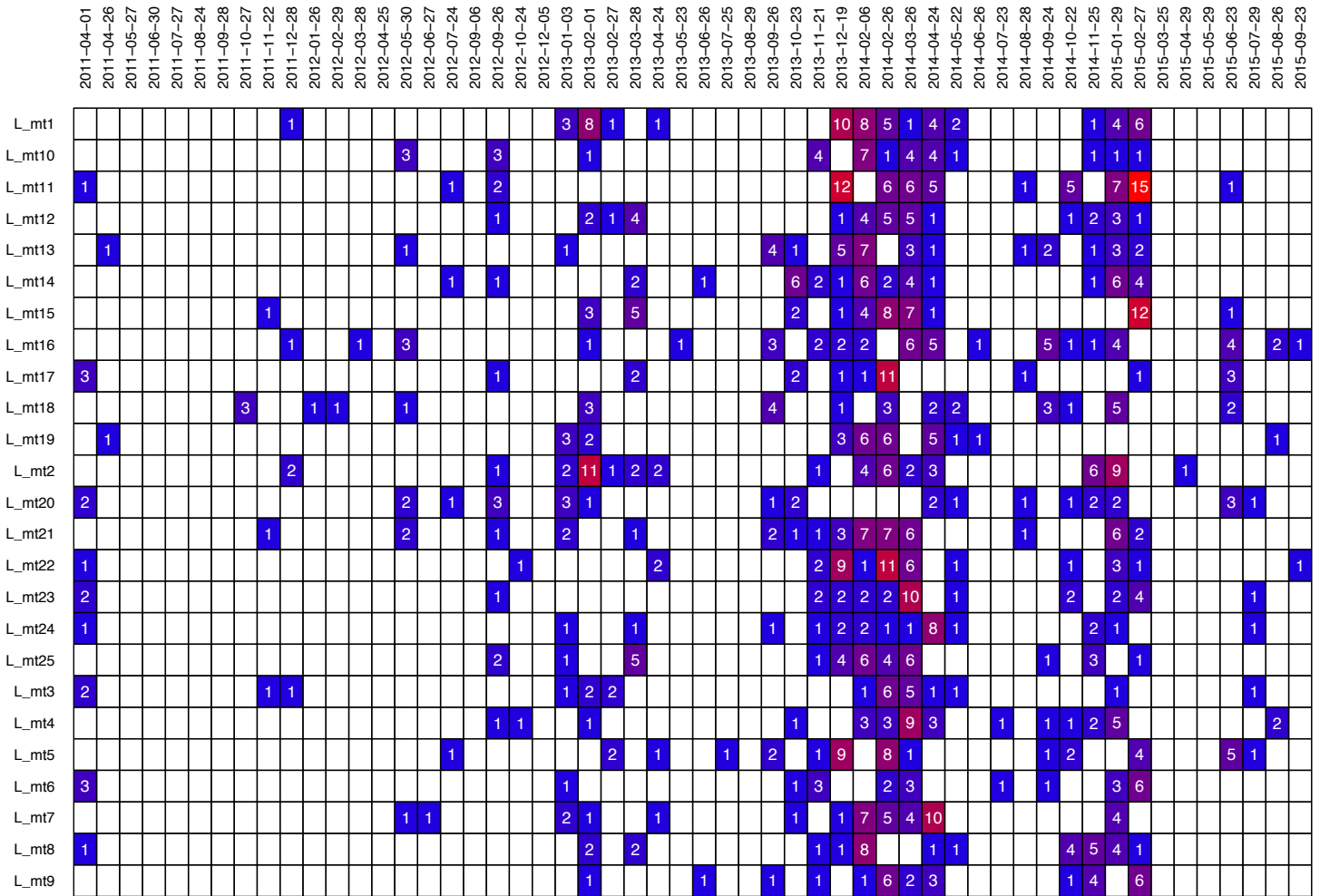


Figure 10. (JELA *L. clamitans*) Number of individuals captured per minnow trap each survey event, year-round data. Rows are minnow traps, columns are survey events. All traps are in a single transect (Transect L). White cells indicate no animals were detected. Otherwise, number of detected individuals ranges from a low of 1 (blue) to a high of 15 (red). Note that this matrix shows data from all surveys, but regression analysis for JELA *L. clamitans* focused on non-breeding season data (October to March of each monitoring year) only.

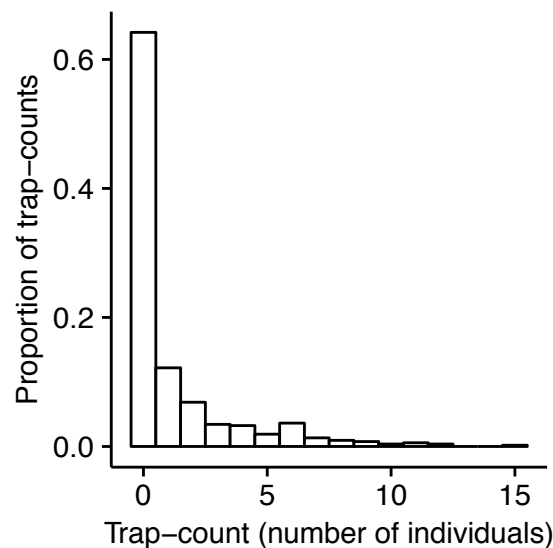


Figure 11. (JELA *L. clamitans*) Density histogram of trap-counts, non-breeding season only. For this site-species, 63% of trap-counts analyzed were zero.

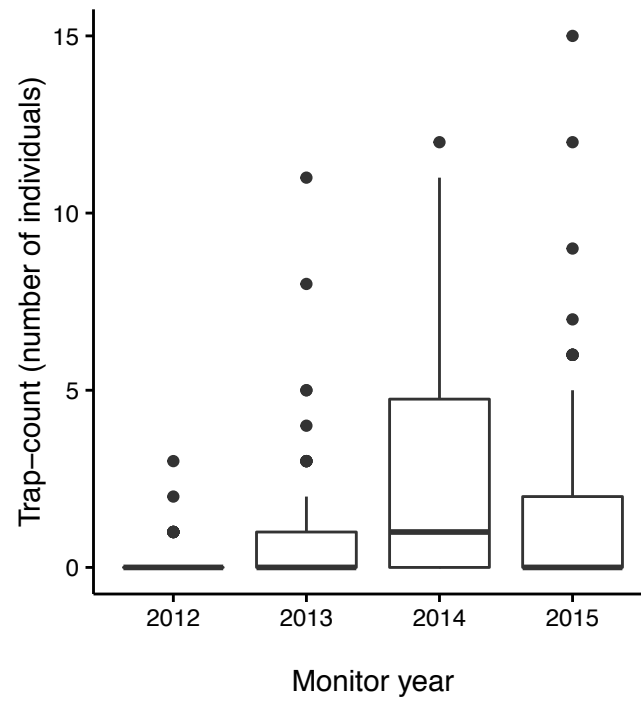


Figure 12. (JELA L. clamitans) Boxplots of detections by monitoring year, non-breeding season only.

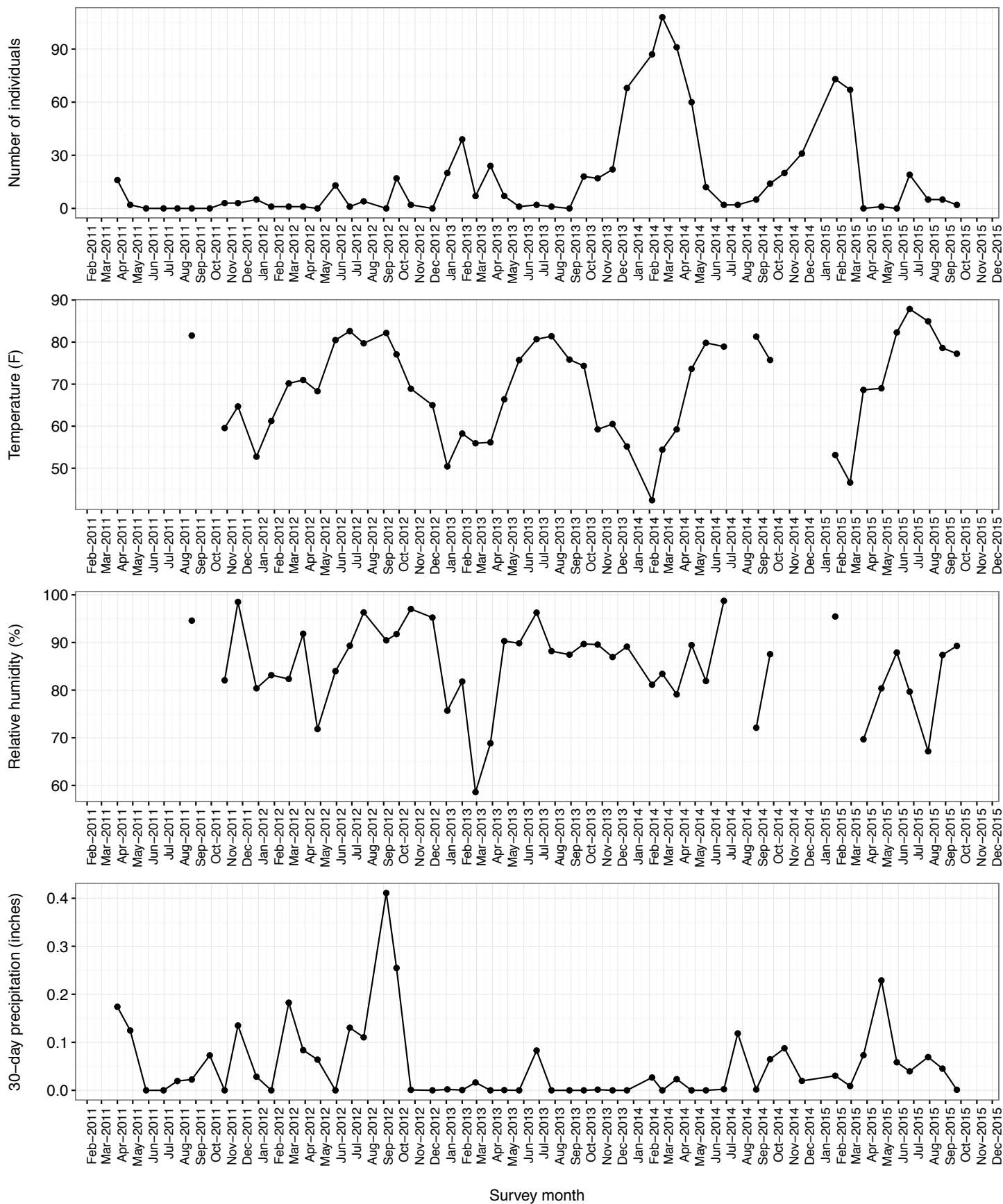


Figure 13. (*JELA L. clamitans*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 30 days prior to each survey event.

Table 12. (*JELA L. clamitans*) Best-supported regression model for *JELA L. clamitans*, non-breeding season only. Transect was not included in the model because all captures occurred on one transect (Transect L). The fixed-effect covariates ‘my.day’ (day-of-year), ‘temp.f’ (temperature), and ‘mean.prior.precip’ (mean monthly precipitation during the prior monitoring year) were centered and scaled so regression coefficients could be presented in standardized units.

```
##
## Call:
## glmnmb(formula = count ~ my.day + I(my.day^2) + temp.f + I(temp.f^2) +
##      mean.prior.precip + (1 | monitor.yr) + (1 | trap.id), data = jela.lc.scale.nb.select.temp,
##      family = "nbinom", link = "log")
##
## AIC: 1250.5
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.2368    0.3268  -0.72    0.469
## my.day          0.3925    0.0921   4.26 2.0e-05 ***
## I(my.day^2)    -0.1734    0.1177  -1.47   0.141
## temp.f         -1.0722    0.1722  -6.23 4.8e-10 ***
## I(temp.f^2)    -0.4564    0.0843  -5.41 6.2e-08 ***
## mean.prior.precip -0.6909    0.3137  -2.20  0.028 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=525, monitor.yr=4, trap.id=25
## Random effect variance(s):
## Group=monitor.yr
##              Variance StdDev
## (Intercept)    0.3278 0.5725
## Group=trap.id
##              Variance StdDev
## (Intercept)    0.04353 0.2086
##
## Negative binomial dispersion parameter: 0.8997 (std. err.: 0.14484)
##
## Log-likelihood: -616.244
```

Table 13. (*JELA L. clamitans*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. Year and trap variance are random effect variances estimated in the model. ‘Dispersion parameter’ is associated with the variance of the negative binomial distribution.

Model	dAIC	Number of parameters	Coherent interannual variation	Trap variance	Dispersion parameter	Difference from global model
mod1	0.00	9	0.33	0.04	0.90	remove precip.30
mod2	0.17	8	0.30	0.05	0.88	remove precip.30 and make my.day linear
mod3	1.06	8	0.81	0.04	0.90	remove precip.30 and mean.prior.precip
mod4	1.19	7	0.75	0.05	0.88	remove precip.30 and mean.prior.precip and make my.day linear

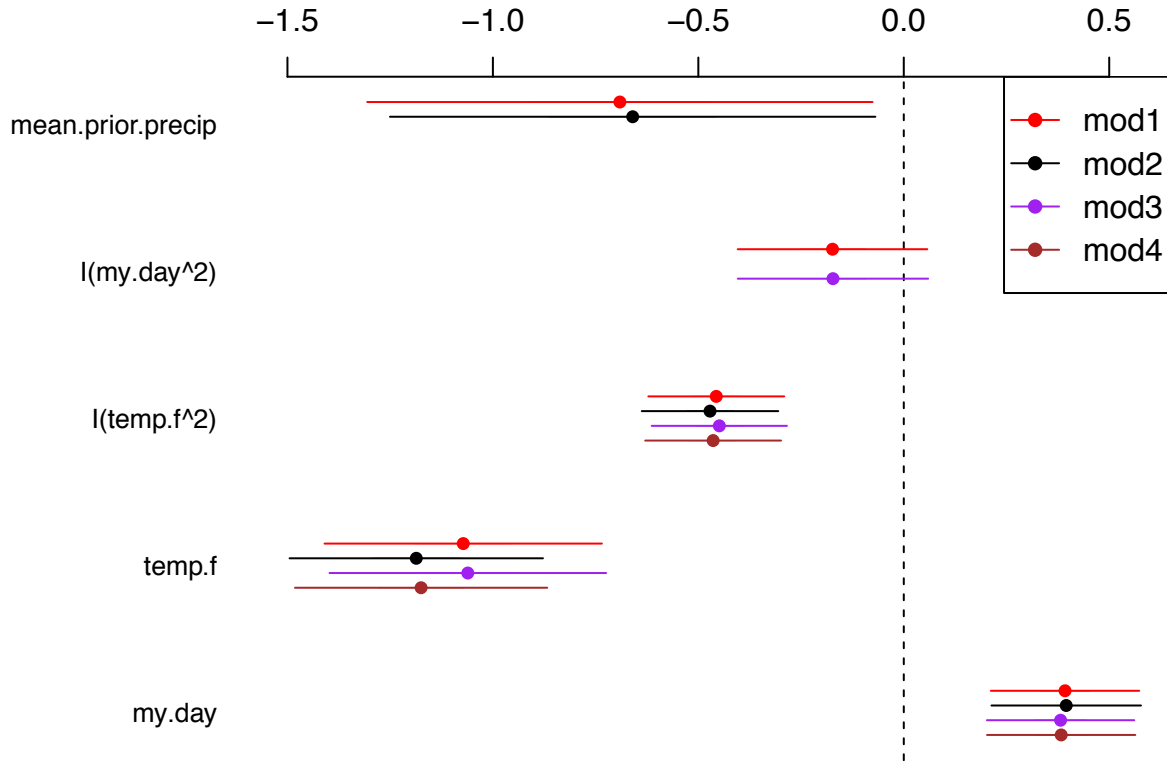


Figure 14. (*JELA L. clamitans*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

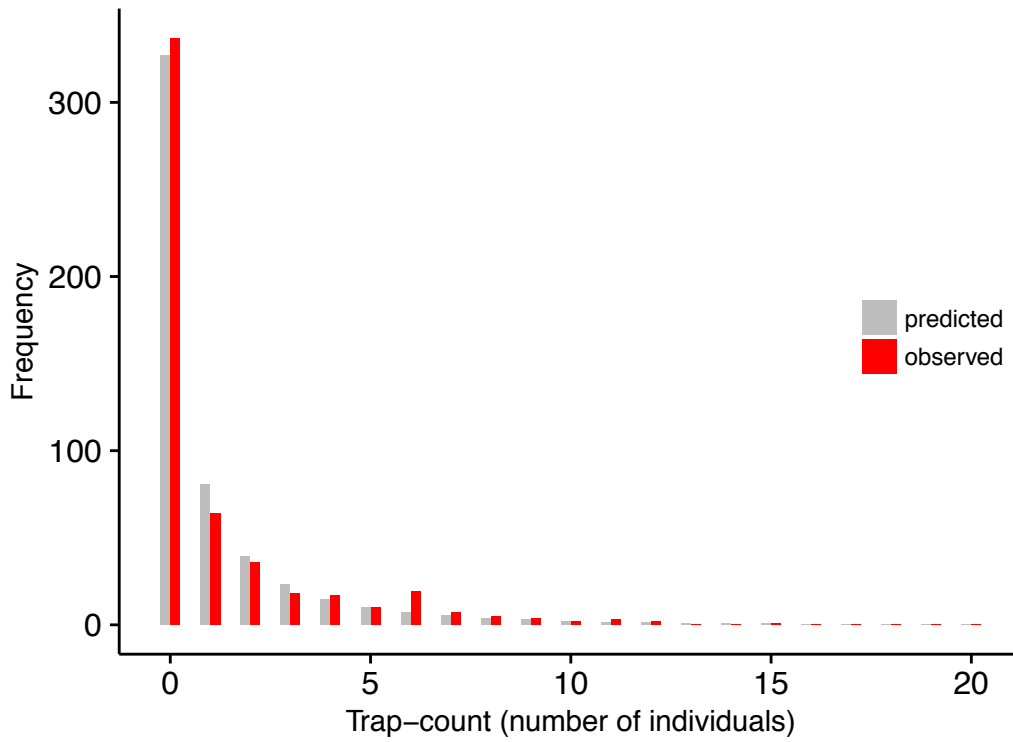


Figure 15. (*JELA L. clamitans*) Frequency histogram of trap-counts for observed data vs. model predictions, non-breeding season only.

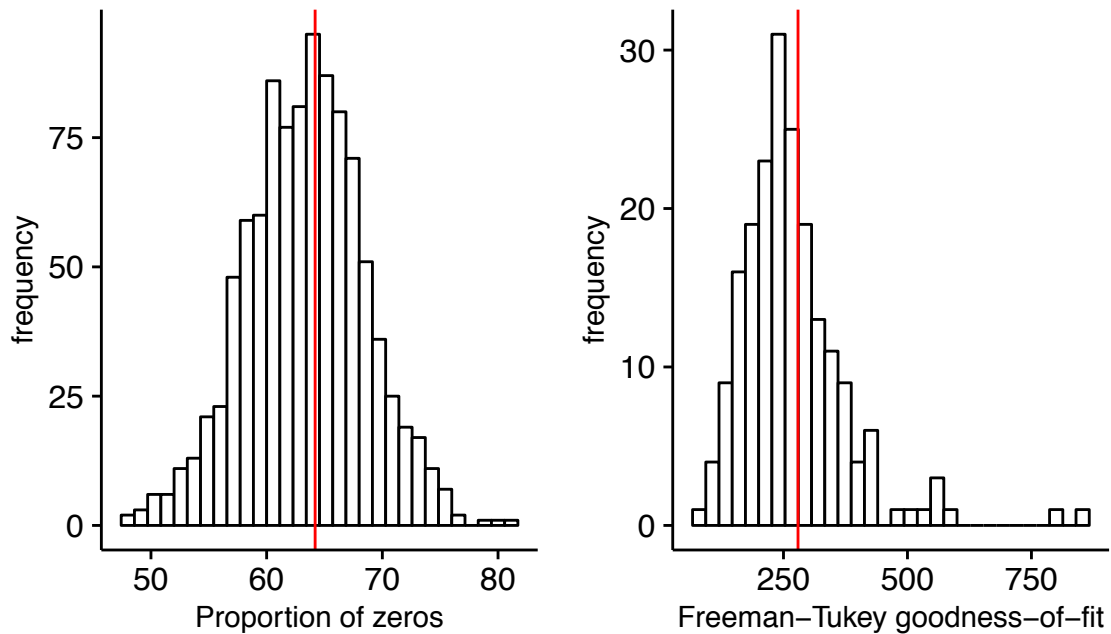


Figure 16. (*JELA L. clamitans*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is good. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

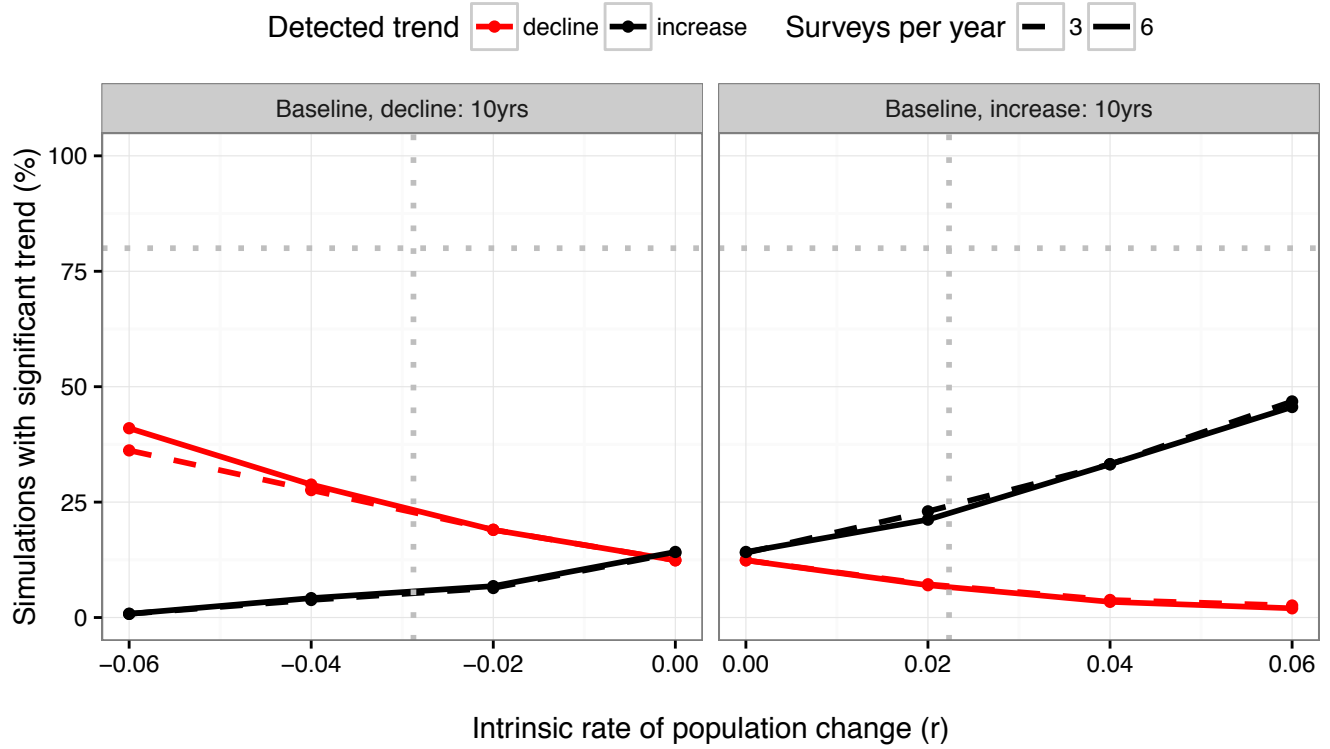


Figure 17. (*JELA L. clamitans*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), non-breeding season only. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 3 surveys per monitoring year (once in each month of November, January, and March). Solid lines represent 6 surveys per monitoring year (monthly from October through March). The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 14. (*JELA L. clamitans*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of α (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate α . Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p-values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for *JELA L. clamitans* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant ($\alpha=0.10$)	% significant ($\alpha=0.20$)
-0.06	Baseline	3	26 / 1	36 / 1
-0.06	Baseline	6	28 / 0	41 / 1
-0.04	Baseline	3	19 / 1	28 / 4
-0.04	Baseline	6	19 / 1	29 / 4
-0.02	Baseline	3	11 / 4	19 / 6
-0.02	Baseline	6	13 / 4	19 / 7
0.00	Baseline	3	7 / 7	12 / 14
0.00	Baseline	6	7 / 7	12 / 14

JELA: *Eleutherodactylus planirostris*

E. planirostris was not captured at JELA prior to September 2012. This species was primarily captured during monitoring years 2012 and 2013. The largest number of captures at a single coverboard in a survey was 6 animals. When only non-breeding season months were considered, 85% of trap-counts were zero-counts (no animals captured).

The best-supported model for JELA *E. planirostris* was Poisson-distributed (i.e., dispersion parameter $K = 0$) and included day-of-year, recent precipitation, and relative humidity. The AIC confidence set included two other models with AIC values within six units of the best-supported model. These models differed from the best-supported model in that they did not include recent precipitation, and the response variable in one model used a negative binomial rather than a Poisson distribution. The confidence set models that excluded recent precipitation had much larger coherent interannual variation (4.92, compared to 3.35 for the best-supported model). In all confidence set models, day-of-year and relative humidity were still important predictors of trap-counts. Regression coefficients and trap variance were similar across all confidence set models.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

Power to detect trends was very low for JELA *E. planirostris*. With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 19% for 3 surveys per year and 20% for 6 surveys per year. The proportion of false trends for these scenarios was 8%–9%. Increasing the number of traps by 50% in existing or new transects improved power by up to 4%.

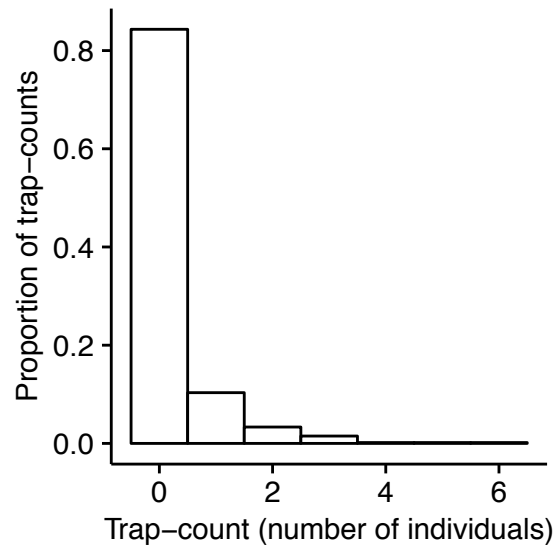


Figure 19. (*JELA E. planirostris*) Density histogram of trap-counts, non-breeding season only. For this site-species, 85% of trap-counts analyzed were zero.

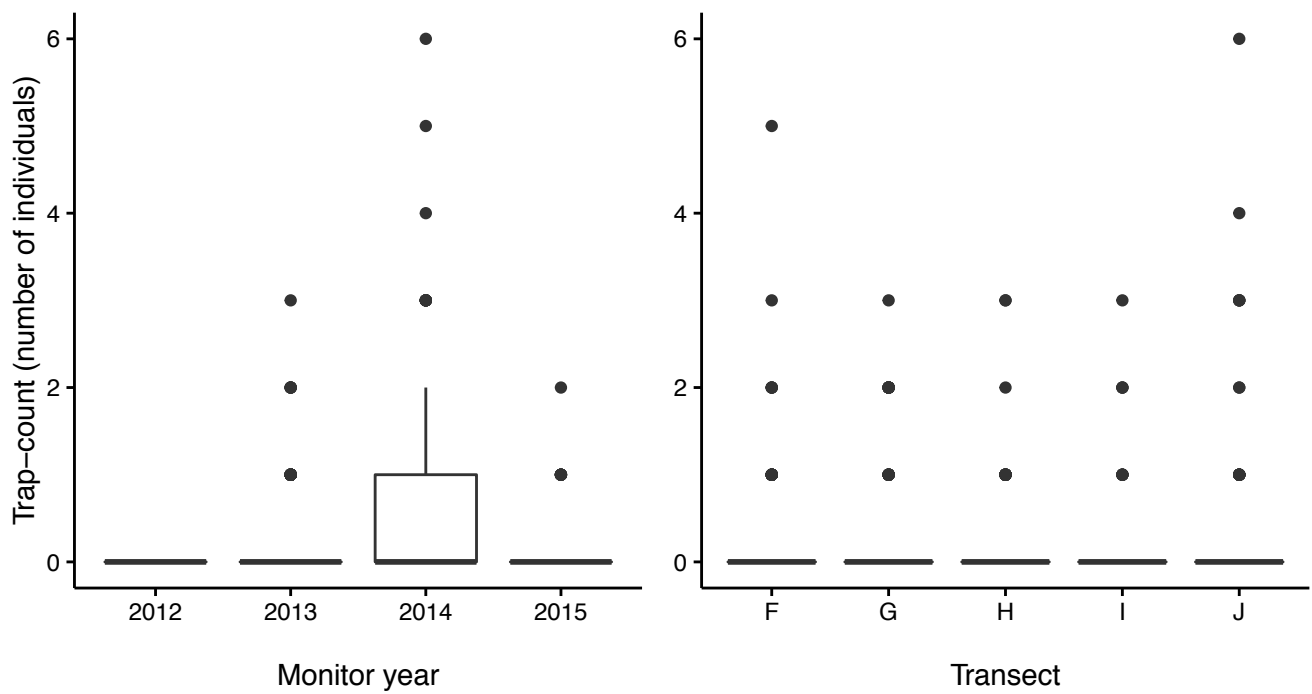


Figure 20. (*JELA E. planirostris*) Boxplots of detections by monitoring year (left) and transect (right), non-breeding season only.

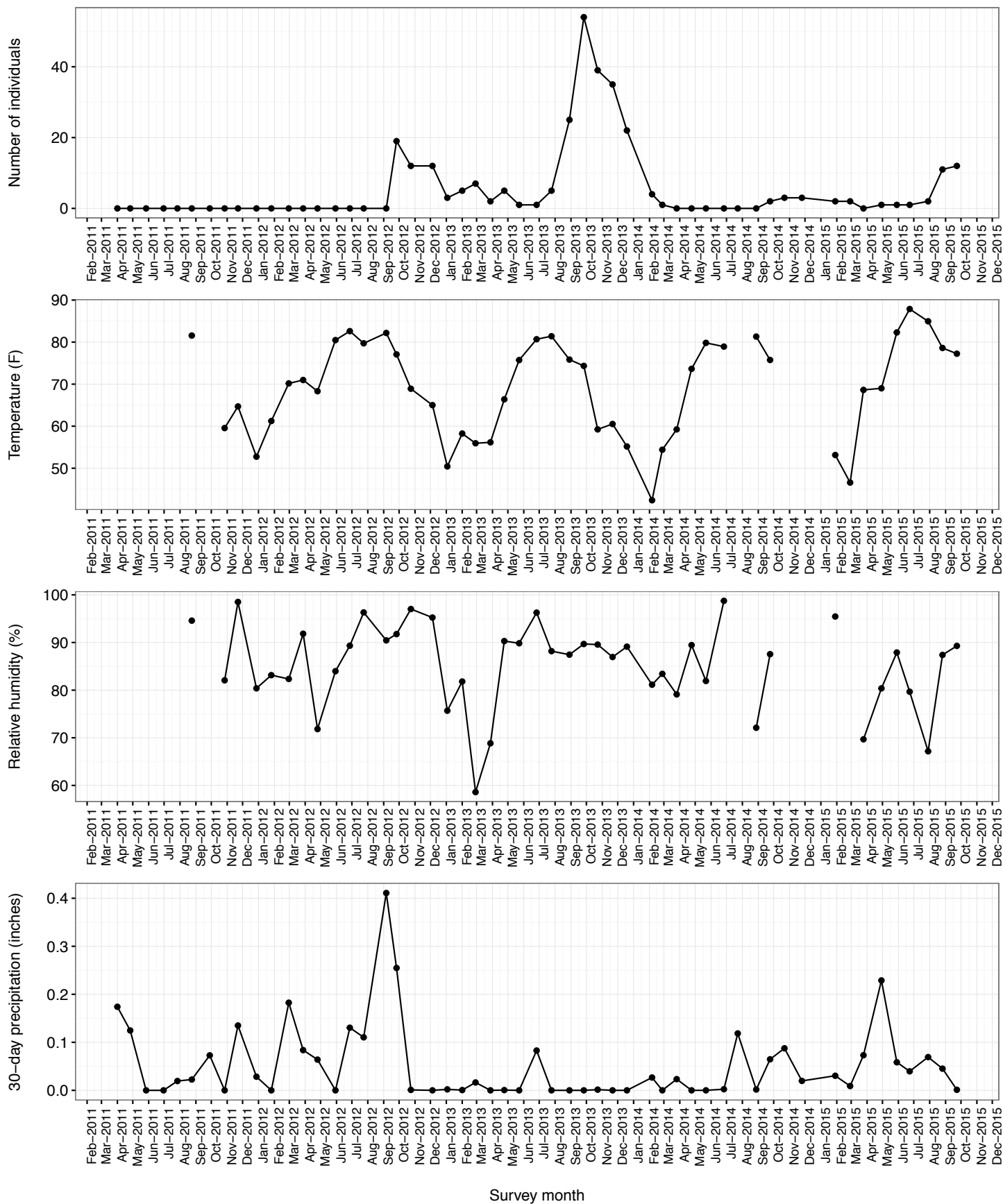


Figure 21. (*JELA E. planirostris*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 30 days prior to each survey event.

Table 15. (*JELA E. planirostris*) Best-supported regression model for *JELA E. planirostris*, non-breeding season only. The fixed-effect covariates ‘my.day’ (day-of-year), ‘precip.30’ (mean daily precipitation in 30 days prior to survey), and ‘rh’ (relative humidity) were centered and scaled so regression coefficients could be presented in standardized units. The best-supported model was Poisson-distributed, not negative binomial-distributed.

```
##
## Call:
## glmmadmb(formula = count ~ transect + my.day + I(my.day^2) +
##      precip.30 + rh + I(rh^2) + (1 | monitor.yr) + (1 | trap.id) -
##      1, data = jela.ep.nb.select.rh, family = "poisson", link = "log")
##
## AIC: 532.8
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## transectJELA_F  2.35e+01  7.56e+00   3.11  0.0019 **
## transectJELA_G  2.38e+01  7.55e+00   3.15  0.0016 **
## transectJELA_H  2.33e+01  7.56e+00   3.09  0.0020 **
## transectJELA_I  2.28e+01  7.56e+00   3.01  0.0026 **
## transectJELA_J  2.35e+01  7.56e+00   3.11  0.0019 **
## my.day          6.45e-03  8.84e-03   0.73  0.4659
## I(my.day^2)     -1.49e-04  5.76e-05  -2.59  0.0096 **
## precip.30       -3.11e+01  2.01e+01  -1.55  0.1219
## rh              -5.87e-01  1.91e-01  -3.08  0.0021 **
## I(rh^2)         3.42e-03  1.20e-03   2.85  0.0043 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=600, monitor.yr=4, trap.id=30
## Random effect variance(s):
## Group=monitor.yr
##              Variance StdDev
## (Intercept)   3.351  1.831
## Group=trap.id
##              Variance StdDev
## (Intercept)   0.2161 0.4649
##
##
## Log-likelihood: -254.397
```

Table 16. (*JELA E. planirostris*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. Year and trap variance are random effect variances estimated in the model. ‘Dispersion parameter’ is associated with the variance of the negative binomial distribution.

Model	dAIC	Number of parameters	Coherent interannual variation	Trap variance	Dispersion parameter	Difference from global model
mod1	0.00	12	3.35	0.22	NA	remove temp and make precip.30 linear; use Poisson
mod2	1.22	11	4.92	0.22	NA	remove temp and precip.30; use Poisson
mod3	2.41	12	4.92	0.21	7.92	remove temp and precip.30

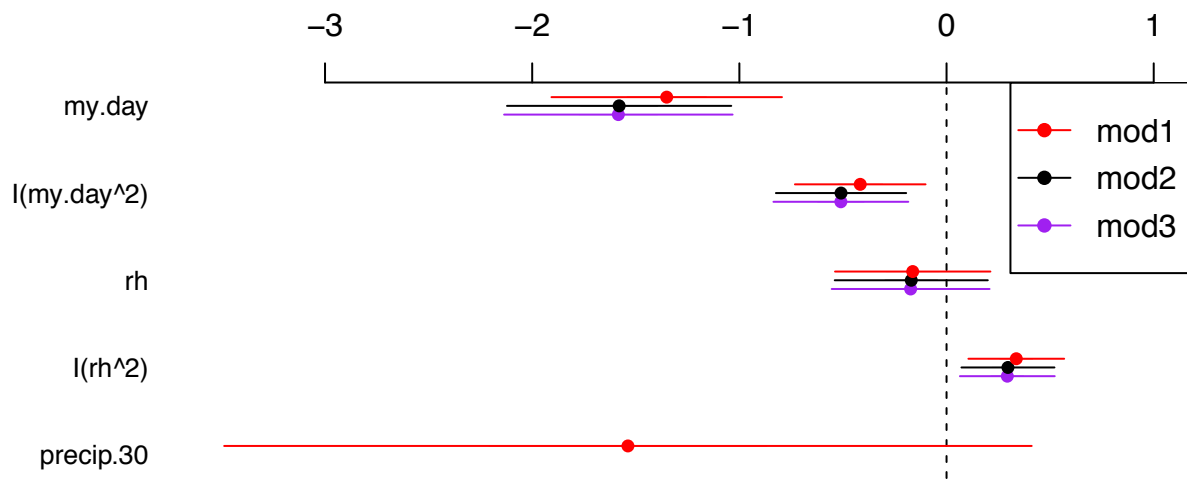


Figure 22. (*JELA E. planirostris*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

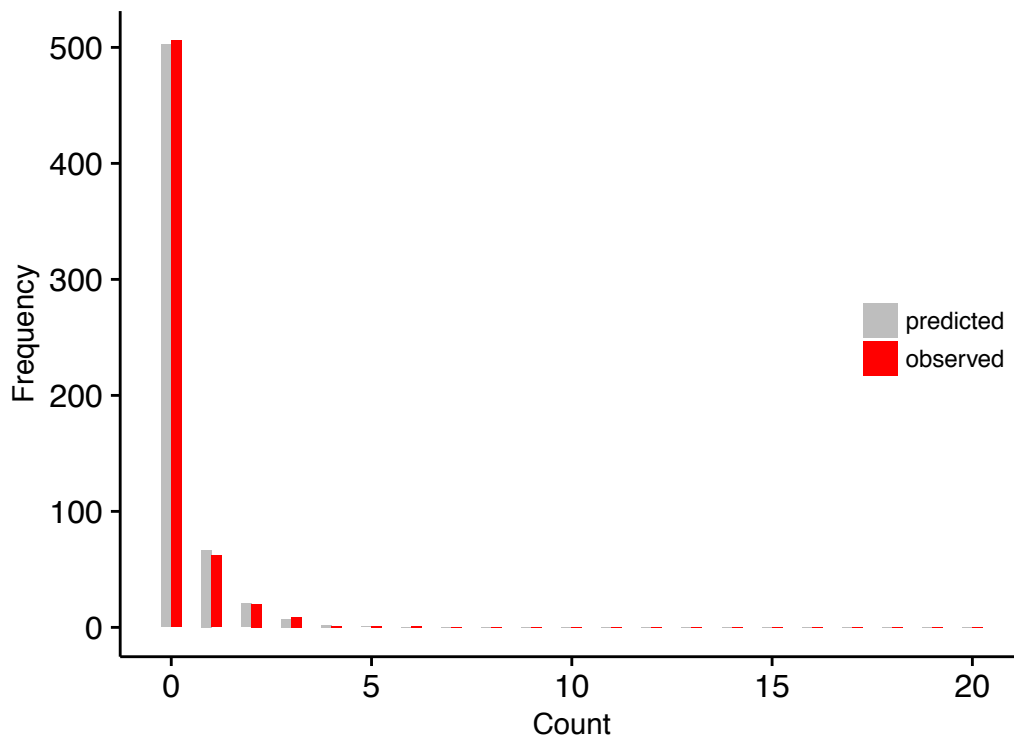


Figure 23. (*JELA E. planirostris*) Frequency histogram of trap-counts for observed data vs. model predictions, non-breeding season only.

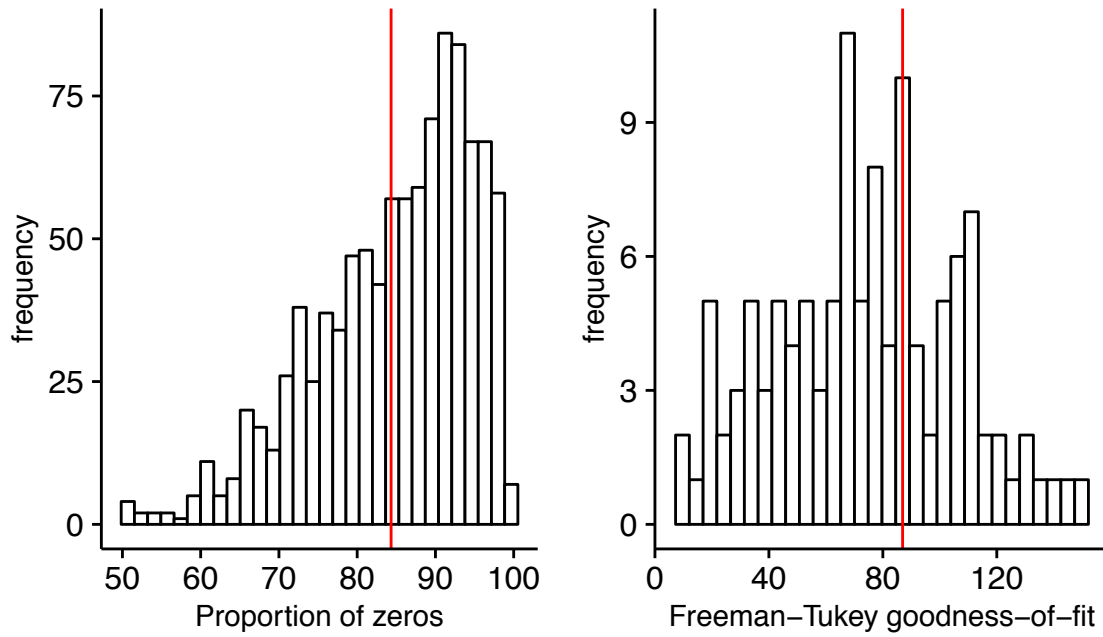


Figure 24. (*JELA E. planirostris*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Proportion of zeros cannot exceed 100%, so the distribution on the left is non-normal. Model fit is adequate. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

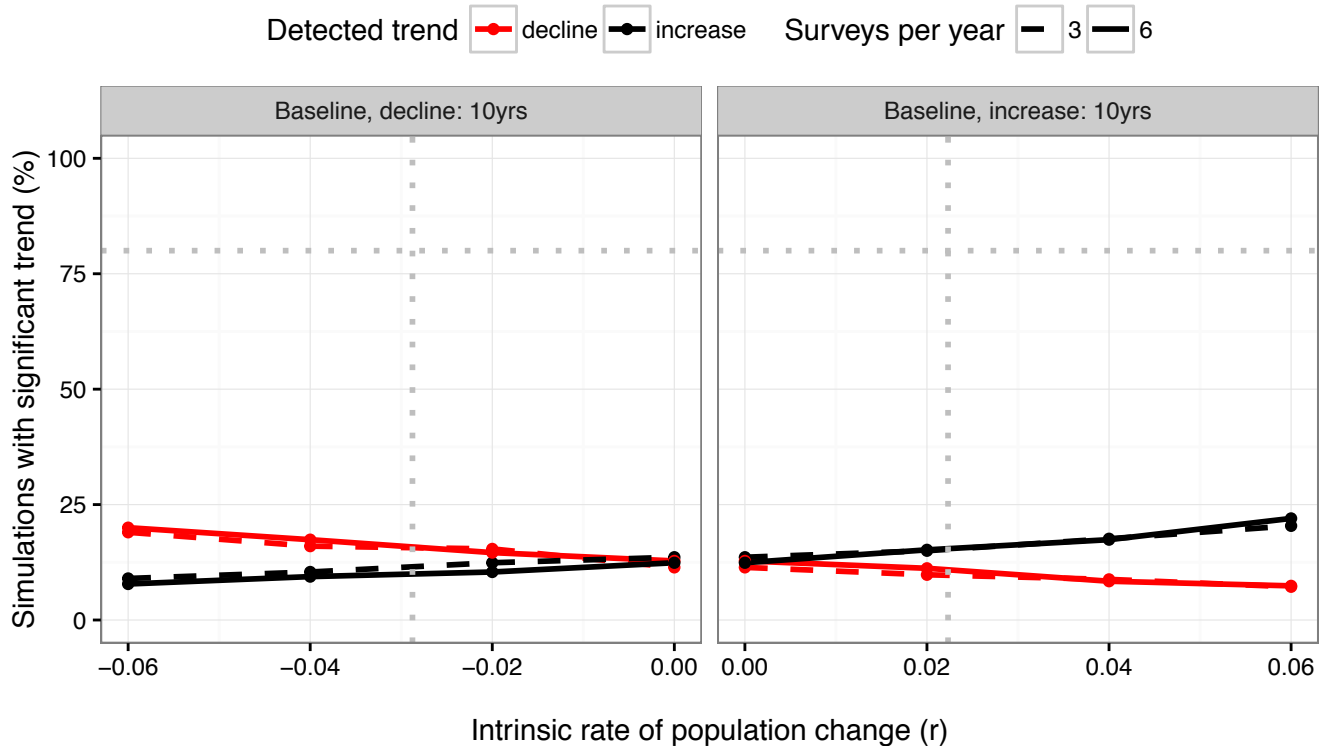


Figure 25. (*JELA E. planirostris*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), non-breeding season only. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 3 surveys per monitoring year (once in each month of November, January, and March). Solid lines represent 6 surveys per monitoring year (monthly from October through March). The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 17. (*JELA E. planirostris*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate alpha. Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p -values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for *JELA E. planirostris* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	11 / 4	19 / 9
-0.06	Baseline	6	13 / 4	20 / 8
-0.06	50% more (old transects)	3	14 / 6	22 / 8
-0.06	50% more (old transects)	6	14 / 6	23 / 8
-0.06	50% more (new transects)	3	14 / 5	22 / 8
-0.06	50% more (new transects)	6	14 / 6	22 / 9
-0.04	Baseline	3	9 / 5	16 / 10
-0.04	Baseline	6	10 / 5	17 / 9
-0.04	50% more (old transects)	3	13 / 6	19 / 10
-0.04	50% more (old transects)	6	12 / 6	20 / 10
-0.04	50% more (new transects)	3	12 / 7	19 / 10
-0.04	50% more (new transects)	6	11 / 7	19 / 9
-0.02	Baseline	3	8 / 6	15 / 12
-0.02	Baseline	6	8 / 7	15 / 10
-0.02	50% more (old transects)	3	10 / 7	16 / 11
-0.02	50% more (old transects)	6	10 / 7	17 / 11
-0.02	50% more (new transects)	3	10 / 8	17 / 11
-0.02	50% more (new transects)	6	10 / 8	16 / 12
0.00	Baseline	3	6 / 8	11 / 14
0.00	Baseline	6	7 / 8	13 / 12
0.00	50% more (old transects)	3	9 / 8	16 / 13
0.00	50% more (old transects)	6	8 / 9	15 / 14
0.00	50% more (new transects)	3	8 / 9	14 / 14
0.00	50% more (new transects)	6	9 / 8	14 / 13

GUIS

We analyzed data for two GUIS species that accounted for 55% of all captures at the site. Data for *Hyla squirella* and *Hyla cinerea* were combined as *Hyla* mix, because these sister species share the same ecological niche and some morphs can be difficult to visually distinguish. The two GUIS species exhibited seasonality in trap-counts, with a large proportion of captures occurring during non-breeding season months (October through March). We therefore focused our GUIS analyses only on non-breeding season data.

Trap-count patterns for the GUIS species were very similar to patterns for these same species at JELA. Trap-counts for *Hyla* mix were highest in monitoring year 2012 and slowly declined in subsequent years. *E. planirostris* peaked in monitoring year 2013. Almost 15% of survey data were omitted from regression analyses for GUIS species due to missing relative humidity and temperature information. Prior year's precipitation was highest for monitoring year 2015 and lowest for 2012, i.e., monitoring year 2014 was an especially wet year and 2011 a somewhat drier year.

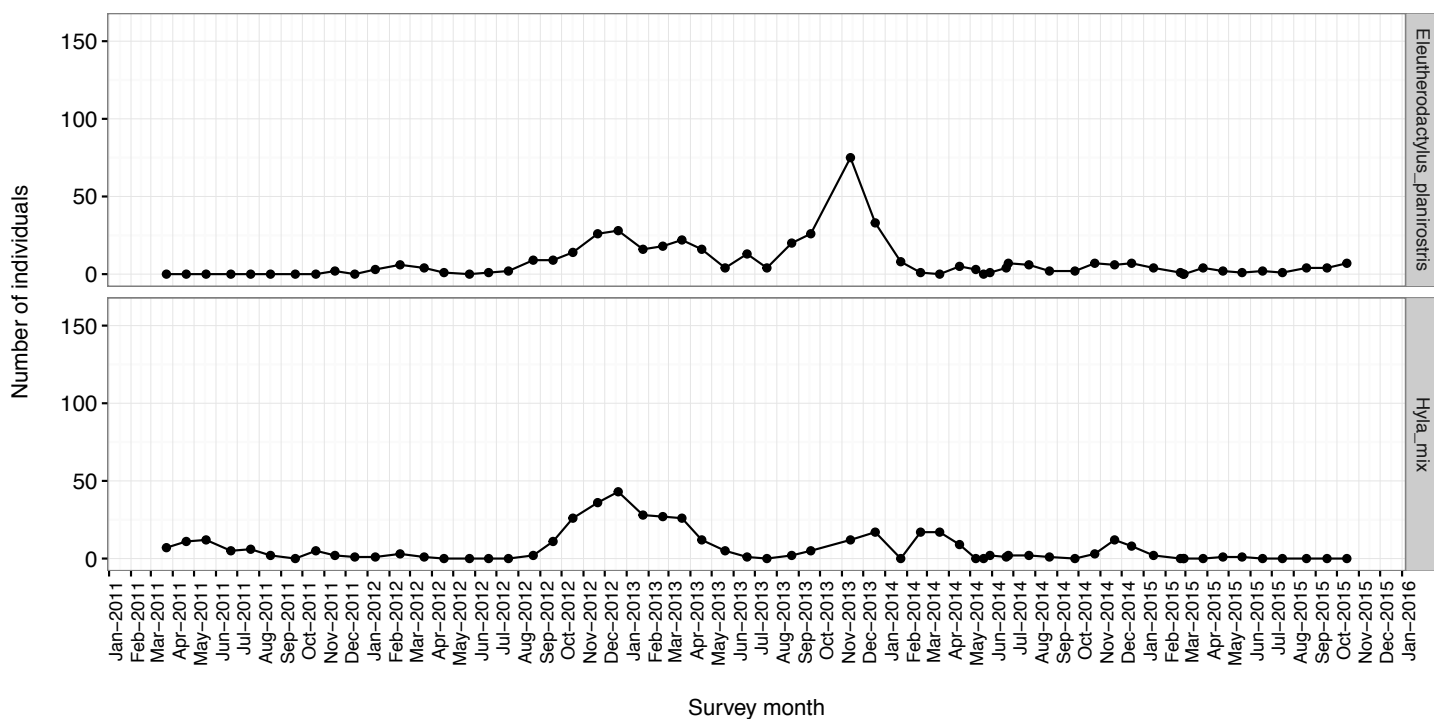


Figure 26. (GUIS) Number of individuals captured each survey event, summed across traps, for two species (horizontal panels).

Table 18. Summary of GUIS covariate data corresponding with survey dates, non-breeding season only. Minimum, median, and maximum values, and % of non-breeding season data missing for covariates: ‘temp.f’ = temperature (Fahrenheit), ‘rh’ = relative humidity (%), ‘precip.7’ = mean daily precipitation in 7 days prior to survey (inches).

Statistic	temp.f	rh	precip.7
Minimum	41.79	63.56	0.00
Median	57.22	80.66	0.01
Maximum	74.34	100.00	0.40
% missing data	14.8	14.8	0

Table 19. (GUIS) For each monitoring year, the mean monthly precipitation (in inches) for the prior monitoring year.

monitor.yr	mean.prior.precip
2011	5.98
2012	4.01
2013	5.88
2014	5.79
2015	7.08
2016	5.10

GUIS: *Eleutherodactylus planirostris*

E. planirostris was first captured at GUIS in November 2011. The largest number of captures at a single coverboard in a survey was 10 animals. When only non-breeding season months were considered, 83% of trap-counts were zero-counts (no animals captured).

The best-supported model for GUIS *E. planirostris* included prior year's precipitation, recent precipitation, and relative humidity. The AIC confidence set included five other models with AIC values within six units of the best-supported model. These models differed from the best-supported model in that they did excluded relative humidity or prior year's precipitation, or included temperature. The confidence set models that excluded prior year's precipitation had much larger coherent interannual variation (0.66–0.88, compared to 0.40 for the best-supported model). In all confidence set models, recent precipitation was still an important predictor of trap-counts. Regression coefficients, trap variance, and the negative binomial dispersion parameter were similar across all confidence set models.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 41% for 3 surveys per year and 43% for 6 surveys per year. The proportion of false trends for these scenarios was 2%–3%.

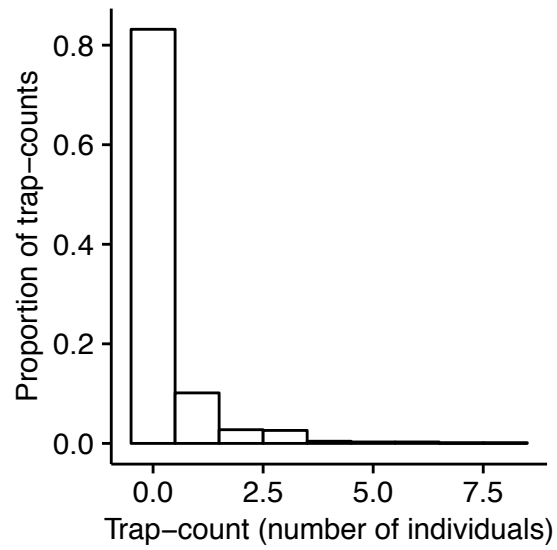


Figure 28. (*GUIS E. planirostris*) Density histogram of trap-counts, non-breeding season only. For this site-species, 83% of trap-counts analyzed were zero.

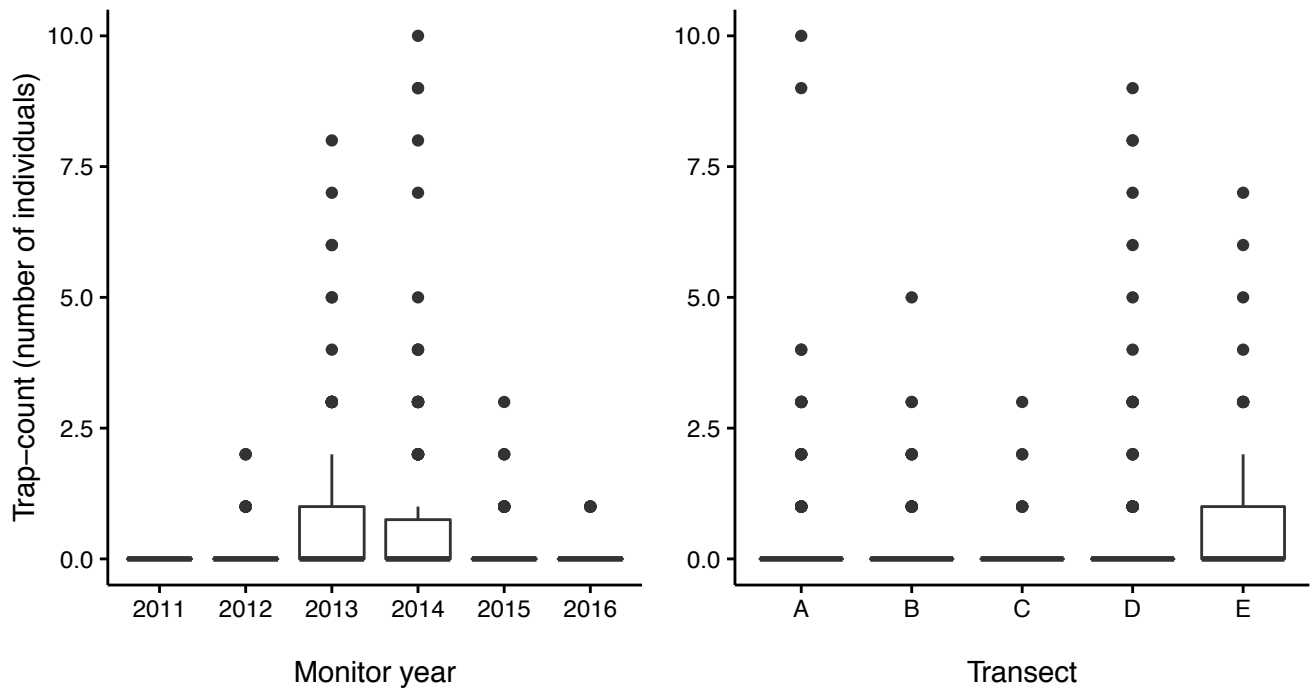


Figure 29. (*GUIS E. planirostris*) Boxplots of detections by monitoring year (left) and transect (right), non-breeding season only.

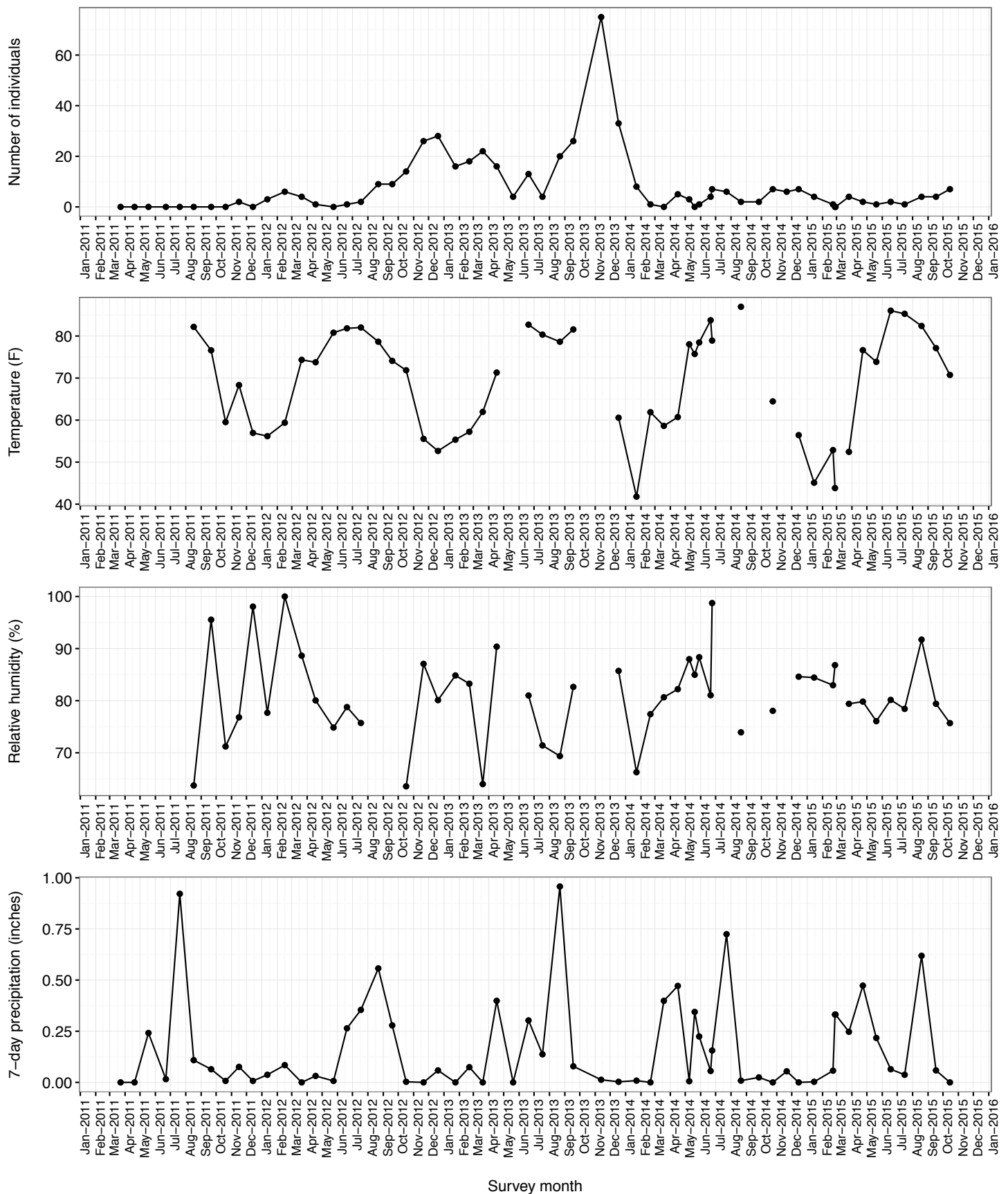


Figure 30. (*GUIS E. planirostris*) Number of individuals captured (summed across traps) and survey-level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 7 days prior to each survey event.

Table 20. (*GUIS E. planirostris*) Best-supported regression model for *GUIS E. planirostris*, non-breeding season only. The fixed-effect covariates ‘mean.prior.precip’ (mean monthly precipitation during the prior monitoring year), ‘precip.7’ (mean daily precipitation in 7 days prior to survey), and ‘rh’ (relative humidity) were centered and scaled so regression coefficients could be presented in standardized units.

```
##
## Call:
## glmnmb(formula = count ~ transect + mean.prior.precip + precip.7 +
##       I(precip.7^2) + rh + (1 | monitor.yr) + (1 | trap.id) - 1,
##       data = guis.ep.scale.nb.select.rh, family = "nbinom", link = "log")
##
## AIC: 756.9
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## transectGUIS_A    -1.8769     0.6652  -2.82   0.0048 **
## transectGUIS_B    -2.4566     0.6803  -3.61   0.0003 ***
## transectGUIS_C    -3.8074     0.8370  -4.55  5.4e-06 ***
## transectGUIS_D    -1.5055     0.6461  -2.33   0.0198 *
## transectGUIS_E    -0.9601     0.6207  -1.55   0.1219
## mean.prior.precip   0.6821     0.3460   1.97   0.0487 *
## precip.7           0.0311     0.2826   0.11   0.9125
## I(precip.7^2)      -0.4329     0.1800  -2.41   0.0162 *
## rh                 0.2300     0.0967   2.38   0.0174 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=690, monitor.yr=5, trap.id=30
## Random effect variance(s):
## Group=monitor.yr
##               Variance StdDev
## (Intercept)   0.4031 0.6349
## Group=trap.id
##               Variance StdDev
## (Intercept)   1.489   1.22
##
## Negative binomial dispersion parameter: 1.565 (std. err.: 0.55424)
##
## Log-likelihood: -366.474
```

Table 21. (GUIS *E. planirostris*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. Year and trap variance are random effect variances estimated in the model. ‘Dispersion parameter’ is associated with the variance of the negative binomial distribution.

Model	dAIC	Number of parameters	Coherent interannual variation	Trap variance	Dispersion parameter	Difference from global model
mod1	0.00	12	0.40	1.49	1.56	remove my.day and temperature and make rh linear
mod2	2.10	13	0.33	1.50	1.51	remove my.day and rh
mod3	3.74	11	0.34	1.49	1.46	remove my.day, temperature and rh
mod4	3.83	14	0.84	1.51	1.61	remove mean.prior.precip and my.day
mod5	5.07	14	0.66	1.51	1.60	remove mean.prior.precip and rh
mod6	5.92	14	0.88	1.50	1.64	remove mean.prior.precip and temperature

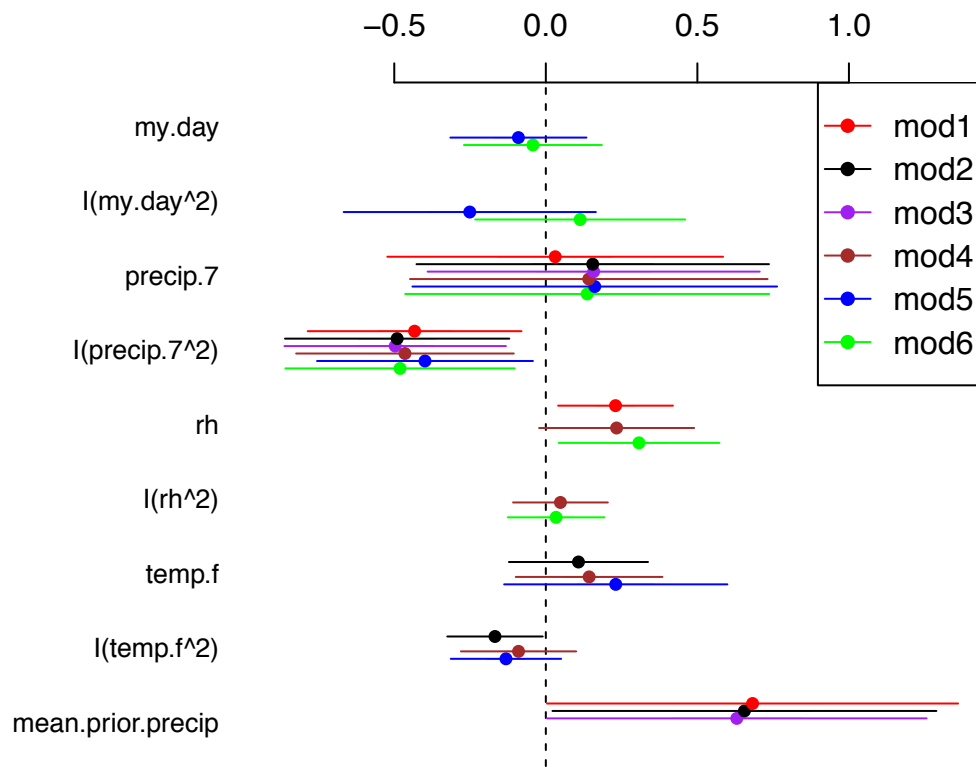


Figure 31. (GUIS *E. planirostris*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

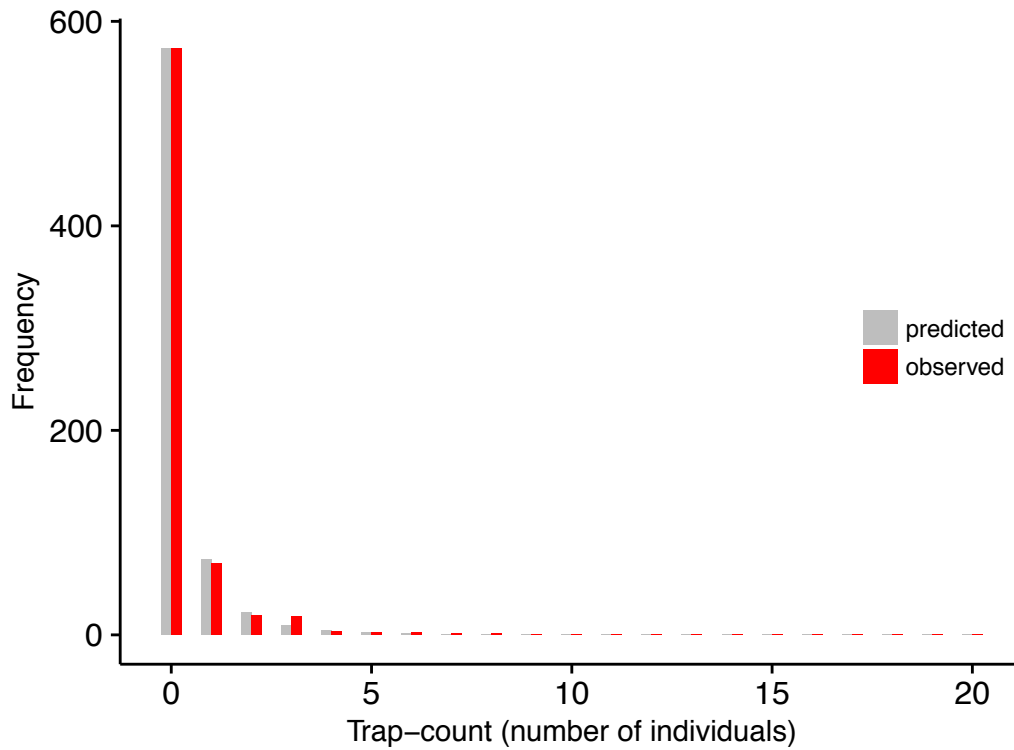


Figure 32. (*GUIS E. planirostris*) Frequency histogram of trap-counts for observed data vs. model predictions, non-breeding season only.

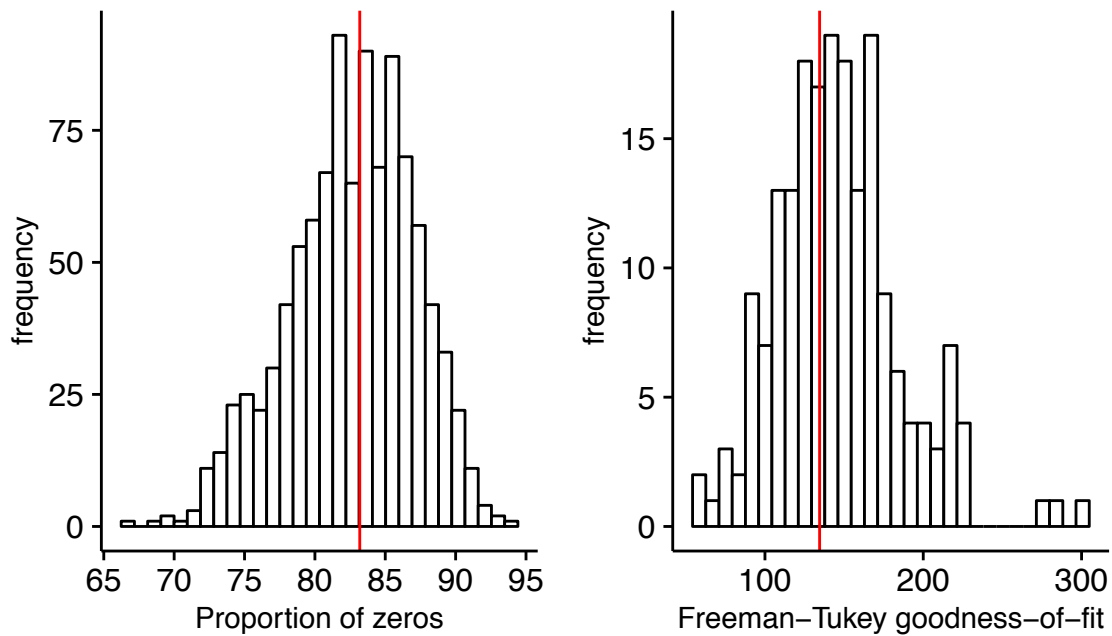


Figure 33. (*GUIS E. planirostris*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is good. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

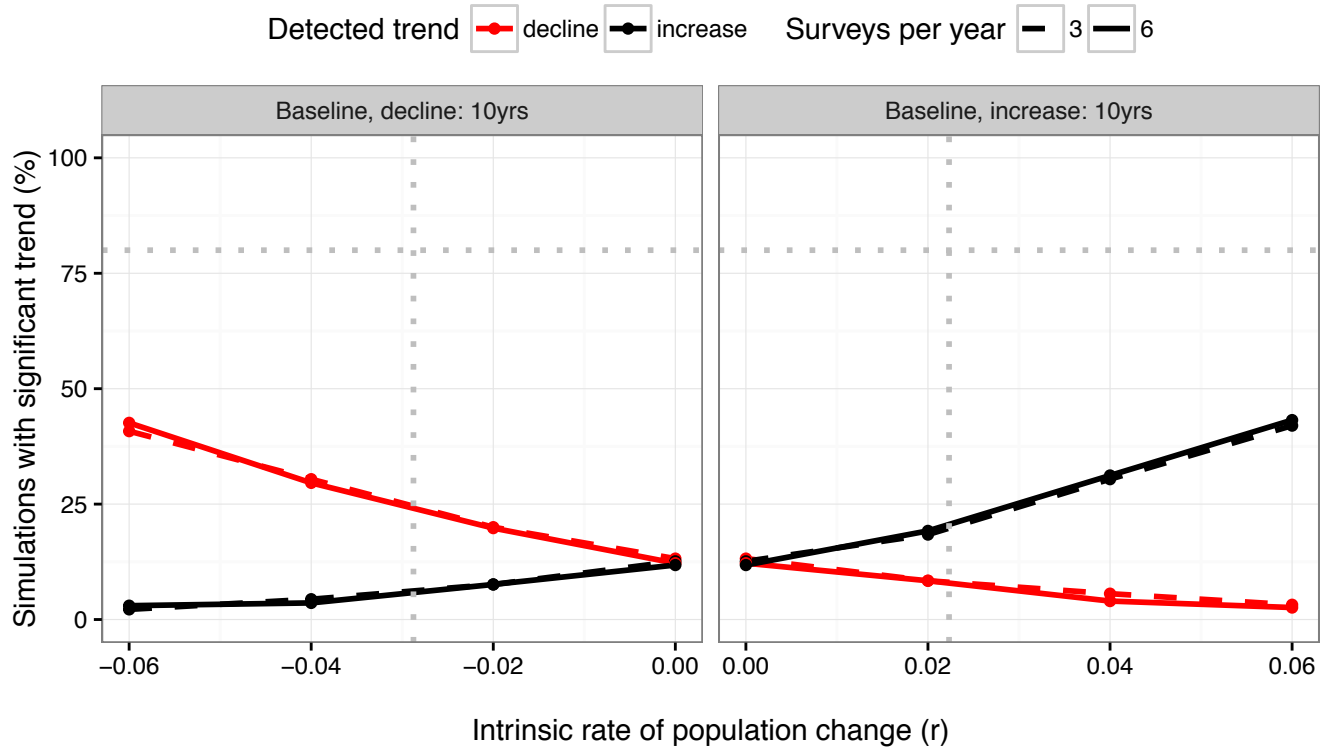


Figure 34. (GUIS *E. planirostris*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), non-breeding season only. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 3 surveys per monitoring year (once in each month of November, January, and March). Solid lines represent 6 surveys per monitoring year (monthly from October through March). The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 22. (GUIS *E. planirostris*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of α (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate α . Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p-values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for GUIS *E. planirostris* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant ($\alpha=0.10$)	% significant ($\alpha=0.20$)
-0.06	Baseline	3	29 / 1	41 / 2
-0.06	Baseline	6	28 / 1	43 / 3
-0.04	Baseline	3	20 / 2	30 / 4
-0.04	Baseline	6	20 / 3	30 / 4
-0.02	Baseline	3	12 / 4	20 / 8
-0.02	Baseline	6	11 / 4	20 / 8
0.00	Baseline	3	9 / 9	13 / 13
0.00	Baseline	6	7 / 8	12 / 12

GUIS: Hyla mix

Similar to JELA *Hyla* mix, trap-counts of GUIS *Hyla* mix peaked between October 2012 and March 2013. The maximum trap-count for this site-species was nine individuals in a single PVC set in a survey. When only non-breeding season months were considered, 76% of trap-counts were zero-counts (no animals captured).

The best-supported model for GUIS *Hyla* mix included day-of-year and temperature. The AIC confidence set included one other model that had an AIC value within six units of the best-supported model. That model differed from the best-supported model in that the response variable used a Poisson rather than a negative binomial distribution. Coherent interannual variation, regression coefficients and trap variance were similar for the two confidence set models.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

Power to detect trends was very low for GUIS *Hyla* mix. With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 21% for 3 surveys per year and 22% for 6 surveys per year. The proportion of false trends for these scenarios was 7%.

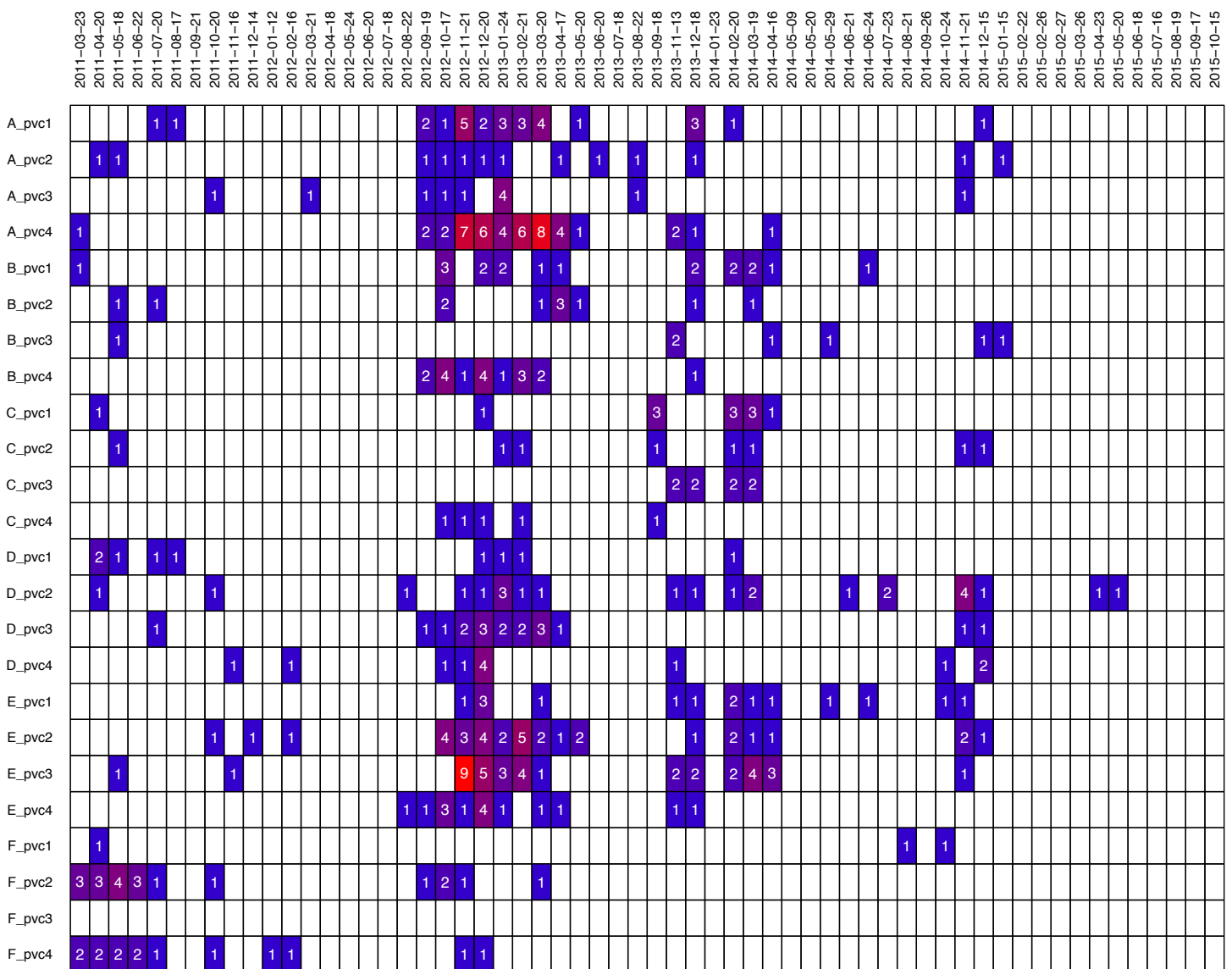


Figure 35. (GUIS *Hyla mix*) Number of individuals captured per PVC trap set each survey event, year-round data. Rows are PVC trap sets, columns are survey events. The initial capital letter (A–F) in each row name indicates the transect. White cells indicate no animals were detected. Otherwise, number of detected individuals ranges from a low of 1 (blue) to a high of 9 (red). Note that this matrix shows data from all surveys, but regression analysis for GUIS *Hyla mix* focused on non-breeding season data (October to March of each monitoring year) only.

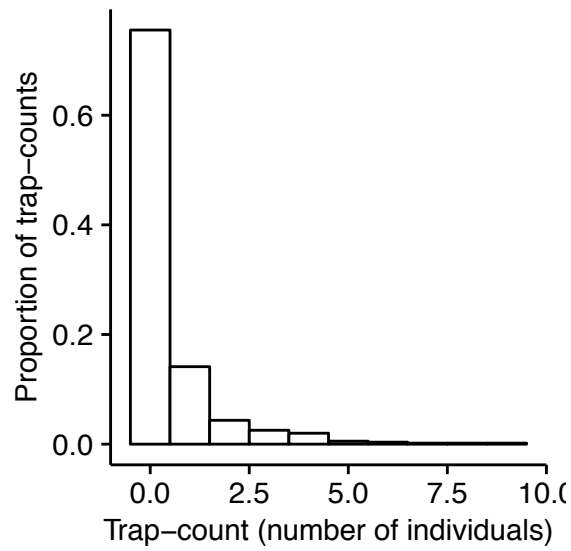


Figure 36. (*GUIS Hyla mix*) Density histogram of trap-counts, non-breeding season only. For this site-species, 76% of trap-counts analyzed were zero.

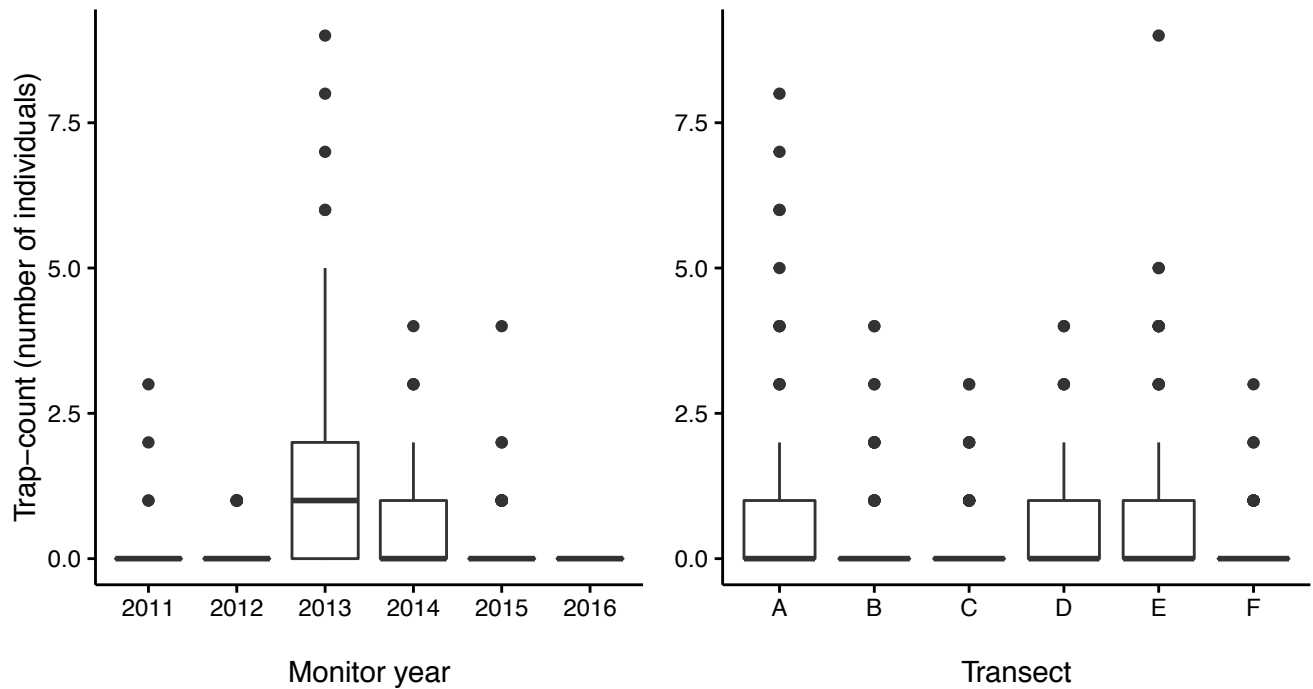


Figure 37. (*GUIS Hyla mix*) Boxplots of detections by monitoring year (left) and transect (right), non-breeding season only.

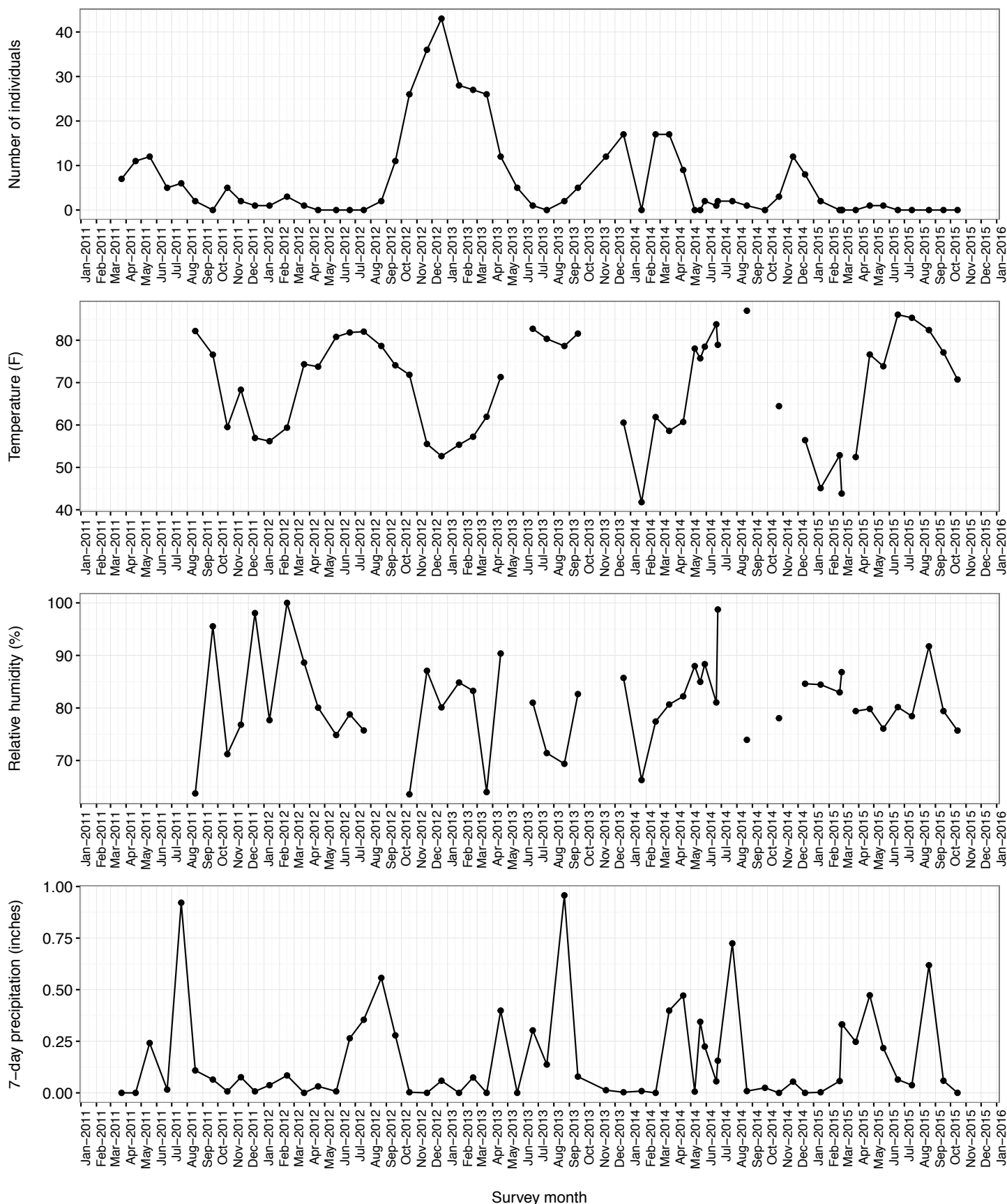


Figure 38. (*GUIS Hyla mix*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 7 days prior to each survey event.

Table 23. (GUIS *Hyla mix*) Best-supported regression model for GUIS *Hyla mix*, non-breeding season only. The fixed-effect covariates ‘my.day’ (day-of-year) and ‘temp.f’ (temperature) were centered and scaled so regression coefficients could be presented in standardized units.

```
##
## Call:
## glmnmb(formula = count ~ transect + my.day + temp.f + I(temp.f^2) +
##       (1 | monitor.yr) + (1 | trap.id) - 1, data = guis.hmix.scale.nb.select.rh,
##       family = "nbinom", link = "log")
##
## AIC: 783.8
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## transectGUIS_A   -1.2140    0.7648   -1.59  0.11246
## transectGUIS_B   -1.7976    0.7753   -2.32  0.02042 *
## transectGUIS_C   -2.2192    0.7840   -2.83  0.00465 **
## transectGUIS_D   -1.5955    0.7699   -2.07  0.03823 *
## transectGUIS_E   -1.0048    0.7620   -1.32  0.18725
## transectGUIS_F   -2.9862    0.8173   -3.65  0.00026 ***
## my.day           -0.2683    0.0869   -3.09  0.00202 **
## temp.f            0.2526    0.1248    2.02  0.04288 *
## I(temp.f^2)      -0.4468    0.1045   -4.28  1.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=552, monitor.yr=5, trap.id=24
## Random effect variance(s):
## Group=monitor.yr
##           Variance StdDev
## (Intercept)  2.179  1.476
## Group=trap.id
##           Variance StdDev
## (Intercept)  0.2128 0.4613
##
## Negative binomial dispersion parameter: 5.0149 (std. err.: 3.3629)
##
## Log-likelihood: -379.886
```

Table 24. (GUIS *Hyla mix*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. Year and trap variance are random effect variances estimated in the model. ‘Dispersion parameter’ is associated with the variance of the negative binomial distribution.

Model	dAIC	Number of parameters	Coherent interannual variation	Trap variance	Dispersion parameter	Difference from global model
mod1	0.00	12	2.18	0.21	5.01	remove mean.prior.precip, rh and precip.7, and make my.day linear
mod2	1.52	11	2.18	0.25	NA	remove mean.prior.precip, rh and precip.7, and make my.day linear; use Poisson

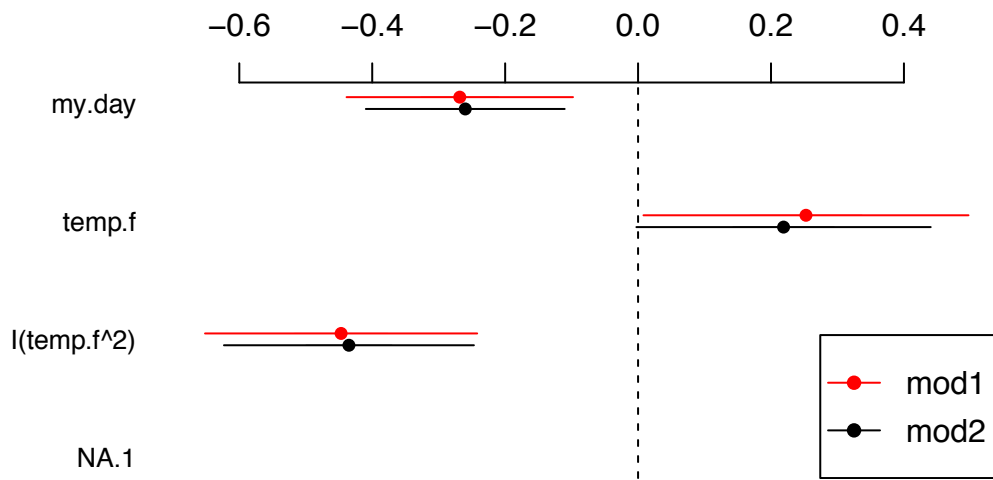


Figure 39. (*GUIS Hyla mix*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

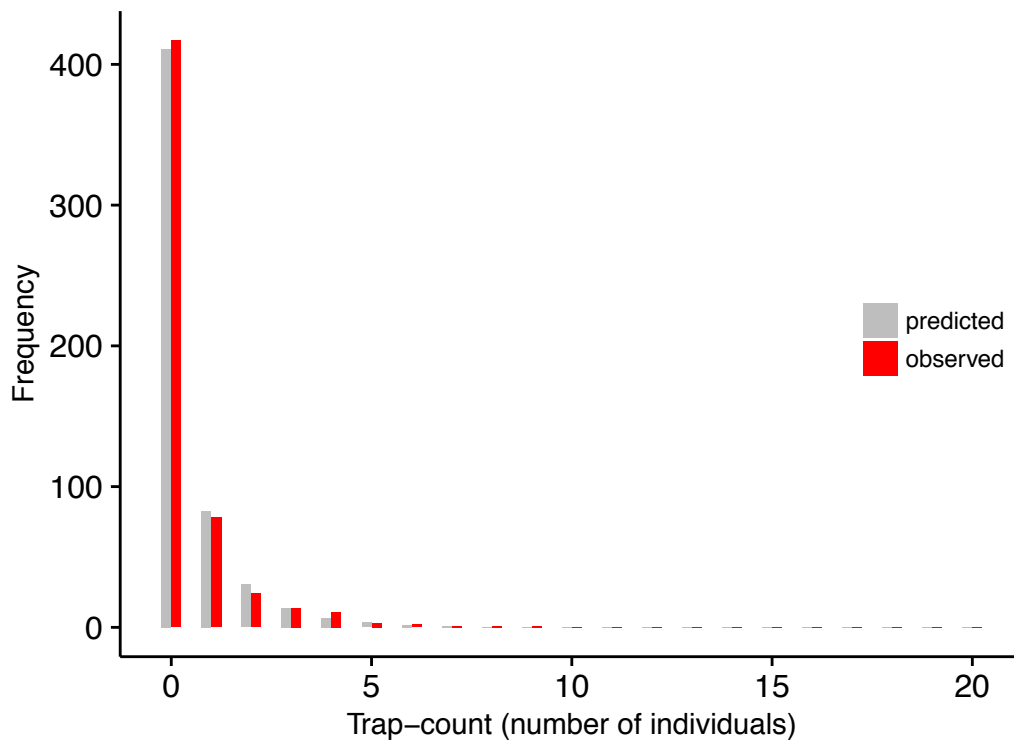


Figure 40. (*GUIS Hyla mix*) Frequency histogram of trap-counts for observed data vs. model predictions, non-breeding season only.

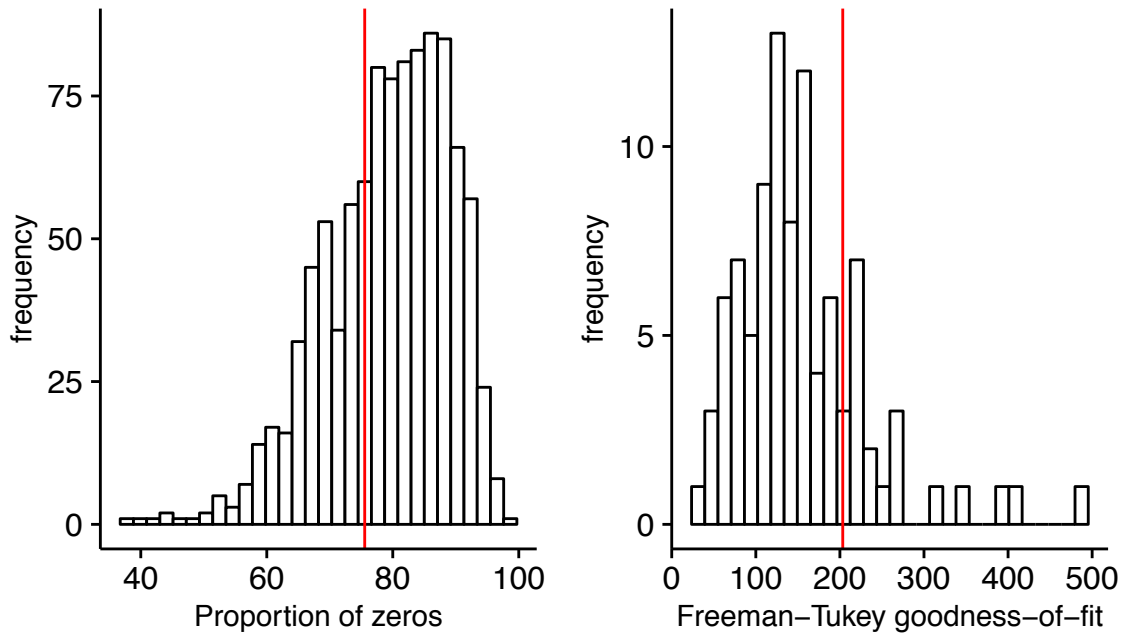


Figure 41. (*GUIS Hyla mix*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is adequate. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

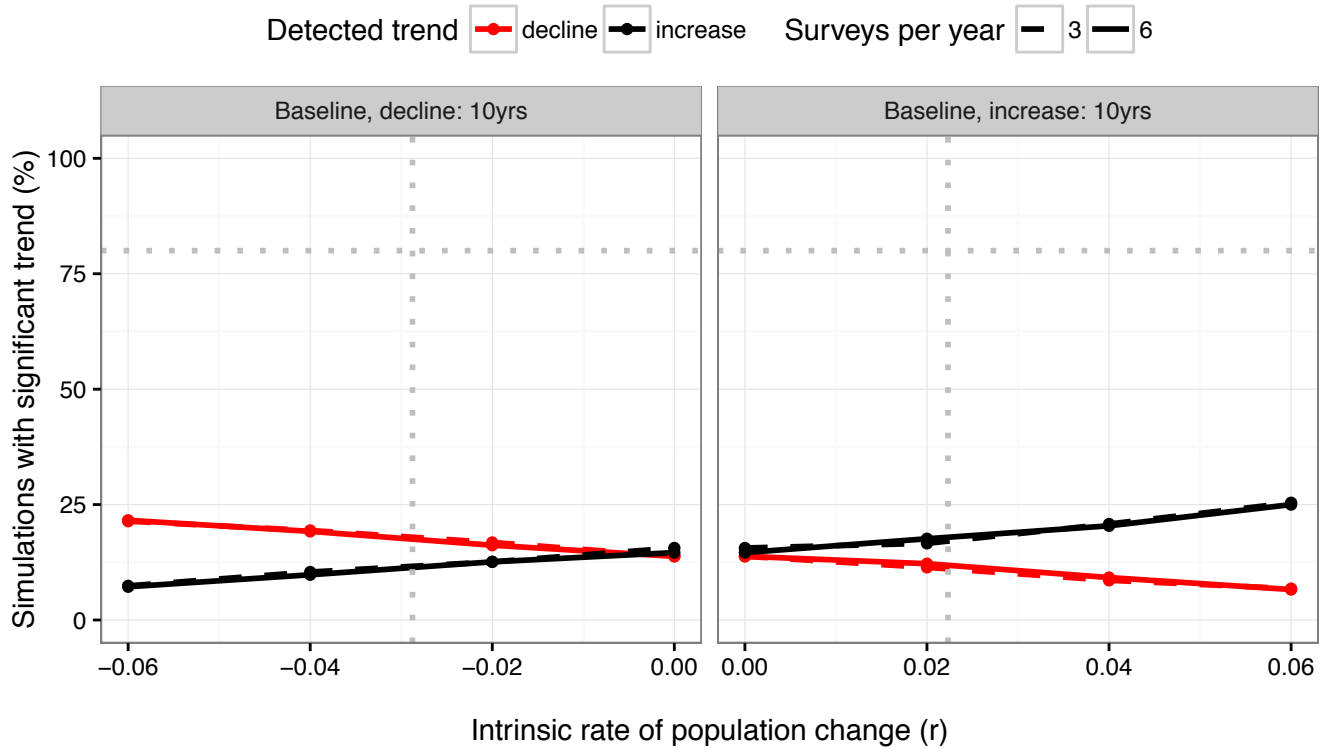


Figure 42. (*GUIS Hyla mix*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), non-breeding season only. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 3 surveys per monitoring year (once in each month of November, January, and March). Solid lines represent 6 surveys per monitoring year (monthly from October through March). The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 25. (GUIS *Hyla mix*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate alpha. Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p -values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for GUIS *Hyla mix* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	14 / 5	21 / 7
-0.06	Baseline	6	14 / 4	22 / 7
-0.04	Baseline	3	13 / 6	19 / 10
-0.04	Baseline	6	13 / 5	19 / 10
-0.02	Baseline	3	10 / 8	17 / 13
-0.02	Baseline	6	10 / 7	16 / 13
0.00	Baseline	3	8 / 10	14 / 16
0.00	Baseline	6	8 / 8	14 / 15

SAAN

We analyzed data for two SAAN species that accounted for 93% of all captures at the site. Due to changes in water management practices beginning in monitoring year 2012, which potentially influenced amphibian captures, we excluded monitoring year 2011 data from SAAN analyses. Temperature and relative humidity data were missing for several years of the pilot survey, so were not included as potential covariates for regression models. SAAN species did not exhibit the strong seasonality in trap-counts observed at JELA and GUI5, so SAAN analyses were based on year-round data.

Trap-counts for *I. nebulier* peaked in monitoring year 2012 and slowly declined in subsequent years. Trap-counts for *E. cystignathoides* were more sporadic. Prior year's precipitation was highest for monitoring year 2011 and lowest for 2012, i.e., monitoring year 2010 was a relatively wet year and 2011 a relatively dry year at this site.

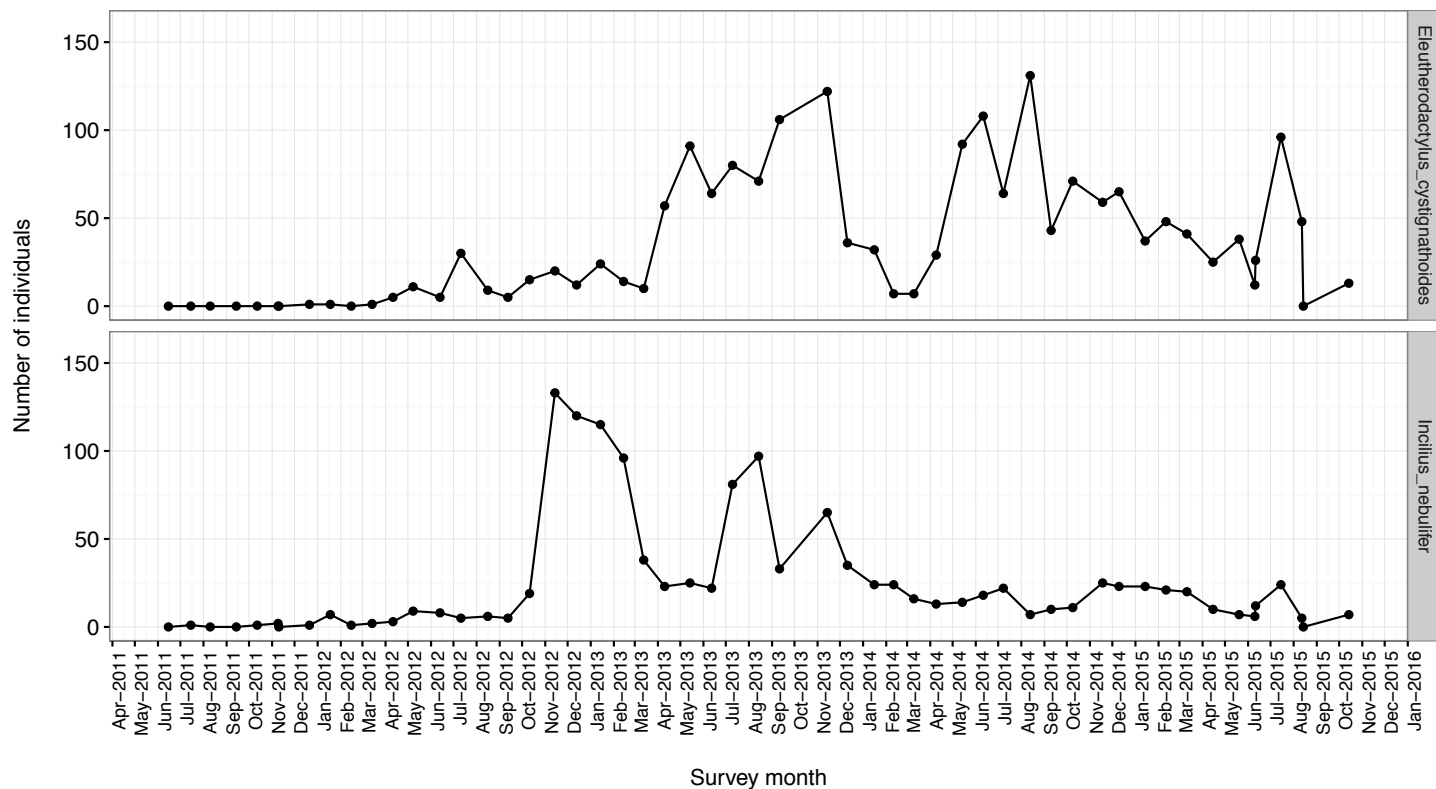


Figure 43. (SAAN) Number of individuals captured each survey event, summed across traps, for two species (horizontal panels).

Table 26. Summary of SAAN covariate data corresponding with survey dates, year-round. Minimum, median, and maximum values, and % of year-round data missing for covariates: ‘precip.7’ = mean daily precipitation in 7 days prior to survey (inches).

Statistic	precip.7
Minimum	0.00
Median	0.00
Maximum	0.76
% missing data	0

Table 27. (SAAN) For each monitoring year, the mean monthly precipitation (in inches) for the prior monitoring year.

monitor.yr	mean.prior.precip
2012	0.76
2013	2.42
2014	1.51
2015	1.51
2016	2.37

SAAN: *Eleutherodactylus cystignathoides*

Trap-counts of SAAN *E. cystignathoides* reached a maximum of 11 individuals at a single coverboard trap in a survey. Approximately 73% of trap-counts were zero-counts (no animals captured). The best-supported model for SAAN *E. cystignathoides* included day-of-year, precipitation, and a binary covariate for breeding season. The AIC confidence set included one other model that had an AIC value within two units of the best-supported model. That model differed from the best-supported model in that it excluded recent precipitation as a predictor. In both confidence set models, day-of-year and the breeding season covariate were still important predictors of trap-counts. Coherent interannual variation, regression coefficients, trap variance, and the negative binomial dispersion parameter were similar for both confidence set models.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 31% for 6 surveys per year and 30% for 12 surveys per year. The proportion of false trends for these scenarios was 3%.

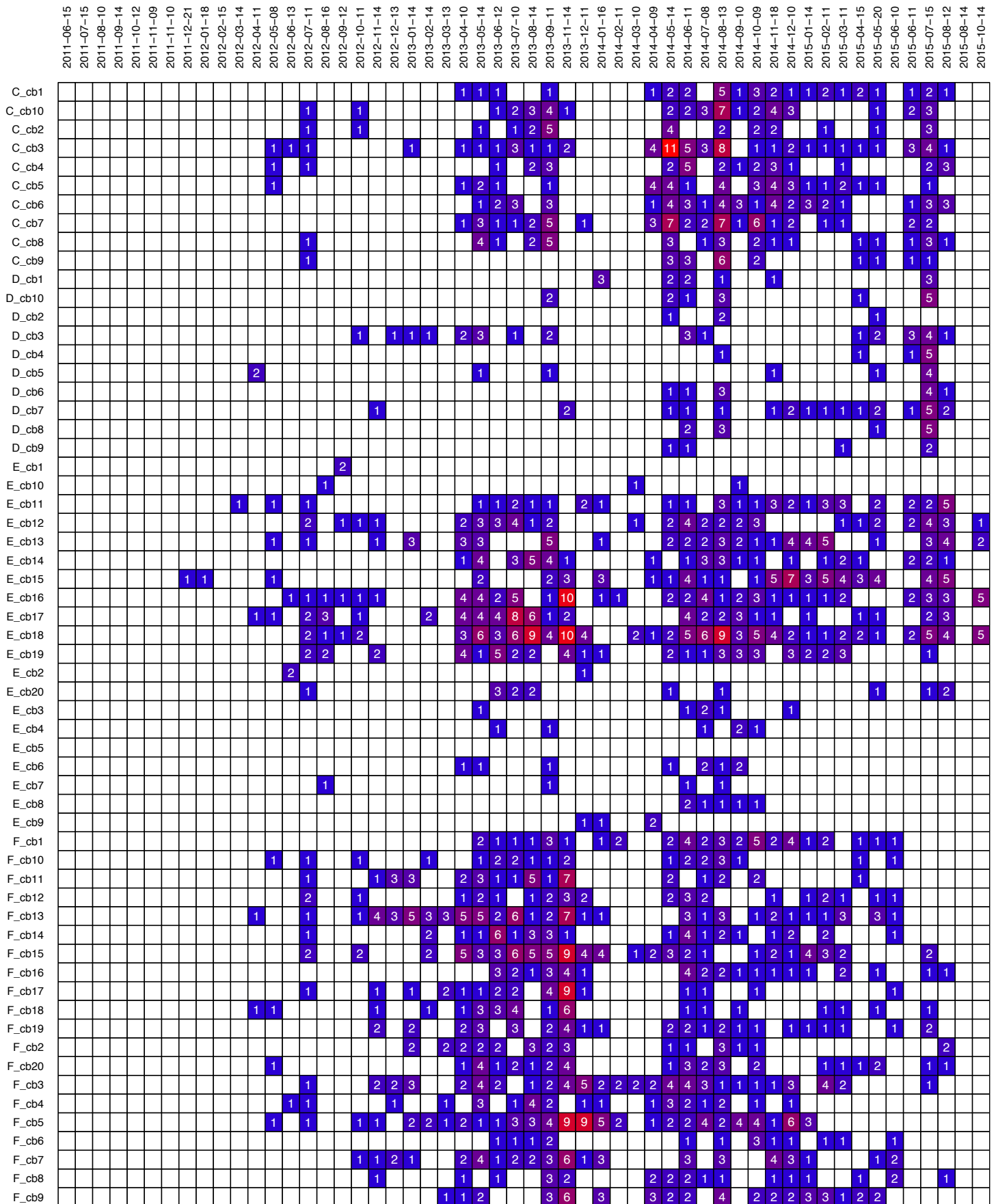


Figure 44. (SAAN *E. cystignathoides*) Number of individuals captured per coverboard trap each survey event, year-round data. Rows are coverboard traps, columns are survey events. The initial capital letter (C–F) in each row name indicates the transect. White cells indicate no animals were detected. Otherwise, number of detected individuals ranges from a low of 1 (blue) to a high of 11 (red).

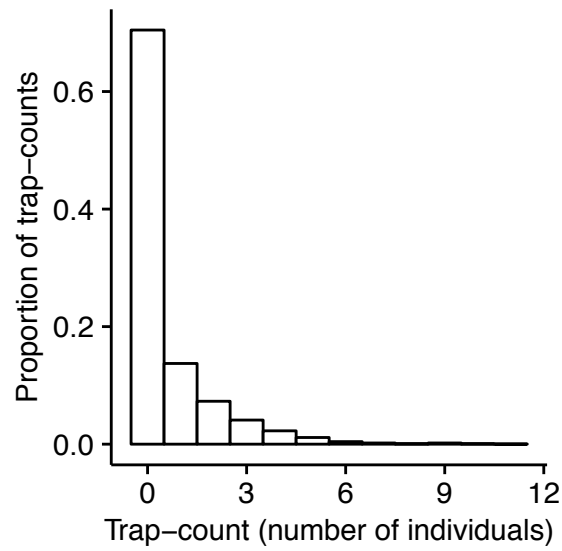


Figure 45. (SAAN *E. cystignathoides*) Density histogram of trap-counts, year-round data. For this site-species, 73% of trap-counts analyzed were zero.

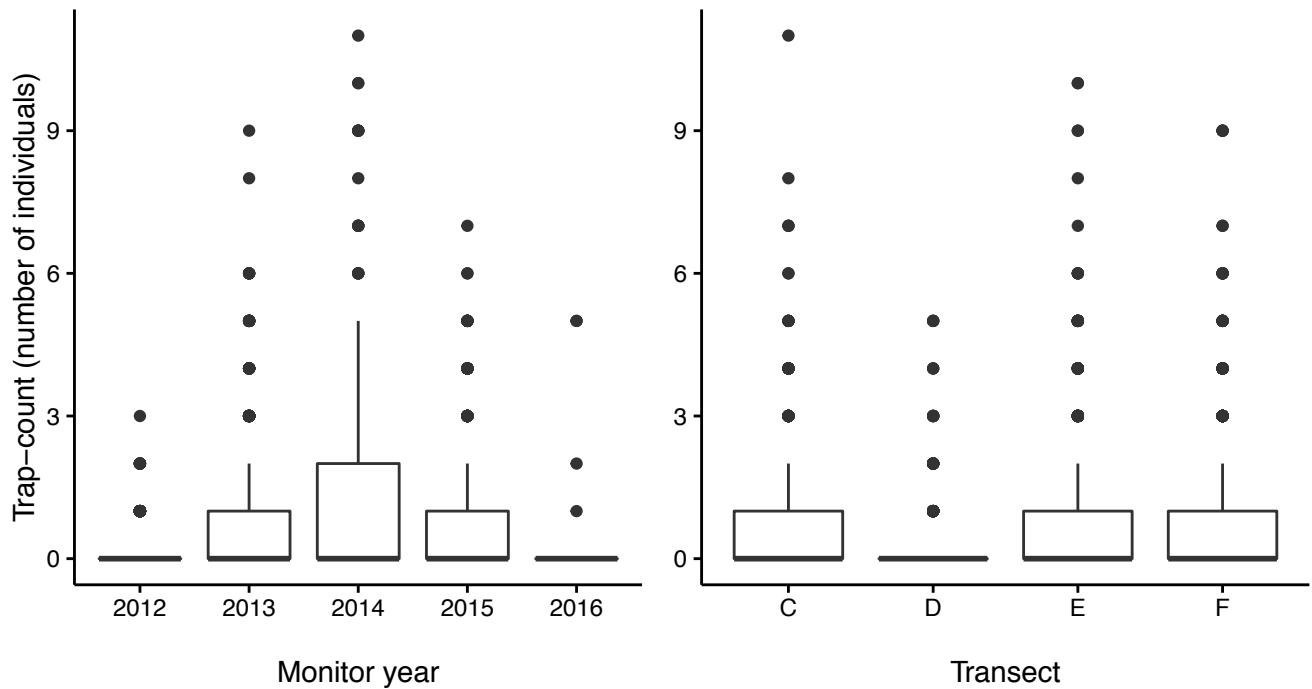


Figure 46. (SAAN *E. cystignathoides*) Boxplots of detections by monitoring year (left) and transect (right), year-round data.

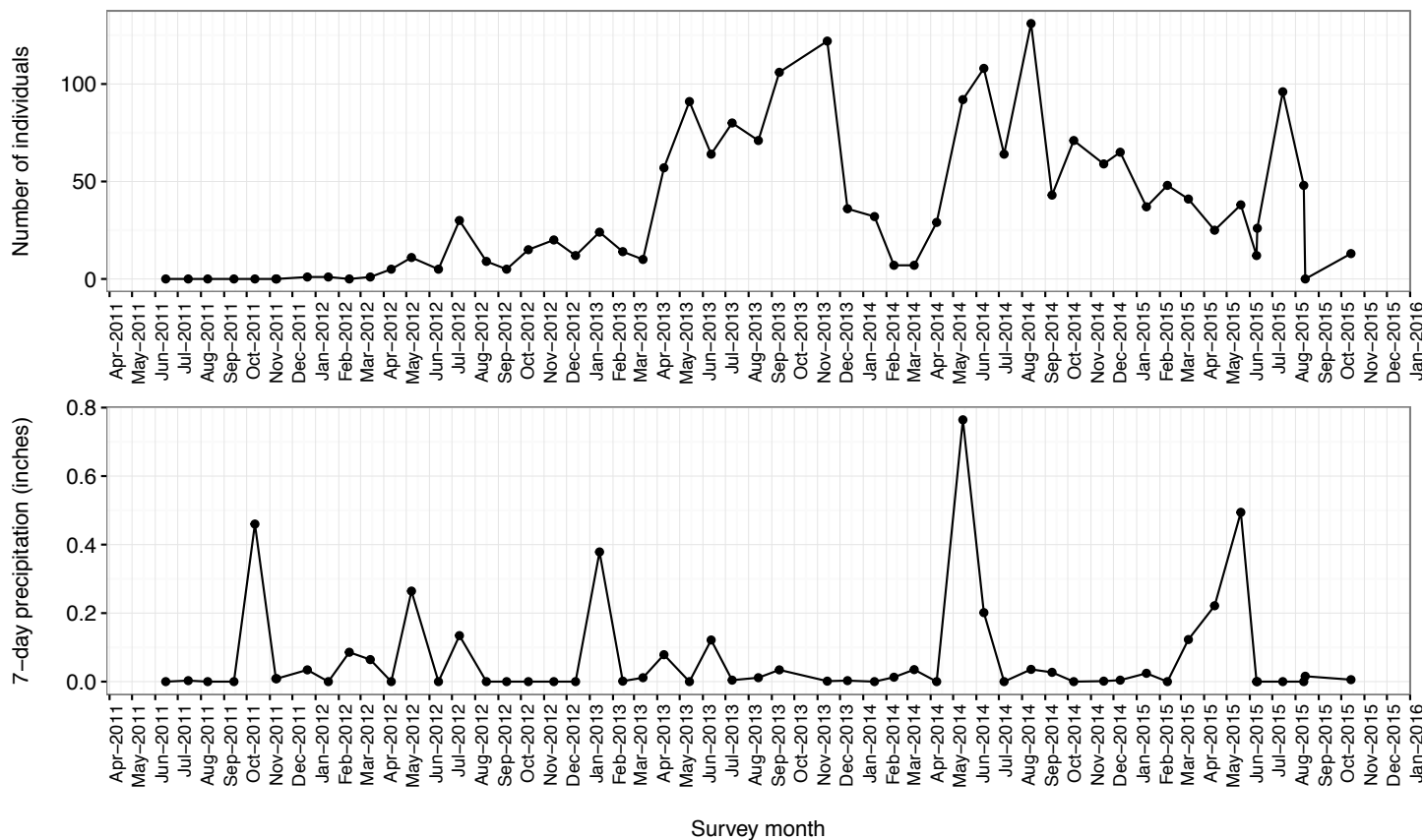


Figure 47. (SAAN *E. cystignathoides*) Number of individuals captured (summed across traps) each survey event and mean daily precipitation for the 7 days prior, year-round data.

Table 28. (SAAN *E. cystignathoides*) Best-supported regression model for SAAN *E. cystignathoides*, year-round data. The fixed-effect covariates ‘my.day’ (day-of-year) and ‘precip.7’ (mean daily precipitation in 7 days prior to survey) were centered and scaled so regression coefficients could be presented in standardized units. ‘nb.01’ refers to a binary predictor indicating if the survey was conducted during a breeding or non-breeding season month. Data from monitoring year 2011 were omitted from analysis because a change in water management operations since 2012 may have influenced amphibian capture rates.

```
##
## Call:
## glmnmb(formula = count ~ transect + my.day + I(my.day^2) +
##       precip.7 + nb.0 + (1 | monitor.yr) + (1 | trap.id) - 1, data = saan.ec.scale.yr,
##       family = "nbinom", link = "log")
##
## AIC: 5445.4
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## transectSAAN_C -1.7294     0.5282  -3.27 0.00106 **
## transectSAAN_D -2.9044     0.5352  -5.43 5.8e-08 ***
## transectSAAN_E -2.4507     0.4977  -4.92 8.5e-07 ***
## transectSAAN_F -1.6257     0.4927  -3.30 0.00097 ***
## my.day         -0.1626     0.0745  -2.18 0.02902 *
## I(my.day^2)      0.2742     0.0420   6.53 6.6e-11 ***
## precip.7         0.0563     0.0332   1.70 0.08960 .
## nb.01           0.9988     0.1561   6.40 1.6e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=3000, monitor.yr=5, trap.id=60
## Random effect variance(s):
## Group=monitor.yr
##           Variance StdDev
## (Intercept)  0.958 0.9788
## Group=trap.id
##           Variance StdDev
## (Intercept)  0.6523 0.8076
##
## Negative binomial dispersion parameter: 1.0616 (std. err.: 0.096527)
##
## Log-likelihood: -2711.68
```

Table 29. (SAAN *E. cystignathoides*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. Year and trap variance are random effect variances estimated in the model. ‘Dispersion parameter’ is associated with the variance of the negative binomial distribution.

Model	dAIC	Number of parameters	Coherent interannual variation	Trap variance	Dispersion parameter	Difference from global model
mod1	0.00	11	0.96	0.65	1.06	remove mean.prior.precip and make precip.7 linear
mod2	0.88	10	0.96	0.65	1.06	remove mean.prior.precip and precip.7

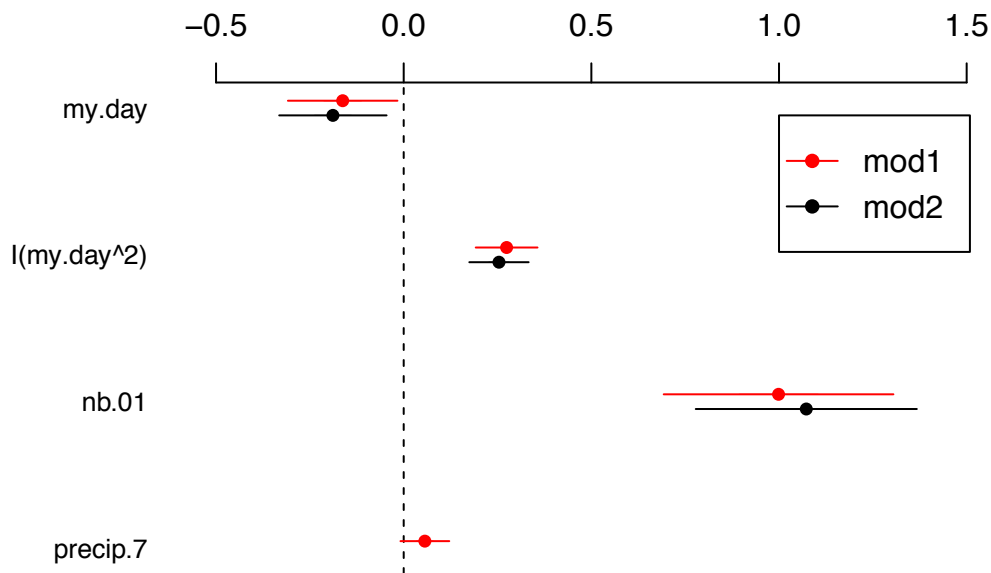


Figure 48. (SAAN *E. cystignathoides*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

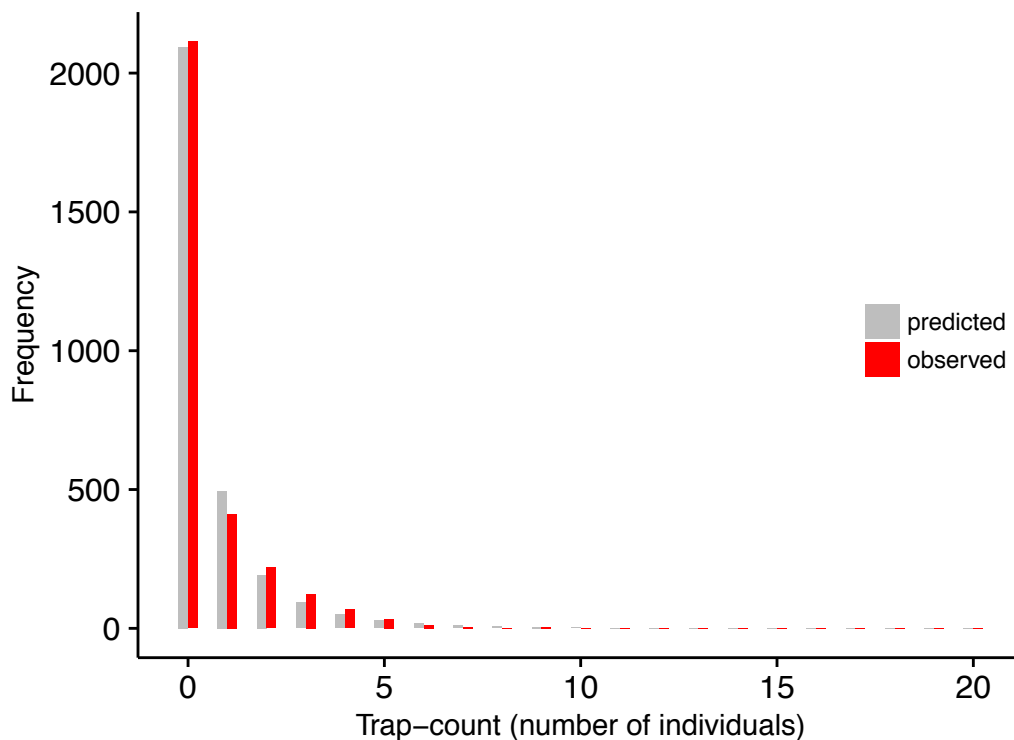


Figure 49. (SAAN *E. cystignathoides*) Frequency histogram of trap-counts for observed data vs. model predictions, year-round data.

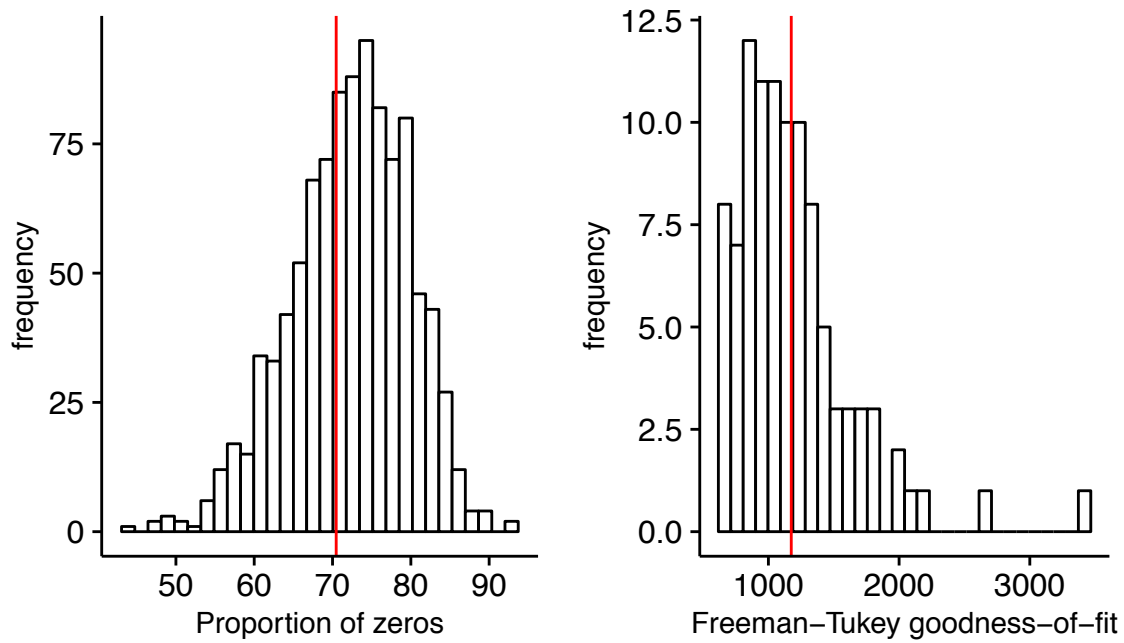


Figure 50. (SAAN *E. cystignathoides*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is good. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

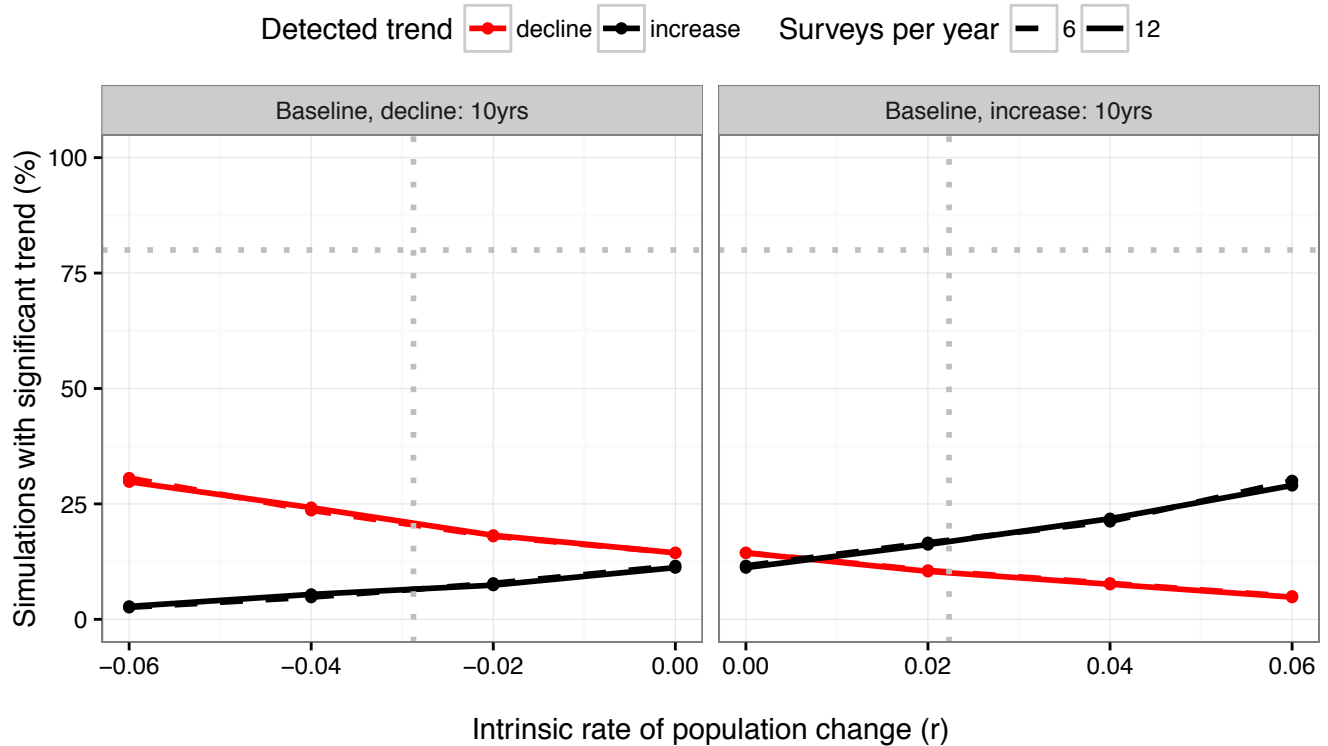


Figure 51. (SAAN *E. cystignathoides*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), with surveys year-round. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 6 surveys per monitoring year (alternate months throughout the year). Solid lines represent year-round monthly surveys. The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 30. (SAAN *E. cystignathoides*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline (first number in column) or a significant increase (second number in column) at two levels of alpha (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate alpha. Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p -values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for SAAN *E. cystignathoides* ‘Baseline, decline: 10yrs’. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	6	21 / 1	31 / 3
-0.06	Baseline	12	21 / 1	30 / 3
-0.04	Baseline	6	16 / 2	24 / 5
-0.04	Baseline	12	15 / 2	24 / 5
-0.02	Baseline	6	12 / 4	18 / 8
-0.02	Baseline	12	12 / 4	18 / 7
0.00	Baseline	6	9 / 6	14 / 12
0.00	Baseline	12	9 / 7	14 / 11

SAAN: *Incilius nebulifer*

Trap-counts of SAAN *I. nebulifer* reached a maximum of 25 individuals at a single coverboard trap in a survey. Approximately 83% of trap-counts were zero-counts (no animals captured). The best-supported model for SAAN *I. nebulifer* included only a binary covariate for breeding season. This was the only model in the AIC confidence set, i.e., no candidate model with likely informative parameters had AIC values within six units of the best-supported model.

Predictions from the best-supported model had a frequency distribution similar to the observed data. In addition, the proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual 6% decline over 10 years was 28% for 6 surveys per year and 29% for 12 surveys per year. The proportion of false trends was 3%–4%.

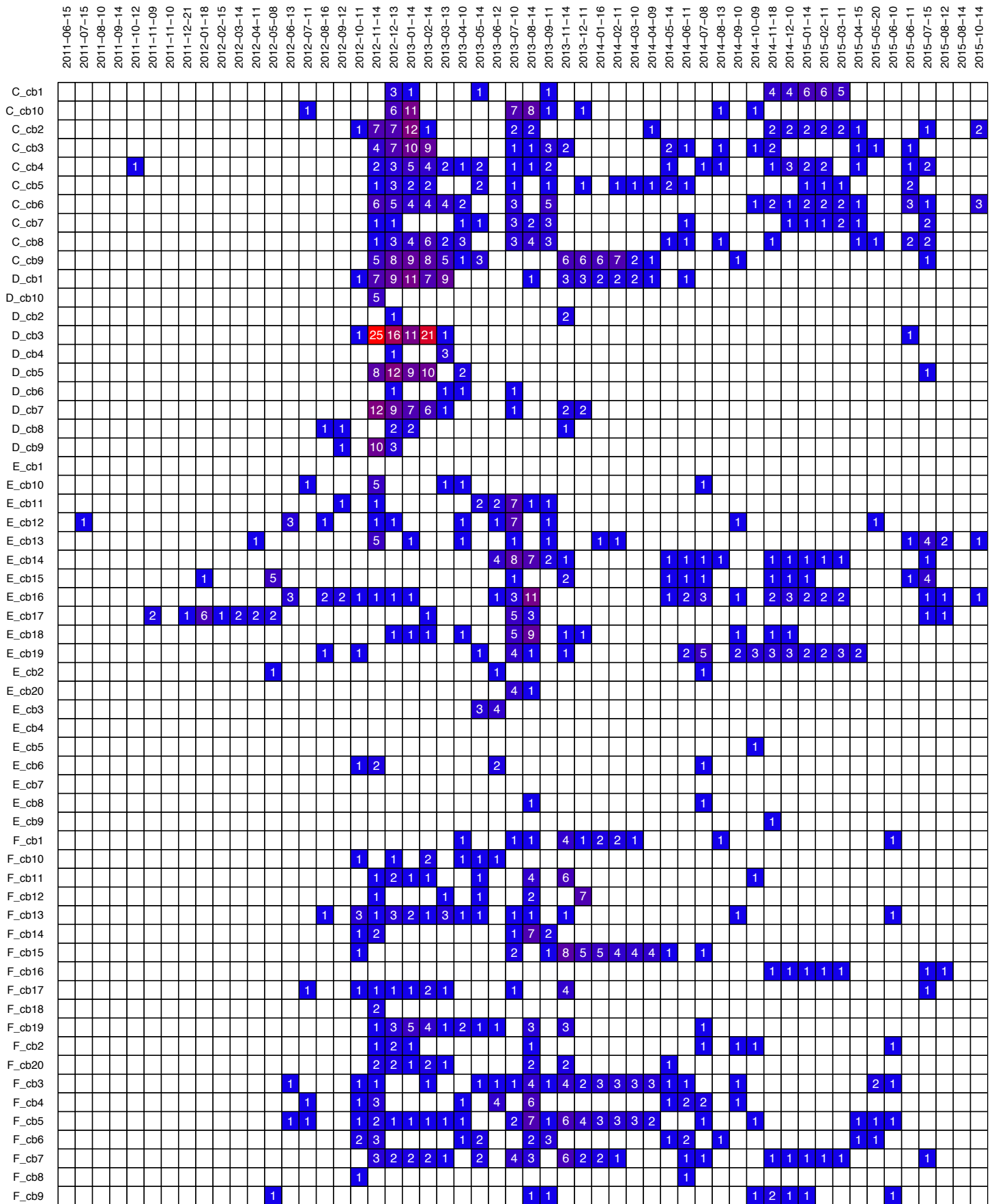


Figure 52. (*SAAN I. nebulifer*) Number of individuals captured per coverboard trap each survey event, year-round data. Rows are coverboard traps, columns are survey events. The initial capital letter (C-F) in each row name indicates the transect. White cells indicate no animals were detected. Otherwise, number of detected individuals ranges from a low of 1 (blue) to a high of 25 (red).

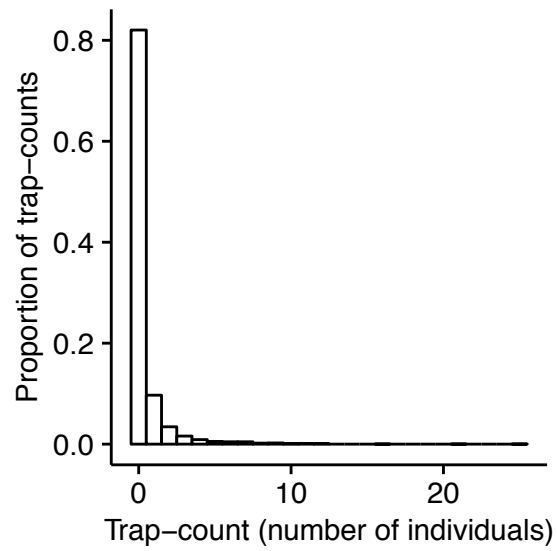


Figure 53. (SAAN *I. nebulifer*) Density histogram of trap-counts, year-round data. For this site-species, 83% of trap-counts analyzed were zero.

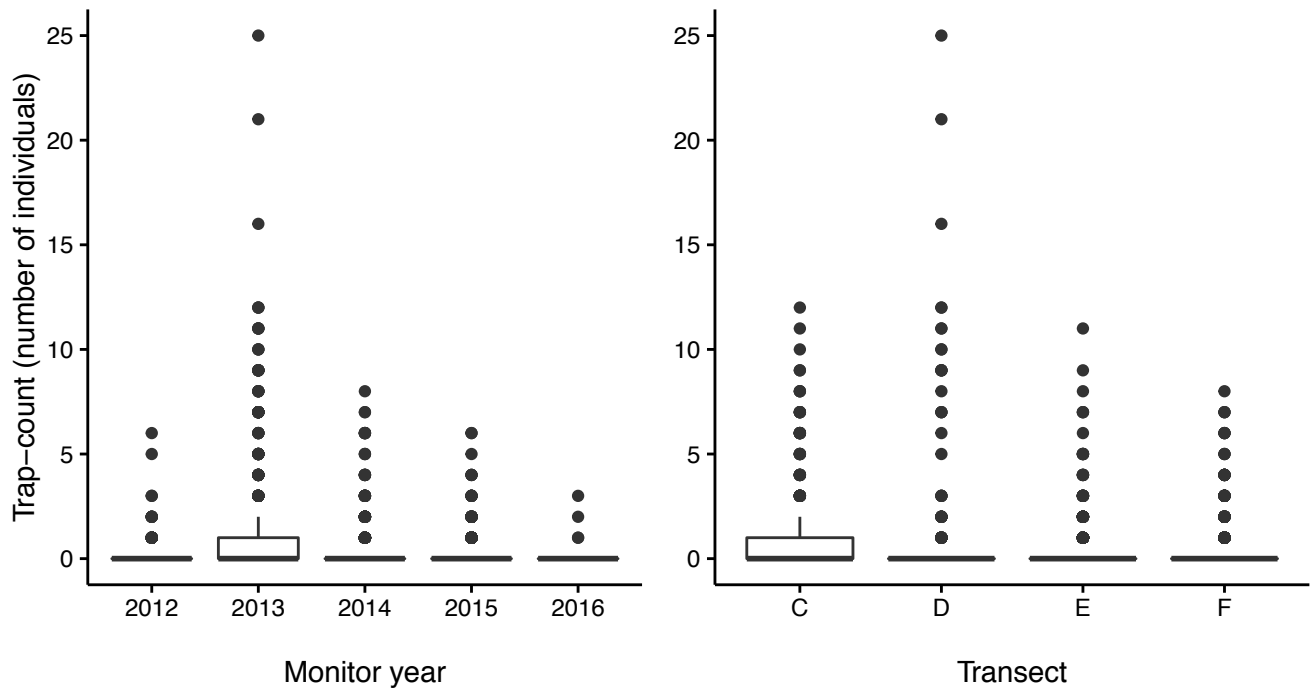


Figure 54. (SAAN *I. nebulifer*) Boxplots of detections by monitoring year (left) and transect (right), year-round data.

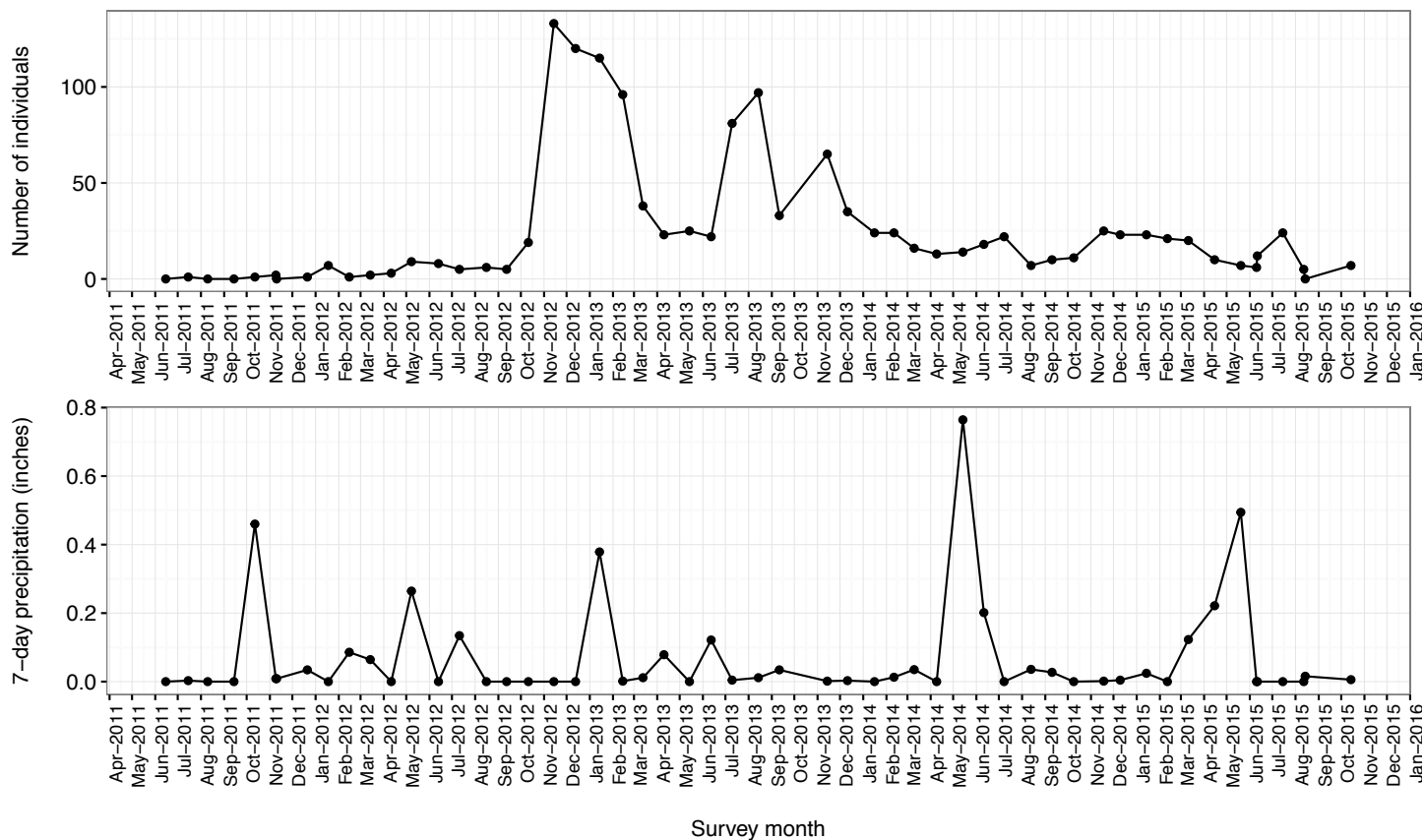


Figure 55. (*SAAN I. nebulifer*) Number of individuals captured (summed across traps) each survey event and mean daily precipitation for the 7 days prior, year-round data.

Table 31. (SAAN I. nebulifer) Best-supported regression model for SAAN I. nebulifer, year-round data. ‘nb.01’ refers to a binary predictor indicating if the survey was conducted during a breeding or non-breeding season month. Data from monitoring year 2011 were omitted from analysis because a change in water management operations since 2012 may have influenced amphibian capture rates.

```
##
## Call:
## glmnadm(formula = count ~ transect + nb.0 + (1 | monitor.yr) +
##      (1 | trap.id) - 1, data = saan.in.scale.yr, family = "nbinom",
##      link = "log")
##
## AIC: 4029.9
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## transectSAAN_C  -0.703    0.582   -1.21  0.22683
## transectSAAN_D  -1.973    0.591   -3.34  0.00085 ***
## transectSAAN_E  -2.123    0.548   -3.88  0.00011 ***
## transectSAAN_F  -1.650    0.540   -3.06  0.00225 **
## nb.01           -0.492    0.104   -4.71  2.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of observations: total=3000, monitor.yr=5, trap.id=60
## Random effect variance(s):
## Group=monitor.yr
##      Variance StdDev
## (Intercept)  1.135  1.065
## Group=trap.id
##      Variance StdDev
## (Intercept)  0.9285 0.9636
##
## Negative binomial dispersion parameter: 0.37452 (std. err.: 0.031496)
##
## Log-likelihood: -2006.96
```

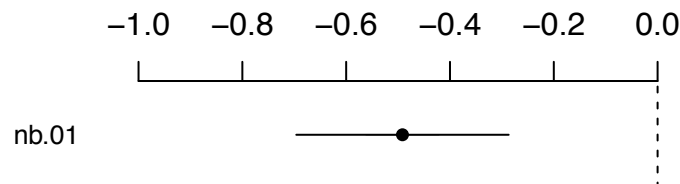


Figure 56. (SAAN I. nebulifer) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For SAAN I. nebulifer, the confidence set only includes the best-supported model.

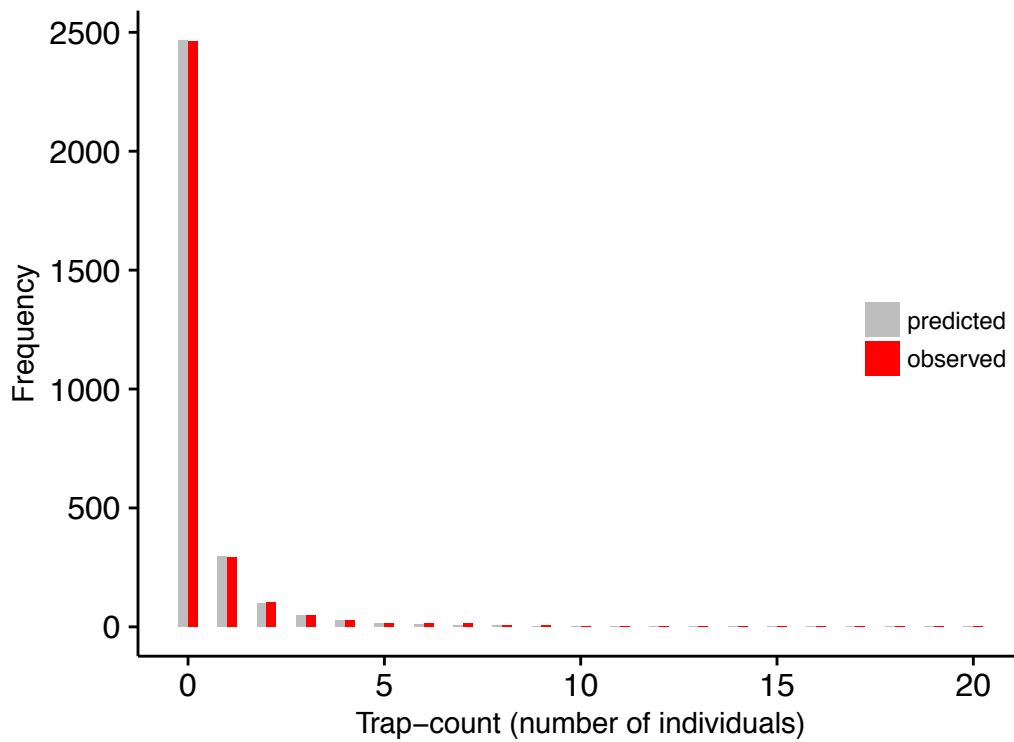


Figure 57. (SAAN *I. nebulifer*) Frequency histogram of trap-counts for observed data vs. model predictions, year-round data.

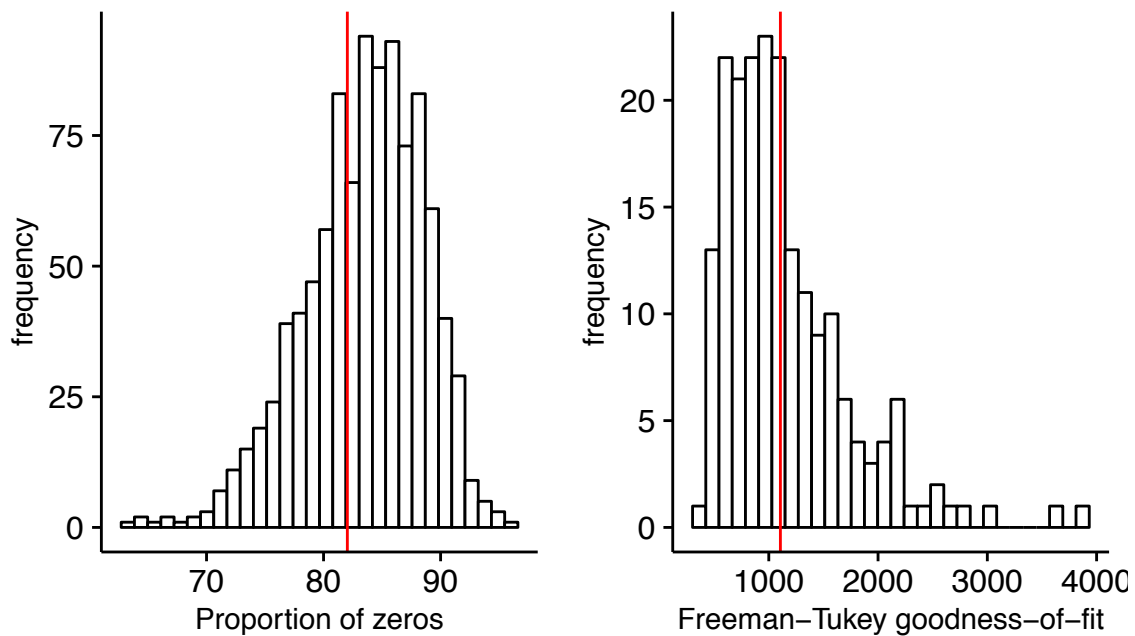


Figure 58. (SAAN *I. nebulifer*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is adequate. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

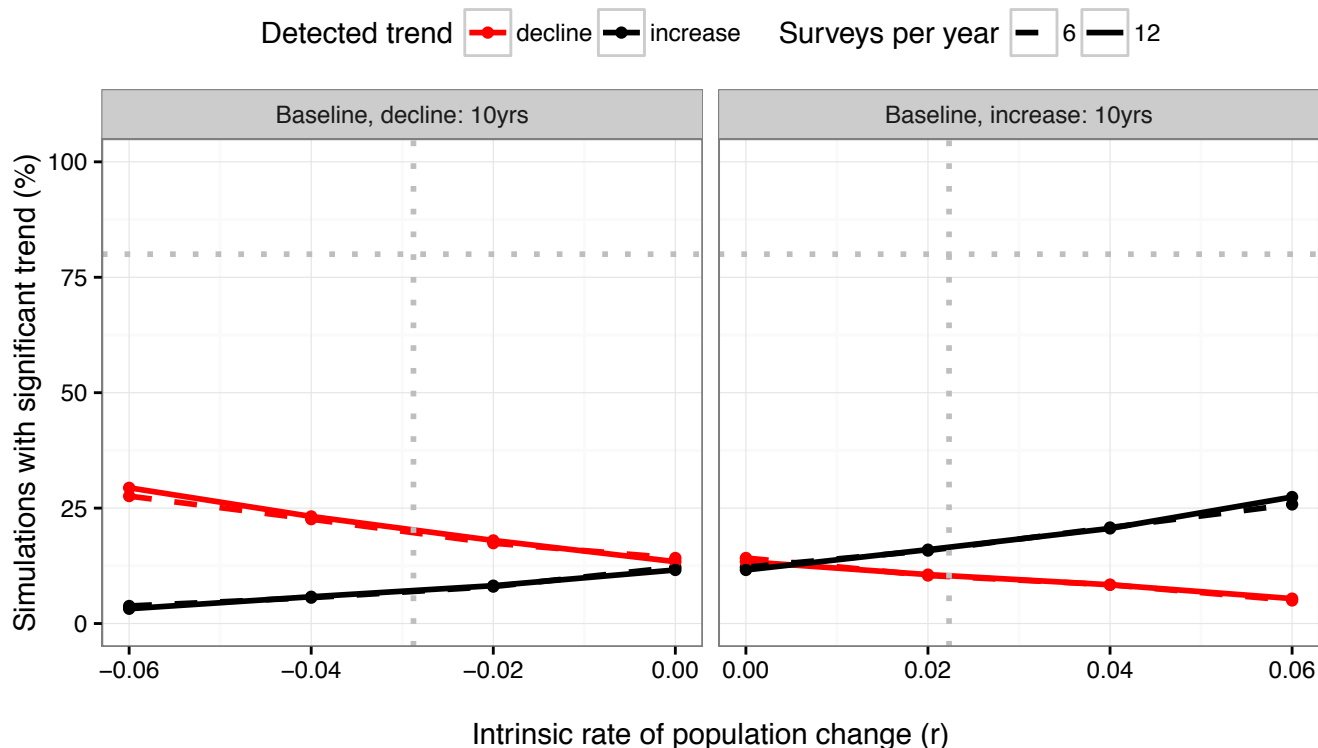


Figure 59. (SAAN *I. nebulifer*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), with surveys year-round. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, population declines (intrinsic population growth of $r = -0.06$ to 0.00) are simulated; right panel, population increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. Dashed lines represent results from conducting 6 surveys per monitoring year (alternate months throughout the year). Solid lines represent year-round monthly surveys. The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses. The dotted gray vertical lines mark the intrinsic rate of population change corresponding with a 25% change over 10 years.

Table 32. (SAAN *I. nebulifer*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline (first number in column) or a significant increase (second number in column) at two levels of α (0.10 and 0.20). Ideally, when $r = 0.00$, the sum of % significant decline and % significant increase should approximate α . Minor discrepancies in our results can be attributed to the use of likelihood ratio tests (instead of more accurate but computationally intensive bootstrapping) to calculate p-values for mixed models, and the relatively small number of simulations (500 per scenario) considered logistically feasible. The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for SAAN *I. nebulifer* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant ($\alpha=0.10$)	% significant ($\alpha=0.20$)
-0.06	Baseline	6	19 / 1	28 / 4
-0.06	Baseline	12	19 / 1	29 / 3
-0.04	Baseline	6	15 / 3	23 / 6
-0.04	Baseline	12	15 / 2	23 / 6
-0.02	Baseline	6	12 / 5	17 / 8
-0.02	Baseline	12	11 / 5	18 / 8
0.00	Baseline	6	9 / 7	14 / 12
0.00	Baseline	12	9 / 7	13 / 12

VICK

We analyzed data for two VICK species that accounted for 38% of all captures at the site. VICK species did not exhibit the strong seasonality in captures observed at JELA and GUIs, so VICK analyses were based on year-round data.

Captures for both VICK species were extremely low. More than a quarter of survey data were omitted from initial regression analyses for VICK species due to missing relative humidity or temperature information. However, neither of these covariates had support in candidate regression models, so the covariates were removed from analyses. Prior year's precipitation was highest for monitoring year 2015 and lowest for 2011, i.e., monitoring year 2014 was a wet year and 2010 a relatively dry year.

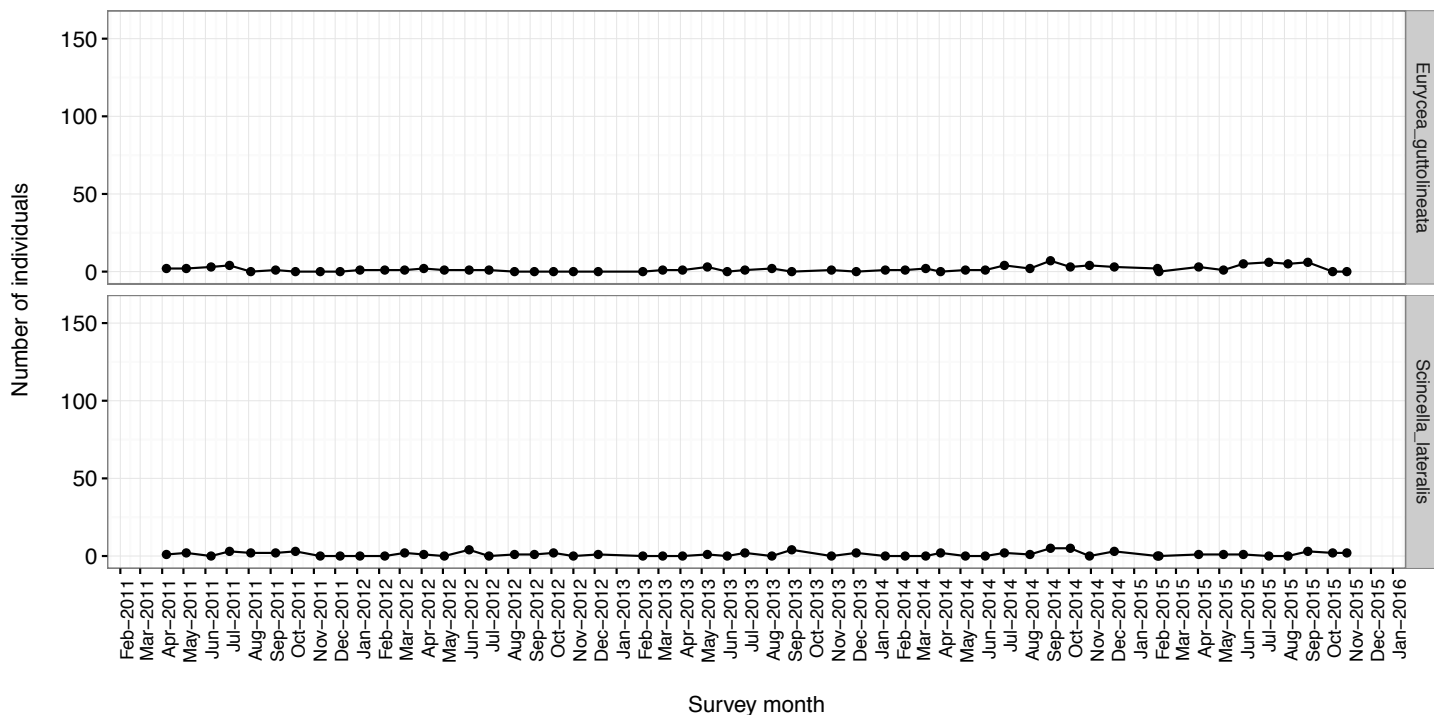


Figure 60. (VICK) Number of individuals captured each survey event, summed across traps, for two species (horizontal panels).

Table 33. Summary of VICK covariate data corresponding with survey dates, year-round. Minimum, median, and maximum values, and % of year-round data missing for covariates: ‘temp.f’ = temperature (Fahrenheit), ‘rh’ = relative humidity (%), ‘precip.7’ = mean daily precipitation in 7 days prior to survey (inches). Due to the high proportion of missing data for temperature and relative humidity, these two covariates were not included in analyses.

Statistic	temp.f	rh	precip.7
Minimum	17.14	19.68	0.00
Median	64.65	91.74	0.08
Maximum	78.90	100.00	1.18
% missing data	18.9	26.4	0

Table 34. (VICK) For each monitoring year, the mean monthly precipitation (in inches) for the prior monitoring year.

monitor.yr	mean.prior.precip
2011	2.95
2012	4.19
2013	4.99
2014	5.27
2015	5.66
2016	4.45

VICK: *Eurycea guttolineata*

Trap-counts of VICK *E. guttolineata* reached a maximum of six individuals at a single coverboard trap in a survey. The species was never captured at 50 of VICK's 60 coverboard traps. Almost all captures of *E. guttolineata* were at two traps (cb1 and cb2 of Transect B) located near seeps. To facilitate model estimation with this very sparse dataset, we analyzed the trap-counts as binary response variables summarized at the transect level (i.e., captured versus not captured on each transect). Even then, 75% of survey-transects were zeros (no animals captured on the transect during that survey).

The best-supported logistic regression model for VICK *E. guttolineata* included day-of-year. The AIC confidence set included two other models with AIC values within six units of the best-supported model. Those models differed from the best-supported model in that they excluded day-of-year, or included prior year's precipitation or the binary breeding season covariate as predictors. Coherent interannual variation was zero for all confidence set models. The regression coefficient for the binary breeding season covariate was considerably different for one confidence set model. The proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual change in probability of occupancy (probability of species capture) of -0.06 on the logit scale, over 10 years, was 49% for 12 surveys per year. The proportion of false trends for this scenario was 0%. Because VICK data were simulated in power analyses as logistic regression models, power analysis results for VICK are not directly comparable with the other GRAMP sites.



Figure 62. (VICK *E. guttolineata*) Bar plots showing the number of survey-transects in which *E. guttolineata* was captured (green) or not (gray), by monitoring year (left) and transect (right), year-round data. The height of each bar plot shows the total number of survey-transects. The species was not captured in 75% of survey-transects.

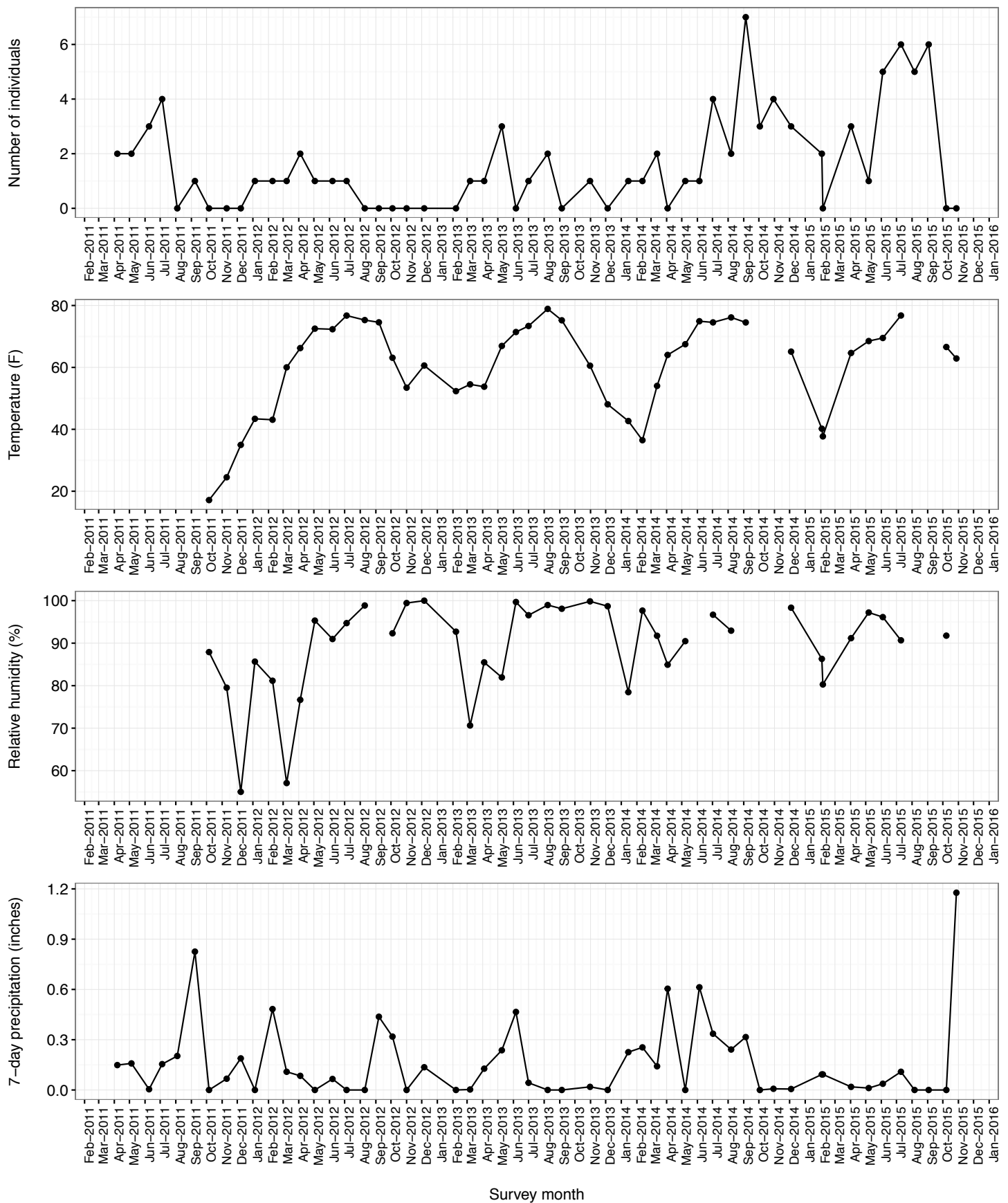


Figure 63. (*VICK E. guttolineata*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 7 days prior to each survey event.

Table 35. (*VICK E. guttolineata*) Best-supported regression model for *VICK E. guttolineata*, year-round data. Due to the high proportion of zeros (98.4%) in *VICK E. guttolineata* trap-level count data, we applied a logistic regression model for modeling species capture at the survey-transect level. The model response, ‘count’, is a binary variable with ‘1’ indicating the species was captured and ‘0’ indicating the species was not captured. Because data were analyzed at the transect—rather than the trap—level, ‘trap ID’ was not included as a random effect variable in the model. The fixed-effect covariate ‘my.day’ (day-of-year) was centered and scaled so its regression coefficient could be presented in a standardized unit.

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: count ~ transect + my.day + I(my.day^2) + (1 | monitor.yr) -
## 1
## Data: vick.egTL.scale.yr
##
##      AIC      BIC    logLik deviance df.resid
##    129.2    147.7    -58.6    117.2     153
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.7743 -0.3526 -0.2768  0.2355  4.4999
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## monitor.yr (Intercept) 1.435e-13 3.788e-07
## Number of obs: 159, groups:  monitor.yr, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## transectVICK_A  -2.1694     0.5694  -3.810 0.000139 ***
## transectVICK_B   1.0029     0.4166   2.407 0.016074 *
## transectVICK_C  -2.1694     0.5694  -3.810 0.000139 ***
## my.day           0.5819     0.2627   2.215 0.026773 *
## I(my.day^2)      -0.5682     0.2913  -1.951 0.051112 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              tVICK_A tVICK_B tVICK_C my.day
## trnscVICK_B  0.237
## trnscVICK_C  0.149  0.237
## my.day       -0.157  0.021 -0.157
## I(my.day^2) -0.346 -0.682 -0.346 -0.042
```

Table 36. (*VICK E. guttolineata*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. For all models, year variance was zero when rounded to two decimal points.

Model	dAIC	Number of parameters	Coherent interannual variation	Difference from global model
mod1	0.00	6	0	remove precip.7, mean.prior.precip and nb
mod2	3.89	6	0	remove precip.7 and my.day
mod3	4.75	9	0	remove mean.prior.precip and my.day

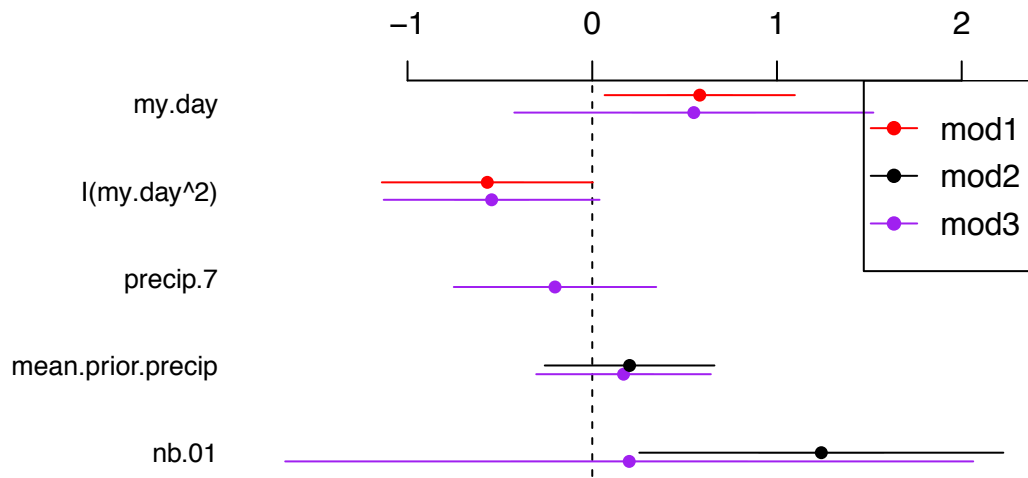


Figure 64. (*VICK E. guttolineata*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

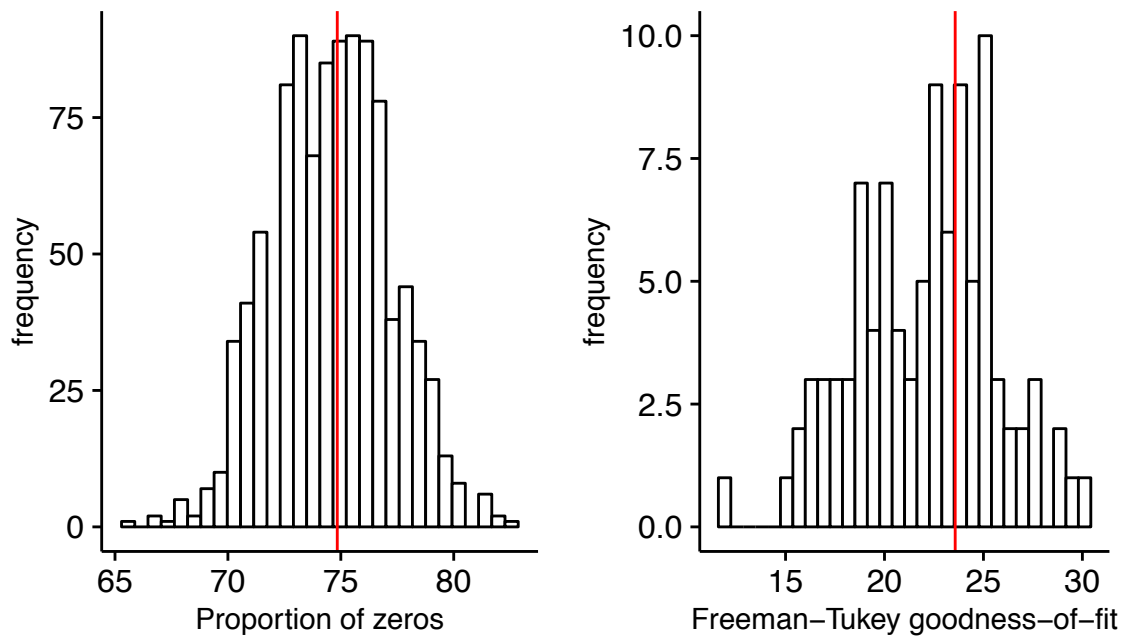


Figure 65. (*VICK E. guttolineata*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is good. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

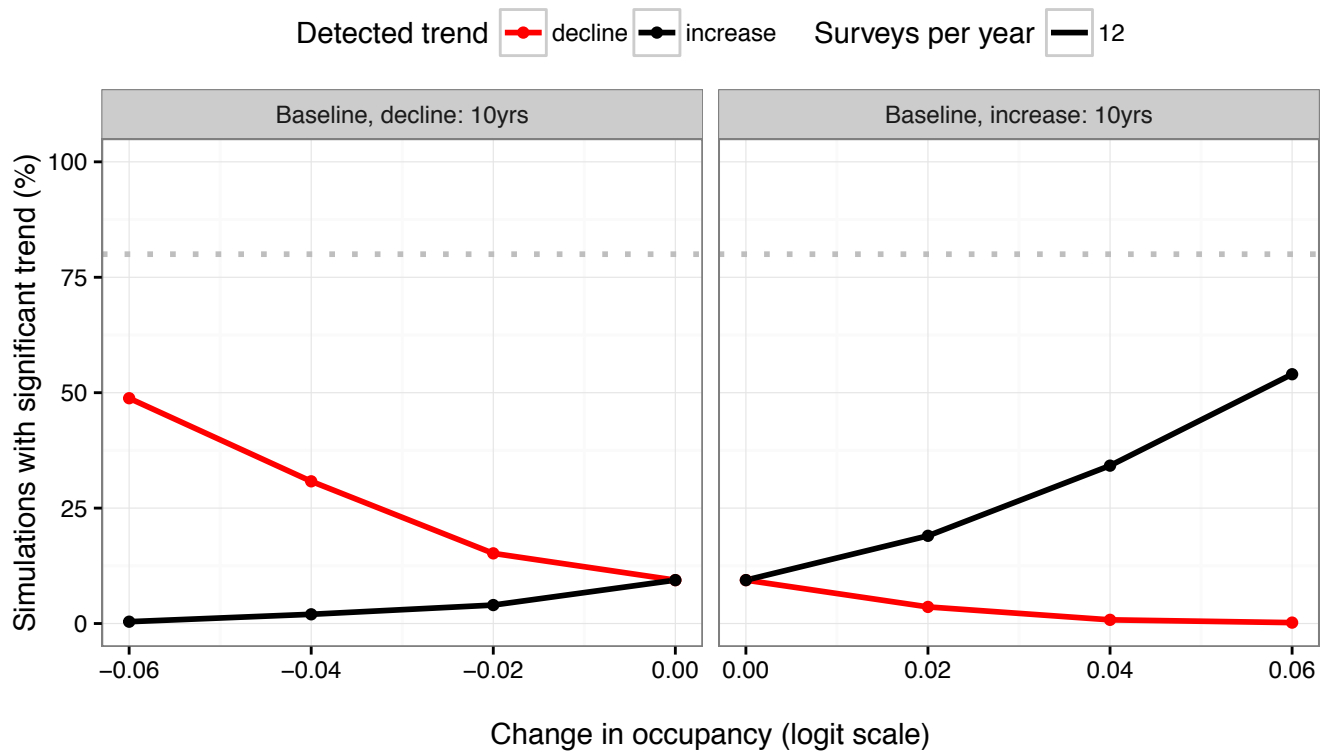


Figure 66. (*VICK E. guttolineata*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), with surveys year-round. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, declines in probability of occupancy at the survey-transect level are simulated; right panel, increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses.

Table 37. (*VICK E. guttolineata*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline (first number in column) or a significant increase (second number in column) at two levels of α (0.10 and 0.20). The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for *VICK E. guttolineata* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant ($\alpha=0.10$)	% significant ($\alpha=0.20$)
-0.06	Baseline	12	31 / 0	49 / 0
-0.04	Baseline	12	18 / 0	31 / 2
-0.02	Baseline	12	8 / 2	15 / 4
0.00	Baseline	12	4 / 4	9 / 9

VICK: *Scincella lateralis*

Trap-counts of VICK *S. lateralis* reached a maximum of 2 individuals at a single coverboard trap in a survey. Compared to VICK *E. guttolineata*, captures of *S. lateralis* were distributed more evenly among traps. However, similar to VICK *E. guttolineata*, trap-counts were very low. To facilitate model estimation with this sparse dataset, we analyzed the trap-counts as binary response variables summarized at the transect level (i.e., captured versus not captured on each transect). Even then, 72% of survey-transects were zeros (no animals captured on the transects).

The best-supported logistic regression model for VICK *S. lateralis* included day-of-year and the binary breeding season covariate. The AIC confidence set included one other model that had an AIC value within six units of the best-supported model. That model differed from the best-supported model in that it excluded the binary breeding season covariate as a predictor. Coherent interannual variation was zero for both confidence set models. The regression coefficient for day-of-year was fairly different between the two models. The proportion of zero-counts and the Freeman-Tukey statistic for the observed data were within the range of expected values calculated from simulations of the best-supported model, suggesting reasonable model fit.

With the current number of traps and $\alpha = 0.20$ for a two-tailed test, power to detect an annual change in probability of occupancy (probability of species capture) of -0.06 on the logit scale, over 10 years, was 61% for 12 surveys per year. The proportion of false trends for this scenario was 0%. Because VICK data were simulated in power analyses as logistic regression models, power analysis results for VICK are not directly comparable with the other GRAMP sites.

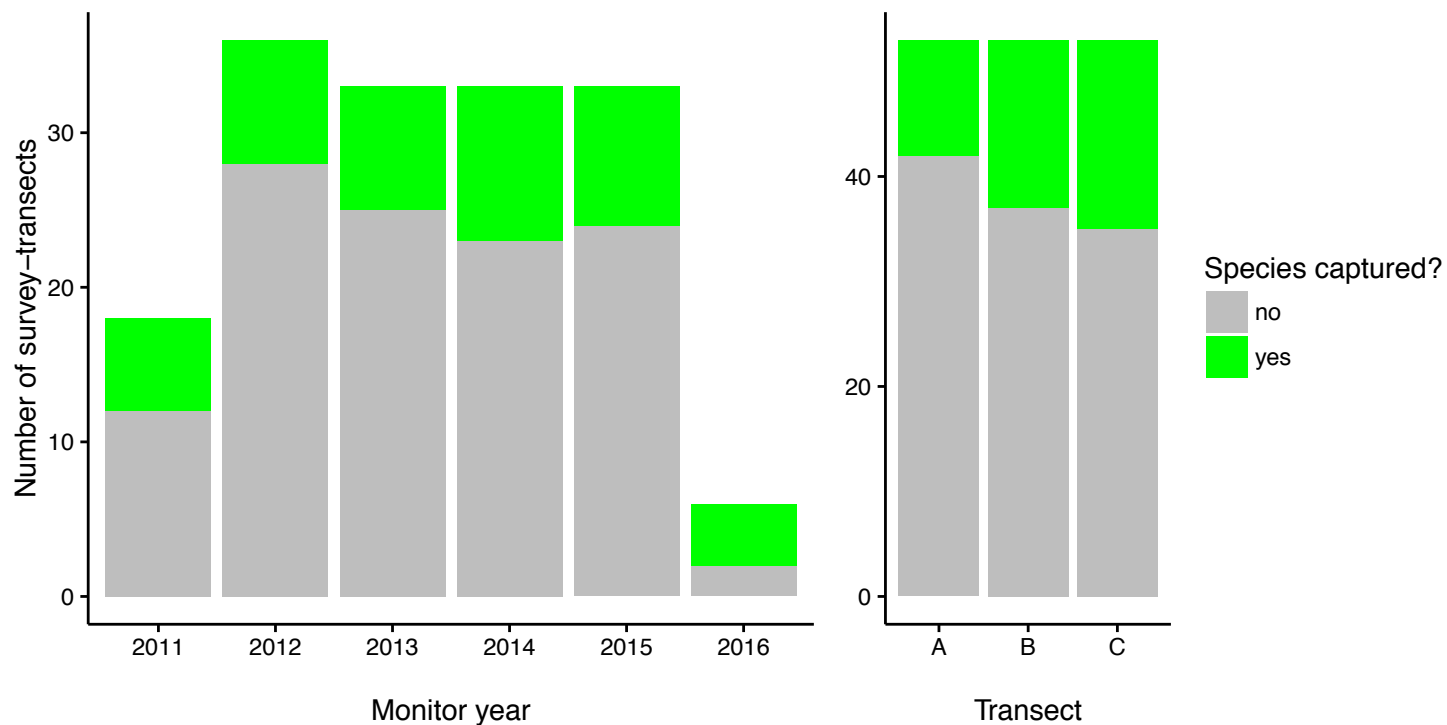


Figure 68. (VICK *S. lateralis*) Bar plots showing the number of survey-transects in which *S. lateralis* was captured (green) or not (gray), by monitoring year (left) and transect (right), year-round data. The height of each bar plot shows the total number of survey-transects. The species was not captured in 72% of survey-transects.

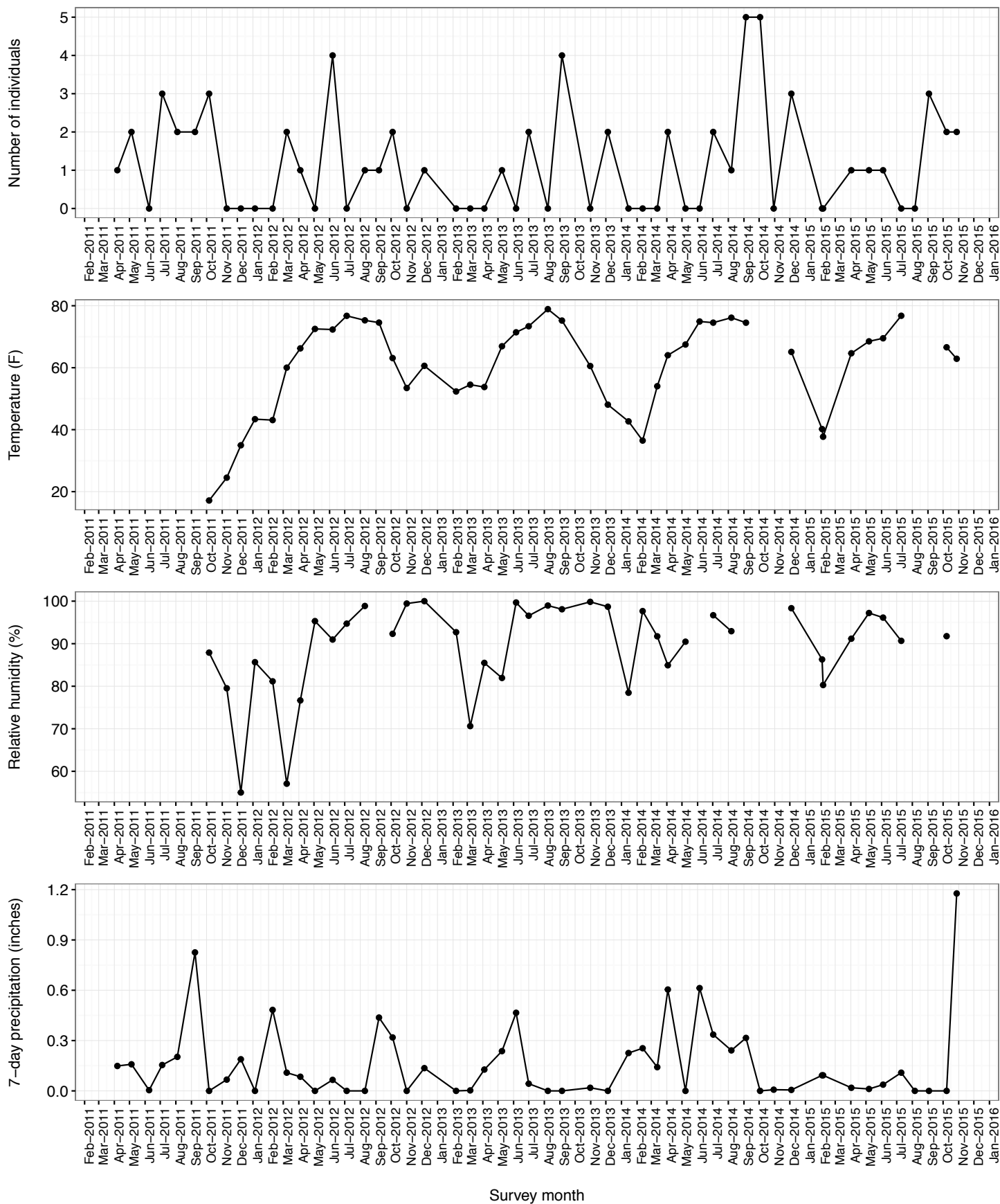


Figure 69. (*VICK S. lateralis*) Number of individuals captured (summed across traps) and survey-level level covariate values for each survey event, year-round data. Disconnected lines indicate missing data. The bottom graph shows mean daily precipitation for the 7 days prior to each survey event.

Table 38. (*VICK S. lateralis*) Best-supported regression model for *VICK S. lateralis*, year-round data. Due to the high proportion of zeros (98.1%) in *VICK S. lateralis* trap-level count data, we applied a logistic regression model for modeling species capture at the survey-transect level. The model response, ‘count’, is a binary variable with ‘1’ indicating the species was captured and ‘0’ indicating the species was not captured. The fixed-effect covariate ‘my.day’ (day-of-year) was centered and scaled so its regression coefficient could be presented in a standardized unit. ‘nb.01’ refers to a binary predictor indicating if the survey was conducted during a breeding or non-breeding season month.

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula:
## count ~ transect + my.day + I(my.day^2) + nb.0 + (1 | monitor.yr) -      1
## Data: vick.slTL.scale.yr
##
##      AIC      BIC    logLik deviance df.resid
##    181.8    203.3    -83.9    167.8     152
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.2468 -0.6662 -0.4564  0.8358  4.1336
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## monitor.yr (Intercept) 0          0
## Number of obs: 159, groups: monitor.yr, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## transectVICK_A  -3.7587     0.9336  -4.026 5.67e-05 ***
## transectVICK_B  -3.2002     0.9052  -3.536 0.000407 ***
## transectVICK_C  -3.0030     0.8955  -3.353 0.000798 ***
## my.day          -0.5635     0.4265  -1.321 0.186458
## I(my.day^2)      0.8990     0.2445   3.677 0.000236 ***
## nb.01           2.2780     1.0616   2.146 0.031891 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              tVICK_A tVICK_B tVICK_C my.day I(.^2)
## trnscVICK_B  0.866
## trnscVICK_C  0.867  0.878
## my.day       0.748  0.767  0.772
## I(my.day^2) -0.693 -0.687 -0.683 -0.345
## nb.01       -0.875 -0.892 -0.896 -0.906  0.494
```

Table 39. (*VICK S. lateralis*) AIC results and model variance estimates for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. ‘dAIC’ is the difference in AIC from the best-supported model. For all models, year variance was zero when rounded to two decimal points.

Model	dAIC	Number of parameters	Coherent interannual variation	Difference from global model
mod1	0.00	7	0	remove precip.7 and mean.prior.precip
mod2	3.55	6	0	remove precip.7, mean.prior.precip and nb

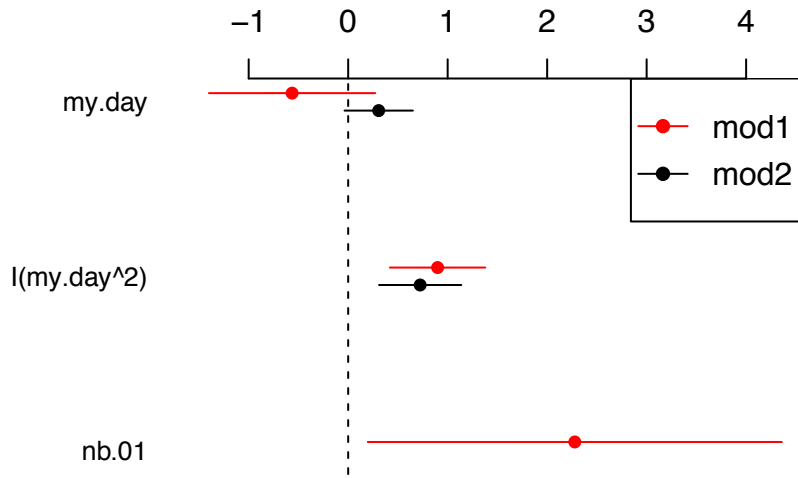


Figure 70. (*VICK S. lateralis*) Regression coefficients with 95% confidence intervals for covariates in the “confidence set” of models. Models within six AIC of the best-supported model and with fewer or an equal number of parameters. Confidence intervals are naive estimates based on Wald standard errors. For this figure, all covariates were centered and scaled for easier comparison of effect sizes.

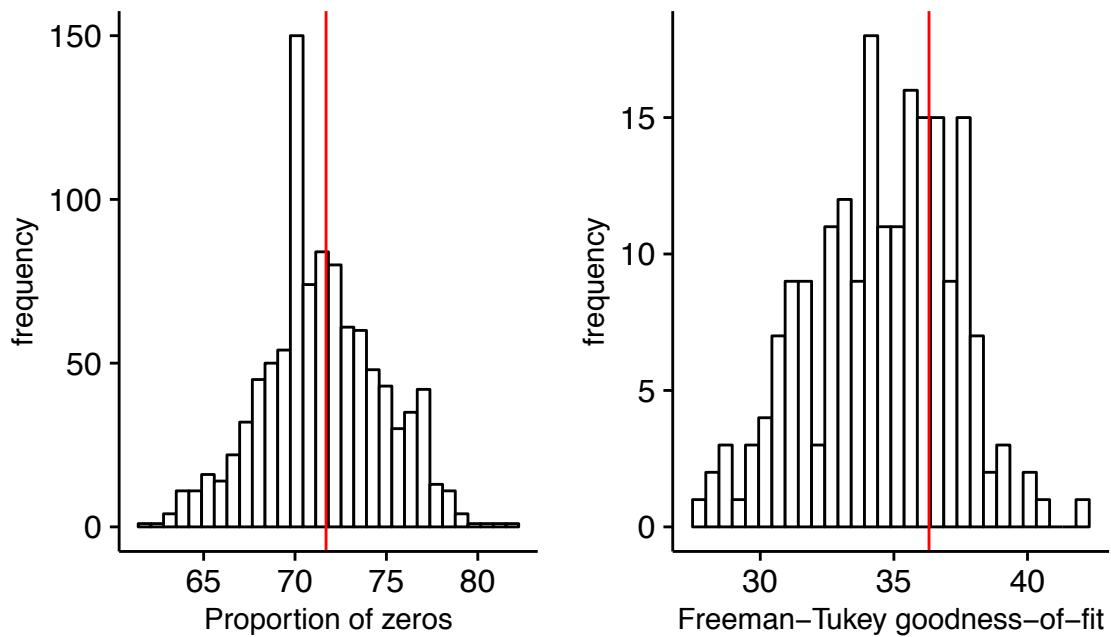


Figure 71. (*VICK S. lateralis*) (left) Simulated versus observed proportion of zeros and (right) Freeman-Tukey goodness-of-fit statistic, non-breeding season only. Model fit is adequate. A red vertical line (representing observed data) in the extreme tails of the distribution would have suggested poor model fit.

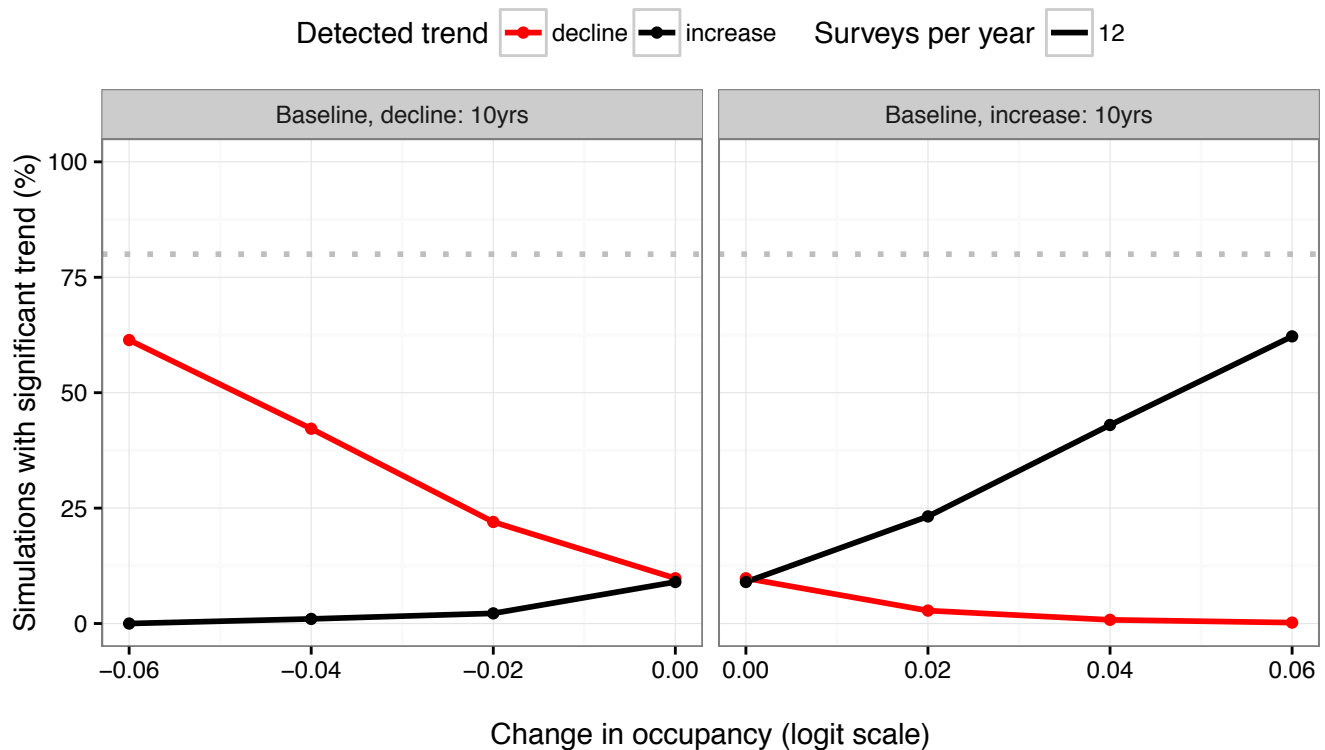


Figure 72. (*VICK S. lateralis*) Statistical power to detect a simulated trend after 10 YEARS OF BASELINE SAMPLING (current number of traps), with surveys year-round. Results are shown for $\alpha = 0.20$ in a two-tailed test. In the left panel, declines in captures at the survey-transect level are simulated; right panel, increases. Red lines indicate the proportion of 500 simulations in which a significant decline was observed; black lines indicate a significant increase. The dotted gray horizontal line is 80% statistical power, a common benchmark for power analyses.

Table 40. (*VICK S. lateralis*) Statistical power to detect SIMULATED DECLINES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline (first number in column) or a significant increase (second number in column) at two levels of α (0.10 and 0.20). The numbers in this table for baseline trap effort at $\alpha = 0.20$ correspond with the power analysis graph for *VICK S. lateralis* 'Baseline, decline: 10yrs'. The power analysis results for other scenario combinations are provided in Appendix A1.

r	Number of traps	Surveys per year	% significant ($\alpha=0.10$)	% significant ($\alpha=0.20$)
-0.06	Baseline	12	45 / 0	61 / 0
-0.04	Baseline	12	26 / 0	42 / 1
-0.02	Baseline	12	12 / 1	22 / 2
0.00	Baseline	12	4 / 4	10 / 9

DISCUSSION

Statistical power and coherent interannual variation

The low statistical power estimated for GRAMP surveys was well within expectations for amphibian studies. Amphibian dynamics are often characterized by high interannual variation, episodes of high recruitment, and complex interactions with environmental covariates (Hayes and Steidl 1997; Marsh 2001; Scherer and Tracey 2011), so distinguishing trends from noise in amphibian data is particularly difficult. In simulations of GRAMP surveys, power to detect a statistically significant decline was negatively correlated with coherent interannual variation (with a simulated decline of 6% per year over 10 years, Pearson's $r = -0.96$, $p < 0.001$; Figure 73). Coherent interannual variation was also an important determinant of false trends, i.e., the proportion of simulations in which an increasing trend was detected, even though a deterministic decline was simulated (with a simulated decline of 6% per year over 10 years, Pearson's $r = 0.95$, $p = 0.001$).

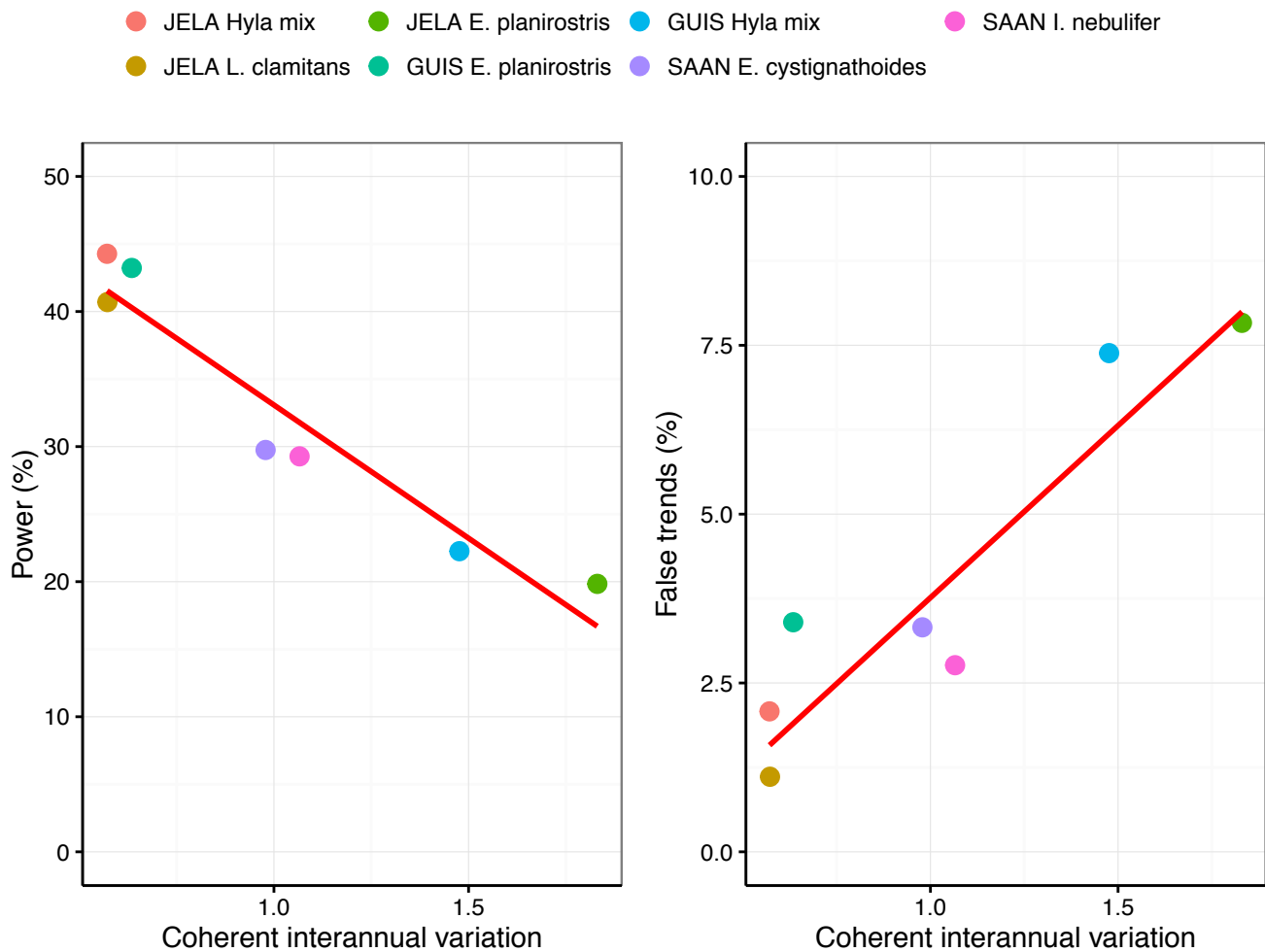


Figure 73. Relationship between coherent interannual variation and power statistics. Results are shown for a simulated decline of 6% per year over 10 years, with monthly surveys (for JELA and GUIS, only non-breeding season months were considered), using the baseline number of traps, and with $\alpha = 0.20$. Each point represents a site-species. ‘Coherent interannual variation’ is estimated as the standard deviation of the ‘year’ random effect (logarithmic scale) from the best-supported model for each site-species. ‘Power’ is the proportion of 500 simulations in which a significant decline was detected. ‘False trends’ is the proportion of 500 simulations in which a significant increase was detected. VICK species are not included because regression models for VICK used a logit instead of a log link function.

When coherent interannual variation is large relative to the magnitude of a deterministic decline, chance can play a large role in the trend that is actually observed, especially for short monitoring durations. A population may be slowly (but noisily) declining over 10 years but if, for example, coherent interannual variation around a deterministic decline generates a “good year” for amphibian captures in Year 8, then the perceived trend may be a statistically significant increase. Generally speaking,

extreme values near the beginning or end of a time series can have high leverage in regression analyses. When coherent interannual variation is large, extreme years are more common, so chances are good that an extreme year early or late in a time series could obscure a deterministic trend or even suggest the opposite trend. The observed relationships in Figure 73 are consistent with these ideas. However, keep in mind that estimates of coherent interannual variation based on only five years of amphibian population counts are highly tentative (Marsh 2001). Therefore, power estimates from the GRAMP pilot data should be considered preliminary.

Statistical power and survey variables

For each site-species we evaluated how power changed under different scenario combinations of monitoring duration, within-year survey frequency, number of traps, trend magnitude, and alpha (Appendix A1). Power to detect simulated increases was roughly similar to that for detecting simulated declines.

The influence of monitoring duration on GRAMP statistical power depends on how the question is asked. NPS was interested in the question “How does increasing monitoring duration influence power to detect an *overall 25% decline* in captures?” To answer this question, we would compare the power to detect (approximately) a 6% decline in captures per year over 5 years, to a 3% decline in captures per year over 10 years (Table 5), and we would find there is almost no impact of monitoring duration. If the question is phrased a little differently—“How does increasing monitoring duration influence power to detect an *annual 6% decline* in captures?”—we would find that for site-species with relatively small coherent interannual variation, increasing monitoring duration increases statistical power. For example, for the three site-species with the smallest coherent interannual variation (JELA *Hyla* mix, JELA *L. clamitans*, and GUIS *E. planirostris*), increasing monitoring duration from 5 to 10 years improved statistical power by 25% to 75%. The probability of false trends also declined considerably with longer monitoring duration for these site-species. For site-species with large coherent interannual variation, increasing monitoring duration had relatively little impact on power or false trends.

For a given rate of annual population change, monitoring duration affects overall population change, so it makes sense that this survey variable could have a large influence on power and false trends. For example, if a population declines 6% per year for 5 years, the population will experience an overall 26% decline from the initial population size. If that same rate of decline (6% per year) is extended an additional 5 years, the population size will experience an overall 45% decline from the initial population size. We would have greater power to detect the 10-year decline primarily because it represents a larger overall population change (45% compared to 26% decline).

Survey frequency (surveying in alternate months versus monthly) had no noticeable effect on power for any site-species. As simulated, surveys in alternate months meant surveying three months per year at JELA and GUIS, and six months per year at SAAN. We did not evaluate power for surveys in alternate months at VICK.

The effects of increasing traps by 50% was evaluated for JELA *Hyla* mix and JELA *E. planirostris*. Increasing traps improved power by less than 5% for the surveyed *Hyla* species. Results were similar whether the additional traps were added to existing or to new transects.

Model selection

Power analysis results were reported for the best-supported model for each site-species. Given the uncertainties surrounding AIC-based model selection for mixed models, we also examined a larger set of “good” models (the confidence set) for each site-species to assess how power estimates may have changed if analyses had been based on other models. Except for GUIS *E. planirostris* (which had six models), the number of models in confidence sets for each site-species was three or fewer. Although parameter combinations sometimes changed, the fixed effects coefficients for included parameters changed little among models in confidence sets. Variance estimates for the trap random effect were also stable among models.

For all site-species, the best-supported model had among the smallest coherent interannual variation of all confidence set models. Because statistical power was negatively correlated with coherent interannual variation, the best-supported model provided relatively high statistical power, compared to other models in the confidence set. For three site-species the coherent interannual variation for some candidate models was much higher than for the best-supported model. In two cases (JELA *L. clamitans* and GUIS *E. planirostris*) the candidate models with larger coherent interannual variation excluded prior year’s precipitation as a predictor, whereas the best-supported models included this variable. In the third case (JELA *E. planirostris*), candidate models with larger coherent interannual variation excluded another precipitation covariate—precipitation within the past 30 days (‘precip.30’)—as a predictor, whereas the best-supported model included this variable.

These findings suggest that power analyses results would have been relatively insensitive to the choice of a best-supported model from confidence sets, except for three site-species. For those three site-species, power analyses based on less supported models in the confidence set would have resulted in lower estimates of statistical power.

Covariate relationships

In regression models with correlated variables, individual coefficient estimates and even their signs (positive or negative) may be influenced by other variables in the model. With these caveats in mind and also considering cautions against drawing conclusions from short time series of amphibian counts, we summarize observed patterns in covariate relationships in GRAMP best-supported models.

Day-of-year and (for analyses of year-round data) the binary breeding season predictor ('nb') were important predictors of captures for most site-species, suggesting seasonal patterns in captures were not fully accounted for with the environmental covariates we analyzed. At JELA and GUIS, captures during the breeding season (April through September) were so low compared to the non-breeding season (October through March) that we focused our analyses of these sites only on data from the non-breeding season. At SAAN and VICK, capture differences between the two seasons were not as extreme, but the breeding season variable was still an important predictor for 3 out of 4 species. We classified breeding season very broadly in our analyses, as a 6-month period that would encompass the breeding activities of most amphibians in most years. Using species- and year-specific breeding dates based on annual field observations of breeding behaviors may further improve model fit and statistical power and additionally provide insights on how climate-related changes in breeding timing may affect future amphibian captures.

Including environmental covariates (temperature, relative humidity, recent precipitation, or prior year's precipitation) improved model fit for most species, reducing coherent interannual variation and residual error. Between one and three environmental covariates were retained in the best-supported models for most of the species analyzed at JELA, GUIS, and SAAN. This finding is consistent with expectations. Amphibians' activity levels, and therefore capture rates, depend on temperature and moisture because amphibians are cold-blooded and their skin must also remain moist for the gas exchange necessary for breathing (Bellis 1962; Reynolds et al. 1994; Roloff et al. 2011; Semlitsch 2002).

For *Hyla* species (tree frogs), temperature was an important covariate. Trap-counts were highest at temperatures between 55F and 60F, but temperature was also highly correlated with day-of-year at all sites (Pearson's $r > 0.70$) and its independent effects are unknown. For *Eleutherodactylus* species, relative humidity and recent precipitation were relevant covariates. *Eleutherodactylus* species do not have an aquatic larval stage. Instead, their young hatch from eggs laid in moist substrates. For these terrestrial species, activity levels may be especially sensitive to atmospheric moisture.

Prior year's precipitation was retained in the best-supported model for two site-species. In both cases, other models included in the confidence set did not include this variable, suggesting that statistical support for this variable may have been weak. Precipitation influences breeding activity and recruitment success for many amphibian species, with consequences for future population size (Reynolds et al. 1998; Semlitsch et al. 1996). But the relationship between precipitation and future abundance is complex, and decades of monitoring data may be required to understand how this yearly variable influences amphibian dynamics. Trenham et al. (2003) reported that frog abundances were correlated with precipitation ranging from one to four years prior. Other studies have suggested that the influence of precipitation on amphibian dynamics is indirect through local hydroperiod, or that the timing of precipitation is more relevant than the amount of precipitation (Marsh 2001; Semlitsch et al. 1996). Not surprisingly, many monitoring studies have been unable to show a strong association between precipitation and amphibian abundance (Marsh 2001; Meyer et al. 1998). In addition to the short duration of data available for our analyses, we obtained precipitation data from off-site weather stations. Longer time series and site-specific precipitation data may be necessary to gain a clearer understanding of how this variable influences amphibian dynamics at GRAMP sites.

The best-supported models for VICK did not include environmental covariates. Data for VICK species were sparse, with only three binary data points (one per transect) from each survey. With such data, it would be difficult to identify anything other than very strong covariate relationships for captures.

Survey considerations

GRAMP surveys appear capable of meeting diverse monitoring objectives. But unfortunately, amphibians are a difficult taxa for monitoring long-term abundance trends with the 80% power often desired in resource management. Based on analyses of the first five years of GRAMP data, we offer the following preliminary considerations to increase understanding of population dynamics for the nine site-species analyzed in this report. These considerations do not address other GRAMP objectives, which may have conflicting survey requirements. We also discuss issues of replication and detectability in amphibian monitoring, which may be applicable to GRAMP objectives.

SURVEY COVARIATES

- Temperature and relative humidity were identified as important predictors of capture for several site-species. Averaged across sites, these covariate data were missing for approximately 16% of survey periods, i.e., they were not available for any of the five mornings centered on a survey date. We excluded from analyses survey dates for which associated covariate data were missing. Although interpolation is sometimes used at the analytical stage to “fill in” missing covariate data, more accurate analyses could be achieved by ensuring temperature and relative humidity data are available for every survey event.
- The temperature and perhaps also relative humidity at a GRAMP site may change considerably from the start of an amphibian survey (typically early morning) to the time a survey is completed several hours later. It may be possible to obtain better coefficient estimates for these covariates, and therefore reduce model error, if the covariates can be included in models at a finer temporal scale such as the transect survey level. For example, if the start and end time for surveying each transect is recorded and data loggers are operating at 10-minute intervals, then captures for each transect can be associated with a temperature and relative humidity that better reflects the conditions at the time each transect was surveyed. These fine-scale covariate relationships may inform understanding of patterns at larger temporal scales.
- The precipitation data used in analyses of GRAMP pilot data were obtained from off-site weather stations. We assumed in analyses that these off-site stations accurately portrayed precipitation patterns at GRAMP sites, but this assumption has not been verified.
- Amphibian detection rates may be influenced by changing activity patterns during the breeding cycle. The exact timing of the breeding cycle for a site-species may differ among years, in accordance with environmental cues. More accurate determination of the annual breeding period for each site-species might improve model estimates and provide insights on how environmental covariates interact with breeding activity to influence capture rates and population trends. Depending on species, mating calls or other signs could be used to estimate the timing of breeding cycles.
- One of the most effective ways to minimize coherent interannual variation in GRAMP models, and therefore increase power to detect long-term trends, may be to incorporate additional covariates that explain some of the year-to-year changes in amphibian counts. Additional relevant covariates may include vegetation structure and type, and availability of freshwater bodies including coverage, proximity, and transience (Grover 1998; Brodman 2009; Smith et al. 2014). These factors may influence amphibian captures from survey-to-survey and from year-to-year by, for example, affecting the availability of natural cover at a site or influencing amphibian movement and breeding activities. Collecting additional covariate data can be time consuming, however, and the potential value of such efforts may differ across sites and by species.

SURVEY EFFORT

- Species at JELA and GUIB exhibited strong seasonal trends in captures, with very few captures during the breeding season months. For the purpose of detecting long-term trends in these site-species, the value of surveying during the breeding season months may be very low. If reducing survey effort is a priority, one possibility would be to discontinue surveys at these sites for the 4-month period of May through August. In the first five years of GRAMP surveys, almost no captures were recorded during these months for the three most common amphibian species. However, the timing of the breeding season may shift as environmental conditions change in the future, possibly requiring adjustments in survey timing.
- Analyses of pilot data suggested that GRAMP survey frequency can be reduced without much impact on power. Many long-term amphibian monitoring programs survey less frequently than suggested in the GRAMP protocol, but the frequency and timing of survey effort varies by program. We simulated a reduction of surveys to every other month instead of monthly. However, a less drastic reduction of survey effort could be implemented, such as once every six weeks. Alternatively, surveys could be conducted monthly for three months during the breeding season and three months during the non-breeding season, as some programs have adopted. If resources are limited, statistical power could be much improved and broader inference gained by allocating more effort to collecting additional relevant covariate data and to increasing the number of GRAMP sites at each park (see below, ‘Site Replication’), rather than by conducting frequent surveys at existing GRAMP sites. Because day-of-year is included as a survey covariate in regression models, changing survey frequency and timing would not preclude use of previously collected GRAMP data in analyses. Scenarios of reduced survey frequency were not evaluated for VICK.
- Increasing the number of traps at a site may have limited effect on power to detect long-term trends. This scenario was only evaluated for JELA.

SITE REPLICATION

- GRAMP surveys are not intended to provide park-wide inference on amphibian trends, but in some cases are intended to support local condition and change assessment at sites representative of particular habitat types. However, even if a site is representative of a habitat, a species decline at a single site would provide little inference about its status or trend at other sites of the same habitat type. Increasing the number of GRAMP sites in a park could incur considerable costs in terms of time and effort. But information on amphibian trends at multiple sites could help distinguish between local versus more widespread changes in amphibian dynamics, and can increase understanding of amphibian metapopulation dynamics, even if sites cannot be randomly selected. For amphibian monitoring, this distinction between local and widespread trends could be important (Alford and Richards 1999). In addition, data from multiple sites could be combined in a single analysis (with ‘site ID’ as an additional parameter) to obtain better covariate estimates and improve the accuracy of model predictions for all sites.

DETECTION PROBABILITIES

- In our analyses of GRAMP data, trap-counts were not adjusted for detection probabilities. Detection probability—the probability of detecting an animal that is present—may vary among species, sites, survey methods, survey events (e.g., due to weather or other survey conditions), and over time. Counts uncorrected for changing detection probabilities may lead to incorrect conclusions about population trends. For example, a change in availability of natural cover at a site may cause amphibians to increase or reduce their use of artificial cover traps. This behavioural response may be misinterpreted as a change in population status. Many amphibian studies have incorporated detection estimates in their survey design and analyses (Corn et al. 2005; Dodd & Dorazio 2004; Gomez-Rodriguez et al. 2010; Hyde & Simons 2001; Smith et al. 2014; Tanadini & Schmidt 2011; Weir et al. 2005). Such studies require repeated surveys over short time periods during which populations are assumed closed to changes in abundance or occupancy. The timespan for population closure varies by species. For many amphibians, this timespan may be one to several weeks (checking coverboards more frequently than once a week may cause too much disturbance; Dodd 2003). A GRAMP protocol that accommodates detection estimation could, for example, suggest every site be surveyed every other month, with each survey event comprising two independent surveys a week apart. Pilot data would be required to determine if such a protocol could improve trend determination in GRAMP studies.

LITERATURE CITED

- Alford, R.A. & Richards, S.J. 1999. Global amphibian declines: A problem in applied ecology. *Annual Review of Ecology and Systematics* 30:133–165.
- Arnold, T.W. 2010. Uninformative parameters and model selection using Akaike’s Information Criterion. *Journal of Wildlife Management* 74:1175–1178.
- Baldwin, R.F., Calhoun, A.J.K. & deMaynadier, P.G. 2006. The significance of hydroperiod and stand maturity for pool-breeding amphibians in forested landscapes. *Canadian Journal of Zoology* 84:1604–15.
- Bellis, E.D. 1962. The influence of humidity on wood frog activity. *The American Midland Naturalist* 68(1): 139-148.
- Bolker, B. & R Development Core Team. 2014. *bbmle: Tools for general maximum likelihood estimation*. R package version 1.0.17. <https://CRAN.R-project.org/package=bbmle>.
- Brodman, R. 2009. A 14-year study of amphibian populations and metacommunities. *Herpetological Conservation and Biology* 4:106–119.
- Burnham, K.P. & Anderson, D.R. 2002. *Model selection and multimodel inference: a practical information theoretic approach*. 2nd edition. Springer-Verlag, New York, New York.
- Corn, P.S., Adams, M.J., Battaglin, W.A., Gallant, A.L., James, D.L., Knutson, M., Langtimm, C.A., Sauer, J.R. 2005. *Amphibian Research and Monitoring Initiative—Concepts and implementation: U.S. Geological Survey Scientific Investigations Report 2005–5015*, 23 p.
- Dochtermann, N. & Jenkins, S. 2010. Developing and evaluating multiple hypotheses in behavioral ecology. *Behavioral Ecology and Sociobiology* 65(1):37-45.
- Dodd, Jr., C.K. 2003. *Monitoring Amphibians in Great Smoky Mountains National Park*. U.S. Geological Survey Circular No. 1258. 117 pp.

- Dodd, Jr., C.K. & Dorazio, R.M. 2004. Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. *Herpetologica* 60(4):468-478.
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A. & Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 27(2):233-249.
- Freeman, M.F. & Tukey, J.W. 1950. Transformations related to the angular and the square root. *Annals of Mathematical Statistics* 21:607-611.
- Gomez-Rodriguez C., Guisan, A. Diaz-Paniagua, C. & Bustamante, J. 2010. Application of detection probabilities to the design of amphibian monitoring programs in temporary ponds. *Annales Zoologici Fennici* 47(5):306-322.
- Grant, B.W., Tucker, A.D., Lovick, J.E., Mills, A.M., Dixon, P.M, and J.W. Gibbons. 1992. The use of coverboards in estimating patterns of reptile and amphibian diversity. Pages 279-403 *in* R. Siegel and N. Scott (eds.). *Wildlife 2001*. Elsevier Science, London, England.
- Green, D.H. 2003. The ecology of extinction: population fluctuation and decline in amphibians. *Biological Conservation* 111:331-343.
- Grover, M.C. 1998. Influence of cover and moisture on abundances of the terrestrial salamanders *Plethodon cinereus* and *Plethodon glutinosus*. *Journal of Herpetology* 32:489-497.
- Hayes, J.P. & Steidl, R.J. 1997. Statistical power analysis and amphibian population trends. *Conservation Biology* 11:273-275.
- Hyde, E.J. & Simons, T.R. 2001. Sampling plethodontid salamanderes-sources of variability. *Journal of Wildlife Management* 65:624-632.
- Johnson, P.C.D., Barry, S.J.E., Fergguson, H.M. & Müller, P. 2015. Power analysis for generalized linear mixed models in ecology and evolution. *Methods in Ecology and Evolution* 6:133-142.
- Joppa, L.N., Williams, K.R., Temple, S.A. & Casper, G.S. 2009. Environmental factors affecting sampling success of artificial cover objects. *Herpetological Conservation and Biology* 5:143-148.
- Maas, C.J.M. & Hox, J.J. 2005. Sufficient sample sizes for multilevel modeling. *Methodology* 1: 86-92.
- Marsh, D.M. 2001. Fluctuations in amphibian populations: a meta-analysis. *Biological Conservation* 101:327-335.
- Meyer, A.H., Schmidt, B.R. & Grossenbacher, K. 1998. Analysis of three amphibian populations with quarter-century long time-series. *Proceedings: Biological Sciences* 265(1395):523-528.
- Paton, P.W.C. & Crouch, W.B. 2002. Using the phenology of pond-breeding amphibians to develop conservation strategies. *Conservation Biology* 16(1): 194-204.
- Pechmann, J.H.K., Scott, D.E., Gibbons, J.W. & Semlitsch, R.D. 1989. Influence of wetland hydroperiod on diversity and abundance of metamorphosing juvenile amphibians. *Wetlands Ecology and Management* 1(1):3-11.
- Pechmann, J.H.K., Scott, D.E., Semlitsch, R.D., Caldwell, J.P., Vitt, L.J. & Gibbons, J.W. 1991. Declining amphibian populations: the problem of separating human impacts from natural fluctuations. *Science* 253:892-895.
- Pechmann, J.H.K. & Wilbur, H.M. 1994. Putting declining amphibian populations in perspective: natural fluctuations and human impacts. *Herpetologica* 50: 65-84.
- Petitot, M., Manceau, N., Geniez, P. & Besnard, A. 2014. Optimizing occupancy surveys by maximizing detection probability: application to amphibian monitoring in the Mediterranean region. *Ecology and Evolution* 4(18):3538-3549.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reynolds, R.P., Crombie, R.I. & R.W. McDiarmid. 1994. Keys to a successful project: associated data and planning. *In* Heyer, W.R., Donnelly, M.A., McDiarmid, R.W., Hayek, L.C. & Foster, M.S. (Eds.), *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians* (pp. 66-71). Washington, DC: Smithsonian Institution Press.
- Rice J. 1995. *Mathematical statistics and data analysis* (2nd ed). Duxbury: Belmont, CA.
- Richards, S.A., Whittingham, M.J. & Stephens, P.A. 2011. Model selection and model averaging in behavioural ecology: the utility of the IT-AIC framework. *Behavioral Ecology and Sociobiology* 65:77-89.
- Roloff, G.J., Grazia, T.E., Millenbah, K.F. & Kroll, A.J. 2011. Factors associated with amphibian detection and occupancy in southern Michigan forests. *Journal of Herpetology* 45(1):15-22.
- Scherer, R.D. & Tracey, J.A. 2011. A power analysis for the use of counts of egg masses to monitor wood frog (*Lithobates sylvaticus*) populations. *Herpetological Conservation & Biology* 6:81-90.

- Semlitsch, R.D. 2002. Critical elements for biologically-based recovery plans for aquatic-breeding amphibians. *Conservation Biology* 16(3): 619-629.
- Semlitsch, R.D. & Pechmann, J.H.K. 1985. Diel pattern of migratory activity for several species of pond-breeding salamanders. *Copeia* 1985:86-91.
- Semlitsch, R.D., Scott, D.E., Pechmann, J.H.K. & Gibbons, J.W. 1996. Structure and dynamics of an amphibian community: Evidence from a 16-yr study of a natural pond. Pages 217–248 *in* M.L. Cody and J.A. Smallwood (eds.). *Long-Term Studies of Vertebrate Communities*. Academic Press, USA.
- Shmueli, G. 2010. To explain or to predict? *Statistical Science* 25(3):289-310.
- Skaug, H., Fournier, D., Bolker, B., Magnusson, A. & Nielsen, A. 2016. Generalized Linear Mixed Models using ‘AD Model Builder’. R package version 0.8.3.3.
- Smith, D.H.V., L. Randall, B.L. Jones, and D.R.C. Prescott. 2014. Difference in detection and occupancy between two anurans: The importance of species-specific monitoring. *Herpetological Conservation and Biology* 9:267–277.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.R. & van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B* 64: 583–616.
- Stephens, P.A., Buskirk, S.W., Hayward, G.D. & Del Rio, C.M. 2007. A call for statistical pluralism answered. *Journal of Applied Ecology* 44: 461–463.
- Symonds, M.R. & Moussalli, A. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike’s information criterion. *Behavioral Ecology and Sociobiology* 65:13–21.
- Tanadini, L.G. & Schmidt, B.R. 2011. Population size influences amphibian detection probability: implications for biodiversity monitoring programs. *PLoS ONE* 6(12):e28244.
- Trenham, P.C., Koenig, W.D., Mossman, M.J., Start, S.L., Jagger, L.A. 2003. Regional dynamics of wetland-breeding frogs and toads: turnover and synchrony. *Ecological Applications* 13(6):1522-1532.
- Urquhart N.S. & Kincaid, T.M. 1999. Designs for detecting trend from repeated surveys of ecological resources. *Journal of Agricultural, Biological, and Environmental Statistics* 4:404–414.
- Urquhart, N.S., Paulsen, S.G. & Larsen, D.P. 1998. Monitoring for policy-relevant regional trends over time. *Ecological Applications* 8(2):246-257.
- Weir, L.A., Royle, J.A., Nanjappa, P. & Jung, R.E. 2005. Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. *Journal of Herpetology* 39(4):627-639.

APPENDIX A1: Additional power analysis results

JELA Hyla mix

Table A1.1. (JELA Hyla mix) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	18 / 6	25 / 11
-0.06	Baseline	6	17 / 6	27 / 10
-0.06	50% more (old transects)	3	21 / 4	28 / 8
-0.06	50% more (old transects)	6	21 / 4	29 / 8
-0.06	50% more (new transects)	3	21 / 5	29 / 7
-0.06	50% more (new transects)	6	27 / 12	32 / 13
-0.04	Baseline	3	14 / 9	22 / 13
-0.04	Baseline	6	14 / 9	22 / 13
-0.04	50% more (old transects)	3	18 / 5	24 / 9
-0.04	50% more (old transects)	6	18 / 6	25 / 9
-0.04	50% more (new transects)	3	18 / 6	24 / 10
-0.04	50% more (new transects)	6	19 / 6	26 / 10
-0.02	Baseline	3	11 / 12	18 / 15
-0.02	Baseline	6	11 / 11	18 / 16
-0.02	50% more (old transects)	3	15 / 8	21 / 14
-0.02	50% more (old transects)	6	14 / 7	22 / 12
-0.02	50% more (new transects)	3	14 / 8	21 / 13
-0.02	50% more (new transects)	6	14 / 7	23 / 12
0.00	Baseline	3	9 / 13	14 / 18
0.00	Baseline	6	9 / 13	14 / 19
0.00	50% more (old transects)	3	11 / 10	19 / 16
0.00	50% more (old transects)	6	10 / 10	19 / 17
0.00	50% more (new transects)	3	11 / 9	20 / 16
0.00	50% more (new transects)	6	10 / 10	19 / 17

Table A1.2. (JELA Hyla mix) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	6 / 8	12 / 12
0.00	Baseline	6	7 / 8	13 / 13
0.00	50% more (old transects)	3	8 / 7	14 / 12
0.00	50% more (old transects)	6	8 / 7	14 / 11
0.00	50% more (new transects)	3	8 / 7	14 / 12
0.00	50% more (new transects)	6	8 / 7	14 / 12
0.02	Baseline	3	3 / 13	7 / 19
0.02	Baseline	6	3 / 13	7 / 20
0.02	50% more (old transects)	3	4 / 12	8 / 21
0.02	50% more (old transects)	6	4 / 13	7 / 20
0.02	50% more (new transects)	3	4 / 13	7 / 20
0.02	50% more (new transects)	6	3 / 13	8 / 21
0.04	Baseline	3	1 / 19	2 / 29
0.04	Baseline	6	1 / 19	2 / 30
0.04	50% more (old transects)	3	1 / 21	3 / 32

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.04	50% more (old transects)	6	1 / 21	3 / 30
0.04	50% more (new transects)	3	1 / 20	3 / 32
0.04	50% more (new transects)	6	1 / 21	3 / 31
0.06	Baseline	3	0 / 31	1 / 45
0.06	Baseline	6	0 / 31	1 / 45
0.06	50% more (old transects)	3	0 / 32	1 / 44
0.06	50% more (old transects)	6	0 / 32	1 / 44
0.06	50% more (new transects)	3	0 / 33	1 / 43
0.06	50% more (new transects)	6	0 / 32	1 / 44

Table A1.3. (JELA *Hyla mix*) Statistical power to detect *SIMULATED INCREASES AFTER 5 YEARS, with surveys during non-breeding season only*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	9 / 13	14 / 18
0.00	Baseline	6	9 / 13	14 / 19
0.00	50% more (old transects)	3	11 / 10	19 / 16
0.00	50% more (old transects)	6	10 / 10	19 / 17
0.00	50% more (new transects)	3	11 / 9	20 / 16
0.00	50% more (new transects)	6	10 / 10	19 / 17
0.02	Baseline	3	7 / 15	12 / 22
0.02	Baseline	6	8 / 15	12 / 22
0.02	50% more (old transects)	3	9 / 14	15 / 19
0.02	50% more (old transects)	6	8 / 13	15 / 19
0.02	50% more (new transects)	3	9 / 13	16 / 19
0.02	50% more (new transects)	6	10 / 15	15 / 22
0.04	Baseline	3	6 / 18	10 / 25
0.04	Baseline	6	6 / 18	10 / 25
0.04	50% more (old transects)	3	6 / 16	11 / 23
0.04	50% more (old transects)	6	7 / 16	12 / 22
0.04	50% more (new transects)	3	6 / 16	11 / 23
0.04	50% more (new transects)	6	7 / 16	11 / 23
0.06	Baseline	3	6 / 21	8 / 29
0.06	Baseline	6	5 / 22	8 / 29
0.06	50% more (old transects)	3	5 / 18	8 / 27
0.06	50% more (old transects)	6	5 / 19	8 / 26
0.06	50% more (new transects)	3	5 / 19	8 / 27
0.06	50% more (new transects)	6	6 / 20	8 / 25

JELA *Lithobates clamitans*

Table A1.4. (*JELA L. clamitans*) Statistical power to detect *SIMULATED DECLINES AFTER 5 YEARS, with surveys during non-breeding season only*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	18 / 4	26 / 9
-0.06	Baseline	6	18 / 5	28 / 8
-0.04	Baseline	3	15 / 6	24 / 11
-0.04	Baseline	6	15 / 8	24 / 11
-0.02	Baseline	3	12 / 9	19 / 13
-0.02	Baseline	6	12 / 10	19 / 14
0.00	Baseline	3	9 / 10	14 / 15
0.00	Baseline	6	9 / 10	16 / 17

Table A1.5. (*JELA L. clamitans*) Statistical power to detect *SIMULATED INCREASES AFTER 10 YEARS, with surveys during non-breeding season only*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	7 / 7	12 / 14
0.00	Baseline	6	7 / 7	12 / 14
0.02	Baseline	3	4 / 13	7 / 23
0.02	Baseline	6	3 / 13	7 / 21
0.04	Baseline	3	3 / 21	4 / 33
0.04	Baseline	6	2 / 22	3 / 33
0.06	Baseline	3	2 / 32	3 / 47
0.06	Baseline	6	2 / 34	2 / 46

Table A1.6. (*JELA L. clamitans*) Statistical power to detect *SIMULATED INCREASES AFTER 5 YEARS, with surveys during non-breeding season only*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	9 / 10	14 / 15
0.00	Baseline	6	9 / 10	16 / 17
0.02	Baseline	3	7 / 12	12 / 16
0.02	Baseline	6	7 / 13	13 / 21
0.04	Baseline	3	6 / 15	10 / 22
0.04	Baseline	6	5 / 15	10 / 24
0.06	Baseline	3	4 / 16	9 / 24
0.06	Baseline	6	4 / 20	7 / 28

JELA Eleutherodactylus planirostris

Table A1.7. (JELA E. planirostris) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	13 / 5	18 / 9
-0.06	Baseline	6	14 / 6	19 / 11
-0.06	50% more (old transects)	3	14 / 7	20 / 13
-0.06	50% more (old transects)	6	13 / 7	20 / 13
-0.06	50% more (new transects)	3	13 / 6	19 / 11
-0.06	50% more (new transects)	6	13 / 6	20 / 13
-0.04	Baseline	3	10 / 5	16 / 10
-0.04	Baseline	6	13 / 7	19 / 13
-0.04	50% more (old transects)	3	14 / 8	21 / 14
-0.04	50% more (old transects)	6	13 / 8	20 / 14
-0.04	50% more (new transects)	3	11 / 7	18 / 14
-0.04	50% more (new transects)	6	12 / 8	18 / 14
-0.02	Baseline	3	10 / 5	17 / 9
-0.02	Baseline	6	12 / 6	19 / 13
-0.02	50% more (old transects)	3	13 / 9	18 / 15
-0.02	50% more (old transects)	6	13 / 9	18 / 15
-0.02	50% more (new transects)	3	11 / 8	17 / 15
-0.02	50% more (new transects)	6	12 / 8	18 / 15
0.00	Baseline	3	10 / 7	16 / 11
0.00	Baseline	6	10 / 7	19 / 14
0.00	50% more (old transects)	3	12 / 10	17 / 16
0.00	50% more (old transects)	6	12 / 9	17 / 17
0.00	50% more (new transects)	3	10 / 8	16 / 16
0.00	50% more (new transects)	6	11 / 10	17 / 16

Table A1.8. (JELA E. planirostris) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	6 / 8	11 / 14
0.00	Baseline	6	7 / 8	13 / 12
0.00	50% more (old transects)	3	9 / 8	16 / 13
0.00	50% more (old transects)	6	8 / 9	15 / 14
0.00	50% more (new transects)	3	8 / 9	14 / 14
0.00	50% more (new transects)	6	9 / 8	14 / 13
0.02	Baseline	3	5 / 10	10 / 15
0.02	Baseline	6	6 / 10	11 / 15
0.02	50% more (old transects)	3	8 / 10	14 / 16
0.02	50% more (old transects)	6	7 / 10	14 / 16
0.02	50% more (new transects)	3	7 / 9	11 / 15
0.02	50% more (new transects)	6	7 / 10	12 / 15
0.04	Baseline	3	4 / 12	9 / 18
0.04	Baseline	6	5 / 12	8 / 17
0.04	50% more (old transects)	3	6 / 12	11 / 18
0.04	50% more (old transects)	6	5 / 12	10 / 18
0.04	50% more (new transects)	3	7 / 11	10 / 17
0.04	50% more (new transects)	6	5 / 12	10 / 17

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.06	Baseline	3	3 / 14	7 / 20
0.06	Baseline	6	4 / 13	7 / 22
0.06	50% more (old transects)	3	6 / 13	9 / 20
0.06	50% more (old transects)	6	4 / 14	9 / 19
0.06	50% more (new transects)	3	5 / 12	8 / 19
0.06	50% more (new transects)	6	5 / 14	8 / 19

Table A1.9. (*JELA E. planirostris*) Statistical power to detect *SIMULATED INCREASES AFTER 5 YEARS*, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	10 / 7	16 / 11
0.00	Baseline	6	10 / 7	19 / 14
0.00	50% more (old transects)	3	12 / 10	17 / 16
0.00	50% more (old transects)	6	12 / 9	17 / 17
0.00	50% more (new transects)	3	10 / 8	16 / 16
0.00	50% more (new transects)	6	11 / 10	17 / 16
0.02	Baseline	3	10 / 8	15 / 13
0.02	Baseline	6	10 / 8	17 / 16
0.02	50% more (old transects)	3	11 / 11	16 / 17
0.02	50% more (old transects)	6	11 / 10	16 / 18
0.02	50% more (new transects)	3	10 / 9	16 / 17
0.02	50% more (new transects)	6	11 / 10	16 / 18
0.04	Baseline	3	9 / 9	13 / 15
0.04	Baseline	6	9 / 10	16 / 17
0.04	50% more (old transects)	3	10 / 11	15 / 18
0.04	50% more (old transects)	6	11 / 11	14 / 19
0.04	50% more (new transects)	3	9 / 10	14 / 17
0.04	50% more (new transects)	6	10 / 11	15 / 20
0.06	Baseline	3	7 / 9	13 / 16
0.06	Baseline	6	9 / 11	14 / 17
0.06	50% more (old transects)	3	10 / 12	14 / 19
0.06	50% more (old transects)	6	11 / 12	15 / 20
0.06	50% more (new transects)	3	10 / 12	14 / 19
0.06	50% more (new transects)	6	10 / 12	15 / 21

GUIS Eleutherodactylus planirostris

Table A1.10. (*GUIS E. planirostris*) *Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys during non-breeding season only.* The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	18 / 3	26 / 6
-0.06	Baseline	6	18 / 3	27 / 5
-0.04	Baseline	3	15 / 5	23 / 7
-0.04	Baseline	6	14 / 3	24 / 7
-0.02	Baseline	3	13 / 5	20 / 10
-0.02	Baseline	6	13 / 5	22 / 10
0.00	Baseline	3	12 / 6	16 / 14
0.00	Baseline	6	10 / 6	18 / 12

Table A1.11. (*GUIS E. planirostris*) *Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys during non-breeding season only.* The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	9 / 9	13 / 13
0.00	Baseline	6	7 / 8	12 / 12
0.02	Baseline	3	5 / 12	8 / 18
0.02	Baseline	6	4 / 12	8 / 19
0.04	Baseline	3	3 / 19	6 / 30
0.04	Baseline	6	2 / 20	4 / 31
0.06	Baseline	3	2 / 28	3 / 42
0.06	Baseline	6	1 / 31	3 / 43

Table A1.12. (*GUIS E. planirostris*) *Statistical power to detect SIMULATED INCREASES AFTER 5 YEARS, with surveys during non-breeding season only.* The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	12 / 6	16 / 14
0.00	Baseline	6	10 / 6	18 / 12
0.02	Baseline	3	9 / 9	14 / 17
0.02	Baseline	6	8 / 8	14 / 15
0.04	Baseline	3	7 / 12	13 / 21
0.04	Baseline	6	7 / 10	12 / 18
0.06	Baseline	3	6 / 15	10 / 26
0.06	Baseline	6	5 / 13	10 / 23

GUIS Hyla mix

Table A1.13. (GUIS Hyla mix) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	3	12 / 8	19 / 13
-0.06	Baseline	6	12 / 9	19 / 12
-0.04	Baseline	3	12 / 9	18 / 14
-0.04	Baseline	6	11 / 9	17 / 13
-0.02	Baseline	3	10 / 9	16 / 14
-0.02	Baseline	6	11 / 9	17 / 15
0.00	Baseline	3	9 / 10	15 / 16
0.00	Baseline	6	10 / 10	15 / 16

Table A1.14. (GUIS Hyla mix) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	8 / 10	14 / 16
0.00	Baseline	6	8 / 8	14 / 15
0.02	Baseline	3	6 / 11	11 / 17
0.02	Baseline	6	6 / 11	12 / 18
0.04	Baseline	3	5 / 14	9 / 21
0.04	Baseline	6	5 / 13	9 / 20
0.06	Baseline	3	5 / 17	7 / 25
0.06	Baseline	6	5 / 17	7 / 25

Table A1.15. (GUIS Hyla mix) Statistical power to detect SIMULATED INCREASES AFTER 5 YEARS, with surveys during non-breeding season only. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	3	9 / 10	15 / 16
0.00	Baseline	6	10 / 10	15 / 16
0.02	Baseline	3	9 / 11	15 / 18
0.02	Baseline	6	9 / 11	14 / 18
0.04	Baseline	3	8 / 12	13 / 19
0.04	Baseline	6	8 / 12	13 / 20
0.06	Baseline	3	7 / 14	12 / 21
0.06	Baseline	6	8 / 13	12 / 22

SAAN *Eleutherodactylus cystignathoides*

Table A1.16. (SAAN *E. cystignathoides*) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	6	18 / 4	27 / 9
-0.06	Baseline	12	19 / 4	27 / 8
-0.04	Baseline	6	15 / 6	24 / 10
-0.04	Baseline	12	15 / 6	25 / 9
-0.02	Baseline	6	14 / 7	21 / 11
-0.02	Baseline	12	14 / 7	22 / 10
0.00	Baseline	6	13 / 8	18 / 13
0.00	Baseline	12	13 / 8	18 / 13

Table A1.17. (SAAN *E. cystignathoides*) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	6	9 / 6	14 / 12
0.00	Baseline	12	9 / 7	14 / 11
0.02	Baseline	6	5 / 10	11 / 17
0.02	Baseline	12	6 / 10	10 / 16
0.04	Baseline	6	4 / 13	8 / 21
0.04	Baseline	12	4 / 13	8 / 22
0.06	Baseline	6	3 / 17	5 / 30
0.06	Baseline	12	3 / 18	5 / 29

Table A1.18. (SAAN *E. cystignathoides*) Statistical power to detect SIMULATED INCREASES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	6	13 / 8	18 / 13
0.00	Baseline	12	13 / 8	18 / 13
0.02	Baseline	6	11 / 9	16 / 16
0.02	Baseline	12	10 / 9	16 / 15
0.04	Baseline	6	8 / 11	14 / 18
0.04	Baseline	12	8 / 10	15 / 18
0.06	Baseline	6	7 / 13	13 / 21
0.06	Baseline	12	7 / 11	12 / 20

SAAN *Incilius nebulifer*

Table A1.19. (SAAN I. nebulifer) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	6	18 / 5	26 / 8
-0.06	Baseline	12	18 / 5	27 / 8
-0.04	Baseline	6	16 / 6	24 / 9
-0.04	Baseline	12	15 / 5	25 / 9
-0.02	Baseline	6	14 / 6	20 / 12
-0.02	Baseline	12	13 / 6	21 / 10
0.00	Baseline	6	11 / 7	18 / 12
0.00	Baseline	12	12 / 8	19 / 12

Table A1.20. (SAAN I. nebulifer) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	6	9 / 7	14 / 12
0.00	Baseline	12	9 / 7	13 / 12
0.02	Baseline	6	6 / 9	10 / 16
0.02	Baseline	12	6 / 9	11 / 16
0.04	Baseline	6	4 / 14	8 / 21
0.04	Baseline	12	4 / 14	8 / 21
0.06	Baseline	6	2 / 18	5 / 26
0.06	Baseline	12	3 / 17	5 / 27

Table A1.21. (SAAN I. nebulifer) Statistical power to detect SIMULATED INCREASES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	6	11 / 7	18 / 12
0.00	Baseline	12	12 / 8	19 / 12
0.02	Baseline	6	10 / 8	17 / 15
0.02	Baseline	12	10 / 8	17 / 14
0.04	Baseline	6	9 / 10	14 / 16
0.04	Baseline	12	9 / 10	14 / 17
0.06	Baseline	6	8 / 11	12 / 18
0.06	Baseline	12	7 / 11	12 / 18

VICK *Eurycea guttolineata*

Table A1.22. (*VICK E. guttolineata*) Statistical power to detect SIMULATED DECLINES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	12	11 / 2	21 / 4
-0.04	Baseline	12	8 / 3	16 / 6
-0.02	Baseline	12	7 / 3	12 / 9
0.00	Baseline	12	4 / 5	10 / 11

Table A1.23. (*VICK E. guttolineata*) Statistical power to detect SIMULATED INCREASES AFTER 10 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	12	4 / 4	9 / 9
0.02	Baseline	12	1 / 11	4 / 19
0.04	Baseline	12	0 / 21	1 / 34
0.06	Baseline	12	0 / 40	0 / 54

Table A1.24. (*VICK E. guttolineata*) Statistical power to detect SIMULATED INCREASES AFTER 5 YEARS, with surveys year-round. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	12	4 / 5	10 / 11
0.02	Baseline	12	2 / 7	7 / 17
0.04	Baseline	12	1 / 11	5 / 21
0.06	Baseline	12	1 / 15	2 / 25

VICK *Scincella lateralis*

Table A1.25. (*VICK S. lateralis*) Statistical power to detect *SIMULATED DECLINES AFTER 5 YEARS, with surveys year-round*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
-0.06	Baseline	12	16 / 1	30 / 2
-0.04	Baseline	12	12 / 2	24 / 4
-0.02	Baseline	12	9 / 2	19 / 6
0.00	Baseline	12	6 / 4	13 / 9

Table A1.26. (*VICK S. lateralis*) Statistical power to detect *SIMULATED INCREASES AFTER 10 YEARS, with surveys year-round*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	12	4 / 4	10 / 9
0.02	Baseline	12	2 / 11	3 / 23
0.04	Baseline	12	0 / 29	1 / 43
0.06	Baseline	12	0 / 49	0 / 62

Table A1.27. (*VICK S. lateralis*) Statistical power to detect *SIMULATED INCREASES AFTER 5 YEARS, with surveys year-round*. The last two columns show the proportion of simulations that detected a significant decline, followed by the proportion that detected a significant increase (the two numbers are separated by a slash '/') at two levels of alpha (0.10 and 0.20).

r	Number of traps	Surveys per year	% significant (alpha=0.10)	% significant (alpha=0.20)
0.00	Baseline	12	6 / 4	13 / 9
0.02	Baseline	12	3 / 7	9 / 12
0.04	Baseline	12	2 / 10	5 / 17
0.06	Baseline	12	1 / 15	3 / 23