Contributed Paper

Determining Optimal Population Monitoring for Rare Butterflies

NICK M. HADDAD,*†† BRIAN HUDGENS,† CHRIS DAMIANI,‡ KEVIN GROSS,§ DANIEL KUEFLER,* AND KEN POLLOCK*

*Department of Zoology, North Carolina State University, P.O. Box 7617, Raleigh, NC 27695-7617, U.S.A.
†Institute for Wildlife Studies, P.O. Box 1104, Arcata, CA 95518, U.S.A.
‡U.S.D.A. Forest Service, Redwood Sciences Laboratory, 1700 Bayview Drive, Arcata, CA 95521, U.S.A.
§Biomath Program, North Carolina State University, P.O. Box 8203, Raleigh, NC 27695-8203, U.S.A.

Abstract: *Determining population viability of rare insects depends on precise, unbiased estimates of population size and other demographic parameters. We used data on the endangered St. Francis' satyr butterfly (*Neonympha mitchellii francisci*) to evaluate 2 approaches (mark–recapture and transect counts) for population analysis of rare butterflies. Mark–recapture analysis provided by far the greatest amount of demographic information, including estimates (and standard errors) of population size, detection, survival, and recruitment probabilities. Mark–recapture analysis can also be used to estimate dispersal and temporal variation in rates, although we did not do this here. Models of seasonal flight phenologies derived from transect counts (Insect Count Analyzer) provided an index of population size and estimates of survival and statistical uncertainty. Pollard–Yates population indices derived from transect counts did not provide estimates of demographic parameters. This index may be highly biased if detection and survival probabilities vary spatially and temporally. In terms of statistical performance, mark–recapture and Pollard–Yates indices were least variable. Mark–recapture estimates were less likely to fail than Insect Count Analyzer, but mark–recapture estimates became less precise as sampling intensity decreased. In general, count-based approaches are less costly and less likely to cause harm to rare insects than mark–recapture. The optimal monitoring approach must reconcile these trade-offs. Thus, mark–recapture should be favored when demographic estimates are needed, when financial resources enable frequent sampling, and when marking does not harm the insect populations. The optimal sampling strategy may use 2 sampling methods together in 1 overall sampling plan: limited mark–recapture sampling to estimate survival and detection probabilities and frequent but less expensive transect counts.*

Keywords: butterflies, distance sampling, endangered butterflies, insects, mark–recapture, rare species, species monitoring, threatened species, transect counts

Determinación del Monitoreo Poblacional Óptimo para Mariposas Raras

Resumen: *La determinación de la viabilidad poblacional de insectos raros depende de estimaciones precisas, sin sesgos, del tamaño poblacional y otros parámetros demográficos. Utilizamos datos de la mariposa en peligro (*Neonympha mitchellii francisci*) para evaluar 2 métodos (marca-recaptura y conteos en transectos) para analizar las poblaciones de mariposas raras. El análisis de marca-recaptura proporcionó, por mucho, la mayor cantidad de información demográfica, incluyendo estimaciones (y error estándar) del tamaño de la población y las probabilidades de detección, supervivencia y reclutamiento. El análisis de marca-recaptura también puede ser utilizado para estimar la dispersión y variaciones temporales en las tasas, aunque no lo hicimos en este artículo. Los modelos de fenologías de vuelo estacional derivados de los conteos en transectos (Analizador de Conteos de Insectos [ACIN]) proporcionaron un índice del tamaño poblacional y estimaciones de la supervivencia y de la incertidumbre estadística. Los índices de la población de*

Pollard–Yates derivados de los conteos en transectos no proporcionaron estimaciones de los parámetros de-
mográficos. Este índice puede estar muy sesgado si las probabilidades de detección y supervivencia varían
espacial y temporalmente. En términos del funcionamiento estadístico, los índices de marca-recaptura y de
Pollard–Yates fueron menos variables. Las estimaciones de marca-recaptura tuvieron menos probabilidad de
fallar que ACIN, pero las estimaciones de marca-recaptura se volvieron menos precisas a medida que decreció
la intensidad de muestreo. En general, los métodos basados en conteos son menos costosos y causan menos
daño a insectos raros que la marca-recaptura. El método óptimo de monitoreo debe reconciliar lo anterior.
Por lo tanto, la marca-recaptura debe ser favorecida cuando se requieren estimaciones demográficas, cuando
los recursos financieros permitan muestreos frecuentes y cuando el marcaje no dañe a las poblaciones de
insectos. La estrategia óptima de muestreo puede utilizar dos métodos de muestreo en un mismo plan de
muestreo: marca-recaptura limitada para estimar las probabilidades de supervivencia y detección y conteos
en transectos frecuentes pero menos costosos.

Palabras Clave: conteos en transectos, especies amenazadas, especies raras, marca-recaptura, mariposas, mari-
posas en peligro, monitoreo de especies, muestreo de distancia

Introduction

Recovery of rare butterfly populations depends on accurate assessments of population trends (Schultz & Hammond 2003), which often rely on sound estimates of population demographics (Morris & Doak 2002). Demographics of rare butterfly populations are usually monitored by 1 of 2 approaches, mark–recapture or transect counts (New 1991). Determining which approach to use involves weighing complex trade-offs among bias, precision, and cost. We used real and simulated data to compare mark–recapture and count-based methods and to evaluate the efficiency and effectiveness of the 2 methods for monitoring adult butterflies.

Mark–recapture is the most common method used to estimate population size in rare butterflies (e.g., Gall 1984; Bergman 2001; Baguette & Schtickzelle 2003). This method produces estimates of daily and total population sizes, recruitment, survival, and detection probabilities. Mark–recapture is the most rigorous approach to population estimation because it incorporates the greatest amount of information into well-developed statistical methods (Williams et al. 2002). Nevertheless, routine use of mark–recapture for butterflies may have several problems. First, mark–recapture can be much more expensive than transect counts because it requires frequent sampling and additional time to mark individuals. Second, estimated demographic parameters could be imprecise if only a few animals can be marked. Third, marking may damage or kill butterflies (Murphy 1987). Fourth, marking may change butterfly behaviors and thus their probability of capture (Singer & Wedlake 1981).

The transect counts most widely applied to monitor rare butterfly populations are Pollard–Yates counts (Pollard & Yates 1993; Seidl 1999; Mattoni et al. 2001; Schultz & Hammond 2003). Pollard–Yates population indices depend on the assumptions that a snapshot count is directly proportional to the population size and that this propor-

tion is constant over space and time. Under these assumptions, indices provide more precise estimates of population change than mark–recapture (Williams et al. 2002). In some cases Pollard–Yates count indices correlate with population size (e.g., Thomas 1983; Collier et al. 2008), but in others they do not (Harker & Shreeve 2007). Because insects may emerge and die between counts, comparisons of indices between sites or over time can require the (sometimes unrealistic) assumption of constant survival. Most quantitative biologists have concluded that it is essential to estimate detection probability and survival rates in monitoring populations (Pollock et al. 2002; Williams et al. 2002).

To improve the statistical performance of transect counts, 2 other approaches to analysis have been developed. One approach that we evaluated fits analytical models to count phenologies for a generation (Manly 1974). Zonneveld (1991) used counts to estimate mortality rates and total insect population sizes over one generation by assuming constant adult mortality and logistically distributed emergence times. This modeling approach has been incorporated into user-friendly, open-access software (http://www.urbanwildlands.org/INCA/), the Insect Count Analyzer (INCA), which requires a time series of counts as input. Because of its simplicity, accessibility, and supposed ability to estimate demographic parameters, INCA could become an important standard for insect population monitoring.

We did not directly evaluate a third approach to transect counts, distance sampling, which overcomes a key limitation of other count-based methods and estimates detection probabilities (Buckland et al. 2001). Although used infrequently for monitoring rare butterflies, we believe its use is increasing. In our study distance sampling was not possible because we could not assume that transect locations were random or that the area in which butterfly densities were uniform was known. In the discussion, we evaluate when distance sampling is appropriate to monitor rare butterflies.

Using the federally endangered St. Francis' satyr butterfly (*Neonympha mitchellii francisci*) as a case study, we determined optimal sampling strategies for adult butterflies. Our primary objective was to compare population estimates and indices among mark–recapture analysis and the count-based methods of INCA and Pollard–Yates. We used several criteria to evaluate these estimates and indices, including the population and demographic information they provide, their statistical performance, and their cost. We based our evaluation on empirical data collected over 4 years and on simulated data generated from hypothetical populations that were demographically similar to St. Francis' satyr. A secondary objective was to determine how estimates and indices depend on sampling intensity. In our experience with St. Francis' satyr, we have found that the issue of sampling intensity is key in designing a monitoring protocol because the performance of different methods depends on the rate of sampling. Yet the literature provides little guidance on this issue.

Methods

Study Species and Site

St. Francis' satyr is a federally endangered butterfly species that is restricted to Ft. Bragg, North Carolina (Murdock 1996; Kuefler et al. 2008). It is a subspecies of the endangered Mitchell's satyr (*N. mitchellii mitchellii*), which occurs in isolated populations in Michigan, Indiana, Virginia, Alabama, and Mississippi. St. Francis' satyr is bivoltine and lives in grassy wetlands along streams separated into distinct patches. Some populations are contained within restricted areas where bombs are detonated and where we could not sample. Outside those areas, we studied 4 subpopulations in detail that covered <5 ha. This is the area we refer to in the rest of the paper.

Wetlands occupied by St. Francis' satyr are created by beavers (*Castor canadensis*) that build and then abandon their dams. In the early stages of succession following abandonment, sedges (*Carex*) grow that are thought to be the butterfly's host plant (Kuefler et al. 2008). The decline in St. Francis' satyr populations may have followed beaver extinction in North Carolina. The last beaver was recorded in the state in 1897 (D. Woodward, personal communication). Beavers were reintroduced near Ft. Bragg in 1939. St. Francis' satyr may have persisted at Ft. Bragg because wetlands were relatively protected within the military installation, beavers were introduced, and bombing created annual fires that may have retarded succession in grassy wetlands (Murdock 1996).

Population Data

From 2002 through 2005 we conducted surveys along fixed transects where the 4 subpopulations occurred.

A system of transects was established in each site with 3.8 × 14.0 cm planks. An attempt was made to position planks across high points in wetlands (dry soil, exposed roots, or rocks) so that researchers would be above the vegetation and would not trample sensitive wetland vegetation. Transects were located to maximize the view of and minimize the damage to suitable wetland habitats. Within a subpopulation, there were multiple transect sections that penetrated into wetlands from the upland forest edge. Transect lengths were standardized to wetland area at approximately 300 m/ha. We attempted to survey transects daily but did not sample in rain or on some weekend days at the beginning and end of flight periods when abundances were low. We sampled between 11:00 and 18:00, when butterfly activity was consistently high (Kuefler et al. 2008).

We conducted both mark–recapture surveys and transect counts during the second flight periods of 2002–2005. Surveys began before the first butterflies emerged and were terminated when no butterflies were seen for 3 consecutive days. During a survey 2 observers walked neighboring transects within a site at a rate of 200 min/ha, with a minimum survey time of 20 min/observer. While walking transects, we attempted to flush butterflies from vegetation within 2 m of transects by swinging a net over and gently brushing plants. Transects were short and when we reached the end before survey time had expired we walked them again, recording only new individuals that flew into view. This approach deviates from Pollard–Yates and other counting methods (Pollard & Yates 1993), but was necessary because of the small number and inactivity of these butterflies. We also deviated from Pollard–Yates counts by attempting not to double count butterflies. When a butterfly was observed, we recorded data on sex, age, and wing condition. After a survey, the 2 observers tallied the total number of butterflies observed, including a minimum count of butterflies that observers saw and recorded at the same time, and a best count, for which observers used information on butterfly morphology (sex, wing condition) and spatial location to determine the number of unique individuals. The minimum and best counts usually differed by zero or one, and we used the best count for all analyses.

Each day we marked all butterflies we captured that had not already been marked. We used a fine-tipped permanent marker to write a unique alphanumeric code on the ventral sides of each hind wing and immediately released marked individuals at the site of capture (Kuefler et al. 2008). To minimize physical captures and maximize resightings from fixed transects, we used binoculars to visually resight butterflies whenever possible. Butterflies that were observed and designated as unmarked but not captured were added to daily population counts. To account for time spent capturing and handling butterflies, we added 3 min to each survey for every butterfly captured. In the second flight period of 2003 and

2004, we conducted 2 surveys each day on 62 occasions, one in which we used transect counts and another in which we used mark–recapture. Counts obtained via the 2 approaches were highly correlated with the slope of a regression near 1 (i.e., equal counts produced by each method, Haddad et al. 2005). Thus, for analyses of count data, we supplemented counts obtained from observations with counts obtained during mark–recapture surveys when necessary to generate time series of nearly daily counts for all analyses.

Analysis of Population Data

We generated 3 types of population estimates for each of the 4 subpopulations in each year. We assembled daily capture data from transect surveys into capture histories and estimated population size on the basis of the assumption that populations were open to births, deaths, and migration. The particular model of most use for butterfly populations is the Schwarz-Arnason version of the Jolly Seber model (Crosbie & Manly 1985; Schwarz & Arnason 1996). We used the POPAN procedure in Program MARK to determine survival probabilities, detection probabilities, and an estimate of population size for the whole flight period (i.e., the superpopulation rather than the daily population size) (White & Burnham 1999). (Program MARK software is available for no cost at http://www.warnercnr.colostate.edu/gwhite/mark/mark.htm.) We adjusted mark–recapture population estimates generated by the POPAN module in MARK to account for individuals that emerged but were never captured before they died. This can be accomplished by multiplying the superpopulation estimates by $[\log(\Phi)/(\Phi\text{-}1)]^{t}$, where Φ is the estimated survival probability and t is the interval between capture periods (Crosbie & Manly 1985).

We generated 2 indices from count data. First, we input a time series of counts for each subpopulation and year into INCA to generate an index of population size and an estimate of daily mortality rate. Second, we used modified Pollard–Yates counts to generate a population index. To do this, we averaged all counts per week and then summed average counts for the entire flight season (Pollard & Yates 1993). For mark–recapture population estimates and INCA population indices, coefficient of variation was computed as standard error of the mean. In some cases, data sets did not contain enough information to estimate all parameters in either mark–recapture or INCA. We called this model failure.

Empirical Analysis of Sampling Intensity

To determine the effect of sampling intensity on precision of butterfly population estimates, we started with the complete data set with nearly daily counts for each subpopulation and year. We then removed days from the full data set to create new data sets to represent sampling at varying intensities. We tested 4 realistic levels of sampling intensity: daily (the complete data set), every other day (semidaily), weekday (Monday–Friday), or every other day plus 5 continuous days surrounding the peak butterfly emergence dates determined by INCA (a hybrid of weekday and every other day). Sampling schemes were maintained regardless of days missed for bad weather.

Analysis of Simulated Data

We evaluated the precision and bias of estimates in simulated data sets. With the empirical data, we were unable to know true parameter values. Simulated data were generated with parameter values that are realistic for St. Francis' satyr (Kuefler et al. 2008), providing a benchmark against which to compare estimates derived from different analysis techniques. We used simulations with known, constant detection and survival probabilities to ask whether there are biases inherent in the model structure even with perfect information about the population. We did this while recognizing that we removed the bias caused by missing information that is typical of empirical studies with indices, in which detection and survival probabilities generally vary over space and time.

We simulated 1000 data sets that assumed daily sampling from a population whose true size was 100. To generate a single data set, we first simulated the dates of emergence and death for each individual butterfly. Emergence day was drawn from a logistic distribution described by peak day of emergence ($\mu = 10$ days) and spread of emergence ($\beta = 1.5$ days). Life span was drawn from an exponential distribution described by daily mortality rate ($\alpha = 0.35$). We then counted the number of butterflies flying on each day. To simulate realistic rates of butterfly detection probability, we sampled the daily population assuming that each adult alive on a given day was "captured" with a probability of 0.30.

Once we created the 1000 simulated data sets, we generated population estimates with mark–recapture analysis and population indices with INCA and Pollard–Yates calculations. We evaluated mean population sizes and mean-squared error, which provides a measure of variability and bias. To evaluate the effects of sampling intensity on performance, we then removed days to create new data sets as described earlier for empirical data, representing daily, weekday, every other day, and every other day plus peak-day sampling intensities.

Results

Comparisons among Techniques

For each site and year, we used the full data set to estimate mark–recapture population size, survival probabilities, detection probabilities, INCA population index and mortality rates, and Pollard–Yates population index (Table 1). The INCA could not generate a population index in 1 site

Table 1. Estimates of population sizes, indices, and demographic parameters generated by mark–recapture, INCA (Insect Count Analyzer), and Pollard–Yates methods for each St. Francis' satyr population sampled daily in every year.

| Year | Site[b] | Mark-recapture[a] | | | | | | INCA | | | | Pollard-Yates |
		population estimate	SE	survivorship	SE	detectability	SE	population index	SE	survivorship[c]	SE	population index
2002	A1	182.1	22.9	0.58	0.06	0.38	0.09	33.3	8.8	0.76	0.07	16.7
	D1	128.3	10.6	0.56	0.06	0.59	0.09	24.6	8.0	0.82	0.07	18.7
	D3	412.2	223.9	0.64	0.16	0.08	0.07	NA	NA	NA	NA	10.9
	E2	82.5	25.4	0.59	0.14	0.29	0.16	4.3	1.5	0.93	0.04	7.7
2003	A1	366.8	35.2	0.61	0.05	0.35	0.07	122.6	85.7	0.57	0.39	30.4
	D1	120.8	13.5	0.54	0.07	0.51	0.11	103.7	190.0	0.45	1.44	18.6
	D3	372.1	56.9	0.62	0.07	0.27	0.08	16.1	4.4	0.90	0.03	20.9
	E2	105.3	12.5	0.75	0.05	0.35	0.08	19.8	6.2	0.80	0.07	12.3
2004	A1	343.9	54.2	0.43	0.08	0.48	0.15	96.2	77.9	0.60	0.41	30.1
	D1	414.3	31.5	0.70	0.03	0.34	0.05	64.4	16.8	0.81	0.05	43.4
	D3	919.0	101.0	0.64	0.04	0.26	0.05	174.9	73.0	0.66	0.18	66.6
	E2	255.6	61.0	0.49	0.11	0.29	0.13	16.2	3.4	0.88	0.03	18.1
2005	A1	151.3	12.5	0.78	0.04	0.40	0.06	35.6	7.7	0.81	0.04	24.5
	D1	294.8	20.6	0.74	0.03	0.36	0.05	73.4	17.8	0.76	0.06	41.7
	D3	444.0	46.3	0.54	0.06	0.42	0.09	37.7	5.6	0.88	0.02	42.4
	E2	60.3	8.7	0.64	0.11	0.54	0.16	18.2	7.7	0.69	0.15	7.1

[a] For mark-recapture analysis, the model with constant detectability and survivorship was supported.
[b] We did not standardize indices to area or transect length. Transect lengths at each site were A1, 196 m; D1, 249 m in 2002-2003 and 324 m in 2004-2005; D3, 215 m; E2, 140 m.
[b] INCA estimates mortality rates. For comparison with mark-recapture, these rates (but not SE) were converted to survival probabilities ($e^{-\alpha}$).

in 1 year. Using the full data set with daily sampling in orthogonal regression, modified Pollard–Yates indices ($r = 0.84$) were more closely correlated with mark-recapture population estimates than were INCA population indices ($r = 0.71$; Fig. 1; Table 1). To ask how the relationship among techniques might improve with additional knowledge gained from mark-recapture studies, we divided Pollard–Yates and INCA indices at each site by corresponding estimates of detectability. Doing this increased the relationship of both indices to mark-recapture estimates (Pollard–Yates $r = 0.95$; INCA $r = 0.84$). Survival probabilities estimated from mark-recapture were consistently lower than those estimated from INCA (Table 1), and mark-recapture and INCA estimates were uncorrelated.

To compare means, variability, and bias among approaches, we used model parameters to recalibrate our estimates to have an expected value of 100. On the basis of model parameters, indices were expected to be lower than the true population size of 100 (INCA = 30 and Pollard–Yates = 12.24) and could be recalibrated for INCA by dividing the index by detection probability (0.3) and for Pollard Yates by dividing the index by (detection probability/α)/7. In real populations, detection and survival probabilities are typically unknown, causing additional bias in indices. Our goal, however, was to evaluate bias even under the most optimistic scenario. Simulations revealed 3 differences among analyses (Fig. 2). Mark-recapture estimates and Pollard-Yates indices were unbiased, whereas INCA tended to overestimate (mean = +13%) butterfly population sizes. The median

for INCA was much closer to the true population size than the mean (+3%). Second, variation in INCA indices was much larger than variation in mark-recapture estimates or Pollard-Yates indices. The interval covering 95% of INCA indices was about 3 times the range covering mark-recapture estimates or Pollard-Yates indices. Third, INCA failed in over 15% of simulations, whereas mark-recapture and Pollard-Yates analyses did not fail.

Effects of Sampling Intensity

As sampling intensity decreased, both mark-recapture and INCA became less reliable. With fewer sampling occasions, variation in mark-recapture estimates increased and the likelihood of model failure increased (Fig. 3a-c). Although variation was higher in INCA, there was no relationship between sampling intensity and population estimates or variation in estimates (Fig. 3d-f). Nevertheless, as in mark-recapture analyses, model failure rates increased with decreased sampling intensity. Pollard–Yates indices of population size did not change with sampling intensity (Fig. 3g) and, unlike mark-recapture and INCA estimates, they did not produce estimates of sampling variation.

After adjusting indices for known values of detection and survival probabilities, results from analyses of simulated data with known parameters revealed similar trends with sampling intensity as in empirical analyses (Figs. 4 & 5). Mean-squared error was low regardless of sampling intensity for Pollard-Yates indices, high regardless of sampling intensity for INCA indices, and rose as sampling
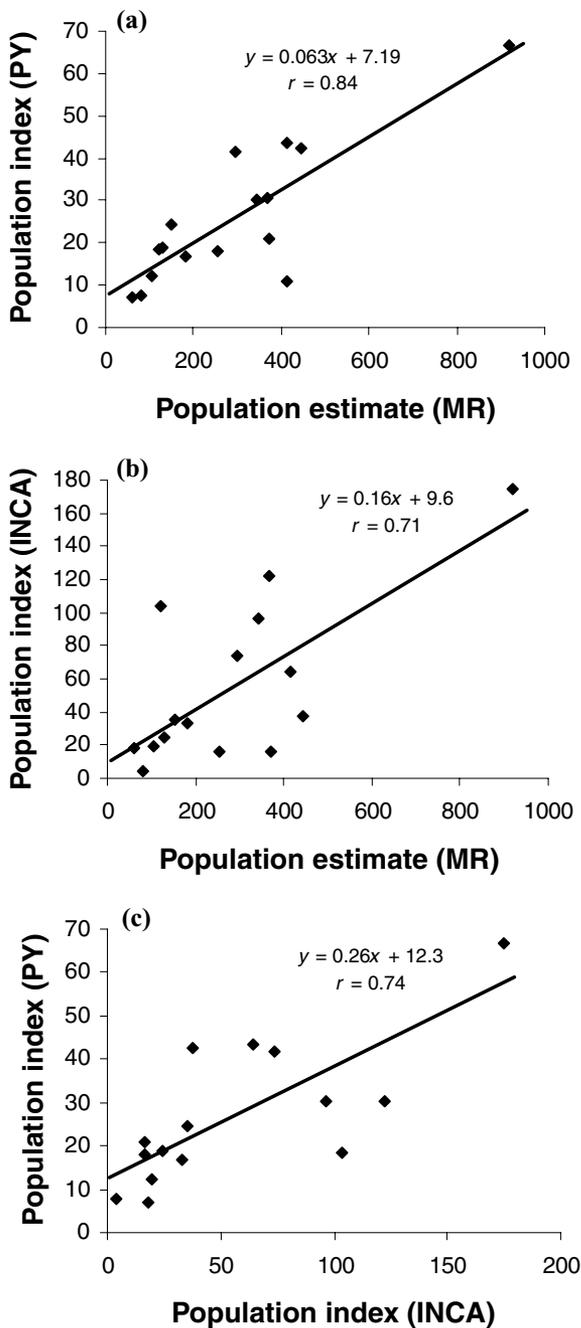
Figure 1. Pairwise relationships between population estimates (mark-recapture, MR) and population indices ([a] Pollard-Yates, PY and [b] Insect Count Analyzer, INCA) and between (c) Pollard-Yates and INCA generated from empirical data. Lines show the best-fit orthogonal regression assuming equal variances, and r *is Pearson's correlation coefficient.*

Figure 2. The distribution of (a) population estimates from mark-recapture analysis, (b) population indices generated by Insect Count Analyzer (INCA), and (c) population indices computed with Pollard-Yates for 1000 simulated analyses of populations with 100 individuals. Filled triangles are the true value, open circles are means, open triangles are medians, and filled diamonds are endpoints that cover 95% of the estimates. To allow comparisons among approaches, we restandardized population indices to 100, removing bias caused by differences in expected values (INCA = 30 and Pollard-Yates = 12.24).

intensity fell for mark–recapture estimates (Fig. 4). For mark-recapture, mean population estimates were unbiased and variability increased as sampling intensity decreased (Fig. 5a). Only analyses of data sampled at the lowest intensity failed, and these failed in only a small proportion (approximately 0.03) of data sets (Fig. 5c). Survival probabilities were biased slightly low, and this bias increased slightly with reduced sampling intensity (Fig. 5b). On the other hand, estimates of detectability were biased high and increased with reduced sampling intensity. In analyses with INCA, the positive bias
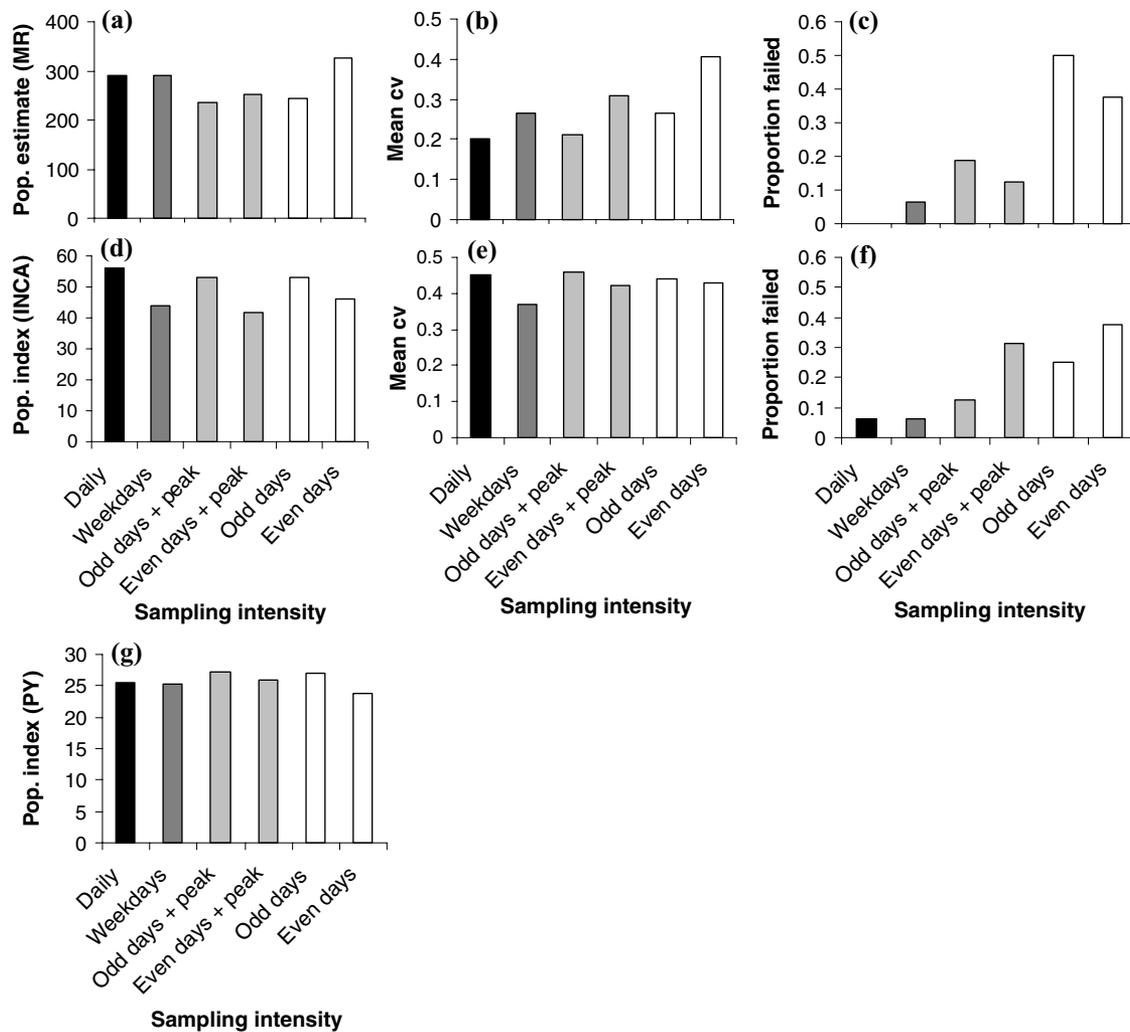
Figure 3. Empirical analysis of St. Francis' satyr populations at 4 sites from 2002 to 2005 to show the influence of sampling intensity: (a-c) mark-recapture (MR), (d-f) Insect Count Analyzer (INCA), and (g) Pollard-Yates (PY). Graphs (a), (d), and (g) show mean estimates for each analysis across all sites and years, (b) and (e) show mean coefficient of variation across all sites and years, and (c) and (f) show proportion of models that failed. Pollard-Yates does not produce estimates of statistical variation and does not fail.

in estimates of population size remained constant regardless of sampling intensity (Fig. 5d). Survival probabilities were biased slightly low, but this bias did not change with sampling intensity (Fig. 5e). The INCA model failed to return parameter estimates in a large fraction of cases, with a minimum of 15% failures with daily sampling, increasing to 24% with semidaily sampling (Fig. 5e). Reduced sampling intensity did not cause bias or model failure in Pollard Yates indices.

Discussion

We evaluated mark-recapture, INCA, and Pollard-Yates with 3 sets of criteria: the information generated, how the model performed, and the method's relative cost (Table

2). None of the approaches was uniformly superior across all criteria. Thus, the proper choice of sampling approach and analysis depends on the weight assigned to each criterion. In many cases, monitoring will involve a hybrid of approaches.

Information Generated

Mark-recapture provided the most demographic information, followed by INCA and then by Pollard-Yates (Table 2). Unlike count-based approaches, mark-recapture generates a population estimate rather than an index, estimates detection probability, and can estimate temporal variation in detection and survival probabilities within a flight period. Although this model was not supported with our St. Francis' satyr data (Kuefler et al. 2008),
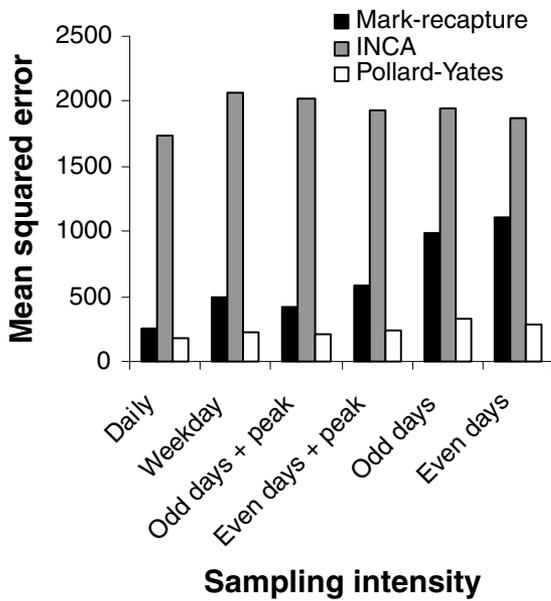
Figure 4. Mean-squared error of population estimates and indices at different sampling intensities. For each case, the expected population size was 100 or indices were adjusted so that their expected values equaled 100, allowing mean-squared error to be evaluated on equal terms across approaches.

variation in detection and survival probabilities across a flight period is likely for most butterflies. The ability of mark–recapture analysis to estimate daily variation increases when local population sizes are high (even when butterflies are globally rare) and when large effort is devoted to increase capture rates. Estimating demographic parameters and their variability is especially important when sites or populations are expected to differ or when individual sites are expected to change over time due to succession or other environmental change. Unlike count-based approaches, mark–recapture is unique because it allows individuals to be tracked during dispersal, which is particularly useful in analysis of the viability of spatially structured populations, such as St. Francis' satyr and many other butterfly populations in fragmented landscapes.

The INCA method shares with mark–recapture the ability to produce estimates of statistical uncertainty and demographic estimates of survivorship. Measures of statistical uncertainty provide context to evaluate the precision of population estimates that may dictate conservation action. Measures of survivorship are central to demographic models used in population viability analysis and are needed to determine total, as opposed to daily, population size. Mark–recapture estimates of survival were less variable than INCA estimates of mortality.
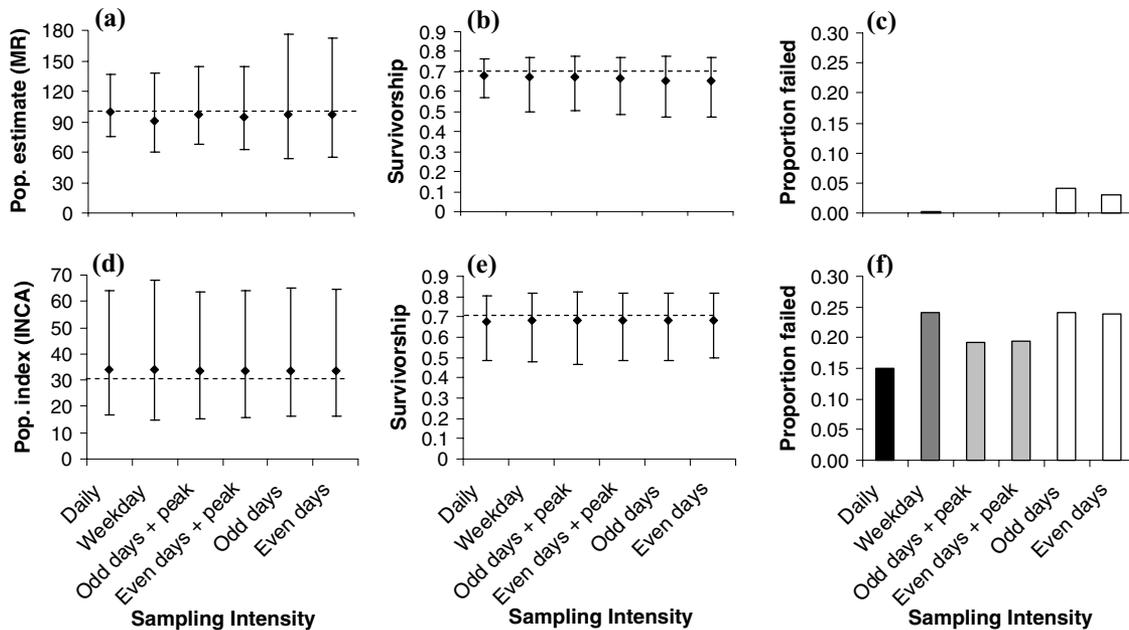

Figure 5. Simulation analysis of model performance with (a, b, c) mark–recapture (MR) and (d, e, f) Insect Count Analyzer (INCA). Mean and 95% coverage intervals for 1000 simulations of population mark–recapture (a) population estimates, (b) survival probabilities, (c) model failure rate and INCA, (d) population indices, (e) survival probabilities, and (f) model failure rates. Graphs (c) and (f) show the proportion of models that were unable to return estimates. The INCA estimates mortality rates, which were converted to survival probabilities ($e^{-\alpha}$) for comparison. No analyses are presented for Pollard–Yates because it did not result in bias, did not estimate survivorship, and did not fail.

Table 2. Comparisons among methods of population analysis.

	Approach			
Criteria	mark-recapture	INCA	Pollard-Yates	distance sampling[a]
Information				
population measure	estimate[b]	index	index	estimate[b]
daily or total population	both[b]	total	total	daily
measures statistical uncertainty	yes[b]	yes[b]	no	yes[b]
estimates survivorship	yes[b]	yes[b]	no	no
estimates detectability	yes[b]	no	no	yes[b]
estimates temporal variation in survival probability	yes[b]	no	no	no
estimates temporal variation in detection probability	yes[b]	no	no	yes[b]
allow measurements of dispersal	yes[b]	no	no	no
Performance				
bias after adjusting for detection and survival probabilities	none if model assumptions met[b]	positive (+13% daily)	none[b]	NA
variation	intermediate	highest	lowest[b]	NA
model failure	lower	higher	none[b]	NA
change with reduced sampling intensity	more variable and more failure	more failure in empirical study	none[b]	NA
Costs				
harm to animals	unknown, greater than counts	none[b]	none[b]	none[b]
relative cost	higher (>25%)	lower[b]	lower[b]	intermediate

[a]Although not analyzed in our case study with St. Francis' satyr, we included distance sampling here to allow comparison with the other most common techniques (NA, not available).
[b]The approach that performed best for each criteria.

The strong assumptions INCA made about demographic process were unrealistic for St. Francis' satyr. The INCA should only be applied when the emergence and death processes conform to INCA's assumptions. With regard to the information they generate, mark-recapture clearly ranked highest, with INCA providing more information than Pollard-Yates when populations conform to INCA's assumptions.

Statistical Performance

Mark-recapture provided the best statistical performance (Table 2). Especially at high sampling frequencies, mark-recapture produced results with low levels of variability and bias. Low bias for mark-recapture was accomplished through an adjustment to estimates obtained from POPAN that accounted for individuals that emerge and die before being captured. A relevant issue in this adjustment is the low demographic contributions of short-lived individuals. We considered them in our analysis to standardize comparisons across approaches, but in practical conservation planning whether or not to consider them will vary among species and depend on their life span relative to sampling intensity. Compared with INCA, the other approach that yielded measures of statistical uncertainty and demographic parameters, mark-recapture performed better. In empirical analysis INCA indices were more variable and less correlated than Pollard-Yates indices to mark-recapture population size, and INCA mor-

tality rates were more variable than and uncorrelated to mark-recapture survival probabilities.

Given the small amount of information used to generate Pollard-Yates indices, they performed remarkably well. In analyses of empirical data, Pollard-Yates indices correlated well with mark-recapture estimates. Simulation results showed that these indices had low variability and, after adjusting for known detection and survival probabilities, low bias. Indices such as Pollard-Yates will only be useful when spatial or temporal variation in these probabilities is low or when those rates can be estimated from other studies. Because of the way Pollard-Yates indices are calculated, they never fail to produce an index of abundance. For small populations, there is some chance that either mark-recapture or INCA will fail, severely limiting their application to the study of rare butterflies.

The poor statistical performance of INCA may have been caused by small St. Francis' satyr population sizes or by the failure of St. Francis' satyr populations to meet model assumptions. St. Francis' satyr population sizes and detection probabilities are low even when compared with other rare butterflies. The INCA performs poorly for small populations because the underlying model cannot successfully estimate mortality rates when the right tail of the distributions of counts over time is too short. There was no association between observed numbers of butterflies and INCA failure in our empirical data. The INCA may also have performed poorly because of the stringent assumption of constant mortality rate through a flight period. Bayesian methods consistent with those

used in INCA improve model performance, but require some previous knowledge of survival and detection probabilities that would most likely come from limited mark–recapture studies (Gross et al. 2007). Regardless, the limitations we found with INCA are likely to apply to many rare butterfly species.

The relative statistical performance of our approaches varied with sampling intensities, shifting the balance over the range of intensities we examined toward count-based approaches as sampling intensity decreased. In analyses of both empirical and simulated data, variability in mark–recapture estimates increased as sampling intensity decreased. Variation in INCA and Pollard–Yates indices did not change as sampling intensity decreased from daily to every other day. In empirical data sets, mark–recapture and INCA had higher failure rates at moderate (every other day) sampling intensity. When sampling intensity was less than every other day (not reported here), mark–recapture analysis would almost surely fail. In simulated data sets, failure rates were highest in INCA, regardless of sampling intensity. For St. Francis' satyr, and for many other rare butterflies, the population size is small and the sampling period is short (on the order of 3 weeks). In practical terms, if a goal of population sampling is to obtain estimates of survival and detection probabilities, our data indicate that minimum mark–recapture sampling intensity for short-lived butterflies should be every weekday. Less frequent sampling (proportional to life span) may be appropriate for longer-lived insects. With higher sampling intensities, our analysis favors mark–recapture, and with lower sampling intensities, our analysis favors Pollard–Yates.

Relative Costs among Approaches

Mark–recapture methods are likely to give better information and estimates of population size and strengthen the knowledge necessary to manage and conserve rare populations. Nevertheless, mark–recapture is relatively expensive because it takes more time per survey and per generation than transect counts. The financial costs of implementing mark–recapture for St. Francis' satyr are marginally higher than transect counts with equal sampling intensity, but that difference grows as sampling intensity is reduced for count-based approaches (Table 2). Mark–recapture also may not be possible with very mobile species that are unlikely to be recaptured, but rare butterflies tend to be patchily distributed and restricted to small areas.

Mark–recapture has the additional disadvantage of potentially damaging the population one is trying to conserve. In our case we found that marking does not harm individual butterflies or butterfly populations (Kuefler et al. 2008). The few other attempts to evaluate effects of marking had variable results. Some researchers found no among-day effects (Gall 1984; Orive & Baughman 1989).

Others found that marking decreases recapture rates due to increasing mortality rates (Morton 1984), increasing migration rates (Singer & Wedlake 1981), or changing activity patterns (Mallet et al. 1987). It may be impossible to mark the smallest and most fragile rare butterflies, such as some Hesperiids and Lycaenids, because of risk of harming them. Under any circumstances, rare butterflies should only be marked by trained researchers after evaluating the effect of marking scientifically and ethically (Murphy 1987). Because the full effects of marking are unknown, we assume transect counts cause less harm.

Distance Sampling

Another approach to estimating butterfly population size that we did not use in our case study and that does not depend on handling butterflies is distance sampling (Table 2; Buckland et al. 2001). Although it has not been applied widely to population studies of rare butterflies, its use is increasing (Brown & Boyce 1998; Powell et al. 2007). By determining the relationship between observation rates and distances from transects, differences in detection probability can be estimated and incorporated into analyses of population density. This method assumes that transects are placed randomly with respect to butterflies (thus assuming a uniform density distribution about the transect); detection is certain at the center of the transect; detection is a decreasing function of distance; there are no measurement errors; and nothing else influences detection.

In our case distance sampling was not possible. We did not locate transects randomly because it was not logistically possible in wetlands that include areas dominated by grasses, thick shrubs, and open water and because we wanted to maximize our chances of seeing individuals in small populations. Random transect placement is not a requirement of mark–recapture analysis and may even cause violation of mark–recapture assumptions that individuals have equal likelihood of capture. Because of habitat variability within habitat patches, we also could not determine over what area we would expect uniform densities of St. Francis' satyr. In studies of butterflies where habitats are uniform and visibility is high, where random transect locations do not harm butterfly habitats, and where butterflies occur in numbers large enough to estimate detection functions, distance sampling is a promising approach for monitoring populations of rare butterflies. Other techniques that overcome limiting assumptions in distance sampling but need development for butterflies are double-observer approaches (Alldredge et al. 2006) and repeated-count methods (Royle & Nichols 2003) that, unlike distance sampling, account for some butterflies being unavailable for detection (e.g., if they are resting). Because distance sampling generates daily population sizes, results must be combined with

mark–recapture survivorship data or with INCA to generate the total butterfly population size across a generation.

Optimal Monitoring Strategy

Considering the costs of mark–recapture and the limitations of transect counts, a hybrid approach may be preferable, whereby limited mark–recapture estimates of survival and detection probabilities are combined with more frequent transect counts to provide accurate indices of population size (Schultz & Dlugosch 1999; Schultz & Hammond 2003). With estimates of detection probabilities, population indices derived from Pollard–Yates indices can be adjusted and are strongly correlated with mark–recapture population estimates. A key question we cannot currently answer completely is, In what proportion of generations or sites must detection probability be estimated? If butterflies are robust to marking, resources are available, and habitats or survey conditions vary greatly across sites or generations, then mark–recapture should be used in each flight period. Where butterflies are fragile and live in stable habitats, perhaps one estimate of detection probability will suffice. One case in which count-based studies were supplemented with a single mark–recapture study is the Fender's blue butterfly (*Icaricia icarioides fenderi*), which lives in stable meadows (Schultz & Hammond 2003).

In contrast, the habitats occupied by St. Francis' satyr are dynamic. After wetland creation by beaver impoundments, successional change increased cover by shrubs and trees that changed visibility, rapidly making sites unsuitable for St. Francis' satyr. This combination of factors causes detection probability for this butterfly to be low and variable, increasing the need for mark–recapture. Mark–recapture or distance sampling should be conducted for a time series that captures changes in survival and detection probabilities in relation to successional changes. Still, spatial variation may require implementation of a design in which detection probability is estimated at a fraction of sites, optimizing the trade-off in costs of surveys and reduction in variation of estimates of detection probability (Pollock et al. 2002).

Our analysis focused on rare butterflies, but it may apply to other rare insects. Butterflies comprise the greatest proportion of threatened insects, including approximately 40% of listed insect species under the U.S. Endangered Species Act, 50% of insect species listed in the British Wildlife and Countryside Act, and 25% of insect species considered threatened in the World Conservation Union Red List. Our results should be applicable to other rare insects that are visually surveyed and can be marked (e.g., Maes et al. 2006). As more becomes known about the rarity of other insects, sampling approaches like those we discuss here will become more widely applied to them. Alternatively, we did not consider 2 important approaches for monitoring rare butterflies. For species that occupy tens to hundreds of patches, an alternative method for trend assessment focuses on patch occupancy (WallisDeVries 2004; MacKenzie et al. 2006; Roy et al. 2007). For a few rare butterfly species, monitoring is best conducted for larvae (Wahlberg et al. 2002), and our discussion regarding transect counts may be applicable.

Optimal monitoring strategies for rare butterflies are determined by information needs, statistical performance, and costs among approaches. In cases where estimates of population demography and variability are needed, where resources permit high sampling intensity, and where potential damage caused by capture is low, then mark–recapture is the best approach for monitoring rare butterfly populations. Where resources do not permit intensive sampling and where butterflies are fragile and cannot be marked, our analysis suggests that Pollard–Yates counts can provide a precise index of population size that, when combined with estimates of detection and survival probabilities that capture their variability over space and time, is also unbiased. Although we did not evaluate it directly, distance sampling may be used in instances when mark–recapture is not possible to provide detection probabilities needed to estimate daily, but not total, population size. The INCA may be applicable in limited circumstances where marking is not possible and demographic information is needed; yet, it should be applied to small populations with caution after making adjustments that may be necessary for small populations (Gross et al. 2007). In most cases, a hybrid of mark–recapture and count-based approaches will provide optimal monitoring data to estimate population viability.

Literature Cited

Alldredge, M. W., K. H. Pollock, and T. R. Simons. 2006. Estimating detection probabilities from multiple-observer point counts. The Auk **123:**1172–1182.

Baguette, M., and N. Schtickzelle. 2003. Local population dynamics are important to the conservation of metapopulations in highly fragmented landscapes. Journal of Applied Ecology **40:**404–412.

Bergman, K. O. 2001. Population dynamics and the importance of habitat management for conservation of the butterfly *Lopinga achine*. Journal of Applied Ecology **38:**1303-1313.

Brown, J. A., and M. S. Boyce. 1998. Line transect sampling of Karner blue butterflies (*Lycaeides melissa samuelis*). Environmental and Ecological Statistics **5:**81-91.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, United Kingdom.

Collier, N., D. A. Mackay, and K. Benkendorff. 2008. Is relative abundance a good indicator of population size? Evidence from fragmented populations of a specialist butterfly (Lepidoptera: Lycaenidae). Population Ecology **50:**17-23.

Crosbie, S. F., and B. F. J. Manly. 1985. Parsimonious modeling of capture-mark-recapture studies. Biometrics **41:**385-398.

Gall, L. F. 1984. The effects of capturing and marking on subsequent activity in *Boloria acrocnema* (Lepidoptera: Nymphalidae), with a comparison of different numerical models that estimate population size. Biological Conservation **28:**139-154.

Gross, K., E. J. Kalendra, B. R. Hudgens, and N. M. Haddad. 2007. Robustness and uncertainty in estimates of butterfly abundance from count data. Population Ecology **49:**191-200.

Haddad, N., D. Kuefler, and B. Hudgens. 2005. Monitoring protocol for St. Francis' satyr: final report for the 2004 field season. North Carolina State University, Raleigh.

Harker, R. J., and T. G. Shreeve. 2008. How accurate are single site transect data for monitoring butterfly trends? Spatial and temporal issues identified in monitoring *Lasiommata megera*. Journal of Insect Conservation **12:**125-133.

Kuefler, D., N. M. Haddad, S. Hall, B. Hudgens, B. Bartel, and E. Hoffman. 2008. Distribution, population structure, and habitat use of the endangered St. Francis' satyr butterfly, *Neonympha mitchellii francisci*. American Midland Naturalist **159:**298-320.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, Boston.

Maes, D., A. Ghesquiere, M. Logie, and D. Bonte. 2006. Habitat use and mobility of two threatened coastal dune insects: Implications for conservation. Journal of Insect Conservation **10:**105-115.

Mallet, J., J. T. Longino, D. Murawski, A. Murawski, and A. S. D. Gamboa. 1987. Handling effects in *Heliconius*: where do all the butterflies go? The Journal of Animal Ecology **56:**377-386.

Manly, B. F. J. 1974. Estimation of stage-specific survival rates and other parameters for insect populations developing through several stages. Oecologia **15:**277-285.

Mattoni, R., T. Longcore, C. Zonneveld, and V. Novotny. 2001. Analysis of transect counts to monitor population size of endangered insects. Journal of Insect Conservation **5:**197-206.

Morris, W. F., and D. F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Sunderland, Massachusetts.

Morton, A. C. 1984. The effects of marking and handling on recapture frequencies of butterflies. Pages 55-61 in R. I. Vane-Wright and P. R. Ackery, editors. The biology of butterflies. Princeton University Press, Princeton, New Jersey.

Murdock, N. A. 1996. Recovery plan for Saint Francis' Satyr (*Neonympha mitchellii francisci*). U.S. Fish and Wildlife Service, Asheville, North Carolina.

Murphy, D. D. 1987. Are we studying our endangered butterflies to death? Journal of Research on the Lepidoptera **26:**236-239.

New, T. R. 1991. Butterfly conservation. Oxford University Press, Oxford, United Kingdom.

Orive, M. E., and J. F. Baughman. 1989. Effects of handling on *Euphydryas edithia* (Nymphalidae). Journal of the Lepidopterists Society **43:**244-247.

Pollard, E., and T. J. Yates. 1993. Monitoring butterflies for ecology and conservation. Chapman & Hall, London.

Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics **13:**105-119.

Powell, A. F. L. A., W. H. Busby, and K. Kindscher. 2007. Status of the regal fritillary (*Speyeria idalia*) and effects of fire management on its abundance in northeastern Kansas, USA. Journal of Insect Conservation **11:**299-308.

Roy, D. B., P. Rothery, and T. Brereton. 2007. Reduced-effort schemes for monitoring butterfly populations. Journal of Applied Ecology. **44:**993-1000.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology **84:**777-790.

Schultz, C. B., and K. M. Dlugosch. 1999. Nectar and hostplant scarcity limit populations of an endangered Oregon butterfly. Oecologia **119:**231-238.

Schultz, C. B., and P. C. Hammond. 2003. Using population viability analysis to develop recovery criteria for endangered insects: case study of the fender's blue butterfly. Conservation Biology **17:**1372-1385.

Schwarz, C. J., and A. N. Arnason. 1996. A capture methodology for the analysis of capture–recapture experiments in open populations. Biometrics **52:**868-873.

Seidl, A. L. 1999. Population demographics and the conservation status of the Uncompahgre Fritillary *Boloria acrocnema* (Nymphalidae). Journal of the Lepidopterists Society **53:**32-36.

Singer, M. C., and P. Wedlake. 1981. Capture does affect probability of recapture in a butterfly species. Ecological Entomology **6:**215-216.

Thomas, J. A. 1983. A quick method for estimating butterfly numbers during surveys. Biological Conservation **27:**195-211.

Wahlberg, N., T. Klemetti, and I. Hanski. 2002. Dynamic populations in a dynamic landscape: the metapopulation structure of the marsh fritillary butterfly. Ecography **25:**224-232.

WallisDeVries, M. F. 2004. A quantitative conservation approach for the endangered butterfly *Maculinea alcon*. Conservation Biology **18:**489.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study **46**(suppl.):120-138.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, New York.

Zonneveld, C. 1991. Estimating death rates from transect counts. Ecological Entomology **16:**115-121.