

Comparing Population Size Estimators for Plethodontid Salamanders

LARISSA L. BAILEY,^{1,2} THEODORE R. SIMONS¹ AND KENNETH H. POLLOCK³

¹*U.S. Geographical Survey, Cooperative Fish and Wildlife Research Unit, Department of Zoology, North Carolina State University, Campus Box 7617, Raleigh, North Carolina 27695-7617, USA*

³*Department of Statistics, Biomathematics, and Zoology, North Carolina State University, Raleigh, North Carolina 27695-8203, USA*

ABSTRACT.—Despite concern over amphibian declines, few studies estimate absolute abundances because of logistic and economic constraints and previously poor estimator performance. Two estimation approaches recommended for amphibian studies are mark-recapture and depletion (or removal) sampling. We compared abundance estimation via various mark-recapture and depletion methods, using data from a three-year study of terrestrial salamanders in Great Smoky Mountains National Park. Our results indicate that short-term closed-population, robust design, and depletion methods estimate surface population of salamanders (i.e., those near the surface and available for capture during a given sampling occasion). In longer duration studies, temporary emigration violates assumptions of both open- and closed-population mark-recapture estimation models. However, if the temporary emigration is completely random, these models should yield unbiased estimates of the total population (superpopulation) of salamanders in the sampled area. We recommend using Pollock's robust design in mark-recapture studies because of its flexibility to incorporate variation in capture probabilities and to estimate temporary emigration probabilities.

Despite concern over amphibian population declines, few amphibian studies estimate population abundances because of logistical and financial costs of obtaining precise parameter estimates (Welsh and Lind, 1992; Jung et al., 2000; Petranka and Murray, 2001; Schmidt, 2003). Rather, most inferences are made using relative abundance indices based on count data (Schmidt, 2003). Such indices assume that there is a constant, linear relationship between the index and population size and that all individuals of a particular species have the same detection probability (equal capture probability; Lancia et al., 1994). These assumptions are unlikely in most wildlife populations including the terrestrial salamander populations we studied in the southern Appalachians (Bailey et al., 2004a,b).

Many methods are available to estimate population size, including change-in-ratio methods, catch-per-unit-effort, removal and depletion methods, and mark-recapture methods (for reviews, see Seber, 1982). Two approaches recommended for amphibian studies are mark-recapture and depletion (or sometimes referred to as removal) sampling (Heyer et al., 1994). Because these methods are labor intensive, they have been recommended only when researchers require detailed knowledge of the target population; otherwise relative abundance indices are

currently preferred (Heyer et al., 1994). It should be noted that, even when the focus is on relative abundance, investigators should provide evidence that detection probabilities are practically equivalent between populations being compared (MacKenzie and Kendall, 2002).

In this paper, we compare abundance estimation via various mark-recapture models and depletion models, using data from a three-year study of terrestrial salamanders in Great Smoky Mountains National Park (GSMNP). Terrestrial salamanders are known to reach high densities in this area and other eastern forests (Burton and Likens, 1975; Howard, 1987; Petranka and Murray, 2001). Moreover, plethodontids are susceptible to a variety of natural and anthropogenic perturbations and have been promoted as excellent indicators of forest ecosystem integrity (see review by Welsh and Droege, 2001). We compare salamander population estimates from depletion models and categories of mark-recapture models including closed-population, open-population, and robust design models to illustrate how different model assumptions can result in wide variation in population estimates. We also consider and clarify which methods estimate surface population versus superpopulation (total population) size.

Population Estimation.—Mark-recapture estimators can be separated into three classes: closed-population; open-population; and robust design models. Closed-population models make three general assumptions: (1) The population is closed to births, deaths, immigrants, and emi-

² Corresponding Author. Present address: USGS Patuxent Wildlife Research Center, 12100 Beach Forest Road, Laurel, Maryland 20708, USA; E-mail: lbailey@usgs.gov

grants during the sampling interval (closure assumption); (2) All animals are equally likely to be captured in each sample (equal capture probability) and no animal has probability of capture equal to zero, $p \neq 0$; and (3) Marks are not lost or overlooked by the observer. The equal capture probability assumption may be relaxed by allowing capture probabilities to vary as a result of time (t), behavioral (or trap) response (b), and heterogeneity (h). Heterogeneity implies that capture probabilities vary among individual salamanders, possibly because of factors such as species type, body size, or sex. Well-known Lincoln-Peterson (for two samples), Schnabel, and Schumacher-Eschmeyer estimators (Seber, 1982) allow time variation in capture probabilities. These methods only require batch marks. However, if animals have been individually marked, the use of program CAPTURE (Otis et al., 1978; Rexstad and Burnham, 1991) to calculate maximum-likelihood estimators is recommended (Pollock et al., 1990). Depletion or removal models (Zippen, 1956; Seber, 1982) are similar to the behavioral model, M_b , in program CAPTURE, because only information from an animal's first capture contributes to population estimates (Pollock et al., 1990). Studies that remove individuals from the population assume a closed population, equal sampling effort, and constant capture probabilities over time (White et al., 1982). The behavioral model, M_b , assumes that marked and unmarked individuals have different capture probabilities, but no temporal variation in capture probabilities (Otis et al., 1978; Pollock et al., 1990). Finally, capture probabilities may vary among individuals (heterogeneity, h) and estimators for heterogeneous populations included Burnham's "jackknife" technique (Burnham and Overton, 1978, 1979), Chao's M_h estimator (Chao, 1988), and finite mixture models (Norris and Pollock, 1996; Pledger, 2000). Program CAPTURE contains estimators for seven of the eight closed population models proposed by Otis et al., (1978; M_{or} , M_b , M_{lh} , M_t , M_{bh} , M_{tb} , M_{th}). CAPTURE also contains a test for closure and a model selection procedure, together with a series of goodness-of-fit tests to aid investigators in model choice. The test for closure and the model selection procedure both have limitations, and biologists should interpret model results carefully (Otis et al., 1978; Menkens and Anderson, 1988; Pollock et al., 1990; Stanley and Burnham, 1999; Pledger, 2000).

In studies where the demographic closure assumption cannot be met, open-population models can estimate population size at each sampling period, and survival and recruitment probabilities between sampling periods (Pollock et al., 1990). The most common model, the Jolly-Seber (JS) model, (Jolly, 1965; Seber, 1965) has

three main assumptions (Pollock et al., 1990): (1) All animals present in the population at time i ($i = 1, 2, \dots k$) are equally likely to be captured (equal capture probability); (2) Every marked animal present in the population at time i has the same probability of surviving from i to $i + 1$; and (3) Marks are not lost or overlooked by the observer. An additional, implicit assumption is that all emigration from the population is permanent; temporary emigration in and out of the population violates assumptions of the Jolly-Seber (JS) model. Seber (1982) and Pollock et al. (1990) presented several variations of the JS model and many are included in programs JOLLY (Pollock et al., 1990) and MARK (White and Burnham, 1999). Despite these additions, the JS model is still limited in that some parameters cannot be estimated (e.g., population size in the first and last periods) and the population size estimator is not robust to variation in capture probabilities (other than time variation; Kendall and Pollock, 1992).

To address these concerns, Pollock developed the "robust design" model (Pollock, 1982). Under this design, primary sampling periods, i ($i = 1, 2, \dots k$) contain l_i secondary sample occasions that are separated by a time interval that is short enough to assume demographic closure (i.e., no birth, death, immigration, or emigration). Primary periods are separated by longer time intervals during which population additions (immigration and birth) and deletions (emigration and death) can occur. Data from secondary samples within each primary period can be analyzed using closed-population models that allow for unequal capture probability (Otis et al., 1978; White et al., 1982). The closed-population models estimate conditional capture probabilities (p), recapture probabilities (c), and population size for each primary period (N). Data within each primary period are pooled to estimate survival probabilities (ϕ), and data from both primary and secondary periods are used to estimate temporary emigration (γ ; Kendall and Nichols, 1995; Kendall et al., 1997; Kendall, 1999). This design is especially useful for estimating temporary emigration probabilities (Kendall and Nichols, 1995; Kendall et al., 1997; Kendall, 1999; Kendall and Hines, 1999).

Previous salamander studies have used a wide variety of closed- and open-population estimation methods. Closed-populations methods include Lincoln-Peterson for two samples (Burton and Likens, 1975; Welsh and Lind, 1992); and for more than two samples: Schabel's (Stewart and Bellis, 1970; Howard, 1987; Smith and Petranka, 2000), Schumacher-Eschmeyer (Semlitsch, 1980; Howard, 1987), and depletion methods (Bruce, 1995; Petranka and Murray, 2001; Salvidio, 2001). We know of only one salamander study (Jung et

al., 2000) that has used the multiple closed-population models discussed in Otis et al. (1978). This is important because none of the other methods allows, or tests for, heterogeneity of capture probabilities between different animals. JS open-population models have been used in salamander studies by Tilley (1980), Welsh and Lind (1992), and Marvin (1996). Most studies were conducted on 1–4 sites (but see Jung et al., 2000; Smith and Petranka, 2000), and they reported low effective capture probabilities ($p < 0.15$) which resulted in imprecise population estimates (Howard, 1987; Welsh and Lind, 1992; Jung et al., 2000). Bailey et al. (2004a,b) represents the most extensive terrestrial salamander study that focuses on population estimation and is the only study to use the robust design.

MATERIALS AND METHODS

Study Area.—Great Smoky Mountains National Park, located along the Tennessee-North Carolina border, is internationally recognized for its rich temperate forest biodiversity. The park's unique geology, along with steep, complex topography, creates temperature and moisture gradients across 205,665 ha of contiguous forest. These gradients contribute to produce high levels of temperate species diversity in many taxa, including salamanders. Approximately 10% of the world's salamander species are found in the southern Appalachian region (Petranka, 1998), with 31 species occurring inside the park's boundaries (Dodd, 2003). Salamanders are a high priority taxon for the park's inventory and monitoring program because of their high diversity, large number of species endemic to the southern Appalachians, and the limited amount of data on the distribution, abundance, and natural history of most species.

Field Methods.—In 1999, we initiated a three-year mark-recapture study on 15 plots (15 × 15 m) in the Roaring Fork Watershed (Mt. LeConte USGS Quadrangle); five additional plots were added in 2000. Each plot was enclosed with a silt fence to inhibit horizontal salamander movement. The bottom edge of the silt fence was buried 10–15 cm into the soil, and the remainder of the fence was raised and stapled to 60-cm tall wooden stakes. The top edge of the fence was draped toward the inside of the plot creating a lip that made it difficult for salamanders to escape. We performed salamander searches outside the plots each year; of over 460 salamanders captured, only six were recaptured having escaped the fences.

Three parallel 15-m transects were established within each plot, including a natural cover transect, a transect of five coverboard arrays and a set of five leaf litter search stations (for details, see Bailey et al., 2004a). Plots were

sampled according to Pollock's robust design with four primary periods, each consisting of 3–4 consecutive sample days (Pollock, 1982). Primary periods were separated by 6–8 days in 1999 and 12 days in 2000 and 2001. Sampling was conducted from 1 April to mid-June each year. During each sampling occasion, the three transects were sampled first; then the remaining natural cover was turned, and the inside edge of the fence was searched. This procedure ensured that all available animals within the plot had a nonzero capture probability during the given primary period. All unmarked salamanders over 18-mm snout-vent length (SVL) were individually marked using fluorescent elastomer (Bailey et al., 2004a).

Following the 30–40 day mark-recapture study, a depletion study was initiated on each plot. Plots were sampled every other day for eight days, resulting in four depletion samples per plot. Captured animals were removed from plots and stored in a refrigerator at 4–5°C. Salamanders can be retained in this manner with minimal care for several weeks (Stewart and Bellis, 1970; DeNardo, 1995; Salvidio, 1998).

In 1999, a complete removal study was conducted on a subset of three plots. All natural cover and leaf litter was removed from these plots using the method presented in Heatwole and Sexton (1966). The material was thoroughly searched for salamanders before it was removed from the plot. Leaf litter was collected and stored in plastic garbage bags outside the plot. All captured salamanders were stored in refrigerators. Following excavation, plots were visited on consecutive nights to remove all available salamanders. Sampling was terminated when no new salamanders were found on nightly surveys, usually after 4–5 nights. It is thought that removing leaf litter and cover yields a complete count of forest-floor herpetofauna within an enclosed quadrat (Jaeger and Inger, 1994; Heatwole, in press). After complete removal sampling was halted, the bagged leaf litter and natural cover were replaced, carefully reconstructing the plot as closely to its original design as possible. Captured salamanders were marked and released at their original capture locations.

Population Estimation.—We used several methods to estimate salamander abundances at each site in each year (site-year). We used all 14–16 sampling occasions (four primary periods × 3–4 secondary samples each) to construct capture history datasets for each site-year and used program CAPTURE to test demographic closure, perform model selection, and estimate population size under a number of closed-population models (Otis et al., 1978; White et al., 1982; Rexstad and Burnham, 1991). In this case, the closure assumption applied across all sampling

TABLE 1. The number of individual salamanders captured (>18-mm snout-vent length) on 15 × 15 m plots in Great Smoky Mountains National Park during mark-recapture, depletion and complete removal studies. *N*-plots were sampled 14–16 times per year. Complete removals were only performed at three plots in 1999.

| Year | <i>N</i> | Mark-recapture | Depletion | | Complete removal | | Total |
|------|----------|----------------|-----------|--------|------------------|--------|--------|
| | | Caught | Unmarked | Marked | Unmarked | Marked | Caught |
| 1999 | 14 | 1417 | 375 | 122 | 172 | 34 | 1964 |
| 2000 | 19 | 2140 | 334 | 189 | — | — | 2474 |
| 2001 | 19 | 1778 | 346 | 194 | — | — | 2124 |

occasions within the season, and we refer to these estimates as full-season closed-population estimates. This analysis was the first exploratory step to determine the types of variation in capture probabilities (heterogeneity, behavioral or time) that characterized each dataset. Data from the four-sample depletion study were analyzed using the behavioral model, M_b , and the generalized removal method, M_{bh} (program CAPTURE, Otis et al., 1978). The generalized removal method allows heterogeneity of capture probabilities.

A third closed-population estimate was obtained using the Lincoln-Peterson method. Because the Lincoln-Peterson estimator uses only two sampling occasions, all individuals caught in 30–40 day mark-recapture study were pooled into a single sampling period, n_1 , and the individuals obtained in the eight-day depletion study were pooled into a single recapture period, n_2 . The number of previously marked animals in the depletion sample became m_2 , and an unbiased estimate of population size was obtained using:

$$\hat{N} = \left[\frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} \right] - 1 \text{ (Chapman, 1951)}$$

Pollock et al. (1990) and Seber (1982) presented detailed exposition on this model and its standard error (not shown here). Closed-population estimates were compared to one another and to the complete removal counts from three plots in 1999.

Open-population estimates were obtained for each data set by pooling secondary samples within each primary period. We considered three standard JS models including the full Time-specific Model (Model A), the Constant Survival Model (Model B), and the Constant Survival and Capture Model (Model D; Pollock et al., 1990). Analysis was performed using program JOLLY (Pollock et al., 1990).

We analyzed the robust design mark-recapture data using program MARK (White and Burnham, 1999). In previous work, we fit a series of models that represented possible salamander responses (time and behavioral) with and without temporary emigration (Bailey et al., 2004a).

Models that contained random temporary emigration were chosen best for over 80% of our datasets (using QAIC criteria; Akaike, 1973; Burnham and Anderson, 2002; Bailey et al., 2004a). In this study, we used a subset of robust design models, each containing constant random temporary emigration, $\gamma(\cdot)$, time-invariant average surface population size, $N(\cdot)$, and either constant, behavioral, or time-specific capture probabilities (RD- M_o , RD- M_b , RD- M_t ; Bailey et al., 2004a). A detailed description of these models is given in Appendix 1. To compare robust design model results to JS open-population estimates, we did not fix apparent survival probability over primary periods ($\phi(\cdot) \neq 1$) but allowed the model to estimate a constant survival parameter, $\phi(\cdot)$, between primary periods. Finally, we analyzed the secondary samples within each primary period using closed-population models M_b and M_h in program CAPTURE. Here the closure assumption only applies to the short, 3–4-day period over which secondary samples were obtained.

RESULTS

A total of 6562 salamanders (>18 mm SVL) and 11 species were captured and marked during our three-year study (Table 1). Of these, 206 salamanders were removed from three plots during the complete removal study in 1999. Most of the captures were members of seven species or species complexes including *Plethodon jordani*, *Plethodon serratus*, *Desmognathus wrighti*, *Eurycea wilderae*, and three species complexes described by Petranka (1998): *Plethodon glutinosus* complex, *Desmognathus fuscus* complex, and *Desmognathus imitator* complex. Captures of *Desmognathus quadramaculatus*, *Desmognathus monticola*, *Gyrinophilus porphyriticus*, and *Pseudotriton ruber* were rare.

The average number of salamanders caught per site (excluding complete removals) varied little between years: 1999 mean \pm 1 SE = 128 \pm 19.89, N = 14 sites; 2000 = 130 \pm 28.02, N = 19 sites; 2001 = 112 \pm 20.32, N = 19 sites.

Despite the use of a silt fence as a barrier to salamander movement, the test for closure was

TABLE 2. Percentage of the datasets for which different closed-population models were selected by program CAPTURE. Datasets contained all 14–16 sampling occasions. N = number of datasets (sites).

| Year | N | Models | | | | | | | |
|-------|-----|--------|-------|-------|----------|-------|----------|-------------|-------------|
| | | M_o | M_h | M_b | M_{bh} | M_t | M_{th} | M_{tb} | M_{tth} |
| 1999 | 14 | 28.6 | 14.3 | 7.1 | 0.0 | 0.0 | 7.1 | 7.1 | 35.7 |
| 2000 | 19 | 10.5 | 5.3 | 5.3 | 0.0 | 26.3 | 5.3 | 26.3 | 21.0 |
| 2001 | 19 | 5.3 | 0.0 | 5.3 | 5.3 | 5.3 | 15.8 | 42.1 | 21.0 |
| Total | 52 | 13.5 | 5.8 | 5.8 | 1.9 | 11.5 | 9.6 | 26.9 | 25.0 |

rejected ($P < 0.05$) for 30 of the 52 datasets containing all 14–16 sample occasions within a sampling season (program CAPTURE). Closed-population models containing time and behavioral variation in capture probabilities fit our data better than other competing models according to model selection procedures in program CAPTURE (Table 2). Time variation (t) was included in 73.0% of selected models, whereas behavioral effects (b) and heterogeneity (h) were included in 59.6% and 42.3%, respectively; constant capture probabilities were suggested for only 13.5% of our datasets. Unfortunately, models with both time and behavioral effects, M_{tbt} , are often difficult to fit with sparse data (few numbers of captured or recaptured individuals), and there is currently no estimator available for model M_{tbt} . Because model selection procedures in program CAPTURE often perform poorly with sparse data (Menkens and Anderson, 1988), we analyzed the same datasets using the closed-captures procedure in program MARK (White and Burnham, 1999) where AIC_c -values could be used to evaluate relative model performance. Models containing heterogeneous capture probabilities are excluded in this assessment, but model selection procedures in program CAPTURE indicated that heterogeneity was the least influential of the three possible capture probability effects (Table 2). A subset of four closed-population models (M_o , M_b , M_t , M_{tb}) were tested, and all models with $\Delta AIC_c < 2.0$ were considered appropriate for a given dataset (Burnham and Anderson, 2002). Model selection results using

AIC_c were compared to the best subset model from program CAPTURE (i.e., the subset model with the value closest to 1.0; Table 3). Neither model selection procedure yielded a dominant capture probability effect but suggested that either time or behavioral or both effects influenced salamander capture probabilities.

Capture probabilities using full-season closed-population models were low regardless of model choice; the highest average capture probability estimate for any dataset was $\hat{p}(\cdot) = 0.06$ using the null model, M_o , and $\hat{p}(\cdot) = 0.04$ using the heterogeneity model, M_h ($N = 52$ datasets). Nearly all data sets suggested a trap-shy behavioral response, where the highest estimates of initial capture probability (p) and recapture probability (c) were $\hat{p}(\cdot) = 0.27$ and $\hat{c}(\cdot) = 0.07$ ($N = 42$ datasets).

Open-population estimation was difficult because pooling secondary samples reduced the number of recaptures. For example, if an animal was recaptured multiple times within a primary period but not recaptured in other primary periods that animal appeared to be captured only once in the open-population analysis. Pooling resulted in a 32% reduction in the number of recaptured animals and thus open-population analyses were possible for only 35 of 52 datasets. All three JS models and the full complement of goodness-of-fit tests could be calculated for only 10 datasets. The goodness-of-fit test comparing JS Constant Survival and Capture Model (Model D) versus full Time-specific Model (Model A) was possible for 31 datasets and the more restricted

TABLE 3. Percentage of the datasets for which different closed-population models were considered appropriate by AIC_c -values (calculated in program MARK) and by model selection procedures in program CAPTURE. Datasets contained all 14–16 sampling occasions. N = number of models considered appropriate for datasets in each year.

| Year | MARK model selection | | | | | CAPTURE model selection | | | | |
|-------|----------------------|-------|-------|-------|----------|-------------------------|-------|-------|-------|----------|
| | N | M_o | M_b | M_t | M_{tb} | N | M_o | M_b | M_t | M_{tb} |
| 1999 | 23 | 17.4 | 34.8 | 39.1 | 8.7 | 14 | 71.4 | 7.14 | 0.0 | 21.4 |
| 2000 | 28 | 10.7 | 25.0 | 32.1 | 32.1 | 19 | 26.3 | 5.3 | 36.8 | 31.6 |
| 2001 | 26 | 15.4 | 30.8 | 26.9 | 26.9 | 19 | 26.3 | 5.3 | 15.8 | 52.6 |
| Total | 77 | 14.3 | 29.9 | 32.5 | 23.4 | 52 | 38.5 | 5.8 | 19.2 | 36.5 |

TABLE 4. Example of population size estimates (\hat{N}) for one site, CG008, in 1999, using closed-population, open-population and robust design models (Appendix 1). Capture probabilities are either equal (*o*), or vary as a result of time (*t*), behavioral (or trap) response (*b*), and heterogeneity (*h*). Full-season closed-population model M_{tbh} was the best model selected by program CAPTURE. Values in parentheses are superpopulation estimates using corresponding robust design models. Day removal equals the number of individuals captured while removing all leaf litter and cover from the plot; Complete removal equals the number of individuals captured during depletion and complete removal studies; and Total Captured equals the number of individuals captured during mark-recapture, depletion, and complete removal studies.

| Model | Surface population | | Model | Superpopulation | |
|--------------------|--------------------|---------------|----------------------|-----------------|--------------|
| | \hat{N} | SE \hat{N} | | \hat{N} | SE \hat{N} |
| Robust Design | | | Full-Season Closed | | |
| RD- M_o | 264 (264) | 60.05 | M_o | 345 | 44.79 |
| RD- M_b^* | 64 (260) | 7.54 | M_b | 183 | 16.11 |
| RD- M_t | 264 (264) | 60.25 | M_t | 341 | 43.77 |
| Depletion | | | Open-population (JS) | | |
| M_b | 41 | 6.92 | Model B | 245 | 159.25 |
| M_{bh} | 33 | 4.73 | Model D | 261 | 128.83 |
| Secondary Samples | | | Lincoln-Peterson | 291 | 48.90 |
| Closed (4 periods) | | | | | |
| | — | — | | | |
| $M_b : M_t$ | 59 : 130 | 4.96 : 3.96 | Day Removal | 27 | |
| | 46 : 86 | 7.39 : 11.25 | Complete Removal | 123 | |
| | 70 : 83 | 41.09 : 10.94 | Total Captured | 237 | |

* Robust design model with the lowest QAIC_c value.

Model D was favored for 81% of these datasets. Population estimates were possible for 47 depletion datasets, using model M_b and 43 datasets, using the generalized removal method M_{bh} . Lincoln-Peterson population estimates were possible for 50 of 52 datasets.

We used estimates from the three complete removal sites in 1999 to illustrate the variation among population size estimates obtained from different types of models (closed-population, open-population, robust design, and depletion models; Tables 4–6). Population estimates obtained from different closed-population models varied widely, whereas estimates from open-population JS models and depletion models were more consistent (Tables 4–6). Despite these differences, a distinct pattern emerged that we believe represents two different populations. Depletion models yielded the lowest population estimates but had good agreement with robust design estimates that included temporary emigration and closed-population estimates from secondary sampling periods only (Tables 4–6). We feel these models estimated the population of salamanders near the surface and available for capture at a given sampling occasion (i.e., surface population). Table 6 represents one of the few exceptions to this generality, but in 1999 the RG012 site had an exceptionally low numbers of recaptured individuals (only 7 of 111 individuals were recaptured during the mark-recapture study) resulting in poor estimates and precision.

Closed-population models over all 15–16 sampling occasions (full-season), JS open-population models, and Lincoln-Peterson estimators all yielded much higher population estimates than depletion or robust design estimates. Population estimates from depletion model M_b were usually 5–20% of Lincoln-Peterson estimates for the same site-years (range: 4.39–40.42%, $N = 47$ datasets). The robust design models listed in each table (4–6) used the same number of sampling occasions and made the same capture probability assumptions as the corresponding closed-population models, but their average population size estimates were much less. We believe the Lincoln-Peterson model, full-season closed-population models, and JS open-population models estimated both surface and subterranean individuals (termed superpopulation, Kendall, 1999). Using the robust design, random temporary emigration probabilities were incorporated to estimate the superpopulation at each site and these estimates resembled those from full-season closed-population models, JS open-population models, and Lincoln-Peterson methods (Tables 4–6).

DISCUSSION

Population size estimation is an important aspect of many ecological studies and wildlife management programs. There are numerous estimation methods available and a method's

TABLE 5. Example of population size estimates (\hat{N}) for one site, CG009, in 1999 using closed-population, open-population and robust design models (Appendix 1). Capture probabilities are either equal (*o*), or vary as a result of time (*t*), behavioral (or trap) response (*b*), and heterogeneity (*h*). Full-season closed-population model M_t was the best model selected by program CAPTURE. Values in parentheses are superpopulation estimates using corresponding robust design models. Day removal equals the number of individuals captured while removing all leaf litter and cover from the plot; Complete removal equals the number of individuals captured during depletion and complete removal studies; and Total Captured equals the number of individuals captured during mark-recapture, depletion, and complete removal studies.

| Model | Surface population | | Model | Superpopulation | |
|--------------------|--------------------|---------------|----------------------|-----------------|--------------|
| | \hat{N} | SE \hat{N} | | \hat{N} | SE \hat{N} |
| Robust Design | | | Full-Season Closed | | |
| RD- M_o | 92 (188) | 20.59 | M_o | 228 | 25.35 |
| RD- M_b | 43 (156) | 11.97 | M_b | not est. | not est. |
| RD- M_t^* | 112 (214) | 26.15 | M_t | 225 | 24.70 |
| Depletion | | | Open-population (JS) | | |
| M_b | 57 | 10.39 | Model B | 262 | 175.19 |
| M_{bh} | 56 | 33.83 | Model D | 280 | 196.82 |
| Secondary Samples | | | Lincoln-Peterson | 419 | 121.54 |
| Closed (4 periods) | — | — | | | |
| $M_b : M_t$ | 106 : 101 | 4.96 : 3.96 | Day Removal | 11 | |
| | 40 : 81 | 7.39 : 11.25 | Complete Removal | 77 | |
| | — : 104 | 41.09 : 10.94 | Total Captured | 179 | |

* Robust design model with the lowest QAIC_c value.

performance depends on its underlying assumptions. This is especially important when mark-recapture data are sparse. An estimator's relative bias and coefficient of variation is expected to increase as population size or

individual capture probability decrease (Mekens and Anderson, 1988). Two estimation methods recommended for amphibian populations are mark-recapture and depletion methods (Heyer et al., 1994). Population size estimates

TABLE 6. Example of population size estimates (\hat{N}) for one site, RG012, in 1999, using closed-population, open-population and robust design models (Appendix 1). Capture probabilities are either equal (*o*), or vary as a result of time (*t*), behavioral (or trap) response (*b*), and heterogeneity (*h*). Full-season closed-population model M_o was the best model selected by program CAPTURE. Values in parentheses are superpopulation estimates using corresponding robust design models. Day removal equals the number of individuals captured while removing all leaf litter and cover from the plot; complete removal equals the number of individuals captured during depletion and complete removal studies; and total captured equals the number of individuals captured during mark-recapture, depletion, and complete removal studies.

| Model | Surface population | | Model | Superpopulation | |
|--------------------|--------------------|--------------|----------------------|-----------------|--------------|
| | \hat{N} | SE \hat{N} | | \hat{N} | SE \hat{N} |
| Robust Design | | | Full-Season Closed | | |
| RD- M_o^* | 412 (840) | 226.65 | M_o | 787 | 262.66 |
| RD- M_b | 281 (846) | 725.40 | M_b | 281 | 145.08 |
| RD- M_t | 435 (769) | 240.76 | M_t | 780 | 258.52 |
| Depletion | | | Open-population (JS) | | |
| M_b | 31 | 3.46 | Model B | 729 | 1326.44 |
| M_{bh} | 29 | 1.70 | Model D | 763 | 1057.13 |
| Secondary Samples | | | Lincoln-Peterson | 463 | 150.40 |
| Closed (4 periods) | 153 : 89 | 299.80:11.45 | | | |
| $M_b : M_t$ | — | — | Day Removal | 31 | |
| | 109 : 73 | 189.49:10.41 | Complete Removal | 93 | |
| | — | — | Total Captured | 192 | |

* Robust design model with the lowest QAIC_c value.

derived from these methods varied widely for populations of terrestrial salamanders on our sites in GSMNP. An examination of the assumptions of each estimation method provides some insight into the sources of variability that can influence abundance estimates.

All closed-population estimators assume the population is closed over sampling occasions and no mark loss or observer bias. Marking studies indicate that elastomer mark loss and misidentification is rare (Davis and Ovaska, 2001; Bailey, 2004); thus it is likely that all model assumptions involving mark retention and identification are met. The closure assumption, which is of fundamental importance, was rejected for 58% of our full-season datasets, despite the fact that the closure test in program CAPTURE is not particularly powerful (Stanley and Burnham, 1999). Survival probabilities are not well known for terrestrial salamanders, but several studies suggest annual rates are above 45% (Organ, 1961; Tilley, 1980; Hairston, 1983). Thus, survival probability during our 6–8-week sampling period should be near 1. Likewise, growth rates are sufficiently low (Petranka, 1998) that we would not expect large numbers of recruits into the population (>18 mm SVL salamanders) during our study. For these reasons, we conclude that movement (immigration and emigration) was the likely cause of the closure violation. Although we constrained horizontal surface movement with silt fencing, we could not control the vertical movement of salamanders in the soil. We believe this vertical movement, surface to subsurface, explains the lack of closure.

In addition to the closure assumption, our results suggest that the equal capture probability assumption was also violated. Model selection procedures in programs CAPTURE and MARK indicate that capture probabilities varied temporally and had a trap-shy behavioral response (higher initial capture probabilities than recapture probabilities). Previous salamander studies have used estimators that allow temporal variation in capture probabilities (i.e., Lincoln-Peterson, Schnabel and Schumacher-Eschmeyer estimators), but few have incorporated a behavioral response (Stewart and Bellis, 1970; Semlitsch, 1980; Howard, 1987; Smith and Petranka, 2000).

Other studies have shown that capture probabilities may vary with sex (Tilley, 1980) or species (Smith and Petranka, 2000; Petranka and Murray, 2001; Bailey et al., 2004b). Failure to account for heterogeneous capture probabilities in closed- and open-population estimators will result in negatively biased estimates of abundance (Pollock et al., 1990). Our results indicate that heterogeneity was the least influential of the

three possible variations among capture probabilities at a given site and year.

Our most dramatic result was the large discrepancy between the depletion and robust design estimators and the open and closed-population estimators (over the full season). We suggest that the estimators actually apply to two different populations: surface and superpopulations. Depletion estimators use data collected during a short period of time where the closure assumption is likely valid; thus depletion population estimates should represent the available surface population of salamanders during the eight-day sampling period. Superpopulation refers to the population of salamanders both near the surface and available for capture and those beneath the surface and unavailable for capture during a given primary sampling period. Temporary emigration violates the closure assumption for the longer mark-recapture study, but if the temporary emigration is completely random, closed-population and open-population models should yield unbiased estimates of the superpopulation in the sampled area. Using robust design models, we were able to incorporate and estimate temporary emigration (Bailey et al., 2004a,b) and confirmed that random temporary emigration probabilities between primary periods were high (average temporary emigration estimate = 0.87 ± 0.01 ; Bailey et al., 2004a). Surface population estimates calculated using the robust design models were similar to depletion model estimates and closed-population estimates from secondary sample periods only. Superpopulation estimates from robust design models were similar to estimates from full-season closed-population models, JS open-population models, and the Lincoln-Peterson method.

Finally, complete removal techniques have been highly successful at providing absolute abundances for a wide variety of herpetofaunal species (Heatwole and Sexton, 1966; Rodda et al., 2001; Heatwole, in press) but the technique has not been validated for fossorial or subterranean species (G. Rodda, pers. comm.). Our results indicate that complete removal procedures described in this paper do not produce good estimates of the superpopulation of terrestrial salamanders.

The purpose of this paper was to (1) compare abundance estimation via various mark-recapture models and depletion models over a large number of datasets (total of 52 possible site-year combinations) and (2) clarify which methods estimate surface population versus superpopulation size. The population investigated included all salamanders at a given site and year. For the purpose of comparing our results to density estimates reported in other studies, we converted superpopulation estimates to density using

a robust design model containing random temporary emigration ($\phi(\cdot) = 1$, $\gamma(\cdot)$, $p(\cdot)$, $c(\cdot)$, $N(\cdot)$, Bailey et al., 2004a). In previous studies, our data supported this model as the most convincing across all sites and years, using two different model selection strategies (Bailey et al., 2004a). Superpopulation densities ranged from (\bar{D}) = 1.56 salamanders/m² (SE (\bar{D}) = 0.58 $N = 10$ site-year datasets) on our lower elevation sites (< 860 m) to (\bar{D}) = 4.16 salamanders/m² (SE (\bar{D}) = 0.75 $N = 9$ site-year datasets) on our high elevation sites (980–1100 m). Density estimates over all sites and years where estimation was possible were (\bar{D}) = 2.35 salamanders/m² (SE (\bar{D}) = 2.64, $N = 48$ datasets). These estimates are slightly higher than those obtained by Petranka and colleagues (Smith and Petranka, 2000; Petranka and Murray, 2001) who reported total salamander densities of 0.51 and 1.84 salamanders/m², respectively, in the southern Appalachians. Part of the reason for the discrepancy may be the models Petranka and colleagues chose for their analysis. We suggest future analysis include model selection criteria and some assessment of model fit (Burnham and Anderson, 2002; Bailey et al., 2004a). Certainly, however, these population and density estimates suggest that terrestrial salamander densities are high in southern Appalachian forests.

We did not focus on species-specific population parameters in this paper because (1) differences in capture probabilities among individual salamanders (heterogeneity, possibly cause by species differences) was the least influential of the three possible capture probability effects and (2) species-specific variations in population parameters were explored in another paper (Bailey et al., 2004b). Most of the salamanders on the three sites we present in detail were small plethodontids (*Plethodon serratus*) or large members of the *glutinosus* complex (both *Plethodon glutinosus* and *Plethodon oconaluftee*), *Plethodon jordani*, and hybrids. Species-specific density estimates, using the same robust design model discussed above, were possible for all three sites for large plethodontids and two sites for small plethodontids (RG012 excluded for sample size reasons discussed previously). Superpopulation density estimates for large plethodontids were 0.37, 0.18, and 1.43 salamanders/m² for large plethodontids and 2.76 and 1.43 salamanders/m² for small plethodontids. Our estimates for large plethodontids were similar to those reported by Smith and Petranka (2000) (0.13/m², *Plethodon jordani*, Howard (1987) (0.50/m², *Plethodon jordani*), Petranka and Murray (2001) (0.15/m², *Plethodon jordani*) and Semlitch (1980; 0.52–0.81/m², *Plethodon glutinosus*). Likewise our estimates for small plethodontids were within the range reported by Jung et al. (2000; 2.8–18.5/

m², *Plethodon cinereus*) and Burton and Likens (1975; 0.16–2.72, *Plethodon cinereus*, based on day counts adjusted for capture probability). Our previous works confirms that there was considerable spatial and temporal variation among population parameters for terrestrial salamanders at our study sites in the southern Appalachians (Bailey et al., 2004a,b).

The choice of an appropriate population abundance estimator is study-specific and investigators must weigh logistical, economic, and biological factors when selecting a sample design. Study objectives and model assumptions must be considered when developing sampling protocols for estimating abundance, even relative abundance. We highly recommend the use of the robust design because of its flexibility to incorporate variation in capture probabilities, and estimate temporary emigration probabilities. The method's assumptions fit the biological processes believed to be occurring in most terrestrial salamander populations. The method is costly in terms of labor and time, but it may be possible to apply the method on a subset of sample sites using the dual sampling design described by Pollock et al. (2002). If intensive sampling is not possible at any sites, depletion methods show good potential to estimate available population size over short sample periods (Bruce, 1995; Salvadio, 2001), while closed- or open-population estimators over longer periods may be unbiased for superpopulation sizes, but only if temporary emigration is a random process.

Acknowledgments.—Funding for this research was provided by the Environmental Protection Agency DISPro Project, the U.S. Geological Survey, and the U.S. National Park Service. We thank the staff of GSMNP, especially K. Langdon for their logistic and administrative assistance. N. Haddad, J. Goodwin, D. Forester, and two anonymous reviewers made valuable suggestions to an earlier draft of the manuscript. Finally, we thank E. J. Hyde and our field crew members for their valuable field assistance.

LITERATURE CITED

- AKAIKE, H. 1973. Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Cazakil (eds.), Proceedings of the Second International Symposium for Information Theory, pp. 267–281. Akademiai Kiado, Budapest, Hungary.
- BAILEY, L. L. 2004. Evaluating elastomer marking and photo identification methods for terrestrial salamanders: marking effects and observer bias. *Herpetological Review* 3:38–41.
- BAILEY, L. L., T. R. SIMONS, AND K. H. POLLOCK. 2004a. Estimating detection probability parameters for plethodon salamanders using the robust capture-

- recapture design. *Journal of Wildlife Management* 68:1–14.
- . 2004b. Spatial and temporal variation in detection probability of plethodon salamanders using the robust capture-recapture design. *Journal of Wildlife Management* 68:14–24.
- BRUCE, R. C. 1995. The use of temporary removal sampling in a study of population dynamics of the salamander *Desmognathus monticola*. *Australian Journal of Ecology* 20:403–412.
- BURNHAM, K. P., AND D. R. ANDERSON. 2002. *Model Selection and Multimodel Inference*. Springer-Verlag, New York.
- BURNHAM, K. P., AND W. S. OVERTON. 1978. Estimation of size of a closed population when capture probabilities vary among animals. *Biometrika* 65:625–633.
- . 1979. Robust estimation of population-size when capture probabilities vary among animals. *Ecology* 60:927–936.
- BURTON, T. M., AND G. E. LIKENS. 1975. Salamander populations and biomass in the Hubbard Brook Experimental Forest, New Hampshire. *Copeia* 1975:541–546.
- CHAO, A. 1988. Estimating animal abundances with capture frequency data. *Journal of Wildlife Management* 52:292–300.
- CHAPMAN, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. *Univ. of California Publications in Statistics* 1:131–160.
- DAVIS, T. M., AND K. OVASKA. 2001. Individual recognition of amphibians: effects of toe clipping and fluorescent tagging on the salamander *Plethodon vehiculum*. *Journal of Herpetology* 35:217–225.
- DENARDO, D. 1995. Amphibians as laboratory animals. *ILAR Journal* 37:173–181.
- DODD, C. K. 2003. *Monitoring amphibians in Great Smoky Mountains National Park*. United States Geological Survey Circular 1258. Tallahassee, FL.
- HAIRSTON, N. G. 1983. Growth, survival and reproduction of *Plethodon jordani*: trade-offs between selective pressures. *Copeia* 1983:1024–1035.
- HEATWOLE, H. In Press. Quadrat sampling. In M. S. Foster (ed.), *Measuring and Monitoring Biological Diversity: Standard Methods for Reptiles*. Smithsonian Institution Press, Washington, DC.
- HEATWOLE, H., AND O. J. SEXTON. 1966. Herpetofaunal comparisons between two climatic zones in Panama. *American Midland Naturalist* 75:45–60.
- HEYER, W. R., M. A. DONNELLY, R. W. MCDIARMID, L.-A. C. HAYEK, AND M. S. FOSTER. 1994. *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*. Smithsonian Institution Press, Washington, DC.
- HOWARD, T. A. 1987. *Population and Biomass Estimates in Four Species of Terrestrial Plethodontid Salamanders*. Unpubl. master's thesis. Appalachian State Univ., Boone, NC.
- JAEGER, R. G., AND R. F. INGER. 1994. Quadrat sampling. In W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster (eds.), *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*, pp. 97–102. Smithsonian Institution Press, Washington, DC.
- JOLLY, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration—stochastic model. *Biometrika* 52:225–247.
- JUNG, R. E., S. DROEGE, J. R. SAUER, AND R. B. LANDY. 2000. Evaluation of terrestrial and streamside salamander monitoring techniques at Shenandoah National Park. *Environmental Monitoring and Assessment* 63:65–79.
- KENDALL, W. L. 1999. Robustness of closed capture-recapture methods to violations of the closure assumption. *Ecology* 80:2517–2525.
- KENDALL, W. L., AND J. E. HINES. 1999. An estimation tool for capture-recapture data collected under Pollock's robust design. *Bird Study* 46:S32–S38.
- KENDALL, W. L., AND J. D. NICHOLS. 1995. On the use of secondary capture-recapture samples to estimate temporary emigration and breeding proportions. *Journal of Applied Statistics* 22:751–762.
- KENDALL, W. L., AND K. H. POLLOCK. 1992. The robust design in capture-recapture studies: a review and evaluation by Monte Carlo simulation. In D. R. McCullough and R. H. Barrett (eds.), *Wildlife 2001: Populations*, pp. 31–43. Elsevier Science Publishers, London.
- KENDALL, W. L., J. D. NICHOLS, AND J. E. HINES. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* 78:563–578.
- LANCIA, R. A., J. D. NICHOLS, AND K. H. POLLOCK. 1994. Estimating the number of animals in wildlife populations. In T. A. Bookout (ed.), *Research and Management Techniques for Wildlife Habitats*. Vol. 5, pp. 215–253. Wildlife Society, Bethesda, MD.
- MACKENZIE, D. I., AND W. L. KENDALL. 2002. How should detection probabilities be incorporated into estimates of relative abundance. *Ecology* 83:2387–2393.
- MARVIN, G. A. 1996. Life history and population characteristics of the salamander *Plethodon kentucki* with a review of plethodon life histories. *American Midland Naturalist* 136:385–400.
- MENKENS, G. E., AND S. H. ANDERSON. 1988. Estimation of small-mammal population size. *Ecology* 69:1952–1959.
- NORRIS, J. L., AND K. H. POLLOCK. 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* 52:639–649.
- ORGAN, J. A. 1961. Studies of the local distribution, life history and population dynamics of the salamander genus *Desmognathus* in Virginia. *Ecological Monographs* 31:189–220.
- OTIS, D. L., K. P. BURNHAM, G. C. WHITE, AND D. R. ANDERSON. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62:1–135.
- PETRANKA, J. W. 1998. *Salamanders of the United States and Canada*. Smithsonian Institution Press, Washington, DC.
- PETRANKA, J. W., AND S. S. MURRAY. 2001. Effectiveness of removal sampling for determining salamander density and biomass: a case study in an Appalachian streamside community. *Journal of Herpetology* 35:36–44.
- PLEDGER, S. 2000. Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics* 56:434–442.

- POLLOCK, K. H. 1982. A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management* 46:757–760.
- POLLOCK, K. H., J. D. NICHOLS, C. BROWNIE, AND J. E. HINES. 1990. Statistical inference for capture-recapture experiments. *Wildlife Monographs* 107:1–97.
- 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.
- REXSTAD, E., AND K. P. BURNHAM. 1991. User's guide for interactive program CAPTURE. Colorado Cooperative Fish and Wildlife Research Unit, Fort Collins.
- RODDA, G. H., E. W. CAMPBELL, AND T. H. FRITTS. 2001. A high validity census technique for herpetofaunal assemblages. *Herpetological Review* 32:24–30.
- SALVIDIO, S. 1998. Estimating abundance and biomass of a *Speleomantes strinatii* (Caudata, Plethodontidae) population by temporary removal sampling. *Amphibia-Reptilia* 19:113–124.
- . 2001. Estimating terrestrial salamander abundance in different habitats: efficiency of temporary removal methods. *Herpetological Review* 32: 21–23.
- SCHMIDT, B. R. 2003. Count data, detection probabilities, and the demography, dynamics, distribution and decline of amphibians. *Comptes Rendus Biologies* 326:S119–S124.
- SEBER, G. A. F. 1965. A note on the multiple recapture census. *Biometrika* 52:249–259.
- . 1982. *The Estimation of Animal Abundance and Related Parameters*. Macmillan, New York.
- SEMLITSCH, R. D. 1980. Geographic and local variation in population parameters of the slimy salamander (*Plethodon glutinosus*). *Herpetologica* 36:6–16.
- SMITH, C. K., AND J. W. PETRANKA. 2000. Monitoring terrestrial salamanders: repeatability and validity of area-constrained cover object searches. *Journal of Herpetology* 34:547–557.
- STANLEY, T. R., AND K. P. BURNHAM. 1999. A closure test for time-specific capture-recapture data. *Environmental and Ecological Statistics* 6:197–209.
- STEWART, G. D., AND E. D. BELLIS. 1970. Dispersion patterns of salamanders along a brook. *Copeia* 1970:86–89.
- TILLEY, S. G. 1980. Life histories and comparative demography of two salamander populations. *Copeia* 1980:806–821.
- WELSH, H. H., AND S. DROEGE. 2001. A case for using plethodontid salamanders for monitoring biodiversity and ecosystem integrity of North American forests. *Conservation Biology* 15:558–569.
- WELSH, H. H., AND A. J. LIND. 1992. Population ecology of two relicual salamanders from the Klamath Mountains of Northwestern California. In D. R. McCullough and R. H. Barrett (eds.), *Wildlife 2001: Populations*, pp. 419–437. Elsevier Science Publishers, London.
- WHITE, G. C., AND K. P. BURNHAM. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* 46:120–138.
- WHITE, G. C., D. R. ANDERSON, K. P. BURNHAM, AND D. L. OTIS. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos, NM.
- ZIPFEN, C. 1956. An evaluation of the removal method of estimating animal populations. *Biometrics* 12:163–189.

Accepted: 16 May 2004.

APPENDIX 1

Closed-Population and Depletion Models.—These models assume the salamander population is closed over a designated time period: 30–40-day mark-recapture period for the full-season and Lincoln-Peterson models, eight days for the depletion models, and four days for the closed-population secondary samples models.

Open-Population Models.—Models considered estimate constant survival probability ϕ , and either time-specific (Model B: $\phi(i)$, $p(i)$) or time-constant capture probability (Model D: $\phi(\cdot)$, $p(\cdot)$). Capture probabilities are assumed to be equal for marked and previously unmarked individuals.

Robust Design Models.—Models considered estimate constant survival probability ϕ , and capture probabilities, which can vary among primary periods, i , or between previously marked, c , and unmarked, p , individuals. RD- M_0 model assumes time-constant and equal capture probabilities ($\phi(\cdot)$, $p(\cdot) = c(\cdot)$, $N(\cdot)$); RD- M_b model assumes time-constant but unequal capture probabilities of previously marked and unmarked individuals ($\phi(\cdot)$, $p(\cdot)$, $c(\cdot)$, $N(\cdot)$); and RD- M_t model assumes time-specific but equal capture probabilities ($\phi(\cdot)$, $p(i) = c(i)$, $N(\cdot)$). See Bailey et al. (2004a) for a more detailed robust design analysis.