FIGURE 1. Palomarin Field Station, Point Reyes National Seashore. Nest-searches and territory-mapping of color-banded birds took place in the four contiguous area marked by solid, rectilinear lines, totaling 36 ha. Constant effort mist-netting was conducted at nets marked with asterisks. Dotted lines enclose Wrentit territorial boundaries for a typical year (1985).

territories between a bird's territory and the nearest net, sex, age, and various measures of reproductive success (date of completion of first clutch, number of young hatched or fledged, the number of clutches or broods), using multiple logistic regression (Hosmer and Lemeshow 2000). Date of clutch completion was transformed to the square of the number of days since 21 March (two days prior to the earliest first-egg date in the sample).

To analyze survival probability and recapture probability (i.e., the probability a bird that has survived to year x is caught in year x), we used the statistical program SURGE (Lebreton et al. 1992, Cooch et al. 1996). All analyses were conducted on the mist-net capture–recapture data from 333 different individuals caught over 11 years, and the results compared with detailed observations on individually color-banded Wrentits (244 different individuals for a total of 523 breeder-years). We first analyzed all captures, stratifying

on territorial status, and then carried out analyses on capture data that pooled all adults.

Statistical analyses were carried out using STATA 5.0 (StataCorp 1997). Results give estimates ± SE, unless otherwise stated, and were considered significant if $P < 0.05$.

RESULTS

INFLUENCES ON CAPTURE PROBABILITY OF ADULTS

Territorial status

Most of the adult Wrentits caught in the study did not hold territories within the study area (Table 1). In general, there were about three times as many non-territory holders as territory holders (means = 26.4 and 8.4, respectively; Table 1), although the

TABLE 1. CAPTURE OF AHY WRENTITS OVER 10 YEARS IN RELATION TO TERRITORIAL STATUS

Year	Number of local breeders	Number of non-territory holders	Percent local breeders
1981	8	35	19
1982	8	33	20
1983	12	30	29
1984	10	9	53
1985	9	28	24
1986	7	47	13
1987	8	32	20
1988	10	35	22
1989	5	14	26
1990	6	12	33
1991	10	16	39

Notes: Local breeders were birds known to hold territories in the study area. Non-territory holders were birds that either did not breed, or bred off the study area.

proportion of territory holders was unusually high in 1984 (53%). The number of non-territory holders varied more markedly among years than did the number of territory holders (Table 1), but the ratio of territory holders to non-territory holders did not vary significantly between years (Likelihood Ratio Statistic [LRS] = 26.94, df = 10, P = 0.076). Results were quite similar when only breeding season captures were considered.

Territory holders and non-territory holders were caught throughout the netting season. The two groups did not differ in mean first capture date (24 May ± 35.3 days [SD] for territory holders vs. 29 May ± 29.6 days [SD] for non-territory holders; ANOVA, P > 0.15). In general, fewer adults were captured in July and August (whether territory holders or non-territory holders).

Territorial status influenced the number of times an individual was recaught in the same season (Table 2). Non-territory holders were usually caught only once during a given year (78%), whereas local breeders were usually caught multiple times (71% more than once; 53% three or more times in the same year).

Over the entire study period, 66% of territory holders were recaptured at least once, whereas 56% were recaught more than once and 31% were recaught six times or more. Only four territory holders (out of 59) had any breaks in their capture–recapture records (i.e., a year in which they were not caught, flanked by one or more years in which they were caught). By contrast, only 20% of non-territory holders were recaptured at least once, and only 5% more than once. The difference in number of total captures

for the two groups was highly significant (P < 0.001, Poisson regression).

Most (74%) non-territory holder birds were first caught in the winter or spring as after-hatching year individuals (i.e., they were neither locally fledged young nor caught in the nets in their first calendar year of life). In contrast, 52% of territory holders caught in nets were locally fledged young or were caught in nets in their first calendar year of life.

Capture probability of territory holders in relation to distance from nets

Over the 11-year study period, 523 breeders were identified on the study grid through intensive observations of color-banded individuals (the same individual was counted multiply if it bred in more than one year). Of these, 93 (17.8%) were captured in mist nets some time during the year, nearly all during the breeding season. By far the most important influence on capture probability was distance between the nest and the nearest mist net. All individuals breeding within 50 m of a net were caught (N = 40), while those breeding more than 200 m from the nearest net were rarely caught (0.8%, N = 389; Fig. 2). In between 50 and 200 m, the proportion of breeders caught in nets declined in a smooth fashion (Fig. 2), ranging from 82% caught among those breeding 50–75 m from a net, to 17% caught among those breeding 175–200 m from a net. The statistical significance of distance to the net in predicting capture of a known local breeder was very high (P < 0.001, logistic regression).
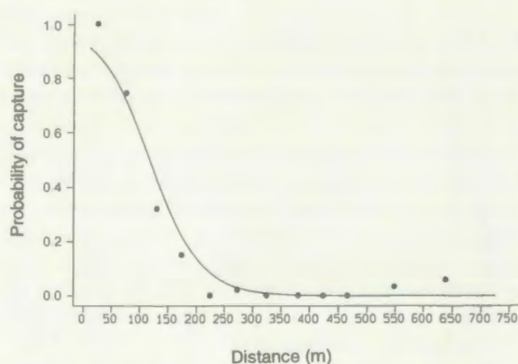

FIGURE 2. Capture probability of Wrentit breeders in relation to distance (m) from the nest to the nearest mist net, 1981–1991. Filled circles show proportion of breeders caught for breeders grouped in 50 m intervals: 0–50 m, 50–100 m, etc. Beyond 500 m, data are shown in 100 m intervals. Solid line gives the best fit to the data using logistic regression.

TABLE 2. FREQUENCY OF CAPTURE AND RECAPTURE OF WRENTITS WITHIN A YEAR, IN RELATION TO TERRITORIAL STATUS. (INCLUDES ONLY INDIVIDUALS CAUGHT AT LEAST ONCE DURING THE BREEDING SEASON)

Local breeders			Non-territory holders		
N times captured	Frequency[a]	Percent	N times captured	Frequency[a]	Percent
1	22	29	1	196	78
2	14	18	2	37	15
3	10	13	3	9	4
4	12	16	4	3	1.2
5	5	7	5	3	1.2
6	3	4	6	1	0.4
7	6	8	7	2	0.8
8–14	4	5	–	–	–
Total	78	100	Total	251	100

[a] Individuals were included more than once if caught in multiple years.

To examine whether territorial boundaries influenced capture probability, we compared proportion captured with respect to the number of territories a Wrentit had to traverse to reach the nearest net (Table 3). This analysis was restricted to birds breeding within 200 m of a net, because we showed above that Wrentits breeding at a greater distance from any net were rarely caught. Where a net was included within a Wrentit's territory, the Wrentit was almost always caught; conversely, Wrentits breeding more than two territories away were never caught (Table 3). Distance to nearest net and number of intervening territories to nearest net had independent and statistically significant effects on capture probability (P = 0.001 and P = 0.011, respectively).

Other factors influencing capture probability of breeders

Date of first clutch completion varied widely in the sample of breeders (minimum, median, and maximum first clutch completion dates were 23 March, 26 April, and 30 June, respectively). Earlier-breeding birds were less likely to be caught than those breeding in the middle or later in the breeding season (Table 4). However, for all breeders whose first clutch was completed from about 21 April on, capture probability was similar, at about 26%. First clutch completion date had a significant effect on capture probability when distance to nearest net was statistically controlled (P = 0.044).

Among breeders, there was a correlation between age and capture probability (one-year old individuals were more likely to be caught than older birds), but this relationship was not significant after controlling

TABLE 3. CAPTURE PROBABILITY IN RELATION TO TERRITORY LOCATION INCLUDING ONLY WRENTITS BREEDING WITHIN 200 M OF THE NEAREST MIST NET

Territory location[a]	Number of birds	Percent caught
0	56	96.4
0.5	24	62.5
1	15	6.7
1.5	7	14.3
2	12	25.0
2.5–3	4	0.0

[a] Coding for N territories: 0 = net was within Wrentit's territory; 0.5 = net was in territorial no-man's land (outside territorial boundary but not within neighbor's territory); 1, 2, 3 = net was one, two or three territories away; 1.5, 2.5 = as with 0.5, but an additional territory or two away.

for distance to nearest net (P > 0.1). Capture probability showed no significant association with the number of young hatched or fledged, the number of clutches or broods, or the sex of the breeder (P > 0.4 in each analysis).

SURVIVAL AND RECAPTURE PROBABILITY

Analyses stratified according to territorial status (territory holder vs. non-territory holder) resulted in estimated survival probabilities of 57% and 38%, respectively (Table 5). Recapture probability was estimated to be 71% for territory holders and 5% for those who were not. The difference in recapture probability between the two groups was significant (LRS = 14.69, P = 0.001), but the difference in survival probability was not (P > 0.3), due to lack of precision regarding the estimate of non-territory holder survival. Low precision was related to the fact that this category of individual was very unlikely to be recaptured the next year.

Annual survival of territory-holding birds caught in mist nets varied from 17–82%, and usually (7 out of 10 years) in a narrower range of 41–78%. Survival of territory holders did not vary significantly with age (LRS = 7.96, P > 0.5) or year (LRS = 8.26, P > 0.5). However, survival estimates showed a tendency

TABLE 4. EFFECT OF BREEDING DATE (DATE FIRST CLUTCH COMPLETED) ON CAPTURE PROBABILITY OF WRENTITS

Date 1st clutch completed	Number of breeders	% caught
Before 11 Apr	65	7.7
11–20 Apr	63	14.3
21–30 Apr	67	26.9
1–10 May	49	26.5
11–21 May	33	21.2
after 22 May	36	27.7

Note: Date categorized into 10-day intervals.

TABLE 5. RESULTS OF SURGE ANALYSIS ON MIST-NET CAPTURES OF WRENTITS, BY TERRITORIAL STATUS

| | Survival | | Recapture | |
	probability	95% CI	probability	95% CI
Local breeders	0.574	0.47 – 0.67	70.8	0.53 – 0.84
Non-territory holders	0.376	0.13 – 0.72	4.8	0.01 – 0.18

to increase with age, consistent with our observations on the color-banded population (Geupel and Ballard 2002).

Analysis of resightings of color-banded birds gave an estimated survival probability of 58.3 ± 2.9% and resighting probability of 91.5 ± 3.1% for females. For males, the estimates were 69.1 ± 2.4%) and 93.4 ± 1.9%. Survival based on resightings differed for the two sexes (likelihood ratio test, P = 0.004), but resighting probability did not. Mean adult survival (averaging values for males and females) based on resighting data was 63.7%, which was somewhat greater than the adult survival estimate obtained from capture data for territory-holding individuals (57.4%), but the confidence intervals of the two estimates overlapped. Thus, survival estimates based on capture–recapture analyses of territory holders caught in mist nets were consistent with those derived from sighting–resighting analyses of color-banded territory holders.

Most investigators running a constant-effort mist-netting program can not distinguish local breeders from non-territory holders. We therefore analyzed data for all mist-net-caught adults, pooling data from territory holders (59 different individuals) and non-territory holders (274 different individuals). The pooled analysis showed no significant variation with year or age, and gave a survival estimate of 30.6% (95% Confidence Interval of 22–41%), vs. 57% for local breeders alone. Recapture probability was estimated at 38.2% (95% CI of 23–56%), as opposed to 71% for known local breeders.

Even though analysis of capture–recapture data gives skewed estimates of survival when non-territory holders are included, it may still provide a reasonable index of annual survival. We investigated whether such an annual index could reliably predict annual survival, by comparing it with survival analyses based on resightings of color-banded birds. There was a trend for the two survival estimates to vary in the same direction (Fig. 3), but the correspondence between the two indices was not significant (R^2 = 0.252, P = 0.14, linear regression). The year 1986 was an outlier, yielding the highest survival estimate of the ten years according to resighting,
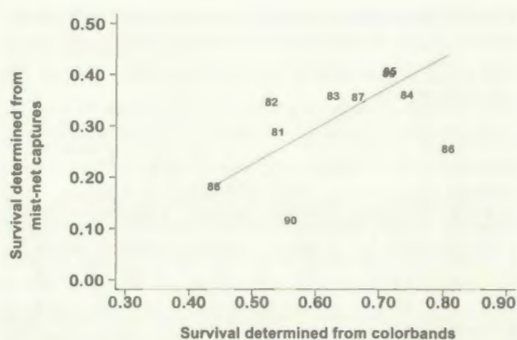

FIGURE 3. Comparison of Wrentit survival indices. Values for each year are shown by two digit codes. 1985 and 1989 are superimposed. On the x-axis is survival as estimated by resighting of color-banded individuals. On the y-axis is the SURGE estimate of survival using the pooled data (not differentiated by breeding status). Because 1986 was an aberrant year (see text), the best least-squares fit to the data excludes 1986.

but a relatively low estimate of survival according to capture–recapture data (third lowest). That year was aberrant in other respects (DeSante and Geupel 1987), and if 1986 was excluded, there was a significant correlation between the two survival indices (R^2 = 0.518, P = 0.029, linear regression).

DISCUSSION

The most important determinant of capture probability for adults in our study was distance from the net. A similar result was obtained for juveniles (see Nur et al. 1995), but the quantitative relationship between distance and capture probability differed for the two classes. For adults, few were captured that bred more than 200 m from the nearest net. Juveniles, however, were caught with a near-constant probability of ~14% beyond 300 m, up to a least 700 m. The catchment area for juveniles was likely more than a kilometer, maybe several. Thus, the populations being sampled by nets were very different for the two age classes. This has implications for the use of estimates of productivity derived by dividing the number of HY birds caught by the sum of AHY + HY captures (as is the practice of the Constant Effort Sites Scheme and MAPS program). This will not pose a serious problem if the numbers of HY birds within 200 m of nets (the area which samples adults) always fluctuate in parallel with numbers of HY birds further from the nets, but this may not be the case, and the subject deserves greater study.

Other than distance from the net, there appeared to be no important factors influencing capture probability of breeders, except that the earliest breeding birds were less likely to be caught. We have no explanation for this result. We speculate that seasonal differences in vegetation (and thus conspicuousness of the nets) may be responsible, but this needs to be examined directly.

The difference in year-to-year variability in number of breeders as opposed to number of transients caught reflected the greater constancy of capture among breeders, presumably because all those breeding close to the nets were caught in every year, whereas all those breeding some distance away were almost never caught. Annual fluctuations in the number of transients is discussed elsewhere (Nur et al. 2000), and is influenced by demographic processes such as last year's production of fledglings and breeding population size.

The most important result of the study was that survival derived from mist net capture–recapture data was underestimated unless local breeders and non-territory holders could be differentiated, due to an almost 18-fold difference between the two groups in recapture probability. If true non-breeders could be distinguished from those transient birds that bred off the study area, then at least non-breeding transients could be excluded from survival analyses. Unfortunately, Wrentit breeders and non-breeders cannot be distinguished in the hand, because both groups commonly display partial brood patches (PRBO, unpubl. data). The same problem is likely to apply to other species as well, such as those in which males do not develop brood patches. Even when the female brood patch is more highly developed among breeders than non-breeders in part of the breeding cycle (e.g., during incubation and the brooding phase), such differences are unlikely to persist throughout the three months or more that constant effort mist netting is conducted. Date of capture might provide some clues as to breeding status, but at least in the Palomarin Wrentit population, breeders and non-breeders cannot be distinguished by this means, and we expect this would also hold true for many other species.

One solution to the problem of differentiating local breeders and transients (whether the latter are breeders or non-breeders) would be to establish the identity of territory-holders within range of mist-nets through the use of unique color-bands or other markings, as in this study. For Wrentits, this identification need be done only within 200 m of the nets, but for other species a greater range would be prudent (perhaps 500 m or more, depending, in part, on territory size). Such an effort would be more time-consuming than the standard mist-netting protocol, but might be justified for a species of high concern.

A second, more expedient solution relies on our observation that non-territory holders were rarely recaptured within the same season, whereas territory holders were usually recaptured (Table 2). Survival could be estimated from only those individuals that had been recaptured in the same season. This would not eliminate the problem of transients, but should definitely reduce its magnitude. Data from some true breeders would be discarded, but at least in Wrentits, only 29% of breeders were not recaught at least once in the same year. An implication of this approach is that, in establishing a constant-effort mist-netting program, one goal would be to maximize the number of adults recaptured, as opposed to number of first captures. Running nets as many days per 10-day period as is feasible would further that goal, but would only be helpful if there was no net-avoidance. The fact that breeding Wrentits were caught so often in the same year, and usually in the breeding season, implies little net-avoidance in this species, even though these birds had ample opportunity to learn where nets were placed. Nets were in permanent locations, and operated at least 3 times/week (daily for more than 6 months of the year).

We applied the within-season recapture criterion to survival analyses of Wilson's Warbler (*Wilsonia pusilla*) capture–recapture data from the Palomarin Field Station (Chase et al. 1997). Individuals were classified as non-transient or transient on the basis of whether they were or were not caught two or more times in the breeding season, at least 7 days apart. Recapture probability for putative transients was only one-fifth that of non-transients (likelihood ratio test, $P < 0.001$). The survival estimate for all individuals pooled was 31%, whereas the estimate exclusive of putative transients was about 46%. True survival in this population was unknown, but is likely to be about 50%.

We have also analyzed data for the Song Sparrow (Nur et al. 2000), with similar results. Territory-holders and non-territory holders had very different recapture probabilities and pooling the two classes of adults resulted in low (biased) survival estimates, whereas distinguishing the two classes of individuals improved survival estimates. One difference between Palomarin Song Sparrows and Wrentits was that for the former, survival estimates for mist-net-caught, known territory holders were still substantially lower than survival as determined from analysis of resightings of color-banded breeders (47% vs. 60%, respectively; Nur et al. 2000). However, for the Song

Sparrow, the double-capture criterion (by which individuals caught twice in the same breeding season are considered non-transients) was very effective in yielding a survival estimate which matched the estimate obtained from capture–recapture analyses of color-banded individuals (both methods yielded estimates of 60% survival for males and females pooled). Thus, the use of the double-capture criterion was substantiated for the Song Sparrow, and that finding supports its use in analyses of Wilson's Warbler survival (Chase et al. 1997). Similar results were obtained by Peach (1993) for several European passerine species.

Even though Wrentit survival estimates were severely skewed when breeders and transients were not distinguished, there may still be value in a survival index based on year-by-year estimates for pooled data. We could not show a significant correlation between the mist-net survival indices and estimates based on individually marked birds, but there was reasonable correspondence between the two survival measures in most years. Any marked temporal trend in survival would probably be detected by the pooled mist-net survival index. We wish to point out, however, that mist-net studies may or may not be able to accurately assess differences in survival between sites. To our knowledge, no validation studies have been carried out to date on this topic.

Since this study, mark–recapture models have been developed to deal specifically with the effect of transients (Pradel et al. 1997). It would be valuable to analyze this data set (where territorial status of individuals is known, not inferred) using Pradel's model, to compare results with those based on color-band resighting data, and to analyze capture–recapture data for known local breeders only.

ACKNOWLEDGMENTS

ESTIMATING ADULT SURVIVAL RATES FROM BETWEEN-YEAR RECAPTURES IN THE BRITISH TRUST FOR ORNITHOLOGY CONSTANT EFFORT SITES SCHEME

WILL J. PEACH AND STEPHEN R BAILLIE

Abstract. Recent developments in the methodology for estimating survival rates from mark–recapture data are summarized. Transient individuals are common in mist-net samples and, unless catered for in the analysis, can cause survival rates to be seriously underestimated. Mark–recapture data from multiple study sites can now be combined analytically to provide regional estimates of survival. Although permanent emigration away from constant effort sites may result in true survival being underestimated, temporal changes in apparent survival may still be useful in highlighting the demographic mechanisms driving population changes.

Key Words: Constant Effort Sites scheme, mark–recapture, survival.

Many passerine species show strong fidelity to breeding sites in successive breeding seasons. This is generally true for long-distance migrants, as well as residents. Consequently, regular captures of marked breeding birds can be an effective means of generating between-year recaptures, and these can be used to estimate apparent survival rates of adult birds. When the sampling effort is standardized across breeding seasons (or at least measured), annual return rates can be estimated with greater precision and less bias.

Birds banded as chicks or as juveniles on constant effort sites have much lower recapture rates in subsequent years than do birds banded as adults (i.e., at least one year old). This is partly because young birds experience higher mortality rates than adult birds, but mainly because many young passerines make their first breeding attempt at sites away from their natal area (Greenwood 1980). For this reason we have not attempted to estimate first-year survival rates using Constant Effort Sites (CES) mark–recapture data (although see Peach et al. 1999). Survival rates of young passerines are probably best estimated using reports of banded birds found dead (Baillie and McCulloch 1993).

In recent years there has been a rapid growth of interest in the application of mark–recapture techniques to the estimation of demographic parameters, in respect to both extensive wildlife monitoring programs and intensive population studies. Comprehensive reviews of the methods and software available for analyzing mark–recapture data have been published (e.g., Pollock et al. 1990, Lebreton et al. 1992, Baillie and North 1999). We have applied these methods to data collected at constant effort sites (Peach et al. 1990, 1995; Peach

1993) and present here a summary of our main findings. For most analyses we recommend the use of program MARK (White and Burnham 1999) combined with program RELEASE (Burnham et al. 1987) for goodness-of-fit tests.

MODELING SURVIVAL RATES

The estimation of survival rates (strictly, return rates) involves the fitting of open population Jolly-Seber models parameterized in terms of survival and recapture rates (Pollock et al. 1990, Lebreton et al. 1992). An important starting point for many analyses is the Cormack-Jolly-Seber (CJS) model in which both survival and recapture rates are time-dependent (Cormack 1964, Jolly 1965, Seber 1965). Following the notation of Lebreton et al. (1992), models are referred to in terms of survival rate (ϕ) and recapture probability (p) with the subscript t denoting time-dependence. The CJS model is therefore referred to as (ϕ_t, p_t).

The CJS model assumes that survival does not vary according to the age of animal, and this is probably reasonable for adults of most short-lived small passerines (e.g., Buckland and Baillie 1987, although see Loery et al. 1987). The choice of a starting model will depend partly on biological knowledge or intuition, and it may be important to consider age-dependent survival models in relatively long-lived species. The goodness-of-fit tests provided by the software RELEASE, now available in the package MARK (White and Burnham 1999), provide explicit tests of the general suitability of the CJS model.

Having decided upon a biologically reasonable starting model that fits the data, the analyst can then

test a series of simpler models, each having fewer parameters than the starting model. For example, an obvious simplification of the CJS model is to constrain recapture probability to be constant over time (ϕ_t p). If the sampling effort has remained relatively constant during each of the sampling periods, as is the case at constant effort sites, then the simpler (ϕ_t p) model should provide a more parsimonious description of the data than the (ϕ_t p_t) model. Estimating fewer parameters from the same data increases the precision of the estimates, although at the risk of introducing bias. Likelihood ratio tests can be used to test specific hypotheses about model structure, and Akaike's Information Criterion (Akaike 1973) can be used to compare large numbers of candidate models without conducting large numbers of statistical tests (Lebreton et al. 1992, Burnham and Anderson 1998). The aim of the modelling procedure is to identify the simplest model that provides an adequate description of the data.

Program MARK has the useful facility of allowing the user to test for linear relationships between time-dependent model parameters and environmental variables (Lebreton et al. 1992, White and Burnham 1999). This can serve both as an aid to model simplification and as a means of testing biological hypotheses about factors affecting survival rates. This facility has, for example, allowed analysts to establish strong relationships between annual survival rates of two long-distance migratory bird species and annual rainfall in the African winter quarters (Kanyamibwa et al. 1990, Peach et al. 1991). It is better to test for relationships between time-dependent survival rates and external variables using the link functions available in MARK, rather than using ordinary least squares approaches, because the latter make no allowance for autocorrelation of successive survival estimates (Lebreton et al. 1992).

If capture effort has varied between sampling periods (e.g., years), then time-dependence in recapture probability can in principal be explained by some measure of capture effort, which can then be incorporated into the model (Clobert et al. 1987). However, in our experience simple measures of capture effort do not always correlate with or explain temporal variation in recapture rates, and we strongly advocate standardization of capture effort whenever this is possible.

THE PROBLEM OF TRANSIENTS IN COHORT SAMPLES

Estimates of survival between the year of capture and the first year of recapture may be biased downwards if cohort samples contain transient individuals

that are unlikely to be retrapped in subsequent years (Buckland 1982). One possible approach to this problem is to restrict analyses to individuals recaptured in years after their first year of capture (equivalent to excluding all first encounters). However, in short-lived species this often involves the loss of a high proportion of the available survival information with a consequent loss of precision (Peach et al. 1990).

Pradel et al. (1997) developed a new approach to the problem of transients that involves estimating the proportion of resident birds in banded samples, in addition to apparent survival and recapture rates. This method has the advantage of removing bias on survival estimates and of providing an estimate of the proportion of transient individuals in cohort samples. A disadvantage of Pradel's method is that for all cohorts the estimate of survival during the year after initial capture remains biased, and this may be a particular problem for short-lived species where a high proportion of all recaptures occur during the first recapture period.

An alternative approach is to use within-year recaptures to identify residents in newly banded cohorts (Buckland and Baillie 1987, Peach et al. 1990). Most constant-effort banding schemes involve repeated sampling of study areas throughout each breeding season, and transient birds should have a lower probability of being retrapped within the same season than resident individuals (note the probability of retrapping transients in subsequent years is, by definition, zero; Pradel et al. 1997). Thus, birds retrapped more than some specified minimum number of days after first being trapped are considered "residents" whereas individuals not retrapped over the same period are considered to be a mixture of "residents" and "transients." We have tended to use 6–10 days as the minimum period that must separate same-year captures of an individual during the first year in which it was encountered for it to be considered resident (Peach 1993). This information is included in the analysis by expanding the encounter histories for each bird to have a single additional encounter period immediately after the first encounter. Modelling then proceeds with a dummy "age" structure that partitions survival and recapture rates between first encounter and subsequent same-year recapture (within-year survival and recapture rates), from same-year recapture and the following breeding season (between-year survival and recapture rates). In this way, individuals not retrapped more than 6–10 days apart in the first year of capture but retrapped in subsequent years, are correctly classified as residents and do contribute to subsequent

estimates of annual survival. Within-year survival and recapture rates can be modelled as constant across years or year-specific. These "within-year" transients models are straightforward to fit using program MARK.

The effect of the "within-year recapture" transients models is to substantially reduce, but not remove, the negative bias on apparent survival estimates caused by the presence of transient individuals in cohort samples (Peach et al. 1990, Peach 1993). However, our approach maximizes the precision of survival estimates by fully using recapture information from the breeding season following initial capture, which often constitute a high proportion of all between-year recaptures for short-lived species. Using within-years recaptures to minimize problems caused by transients will be most appropriate where precision may be limiting statistical inference and small amounts of bias can be tolerated, as is often the case in studies attempting to detect temporal changes in survival rates. The approach is less suited to studies whose primary focus is absolute levels of survival, such as comparative life histories.

COMBINING SURVIVAL INFORMATION ACROSS STUDY SITES

An analytical development of particular importance to the CES Scheme was the capability of programs like SURGE and MARK to handle multiple groups of marked animals within a single analysis (Pradel et al. 1990, White and Burnham 1999). This facility allows the analyst to test for differences in survival or recapture probabilities between groups, which in the CES-context might include study site or sex. The general approach is to fit starting models in which parameters differ between groups, and then to constrain parameters to be identical or even additive across groups (Pradel et al. 1990, Lebreton et al. 1992). In the CES context, this allows the analyst to check whether apparent survival rates differ between sites and, if they do not, to pool survival information across sites to provide more precise regional survival estimates.

The modelling framework for multiple-site analyses is analogous to that used in analysis of variance. An approach that we have adopted in the analysis of CES data is to have a starting model in which survival and recapture probabilities are both year- and site-specific, with an interaction between year and site (Peach 1993). We then attempt to simplify recapture probability, initially by dropping the year-site interaction term, and then by removing the time-dependence. These simplifications are usually parsimonious because of the constant sampling effort maintained at CE sites. Recapture probabilities often differ significantly between sites, which probably reflects the differing numbers and densities of nets at different sites. We then attempt to simplify the survival side of the model, first by seeking to remove the interaction term, and then by removing the site term. We have used this modelling approach to combine mark–recapture data from up to 10 CE sites to provide regional estimates of annual adult survival rates for the Willow Warbler (Peach et al. 1995; scientific names in Table 1).

We compared estimates of adult survival for five passerine species derived from multiple-site CES mark–recapture data (Table 1) with independent estimates from national BTO band recovery data relating to birds found dead (taken from Peach 1993 and Baillie and McCulloch 1993). Survival estimates from recaptures were generally lower than those based on recovery data (Table 1). Although these differences could be a consequence of the differing time periods covered by the two sets of analyses, they are probably partly caused by the permanent emigration of some birds away from constant effort sites (Peach et al. 1990, Cilimburg et al. 2002). Even if apparent survival rates estimated from mark–recapture data do underestimate true survival, they may still constitute a useful index of temporal changes in true survival rates. Moreover, the precision of the survival rates of small passerines estimated from the CES data is generally comparable to or better than that attainable from the national United Kingdom band recovery data, because these species have low band reporting rates (Table 1).

DISCUSSION

Despite the potential problems of negative bias affecting apparent survival estimates, recent developments in both theory and software for analysing mark–recapture data make this aspect of the CES data an exciting prospect for the future. The main application of the CES mark–recapture data will be in the detection of long-term temporal trends in the apparent survival rates of adult passerines, and the testing of relationships between survival and environmental variables such as rainfall in the winter quarters of migrants (e.g., Peach et al. 1991). Knowledge of apparent survival rates is likely to be an important factor affecting the population dynamics of small passerines (Baillie and Peach 1992), and may be critical to our understanding of the mechanisms leading to wide-scale population changes

TABLE 1. ESTIMATES OF AVERAGE ANNUAL ADULT SURVIVAL RATES (AND ASSOCIATED STANDARD ERRORS) DERIVED FROM POOLED MARK–RECAPTURE DATA FROM MULTIPLE CONSTANT-EFFORT BANDING SITES AND FROM RECOVERIES OF DEAD BIRDS BANDED IN BRITAIN AND IRELAND

| | Mark–recapture analyses (1983–1991) | | | | Recovery analyses (1985–1990) | | |
| | | | Survival rate | | | Survival rate | |
Species	Number of combined	Number of individuals retrapped	ϕ	SE	Number of recoveries	S	SE
Willow Warbler (*Phylloscopus trochilus*)	7	183	0.371	0.025	385[a]	0.554	0.056
Blackbird (*Turdus merula*)	4	113	0.566	0.036	1307	0.668	0.020
	4	165	0.581	0.029			
Blackcap (*Sylvia atricapilla*)	3	51	0.443	0.057	197	0.534	0.128
Reed Warbler (*Acrocephalus scirpaceus*)	3	168	0.496	0.026	957[b]	0.558	0.023
Wren (*Troglodytes troglodytes*)	2	29	0.318	0.068	–	–	–
Dunnock (*Prunella modularis*)	1	62	0.422	0.040	265	0.447	0.050

Notes: All survival estimates were derived from time-independent models that fit the data. CE survival estimates are from Peach (1993) and estimates from recoveries are from Baillie and McCulloch (1993).

[a] Estimated from recaptures >5 km from the place of banding.

[b] Estimates based on recoveries of dead birds were supplemented by recaptures >5 km from the place of banding.

(Peach et al. 1999). Other potential applications are the comparisons of apparent annual survival rates between sites (perhaps treatment and control sites), between males and females (e.g., Pratt and Peach 1991), and between different habitats, regions (e.g., Peach et al. 1995), or latitudes.

ACKNOWLEDGMENTS

A EUROPEAN EXAMPLE OF STANDARDIZED MIST NETTING IN POPULATION STUDIES OF BIRDS

ANDREAS KAISER AND PETER BERTHOLD

Abstract. The "MRI-program" is a standardized long-term bird trapping program that has been in existence since 1974. Three central European stations are run daily during the entire autumn migratory period from June through November. Three other stations follow the same highly standardized protocol. In this paper, the field methods are described and standardization is discussed. Advantages of standardization include improved accuracy of capture–recapture estimates of population size and other parameters.

Key Words: capture–recapture, migrant, mist net, monitoring, MRI-program, passerine, standardization.

The standardized study of many different species and populations of birds at the same time, over broad geographic scales, offers valuable opportunities to monitor bird populations and at the same time study factors affecting population dynamics. Two examples of such projects that involve mist netting to capture birds are migration and stopover studies (Bairlein 1998, Bairlein and Giessing 1997, Bairlein et al. 1994), and productivity and survival studies (DeSante 1992, DeSante et al. *this volume*, Peach and Baillie *this volume*). Each of these programs uses highly standardized methods, both to reduce bias in sampling and to facilitate strong statistical analysis. Another example, described here, is the "MRI-program," which currently consists of up to six trapping sites in operation during fall (Fig. 1).

Long-term research programs were begun at three inland stations: the Mettnau peninsula in south Germany, the nature reserve "Die Reit" in north Germany near Hamburg, and in east Austria in the nature reserve at the eastern shore of Lake Neusiedl near Illmitz. Preliminary work was done in 1972 and 1973, and these sites have been run under standard conditions since 1974. Later additions included a banding site at lake Galenbeck in northeastern Germany, and two coastal sites, the Ebro-Delta banding site in Spain and Rybachy at the Kurish Split in Russia (Fig. 1). The latter two sites collaborate closely with the Vogelwarte Radolfzell.

Sites were chosen according to four criteria: (1) at least one site should sample each of the autumn migratory populations of central, northern, western, or eastern Europe, as shown by the atlas of songbird migration (Zink 1973–1985); (2) the stations should be situated in protected areas that would not be disturbed during long-term studies; (3) the areas should have a high degree of climax vegetation and thus show relatively few changes over the long term; and

(4) the areas should be excellent bird conservation areas with rich bird life during the breeding season as well as the migration period. In addition to these considerations, the suitability of the areas was tested by sample trapping during the pilot years.

The program was designed so that a number of questions could be answered, including five main topics:

(1) *Population dynamics and demography*: Short-term and medium-term fluctuations in numbers of migrants, as well as long-term population

FIGURE 1. Banding sites of the Mettnau-Reit-Illmitz-program in Europe and sites in cooperation with the "Vogelwarte Radolfzell" (German bird-banding office). M=Mettnau (Lake Constance), RE=Reit (Hamburg), I=Illmitz (Lake Neusiedl, Austria), G=Galenbeck (Lake Galenbeck), RY=Rybachy (Rossitten, Russia), E=Ebro-Delta (Tarragona, Spain).

changes and their magnitudes, with a special focus on decline of small birds (Marchant 1992, Berthold et al. 1993, Böhning-Gaese 1995, Kaiser and Berthold 1995). Demographic studies were to look at age and sex differences and their role in migratory and stopover behavior, habitat preference, nutrition, and many other topics.

(2) *Migration*: Phenology of migration, migration routes, and strategies of migration and stopover (Berthold 1996, 2001). Also studied are the dependence of these features on sex, age, the breeding area and range of the populations, and seasonal and climatic factors. Finally, questions are investigated on migratory physiology, such as fat deposition, the control of migration, stopover behavior, and the interplay of molt, migration, and energy balance (Berthold et al. 1991; Kaiser 1992, 1993b, 1996).

(3) *Biorhythmicity*: Special attention is given to daily activity patterns of staging individuals, molt (Kasparek 1981), and to the variation in migration patterns from year to year (Bairlein 1981, Brensing 1989).

(4) *Ecosystem research*: Resource partitioning and utilization of stopover sites are of interest, including the role of habitat (Streif 1991), nutritional preferences (Brensing 1977, Grosch 1995), mobility (Bastian 1992), stopover period (Kaiser 1993b), population size and turnover (Kaiser 1995), and competition. Other studies investigate the carrying capacity of a stopover area for small birds and how such an area can be made optimal.

(5) *Methodological research:* Repeatability and observer bias in wing length measurements, fat scoring, and ageing techniques were studied (Berthold and Friedrich 1979; Kaiser 1993a, 1993b). Capture–recapture and other counting methods look at frequency and intensity of operations of the nets required to gain an adequate sample size.

In this paper, we discuss features of the MRI-program that are particularly relevant to population monitoring.

METHODS

The trapping site at Mettnau is typical of the operation of a single large-scale netting station in the MRI-program, and is described as an example. This site is an area of approximately 1 km², situated on the Mettnau Peninsula nature reserve east of Radolfzell at Lake Constance (Berthold et al. 1991). There are 52 mist nets in use, placed in a single transect through a *Phragmites* reed swamp, but sampling all habitat types characteristic of the peninsula (Streif 1991). Distance between nets and release (banding) site range from 55 to 360 m. Operations are run daily through the fall season (30 June–6 November). Nets are open 24 h. Nets are

checked at fixed intervals over the entire day (hourly, except half-hourly in poor weather). No activity is permitted near nets between net checks, and all captures are passive (no chasing or tape lures).

As noted by Bibby et al. (1992), standardization in capture and census methods is needed to reduce bias, and all MRI procedures are highly standardized (Berthold and Schlenker 1975). The number of nets, net locations, hours of operation, timing of net rounds, sequence of checking nets, and height of shelf strings on each net pole are all constant from year to year. Also standardized are all instructions and materials (bands, color rings, balance, tools, rulers). Vegetation is cut back in the off-season to keep habitat and vegetation structure as stable as possible.

Data recorded in the MRI-program include the following:

(1) *Trapping status*: first capture, within-site retrap from the same season, retrap from previous years, or banded elsewhere in the same or an earlier season (foreign retrap). Retraps are handled like first traps except that retraps from the same season do not have wing and foot remeasured, and for same-day retraps, sex, age, and molt are skipped.

(2) *Band number*

(3) *Date*

(4) *Capture time*: time when the net was checked and the bird removed from net.

(5) *Program status*: indicates whether species are study targets (full data collected) or non-target (full data collected only if there is time). Up to 41 species are targets at each site, whereas there are up to 100 non-target species.

(6) *Species-code*: German or Latin abbreviation, or species number

(7) *Net, shelf, and side of net in which the bird was trapped*: net shelves counted from ground upwards, 1 to 4. Left or right of nets are marked by signs at each nets. Data are recorded on a slip of paper placed into the carrying bag for each bird.

(8) *Sex*: recorded only when accurately determinable; otherwise coded as undetermined

(9) *Age*: two age classes are defined: this-year birds (juveniles, yearlings) and adults (older birds, born in the previous calendar year or earlier). Age is recorded only if accurately determinable, for example, by skull pneumatization or by molt limits in the wing (Jenni and Winkler 1994); otherwise coded as unknown age.

(10) *Molt*: body molt is recorded using methods described by Berthold et al. (1970), whereas wing feather molt follows Berthold et al. (1991). Tail feather molt is not recorded.

(11) *Length of the third primary*: length of third wing feather (counting from the outside) gives a relative wing measure that is more convenient to measure than wing chord (Berthold and Friedrich 1979, Svensson 1992).

(12) *Special data for species identification*: notch of the second primary and foot span are measured, to allow discrimination of *Acrocephalus* species.

(13) *Fat class*: using methods of Kaiser (1993a).

(14) *Body mass*: weighed to the nearest 0.1 g within 1 h of capture, using an electronic balance.

RESULTS AND DISCUSSION

Here we discuss some results from the MRI-program that are relevant to the use of mist nets for monitoring, and that illustrate the value of standardization.

STANDARDIZATION

Our results have shown that different species, and different numbers of each species, are caught in different habitats (Bairlein 1981, Streif 1991, Mädlow 1994). Therefore, moving or changing the total number of nets within or between seasons will alter numbers captured and affect annual indices of abundance. In capture–recapture studies, more birds may be recaptured if nets are relocated frequently (see below), but this would alter the probability of capture and recapture in complex ways that would be very difficult to model in analyses. Only in standardized capture–recapture studies are basic model assumptions met and resulting estimates precise (Otis

et al. 1978). We therefore recommend that a station should run with the same number of nets in exactly the same positions each year. For the same reason, it is important to prevent habitat change at the net sites, because habitat change affects capture–recapture probabilities in a manner analogous to moving nets among habitats.

With standard net locations, some species will have low capture probability because relatively little of their specialized habitat is sampled (e.g., Lesser Whitethroat, *Sylvia curruca*; Kaiser 1993b). It is therefore important to determine which species are the targets of study before determining where nets should be placed.

At the main MRI study site, the frequency of all first captures differed among habitats, but the proportions were fairly constant from year to year over a 22-year period (Fig. 2). However, capture indices decreased slightly in the four bushy habitats and increased in reed habitat C. To examine the effect of habitat change, we calculated species-specific long-term population trends separately for the birds
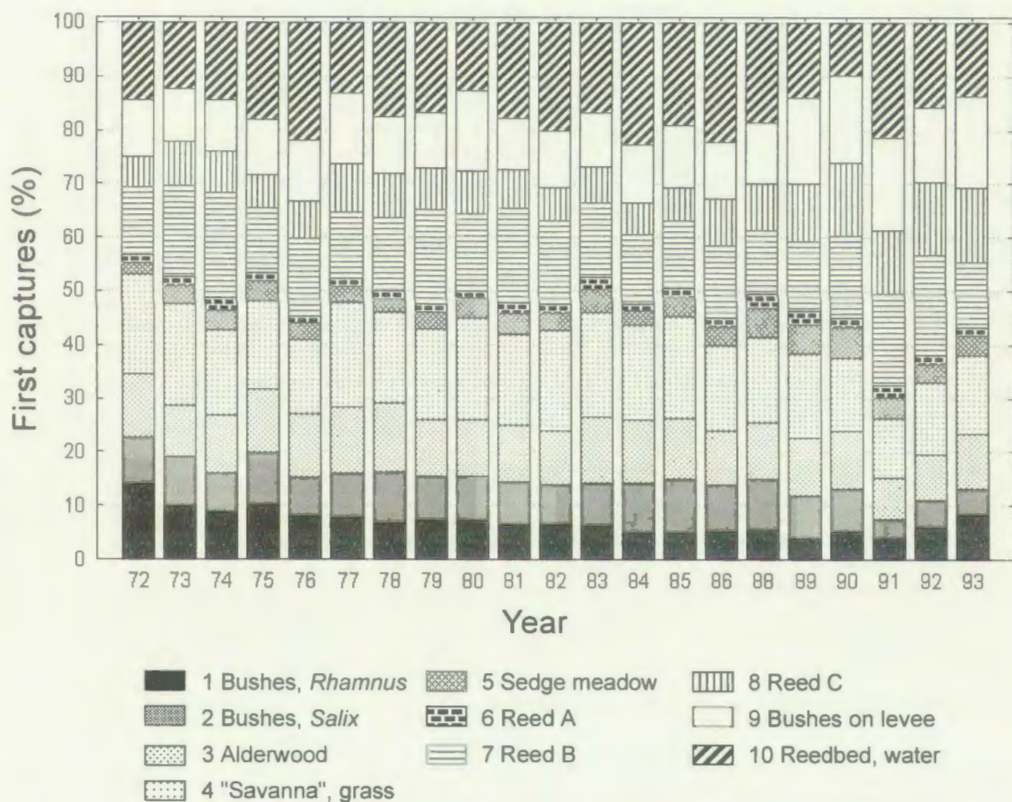

FIGURE 2. Percent of birds captured during June to November at Mettnau in different habitats and years. 1987 missing due to flood.

TABLE 1. LONG-TERM POPULATION TRENDS (1972–1993) FOR FIRST-CAPTURES IN DIFFERENT HABITATS AT METTNAU, GERMANY (KAISER AND BERTHOLD 1995)

Species	Habitat										All habitats	N
	1	2	3	4	5	6	7	8	9	10		
Wryneck (*Jynx torquilla*)	–	ns	–	ns	–	–	–	ns	ns	–	-0.65**	153
Wren (*Troglodytes troglodytes*)	-0.51*	-0.46*	ns	ns	ns	–	ns	ns	ns	-0.49*	ns	1,280
Dunnock (*Prunella modularis*)	ns	ns	-0.50*	ns	ns	–	ns	0.70*	0.62*	ns	ns	1,331
Robin (*Erithacus rubecula*)	-0.62*	ns	ns	0.56*	0.54*	0.60*	0.58*	ns	ns	ns	ns	9,619
Nightingale (*Luscinia megarhynchos*)	ns	ns	ns	ns	–	–	–	–	ns	–	0.52*	219
Bluethroat (*L. svecica*)	–	–	–	–	ns	ns	ns	ns	–	ns	-0.44*	177
Black Redstart (*Phoenicurus ochruros*)	ns	ns	0.45*	ns	ns	–	ns	ns	0.48*	ns	ns	516
Redstart (*P. phoenicurus*)	-0.67*	-0.60*	-0.70**	-0.78**	ns	–	ns	ns	-0.56*	-0.54*	-0.77**	939
Whinchat (*Saxicola rubetra*)	–	–	–	–	-0.64*	ns	ns	ns	–	ns	-0.47*	277
Blackbird (*Turdus merula*)	ns	ns	ns	ns	ns	–	ns	-0.64*	ns	ns	ns	2,379
Song Thrush (*T. philomelos*)	0.46*	ns	ns	ns	ns	–	ns	ns	ns	ns	ns	1,654
Grasshopper Warbler (*Locustella naevia*)	-0.52*	-0.47*	ns	ns	ns	ns	-0.59*	ns	ns	ns	-0.69**	1,087
Savi's Warbler (*L. luscinioides*)	–	–	–	–	–	–	ns	ns	–	ns	ns	144
Aquatic Warbler (*Acrocephalus paludicola*)	–	–	–	–	ns	–	ns	–	–	–	-0.72**	49
Sedge Warbler (*A. schoenobaenus*)	–	–	–	–	ns	ns	-0.66*	ns	ns	ns	-0.49*	1,333
Marsh Warbler (*A. palustris*)	0.89*	ns	-0.80**	-0.54*	ns	ns	-0.84**	ns	-0.49*	-0.56*	-0.77**	1,435
Reed Warbler (*A. scirpaceus*)	-0.73**	-0.58*	-0.51*	-0.72*	ns	0.49*	-0.45*	0.74**	ns	ns	ns	39,309
Great Reed Warbler (*A. arundinaceus*)	–	–	–	–	–	–	-0.65*	ns	ns	-0.64*	-0.75**	329
Icterine Warbler (*Hippolais icterina*)	-0.60*	ns	ns	ns	–	–	ns	–	ns	–	ns	580
Lesser Whitethroat (*Sylvia curruca*)	-0.65*	-0.55*	-0.60*	-0.75**	–	–	ns	ns	-0.47*	ns	-0.79**	2,235
Whitethroat (*S. communis*)	ns	ns	ns	ns	–	ns	ns	ns	ns	ns	ns	456
Garden Warbler (*S. borin*)	-0.68**	-0.56**	-0.46*	ns	0.49*	–	0.67**	ns	ns	0.49*	ns	9,579
Blackcap (*S. atricapilla*)	-0.45*	ns	ns	ns	0.86**	ns	0.70**	0.60*	0.71**	0.50*	0.43*	13,615
Wood Warbler (*Phylloscopus sibilatrix*)	–	–	–	ns	–	–	–	–	ns	–	-0.62**	72
Chiffchaff (*P. collybita*)	-0.72**	ns	-0.43*	-0.60*	ns	ns	0.46*	ns	ns	ns	ns	17,608
Willow Warbler (*P. trochilus*)	-0.86**	-0.82**	-0.64*	-0.83**	ns	ns	ns	ns	ns	ns	-0.81**	7,419
Goldcrest (*Regulus regulus*)	ns	0.55*	ns	0.53*	–	–	–	–	–	–	ns	451
Firecrest (*R. ignicapillus*)	ns	ns	ns	ns	–	–	–	–	–	–	ns	184
Spotted Flycatcher (*Muscicapa striata*)	ns	ns	ns	-0.60*	ns	–	ns	ns	ns	ns	-0.60**	807
Pied Flycatcher (*Ficedula hypoleuca*)	ns	ns	ns	ns	ns	–	–	–	ns	–	ns	656
Blue Tit (*Parus caeruleus*)	-0.60**	ns	ns	-0.52**	ns	–	ns	ns	ns	ns	-0.44*	5,563
Red-backed Shrike (*Lanius collurio*)	–	–	–	ns	–	–	ns	ns	ns	–	-0.49*	153
Goldfinch (*Carduelis carduelis*)	-0.66*	ns	ns	ns	–	–	–	ns	ns	–	ns	481
Bullfinch (*Pyrrhula pyrrhula*)	-0.56*	ns	ns	0.52*	–	–	–	–	–	–	ns	999
Reed Bunting (*Emberiza schoeniclus*)	ns	ns	ns	-0.52*	-0.45*	ns	-0.48*	ns	ns	ns	ns	7,960
All species combined	-0.76**	-0.60*	-0.57*	-0.55*	ns	ns	ns	0.50*	ns	ns		130,478

Notes: Trends are coefficients of annual capture totals linear regressed on year. See Fig. 2 for definition of habitats.

* denotes P < 0.05; ** denotes P < 0.001; ns denotes P > 0.05.

captured in each habitat. Differences in trend among habitats would suggest that habitat change has been taking place over time. For the most part, the trends were very consistent within species among habitats (Table 1). However, in the habitat with dense bushes of buckthorn (habitat 1), the Garden Warbler (*Sylvia borin*), Blackcap (*S. atricapilla*), Robin (*Erithacus rubecula*), and Bullfinch (*Pyrrhula pyrrhula*) were decreasing and the Marsh Warbler (*Acrocephalus palustris*) was increasing, whereas in other habitats population trends of these species were in opposite directions. The Chiffchaff (*Phylloscopus collybita*) showed negative trends in some habitats, but a positive trend in reed B (Table 1). Other species were also captured in remarkably high numbers in later years in reed habitats, and this may be related to an increase in the number of buckthorn bushes within the reed. These results illustrate the importance of maintaining habitat at the same stage over time.

Another possible reason for change in the numbers of birds captured in each habitat could be changes in food abundance, such as fruit patterns related to the height of mist nets or outbreaks of insects in particular habitat types. This kind of variation cannot be controlled with habitat management, but food abundance is not expected to change in a systematic way over time, so long-term trends should be unbiased by this variation.

Timing of operations should be standardized, as well as number and location of nets. Data collected both during migration (Brensing 1989) and during the breeding period (A. Kaiser, unpubl. data) show a strong peak in the number of captures early in the morning, and a second (much lower) peak before dusk. Equal net-hours each day are not equivalent, therefore, unless those net-hours are from the same portion of each day (Karr 1981a). Expressing total number of birds captured as birds/net-h is therefore an ineffective way of controlling for variation in effort, and the schedule of netting operations should instead be standardized.

RELATIVE ABUNDANCE

To test the efficiency and accuracy of mist nets for species inventory and estimates of relative abundance, we compared mist-net counts with different counting methods during the main breeding period from May to July. During this period, population size of adults can be assumed to be relatively constant. At an isolated study plot in south Germany near Espasingen we used a net density of 35-m net/ha in a 9-ha site (and 45-m net/ha in a nearby site of 3 ha), and achieved high capture (and recapture)

probabilities. The correlation between number of all species of breeding birds detected by mist-net captures (first captures only) and point counts was strongly positive ($r = 0.83$, $P < 0.001$), but netting totals were nearly always higher than point count totals (Fig. 3; Kaiser and Bauer 1994). The study suggested that netting can be used to sample a consistent percent of a population (although that percent may differ widely among species). Mist-net captures may therefore be a particularly good means of sampling migrants, because it takes place over many hours (unlike transect or point counts) and does not require birds to be singing for them to be detected.

POPULATION TRENDS

The length of a long-term population monitoring project should be at least 15–20 years to cover natural population fluctuations (Berthold and Querner 1978, Tucker and Heath 1994). Analyses of first capture data from the MRI-program for long term trends have been published regularly (Berthold et al. 1993, Kaiser and Berthold 1995, Berthold *this volume*). Böhning-Gaese (1995) determined that species with similar year-to-year population fluctuations do not necessarily have similar long-term trends. Moreover, results of small-scale study on migration season population trends cannot be taken to represent population change on larger spatial scales in the absence of information on which breeding population is being sampled at the migration station (see Dunn and Hussell 1995).
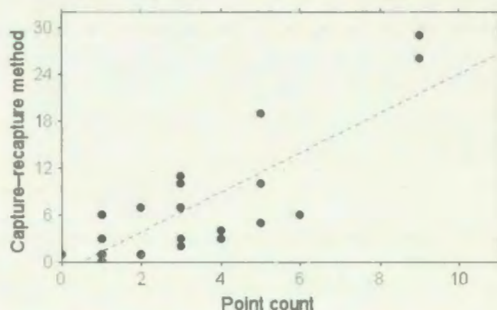

FIGURE 3. Number of local breeders captured (calculated from the number of adult first captures divided by 2 for an estimate of "pairs") compared to point count estimates at a woodland near Espasingen, Germany during the breeding period 1992 (Kaiser and Bauer 1994). Each point represents one species. Linear regression, $r = 0.83$, $R^2 = 0.70$, $P < 0.001$, $N = 29$.

CAPTURE–RECAPTURE STUDIES: BREEDING SEASON

Capture–recapture data are affected by net avoidance by birds that have already been captured once (Kaiser 1995). Recapture rate is generally much lower than expected when trapping is frequent (Buckland and Hereward 1982), although some species do not change their behavior drastically after the first catch. The extent of bias can sometimes be tested using mathematical models. We suggest two types of behavioral response to mist netting: (1) if many nets are used in comparison to the size of the study site, most birds learn to avoid the nets; and (2) intensive netting can cause too much direct human disturbance, causing birds to leave the area. These predictions have to be tested further, for example, in combined capture–recapture and telemetry studies.

In the breeding season, leaving up to 6 days between netting sessions increased capture and recapture rates (Dorsch 1998). One strategy for reducing net avoidance (other than reducing netting frequency) is to change net locations, but this compromises standardization (see above). Despite the problem of net avoidance, the MRI-program continues with daily netting in fixed locations, in part because net avoidance is a smaller problem with migrating birds (see below), and because our main objective is to analyze patterns of first captures under standard conditions.

Mist-net samples do not capture all the birds present, and capture–recapture models can be used to determine total population size. For example, in a study of a Reed Warbler (*Acrocephalus scirpaceus*) population at Lake Galenbeck, 254 adult Reed Warblers were caught at least once, with a total of 106 retraps (Fig. 4). Program CAPTURE (Otis et al. 1978) was used to estimate population size. The appropriate time effects and behavioral response model (White et al. 1982) estimated a population size of 500 birds, and the average estimate of all models was 430 (Fig. 5).

CAPTURE–RECAPTURE STUDIES: MIGRATION SEASON

Population size estimates during the migration period are more difficult to calculate than for breeding populations, because a set of well-defined assumptions of models for open population are violated and recapture numbers are not high (Kaiser 1995). To optimize sampling, density and distribution of nets is important. To obtain more recaptures, their density and distribution has to be adapted to the behavior of passerines stopping over. The interaction between capture behavior, recapture probability, disturbance,
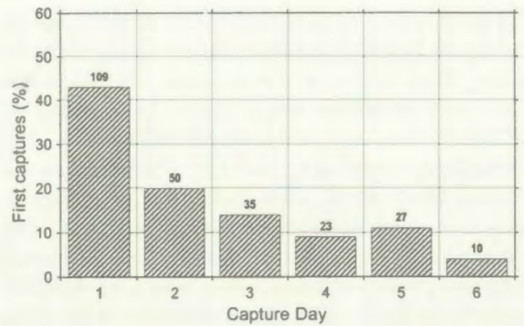

FIGURE 4. Decline in number of first captured Reed Warblers during the first 6 d of the MRI-program, i.e., at the end of the breeding period, at Lake Galenbeck, Germany.

and other biases (Pollock et al. 1990) was discussed by Kaiser (1993b, 1995).

During migration seasons, there is high turnover in individuals present (as shown by the low proportion of retraps), so number of first-time captures is increased by daily netting, and there are few birds stopping over that will develop net shyness (Kaiser 1993b). Nonetheless, Dorsch (1998) has shown that net avoidance may also be an issue with birds that are spending many days at a stopover site. Recapture probabilities during migration must be especially high (>0.2) to estimate other parameters, such as body mass change in relation to capture behavior. At some sites this is feasible, as shown by the 36% retrap rate obtained during 1988–1989 at the Mettnau Peninsula (Kaiser 1995).
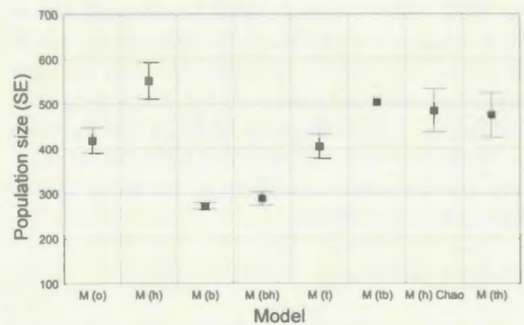

FIGURE 5. Number of breeding Reed Warblers estimated with different models of program CAPTURE (Otis et al. 1978). Capture probabilities are constant in model M_o, or vary by time (M_t), due to behavioral response (M_b), by individual birds (M_h), or by two sources of variation in its capture probabilities (M_{tb}, M_{th}, M_{bh}). Data from Lake Galenbeck, Germany, 1991 and 1992. Point estimates (means) with standard error. Number of first captures was 254, and mean populations size of all models 430.

Mobility of stopover populations was studied by examining the exchange rate of individual birds captured at five banding sites at the Mettnau peninsula during the migration period (Kaiser 1995). With knowledge of the exchange rate, an estimate of the size of the stopover population in the isolated nature reserve was derived from Jolly-Seber estimates. True average stopover time was estimated at 16 days, and it was shown that there were temporal behavioral responses to mist netting and ringing. Nonetheless, variation in capture probability was detected in birds according to differences in body condition, molt, mobility, and behavioral response to mist netting (Kaiser 1993b, 1995). The release of birds at the processing site, up to 500 m away from the trapping site, might affect retrap probabilities by causing the bird to shift its center of activity. Lastly, social interactions, like territorial defense, have an influence on recapture probabilities. All these potential problems should be investigated in further studies. Nonetheless, the capture design chosen in the MRI-program has given clear results for questions of migration patterns, habitat use, and condition of first captures (Berthold et al. 1991, Kaiser 1996).

ACKNOWLEDGMENTS

DETERMINING PRODUCTIVITY INDICES FROM AGE COMPOSITION OF MIGRANTS CAPTURED FOR BANDING: PROBLEMS AND POSSIBLE SOLUTIONS

DAVID J. T. HUSSELL

Abstract. Year, location, day, moon phase, and weather all influenced the daily proportion of young birds captured in nine species of passerines in fall migration at three stations on Long Point, Ontario, in 1961–1988. The proportion of young tended to be higher on days following nights when conditions for nocturnal migration were good. Annual proportions of young may be inconsistent indices of productivity, unless they are adjusted for the daily effects of confounding variables. For single species, correlations between annual proportions of young (adjusted and raw) and fall/spring population ratios were usually low and non-significant. In most species, the annual proportion of young did not explain significant amounts of variation in trend analyses of annual population indices. Nevertheless, adjusted proportions of young performed better than raw proportions in these analyses, suggesting that the proportion of young in populations of migrants does contain useful information about productivity. However, the assumption that proportions of young reflect productivity should not be accepted uncritically. More research is needed to determine how best to use information on ages of fall migrants to elucidate their demography.

Key Words: age proportions, fall migration, Long Point, Ontario, passerine migrants, productivity indices.

The age composition of migrants captured for banding is widely assumed to provide information on the productivity of the preceding breeding season (e.g., Ralph et al. 1993). However, it is not clear whether a direct relationship exists between productivity and the proportion of young birds captured in fall migration. The proportion of young captured over a single migration season at a single station could be influenced by many confounding factors, including differing vulnerability to capture, differing timing and speed of migration of age classes, habitat and coastal effects (e.g., Murray 1966; Ralph 1971, 1981; Hussell 1982, 1991; Dunn et al. *this volume* b), and perhaps by weather. Very little research has been done to determine what effects, if any, these factors have on the daily and annual proportions of young captured during migration and consequently upon annual measures of productivity.

Weather has profound effects on numbers of birds migrating and on the numbers occurring (and therefore available for capture) at a station (e.g., Richardson 1978). If the effects of weather differ among age classes, then even consistently collected data on the number of young and adult birds captured at a single station could be biased by year-to-year variations in weather. Given these potential biases, can we derive a consistent annual index of productivity from migrant age data and can we test that such an index does in fact reflect productivity?

I used regression analysis to examine the effects of weather, moon phase, date, and station on the daily proportion of young birds of nine species captured by personnel of the Long Point Bird Observatory at three stations on Long Point during the autumn migrations of 1961–1988. I calculated two annual indices of productivity: (1) the proportion of young birds (hatched in the preceding breeding season) captured over the entire migration at all stations; and (2) an adjusted proportion of young birds, derived from the regression analysis.

Indices of population size for spring and fall migrations at Long Point were also available (calculated by methods similar to those described by Hussell et al. 1992). If age proportions indicate productivity, and if the population size indices reflect population change, then the annual ratio of the fall to spring population index should be positively correlated with the annual proportion of young birds. This is the case because a high proportion of young birds in the fall population should usually be associated with a high fall population relative to that of the previous spring. The strength of this correlation should provide an independent means of evaluating the effectiveness of methods of calculating population and productivity indices.

If productivity fluctuates from year to year, and if age proportions reflect productivity, we might expect deviations of annual fall population size indices from their general trend to be positively correlated with the proportion of young birds captured in the fall.

Therefore, I tested whether the proportion of young explained additional variability in regression analyses of trends in fall population indices.

METHODS

STUDY LOCATION, SPECIES, AGING

I examined age proportions of nine species captured commonly in fall migration at Long Point, Ontario (approximately 42°33'N, 80°10'N): Swainson's Thrush (*Catharus ustulatus*), Red-eyed Vireo (*Vireo olivaceus*), Tennessee Warbler (*Vermivora peregrina*), Magnolia Warbler (*Dendroica magnolia*), Yellow-rumped Warbler (*D. coronata*), Blackpoll Warbler (*D. striata*), American Redstart (*Setophaga ruticilla*), White-throated Sparrow (*Zonotrichia albicollis*), and White-crowned Sparrow (*Z. leucophrys*).

Data were recorded at three stations: Station 1 at the eastern tip of Long Point; Station 2, 19 km west of Station 1; and Station 3, 9 km west of Station 2. Nearly all the data from Station 2 were collected after 1974 and nearly all from Station 3 after 1983. Nearly all birds were captured in mist nets or Heligoland traps (Woodford and Hussell 1961), but a few were taken in other types of baited ground traps. Trapping and netting effort (including numbers, types, and locations of traps and nets) varied both from year-to-year and day-to-day. I excluded birds captured or killed during nocturnal migration when they were attracted to the lighthouse at Station 1.

Red-eyed Vireos and White-crowned Sparrows were aged as either young (hatched in the current year) or adult (hatched earlier) primarily by eye color and plumage differences, respectively, and I analyzed all data from 1961–1988. Other species were aged mainly by the degree of skull pneumatization (birds with incompletely pneumatized skulls were aged as young) or by obvious plumage characteristics (e.g., adult male American Redstarts and some young Swainson's Thrushes), and I used data only from 1966–1988, because skull examination was not used at Long Point prior to 1966.

For each species, a fall migration period ("migration window") was selected that was identical to that used previously for analysis of migration counts (Hussell et al. 1992). Individuals occurring outside the migration window were excluded from all analyses.

EFFECTS OF WEATHER, MOON, DAY, AND STATION ON AGE PROPORTION

Daily proportion of young was defined for each species, based on numbers of newly captured (unbanded) birds for each day that at least one bird was captured and aged, as: proportion of young = (number of young birds)/(number of young birds + number of adult birds).

I used multiple regression to examine effects of various potential predictor variables on daily proportion of young. The dependent variable was the arcsine (square root

(daily proportion of young)). Proportions of 0 and 1 were counted as $1/4n$ and $(n - ¼)/n$, respectively, where n was the sample size (i.e., the number of young + adults), before transforming to the angular scale (Snedecor and Cochran 1967:327–328). Cases were weighted by $C \times n/N$, where C was the total number of cases (i.e., station-days), n was the sample size for that case (i.e., number of young + number of adults), and N was the sum of n over all cases. This weights in proportion to sample size, and makes the sum of the weights equal to the number of cases. The analysis was otherwise similar to that used for determining indices of abundance (Hussell et al. 1992).

Station-days with captures of aged birds varied from 373 in the White-crowned Sparrow to 942 in the Swainson's Thrush. However, captures and days with captures were not uniformly distributed among stations. If the sum of the case weights for a station was less than 90, it was judged that the coefficients of variables specific to that station could not be adequately estimated and data from that station were excluded from the multiple regression analyses. This criterion excluded Tennessee Warbler, Blackpoll Warbler, and White-crowned Sparrow at Station 3, and White-throated and White-crowned sparrows at Station 2.

I assumed that productivity effects, if they existed, would be associated with year, and would occur across all stations, days of the year (hereafter, "day"), and other conditions. Therefore I included dummy variables for year, as predictor variables in the regression model without interactions with station or any other variables. On the other hand, I assumed that day, weather, and moon effects might be station-specific. Therefore, I designed the regression model to accommodate this assumption by including predictor variables for day, weather, and moon only as interactions with each station.

Age proportion differences between two of the stations were already known to occur in warblers (Dunn and Nol 1980) and preliminary analyses indicated that age proportions change with day of the year, as expected from other research (e.g., Murray 1966, Hall 1981; Hussell 1982, 1991). Therefore, I included dummy variables for station and station-day interaction variables (1st, 2nd, and 3rd order terms in day, D, D^2, and D^3, respectively, where day D was the day of the year, set to zero on a day near the middle of each species' migration window) in the regression model. Inclusion of these predictor variables enables the regression analysis to detect both consistent station effects and different seasonal patterns of change in proportion of young at each station, if they exist in the data.

Moon phase variables were days from new moon (M, or "moonday") and the square of moonday (M^2). These variables enable the analysis to detect an unequal pattern of increase in proportion of young prior to new moon and decrease following full moon, or vice versa, with the possibility of a discontinuity in the proportion of young occurring at full moon. (The sky is moonless late in the night prior to full moon and early in the night following full moon, so the effects of moonlight are likely to by asymmetrical relative to full moon.)

Weather data were from Erie, Pennsylvania (about

50 km south of Long Point on the south shore of Lake Erie) and the variables were identical to those used by Hussell et al. (1992). I used eight variables representing east wind speed, south-east wind speed, south wind speed, south-west wind speed, temperature differences from normal, square root of horizontal visibility, cloud cover, and precipitation. All positive wind speeds indicated direction the wind was coming from, and negative values represented the opposite direction (e.g. a negative south wind speed was the speed of the wind from the north). I reduced the eight weather variables to six weather factors by principal components analysis, followed by varimax rotation. The six weather factors retained 86.2% of the variance of the original eight weather variables. Because the original four wind direction/ speed variables were essentially uncorrelated, they loaded heavily on four factors (referred to as the E, SE, S, and SW wind factors for the wind directions involved). Visibility and temperature loaded heavily on the fifth factor (called "Visibility/Temperature"). Precipitation loaded heavily and cloud loaded moderately on the sixth factor (called "Rain/Cloud").

Predictor variables for weather were formed as interactions between station and the factor scores for the six rotated principal components, enabling the regression model to detect station-specific weather effects. By using factors instead of the original weather variables, the number of station-weather interaction variables was reduced from 24 to 18 at a cost of losing 13.8% of the variance in the original eight weather variables.

In summary, the multiple regression contained up to 63 predictor variables, consisting of up to 28 dummy variables for year, two dummy variables for station, nine station–day interaction variables, six station–moon phase interaction variables, and 18 station–weather factor interaction variables.

PROPORTION OF YOUNG INDICES

I calculated an annual raw proportion of young index as (number of young birds)/(number of young birds + number of adult birds), where numbers were the sums of newly captured birds accumulated from all of the stations over each species' autumn migration window. In addition, I calculated an adjusted annual proportion of young index for each of the nine species from the results of the multiple regressions described above. The adjusted annual proportion of young index was the back-transformed adjusted mean for each year. It is an estimate of what the young proportion would have been in a given year, if the values of the regression variables representing weather, dates, and locations of capture had been the same in all years, and were equal to the average values of those variables recorded in the data.

SPRING AND FALL POPULATION INDICES

Spring and fall population indices for each species counted in migration at Long Point in 1961–1988 were calculated as back-transformed adjusted means for year, from a regression analysis in which the dependent variable was

log (daily count + 1). The "daily count" was an estimated total of number of birds of each species occurring in or passing through a defined count area at each station. The estimate was based on a consistent procedure involving a count along a transect route, unstandardized trapping and netting (as described above), and incidental observations by all observers and banders present at the station (Hussell 1981, Hussell et al. 1992). Indices were calculated in the same way as described elsewhere (Hussell et al. 1992), except as indicated below. Three different sets of indices were calculated using the full data set. I had two reasons for using the full data set, instead of data reduced after an initial regression to remove cases with low predicted values (Hussell et al. 1992): (1) it enabled me to use exactly the same data sets for all three sets of indices, and (2) other analyses indicated that trends in annual indices calculated from the full data sets corresponded more closely to trends in Breeding Bird Survey counts in Ontario than trends based on indices calculated with reduced data sets (D. Hussell and L. Brown, unpublished). The three sets of annual population indices differed in the predictor variables used in the regression analyses. Dummy variables for year were included as predictor variables in all regressions, so that adjusted mean for year could be calculated. Index 1 was based on the full model with station, station–day, station–moon phase, and station–weather variables included as predictor variables (as in Hussell et al. 1992). Index 2 used a reduced model with station and station–day predictor variables. Index 3 was based on a model with dummy variables for station as the only predictor variables (in addition to the year dummy variables). I expected that index 1 would best reflect population size, because effects of variation in weather and moon phase are assigned to those variables. Index 3 would likely be the least satisfactory index of population size.

TESTS OF CONSISTENCY OF PROPORTION OF YOUNG AND POPULATION INDICES

The spring population consists of only adult birds, while the fall population has both young and adult birds. If we assume that the mortality rate of adult birds between spring and fall migrations does not vary importantly among years, then the population ratio = (fall population size index)/(spring population size index) should vary in parallel with fall proportion of young. Therefore, I calculated annual population ratios (population ratio 1, population ratio 2 and population ratio 3) based on each of the three population indices (index 1, index 2, and index 3 for spring and fall) for each of the nine species, and correlated them with annual raw and adjusted proportion of young. If adjustments of proportion of young and population indices were effective, we would expect the highest positive correlation to be between adjusted proportion of young and the population ratio for population index 1.

Rates of change in spring and fall migration indices of 42 species in the period 1967–1987 were positively correlated, as expected if spring and fall indices represent the same source population (Hussell et al. 1992). Fall indices,

however, generally showed greater variability around the trend than did spring indices (D. Hussell, unpubl. analyses). This may reflect variability in proportion of young in fall populations. If so, proportion of young may explain additional variability in the trend analysis and allow more precise estimation of trends.

I tested for the effects of age proportion on trend in fall population index 1 of each species with the following model:

$$\mathrm{Ln} I_j = a + bY_j + c\mathrm{Ln}H_j + e_j \qquad \text{(Eq. 1)}$$

where I_j was index 1 in year j, Y_j was year j, H_j was either raw or adjusted proportion of young, and e_j was an error term, and a, b, and c were coefficients estimated by the regression analysis.

In addition, I tested the effect of age proportion on combined spring and fall trend in each species with the following model:

$$\mathrm{Ln} I_{jk} = a + bY_j + cS_k\mathrm{Ln}H_j + dS_k + e_{jk} \qquad \text{(Eq. 2)}$$

where I_{jk} was index 1 in year j and season k (spring or fall), Y_j was year j, S_k was a dummy variable for season ($S_k = 0$ for spring, $S_k = 1$ for fall), H_j was proportion of young in year j, $S_k\mathrm{Ln}H_j$ was an interaction term (formed by multiplying S_k by $\mathrm{Ln}H_j$), e_{jk} was an error term, and a, b, c, and d were coefficients estimated by the regression analysis. This model assumed a common trend b for spring and fall indices and tested whether fall proportion of young index H_j had a significant additional influence on the fall indices.

In both models, c was expected to be positive (i.e., the greater the proportion of young birds, the higher the annual fall population). In both analyses, cases were weighted by $C \times n_j / N$ where C was the total number of cases, n_j was the number of station-days of observations in year j used in calculating index I_j, and N was the sum of n_j for all cases. I tested for second and third order effects in year (with predictor variables Y_j^2 and Y_j^3) and, in the second model, for season-trend interactions (S_kY_j, $S_kY_j^2$ and $S_kY_j^3$) using a

stepwise procedure. Because this involved many tests and the number of variables was large relative to the number of cases, these effects were considered important enough to be included in the model only if they were significant at the 0.01 level.

I used a sign test on the probabilities (P) associated with c in equations 1 and 2 to determine whether the adjusted proportion of young indices were more effective than raw proportion of young indices as predictors of fall population indices. Because low P values with positive estimates of c indicate good prediction and low P values with negative estimates of c indicate poor prediction, I scored P values associated with negative estimates of c as 2 - P for use in the sign test.

In all tests in this section, I used population ratios based on at least 25 station-days of observations in both spring and fall. Population ratios were excluded if either the spring or the fall index (or both) did not meet the criterion. Adjusted and raw proportion of young indices were used only if captures of aged individuals occurred on at least seven days and at least 50 individuals were aged in that year.

RESULTS

Effects of Weather, Moon, Day, and Station on Age Proportions

Samples of aged birds ranged from 1,328 in the Red-eyed Vireo to 5,414 in the Yellow-rumped Warbler (Table 1). Overall proportion of young varied from 0.549 in the White-crowned Sparrow to 0.916 in the Yellow-rumped Warbler. Except for the Blackpoll Warbler, warblers had proportions of young near 0.90, as reported previously (Dunn and Nol 1980).

Predictor variables in multiple regression analyses accounted for a significant proportion of the variation in transformed proportion of young in all species, with R^2 varying from 0.290 in the White-

TABLE 1. Summary of age data and regression results for nine species captured at Long Point, Ontario

Species	Number of first captures		Proportion of young [a]	Number of station-days [b]	R^2 [c]
	Adult	Young			
Swainson's Thrush	937	3,245	0.776	942	0.300
Red-eyed Vireo	172	1,156	0.870	571	0.391
Tennessee Warbler	191	2,006	0.913	530	0.501
Magnolia Warbler	405	3,225	0.888	831	0.515
Yellow-rumped Warbler	453	4,961	0.916	683	0.393
Blackpoll Warbler	1,061	2,173	0.672	561	0.491
American Redstart	160	1,340	0.893	604	0.289
White-throated Sparrow	669	2,133	0.761	583	0.290
White-crowned Sparrow	1,037	1,260	0.549	372	0.484

[a] Proportion of young for the entire sample = (number of young)/(number of young + number of adults).

[b] Number of station-days for which aged birds were available during the species-specific migration window, over all years used in the analyses (1961–1988 for Red-eyed Vireo and White-crowned Sparrow, 1966–1988 for all other species).

[c] R^2 for the multiple regression of arcsine (square root (proportion of young)), on year, station, station-day, station-moonday, and station-weather predictor variables.

throated Sparrow to 0.515 in the Magnolia Warbler (Table 1). A high R^2 may reflect high year-to-year variability in the proportion of young (variance assigned to the year dummy variables), important effects of other variables, or both.

Interpretation of the effects of independent variables in multiple regressions presents some difficulties, both because some variables are correlated with each other and because effects of individual variables do not occur in isolation from those of other variables (especially where there are higher order terms in the same variable). Nevertheless major effects can be discerned. To summarise the effects of variables (other than dummy variables for station and year), I tabulated the number of times (called "cases" below) that a variable had a significant or near significant ($P \leq 0.1$) positive or negative effect on the proportion of young of a species at a station. In addition, I assessed the importance of positive and negative effects of each variable by summing scores (ordered in accordance with significance level) for each positive and negative effect (Table 2).

Station

The station dummy variables for Stations 2 and 3 always had significant or near significant ($P \leq 0.10$) positive effects on the proportion of young. For nine of the 13 dummy variables (in the regressions for nine species) the effect was significant at $P \leq 0.01$. This indicates a strong tendency for there to be a higher proportion of young birds at Stations 2 and 3 than at Station 1, as previously reported for warblers at Station 2 vs. 1 (Dunn and Nol 1980). In addition to warblers, the effect was also strong in Swainson's Thrush ($P < 0.01$ for both stations) and White-throated Sparrow ($P < 0.01$ for Station 3), but relatively weak in Red-eyed Vireo ($0.05 < P \leq 0.10$ at both sites).

Day

Day of the year (D) had significant effects ($P \leq 0.01$) on proportion of young in 17 of 22 station-species cases (Table 2), including one or more stations in all species. The direction of significant effects was always consistent among stations within species, but was not consistent among species. In most species the effect was negative, indicating that proportion of young tended to decline as the season progressed, but Red-eyed Vireo and Yellow-rumped Warbler showed strong and Swainson's Thrush and White-crowned Sparrow showed weak tendencies in the opposite direction. These effects indicate that the timing of peak migration differs among age classes.

The second order term in day (D^2) had significant effects in 13 of 22 cases (Table 2), but the direction of the effect varied. Negative effects predominated. Because day zero was set near the middle of the species' migration window, a negative second-order effect indicates a tendency for the proportion of young to be higher at the middle of the season than at the

TABLE 2. EFFECTS OF DAY, MOONDAY, AND SIX WEATHER FACTORS FROM PRINCIPAL COMPONENTS ANALYSIS ON PROPORTION OF YOUNG CAPTURED

	Significant and near-significant effects ($P \leq 0.10$)			
	No. of species-stations with significant effects [a]		Score total of strength of effects [b]	
Predictor variable	Positive	Negative	Positive	Negative
Day	7	10	15	26
Day^2	4	9	10	23
Day^3	10	0	25	0
Moonday	3	4	5	8
Moonday^2	5	1	12	2
Visibility/Temperature	5	2	12	4
Rain/Cloud	1	6	3	14
E Wind	2	1	3	2
SE Wind	1	4	3	7
S Wind	3	3	5	7
SW Wind	0	3	0	7

[a] Number of species-station combinations (out of 22) that show positive or negative significant effects of the indicated variable.
[b] Score total indicates the strength of positive and negative effects of variables. Score total = sum of scores assigned to each species-station combination according to the significance level of the effect of the variable. Scores were as follows: score = 1 if $0.05 < P \leq 0.10$, score = 2 if $0.01 < P \leq 0.05$, score = 3 if $P \leq 0.01$. Maximum possible score total is 66.

beginning and end, although this may be modified or reversed (at one end of the season) by the direction and magnitude of the first- and third-order effects. Swainso's Thrush showed a strong tendency for the mid-season proportion of young to be high (P < 0.01 at all three stations). Tennessee Warbler at Station 2, and White-throated and White-crowned sparrows at Station 1, showed strong (P < 0.01) tendencies in the opposite direction: proportion of young tended to be lowest in the middle of the season. These results may indicate that adult Swainson's Thrushes have a long migration period relative to young birds, whereas the opposite is true for Tennessee Warblers, White-throated and White-crowned sparrows.

When present, the effects of the third order term in day (D^3) were consistently positive (Table 2). This indicates that the proportion of young tended to be relatively low near the start of the season and high near the end of the season. However, these effects usually occurred in combination with negative first-order effects, indicating that proportion of young started at a high plateau, declined during the course of the season, then levelled off again near the end of the migration. Such a pattern would be expected if there was a substantial average difference, but much overlap, in the timing of migration of the two age classes. Species showing this pattern strongly at all stations were Tennessee Warbler and Magnolia Warbler.

Moon

First order effects of the number of days from new moon (*M*) occurred in seven of 22 cases and the results were somewhat equivocal (Table 2). Negative effects in four species all occurred at Station 2, indicating a tendency for the proportion of young to be lower in the days before full moon, when the sky is moonless late in the night, than in the days following full moon, when the moon is above the horizon late in the night. In one of those species (Blackpoll Warbler) a strong opposite effect (P < 0.01) occurred at Station 1.

The important result with respect to moon phase was the strong tendency for second-order effects to be positive (Table 2). In four of five cases, these effects occurred at Station 1, where the presence of a lighthouse may magnify the effect (Dunn and Nol 1980). This result indicates that the proportion of young tends to be lower near new moon than near full moon, when both the size of the illuminated lunar disk and the number of nocturnal hours that it is above the horizon are near their maximum values. Species showing this pattern strongly at Station 1

were Swainson's Thrush, Blackpoll Warbler, White-crowned Sparrow (all P < 0.01), and American Redstart (P < 0.05). Red-eyed Vireo showed a tendency to have a higher proportion of young near full moon at Station 2 (P < 0.10) and a lower proportion of young near full moon at Station 3 (P < 0.05).

Weather

Only 23% (31/132) of weather-station interactions had significant (P ≤ 0.10) effects on the proportion of young. Nevertheless some patterns could be detected. High horizontal visibility and warm temperatures usually had positive effects on the proportion of young (Table 2). In the Red-eyed Vireo, however, the effect was strongly positive at Station 1 but strongly negative at Station 3 (P < 0.01 in both cases). Rain and cloud tended to have negative effects on the proportion of young (Table 2). The single exception was Swainson's Thrush at Station 3, where the effect was positive (P < 0.01).

Effects of wind variables were more erratic. Easterly and westerly winds had little effect. Winds with a southerly component tended to have a negative effect on the proportion of young, scoring 21 negative points versus eight positive points (Table 2).

ANNUAL PROPORTIONS OF YOUNG AND FALL/SPRING POPULATION RATIOS

Annual raw and adjusted proportions of young for each species are shown in Figure 1 (left side). In some species (e.g., Swainson's Thrush, Yellow-rumped Warbler) differences between raw and adjusted proportions were small; in others (e.g., Tennessee Warbler, Magnolia Warbler), there were large discrepancies in some years between adjusted and raw proportions of young. Proportions of young showed substantial year to year fluctuations. There were no very obvious trends, although the proportion of young in White-throated Sparrows was generally higher from 1975 to 1988 than between 1966 and 1972, and there was a tendency for Tennessee Warbler proportions of young to decline between 1975 and 1987.

Fall/spring population ratios also fluctuated (Fig. 1, right side). Again there were few obvious trends. Red-eyed Vireo population ratios tended to decline from 1966 to 1988. Population ratios of Tennessee, Magnolia, and Yellow-rumped warblers were high in the 1975–1980 period, corresponding with a spruce budworm (*Choristoneura fumiferana*) outbreak that peaked in Ontario in 1980 (Hussell et al. 1992).
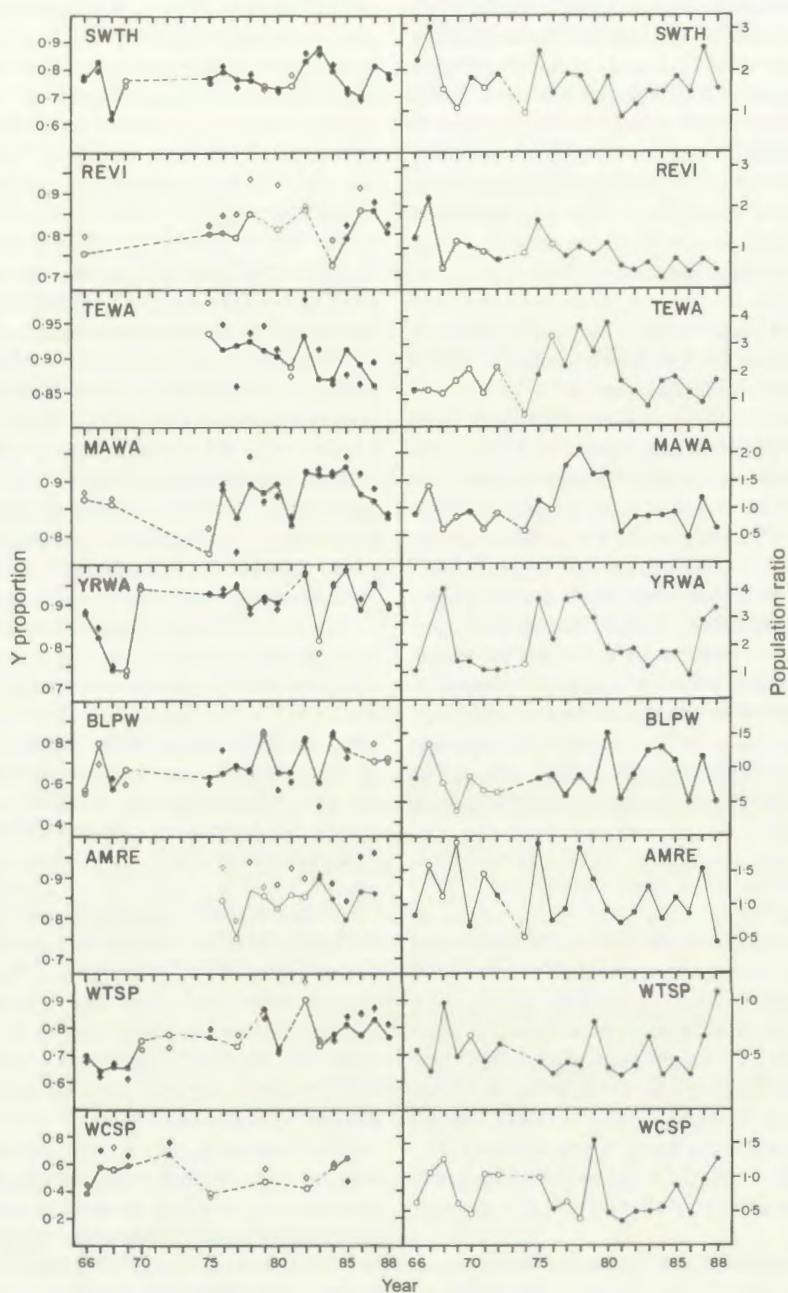
FIGURE 1. Proportion of young (Y proportion, left panels) and fall:spring ratios for population index 1 (right panels) for nine species. Left panels: circles = adjusted proportion of young, diamonds = raw proportion of young; open and closed symbols indicate proportions based on 50–99 and 100+ aged birds, respectively; lines join annual adjusted proportion of young; broken lines span years with missing data. Right panels: closed circles indicate population ratios derived from indices both of which were based on 50+ station-days of observations; open circles indicate ratios calculated from indices at least one of which was based on <50 (25–49) station-days; lines join annual ratios; broken lines span years with missing data. Species Codes: SWTH = Swainson's Thrush, REVI = Red-eyed Vireo, TEWA = Tennessee Warbler, MAWA = Magnolia Warbler, YRWA = Yellow-rumped Warbler, BLPW = Blackpoll Warbler, AMRE = American Redstart, WTSP = White-throated Sparrow, WCSP = White-crowned Sparrow.

For individual species, the only significant correlations between proportions of young and population ratios were those between raw proportion of young and all three population ratios in the White-crowned Sparrow ($r_s = 0.536$ for population ratio 1, $r_s = 0.573$ for population ratio 2, $r_s = 0.664$ for population ratio 3; $N = 11$, $P \leq 0.05$). If proportions of young and population ratios are positively related, however, then the mean correlation coefficient for the nine species should be positive. Mean correlation coefficient between adjusted proportion of young and population ratio 1 was 0.161, which was significantly greater than zero (Table 3). Means of correlations between all other combinations of methods of determining proportion of young and population ratio were non-significant and close to zero. Ranges and standard deviations of the correlation coefficients were lowest when population ratio 1 was used.

The effect of adjusted proportion of young in the trend analyses was positive (in accordance with expectation) in six of nine species for both the fall trend alone and for the combined spring and fall trend (Table 4). Significant or near-significant effects occurred in three and four species for the fall and spring/fall analyses, respectively (Table 4). Raw proportions of young had positive effects in four of nine species in the fall and three of nine species in the spring/fall analyses, with none of the effects significant or near significant. A sign test on the probabilities associated with the effect of proportion of young showed that adjusted proportion of young index was a marginally non-significantly better predictor of fall population indices than raw proportion of young index in the fall trend analyses (seven positive, two negative differences, one-tailed $P = 0.090$). In the spring/fall trend analyses, adjusted proportion of young was a significantly better predictor of the fall population index than the raw proportion of young (eight positive, one negative differences, one-tailed $P = 0.020$).

DISCUSSION

Daily proportion of young was influenced by year, station, date, moon phase, and weather (Table 2). As far as I am aware, this is the first demonstration of effects of weather on the proportion of young captured during fall migration. Although there was considerable variation among species and stations, it appears that there was a general tendency for the proportion of young to be higher when conditions were good for migration than when they were poor. The proportion of young tended to be higher near full moon than near new moon, higher when horizontal visibility was good than when it was poor, higher when there was no rain than when it was raining, and higher when there were tail-winds (northerly component) than when there were head-winds. This indicates that a greater proportion of adult birds land on Long Point when conditions for migration are poor than when they are good. This, perhaps, reflects the relative inexperience of young birds, which are less likely to overfly Long Point when conditions are good.

These effects were detectable despite the fact that capture methods at Long Point were not standardized and varied from day to day and year to year. We do not know to what extent consistent use of the same methods would have improved the precision of the productivity estimates. If the proportion of young captured is influenced by types and siting of traps and nets, then it is likely that the effects of environmental variables, such as station, moon phase and weather, would have been detected even more readily had the data collection been more standardized.

The proportion of young birds varied among the three Long Point stations, with more young recorded at Stations 2 and 3 than at Station 1. Therefore my annual raw proportion of young index (based on numbers of adults and young accumulated over all three stations) is certain to be biased by annual

TABLE 3. SPEARMAN CORRELATION COEFFICIENTS (r_s) BETWEEN PROPORTION OF YOUNG AND POPULATION RATIO FOR NINE SPECIES

Proportion of young		Population ratio model		
		Full	Station + station-day	Site only
Raw	mean	0.072	0.012	-0.020
	min, max	-0.056, 0.434	-0.302, 0.462	-0.496, 0.531
	SD	(0.164)	(0.222)	(0.296)
Adjusted	mean	0.161*	0.070	0.065
	min, max	-0.212, 0.467	-0.212, 0.420	-0.441, 0.420
	SD	(0.202)	(0.242)	(0.314)

* denotes $P \leq 0.05$ for one-tailed t-test for H: mean $r_s = 0$.

variations in the proportion of the total captures at each station. This problem exists only if there are significant differences in the proportions of young birds captured among different stations used to calculate a combined raw proportion of young index. If this is so, then the combined proportion of young index should be standardized such that each station is represented in the same proportion in the total index each year.

Overall, the results presented here imply that we should not assume that raw annual age proportions are reliable and consistent indices of productivity (see also Dunn et al. *this volume* b). It may be necessary to make adjustments for the confounding effects of station, day, moon phase, and weather. This conclusion was supported to a limited extent by my tests of consistency of proportion of young indices and population indices.

The annual proportion of young indices that were adjusted for the confounding effects of station, day, moon phase, and weather sometimes differed substantially from raw proportion of young (Fig. 1). The only significant correlation between proportion of young and population ratio was the one that matched adjusted proportion of young with population ratio 1, which was also fully adjusted for effects of day, moon, and weather (Table 3). Adjusted proportion of young was also more effective than raw proportion of young in accounting for variability in trend analyses (Table 4). All of these results suggest that adjusted proportions of young perform better as indices of productivity than do raw proportions of young.

My attempts to validate proportions of young as productivity indices were disappointing, however, in that most of the single-species correlations between proportion of young and population ratios were low and non-significant (Table 3) and, in most species, the effects of proportion of young in the trend analyses were also not significant (Table 4). My analysis is consistent with the view that both age proportions and population ratios contain information about productivity, but one or both of these measures lack precision. Given small sample sizes of aged birds in some years (particularly adults), and variability in migratory populations, it is likely that both age proportions and population ratios lack precision. Nevertheless, my results indicated that adjusted proportions of young performed better than raw proportions of young. Moreover, fully adjusted population indices outperformed other population indices, as is expected because the adjustments are designed to reduce variability that is not attributable to population size.

Inclusion of proportion of young as a predictor variable in trend analyses may enhance precision of estimates of trends in fall populations of some migrants. In several species proportion of young was not significant and it had little effect on the estimate of trend. In other species proportion of young was significant and its inclusion resulted in a relatively large reduction in residual variance, which would allow earlier detection of a trend, if it exists. For example, explained variation (R^2) in fall abundance indices increased by 18.6% in the Tennessee Warbler and 25.4% in the Magnolia Warbler when proportion of young was included as a predictor variable.

These results indicate that either proportions of young or population ratios or both may be useful for

TABLE 4. Effect of proportion of young in trend analyses of annual population indices [a]

Species	Sample sizes		Adjusted proportion of young		Raw proportion of young	
	Spring	Fall	Fall	Spring/fall	Fall	Spring/fall
Swainson's Thrush	22	18	+	+	−	−
Red-eyed Vireo [c]	22	12	+*	+**	+	+
Tennessee Warbler	22	13	+(*)	+**	−	−
Magnolia Warbler	22	16	+*	+**	+	+
Yellow-rumped Warbler	22	18	−	−	−	−
Blackpoll Warbler	22	17	−	−	−	−
American Redstart	22	12	−	−	−	−
White-throated Sparrow	22	18	+	+	+	−
White-crowned Sparrow	25	13	+	+*	+	+

[a] Except as indicated in footnote c, trends were linear as in equations 1 and 2; that is, there were no higher order or season interaction terms.

[b] (*) = P ≤ 0.10, * = P ≤ 0.05, ** = P ≤ 0.1 in one-tailed test of significance of coefficient c in equation (1) for fall and equation (2) for spring and fall (see METHODS).

[c] Spring and fall linear trends of Red-eyed Vireo differed significantly (P < 0.01). Therefore a season interaction term was included in the spring/fall regression model.

detecting productivity changes in songbird populations, but appropriate adjustments may be necessary to account for effects of confounding variables. Small sample sizes and sampling errors are likely to result in imprecise annual estimates, but long-term trends in productivity should be detectable.

It was notable that in the nine common species selected for analysis here, annual samples of aged birds were often fewer than the 50 that I judged was the minimum acceptable for inclusion in the analyses. Most banding stations probably do not capture large enough samples of more than a few species to be useful for estimating age proportions, unless the data are combined with those from other stations (with appropriate adjustments for station effects). Possibly, much larger samples than this will be needed to obtain precise indices of productivity. Alternatively or additionally, inland stations where higher proportions of adult birds are captured than at Long Point may give more precise estimates and may be less affected by confounding variables.

Validation of productivity indices for small landbird migrants is a difficult problem because reliable benchmark measures of productivity are generally not available. More research is needed to determine the reliability of indices and required sample sizes and to examine the effects of confounding variables at different stations (inland versus coastal). We need more information on whether different capture methods have an important influence on the proportion of young birds captured. Alternative methods of analysis should also be explored.

ACKNOWLEDGMENTS

AN INVESTIGATION OF PRODUCTIVITY INDICES DERIVED FROM BANDING OF FALL MIGRANTS

ERICA H. DUNN, DAVID J. T. HUSSELL, AND RAYMOND J. ADAMS

Abstract. Indices of productivity were estimated for seven species of birds captured during fall migration at two mist-netting stations less than 1 km apart, in Kalamazoo, Michigan, where those species occur only as migrants. The indices were proportion of hatch-year birds in the fall migration catch, and abundance of hatch-year birds. These values were positively correlated. Within species, mean annual hatch-year abundance often differed in magnitude between the two stations, and in some species annual abundance indices showed long-term trends in opposite directions. Nonetheless, there was evidence of parallel annual fluctuation of both productivity indices, both within and between stations. Fall migration productivity indices will rarely be useful for tracking reproductive success of specific breeding populations, because the areas from which fall migrants originate are large and poorly delineated, but such indices should be useful for other purposes (e.g., comparing regional productivity in wet and dry years). More work is needed to test the effect on fall productivity indices of habitat, net location, and frequency of sampling. Also needed are more comparisons of productivity indices among a larger number of stations, and better validation through comparison with independently derived productivity estimates.

Key Words: age ratio, banding, migration, productivity indices, validation.

Annual productivity is a key component of integrated monitoring (Baillie 1990). At local scales, intensive nest searching can provide data on reproductive success, but most such studies focus on a single species and station, often for just a few years. At slightly broader scales, constant-effort mist netting spanning the post-fledging, pre-migration period has been shown to give estimates of breeding success that correspond well with nest studies, at least in some species (du Feu and McMeeking 1991, Nur and Geupel 1992). Cooperative programs such as the Monitoring Avian Productivity and Survival program (MAPS) in North America (DeSante et al. 1995) and Constant Effort Sites (CES) in Great Britain (Peach et al. 1996) depend on many contributors to track productivity on regional scales. These programs have provided further evidence that summer mist netting reflects true levels of productivity: productivity indices may fluctuate in parallel among stations (Baillie et al. 1986), long runs of data sometimes show patterns and periodic anomalies that correspond well to suspected causal events (e.g., DeSante and Geupel 1987), and large drops in productivity indices may precede declines in breeding populations the next year (DeSante et al. 1998).

Data on birds captured during migration may provide another valuable source of productivity data. In particular, productivity measures from migrants could provide information on species whose breeding ranges are largely inaccessible for other kinds of survey, such as boreal-nesting songbirds. Migrants captured at a single station can come from broad areas of breeding range (Brewer et al. 2000, Wassenaar and Hobson 2001), so it may take only a few stations to provide results representative of a broad geographic area. Finally, fall banding produces relatively large sample sizes compared to MAPS and CES, which may contribute to making productivity indices more robust. However, although there is widespread belief that age data from the migration season reflect annual reproductive success (e.g., Ralph et al. 1993), there are no studies comparable to those for MAPS and CES that have attempted to demonstrate the validity of fall migration productivity indices.

Here we examine two productivity indices for fall migrants captured at two neighboring stations in southern Michigan: the proportion of young birds in the total sample, and an index representing abundance of young birds. Although abundance of young will vary with population size, a portion of the annual fluctuation in numbers of young should reflect variation in productivity. We compare the two indices with each other both within and between stations and to data from the Breeding Bird Survey, and outline needs for further validation.

METHODS

We analyzed age data for 1979–1991 from two banding stations that are about 0.75 km apart, located at Kalamazoo,

in southern Michigan. The "River" station had 30–35 12 m, 30 mm-mesh nets in second growth, open riparian woodland, and marsh shrub. The "Marsh" station had 15–20 similar nets in shrub vegetation bordering a marsh and woodland. Mist nets were operated daily (weather permitting) from early August to mid-November, from shortly after dawn until early afternoon. More than 80% of days in the fall migration period were covered annually.

Species chosen for analysis were Gray-cheeked Thrush (*Catharus minimus*), Hermit Thrush (*C. guttatus*), Swainson's Thrush (*C. ustulatus*), Magnolia Warbler (*Dendroica magnolia*), Yellow-rumped Warbler (*D. coronata*), Dark-eyed Junco (*Junco hyemalis*), and White-throated Sparrow (*Zonotrichia albicollis*). None of these species breeds as far south as the study station, so capture of local residents and dispersing juveniles was not a complicating factor in the analyses.

Data were restricted to first captures only, in species-specific migration "windows" (as defined at Long Point, Ontario, 650 km east of Kalamazoo; Hussell et al. 1992). A species was analyzed only if at least 0.2 adult birds/day were captured (on average) within the appropriate migration window, so that results would not be affected by chance variation in low numbers of adults. All birds were aged by the degree of skull pneumatization, and all species chosen for analysis can be aged by this method through the entire migration period. Unaged birds were excluded from the study, and did not exceed 0.3% of the totals for any species analyzed.

We constructed three indices of annual productivity from the number of birds captured, which we term "Raw HY Proportion" (where HY = hatching year birds), "HY Abundance," and "Adjusted HY Proportion." The first index was calculated for each species for each area as (N of HY birds)/(total of HY + AHY birds). To construct the other two indices, annual estimates of abundance for all birds, and for HY birds alone, were calculated from multiple regressions designed to assign variability in daily numbers to date, weather, moon phase, and year. Analysis was identical to that detailed in Dunn et al. (1997). These abundance indices represent the number of all birds, or of HY birds alone, that would be expected in a given year on an average date, under average conditions of moon phase and weather. HY Abundance was simply the abundance index

for HY birds estimated from the regressions. Adjusted HY Proportion was HY Abundance divided by the abundance of all birds, as estimated from the regressions. This figure differed from the Raw HY Proportion in that it was corrected for any variation that may have been caused by weather, moon phase, or date in the season.

Trends in breeding populations for Ontario and Michigan, according to the Breeding Bird Survey (BBS), were obtained from Sauer et al. (2000). Other evidence suggests that migrants at the study stations come from both these areas (Dunn et al. 1997). Trends in HY Abundance were calculated as the slope of the log-transformed annual indices regressed on year, producing an estimated annual percent rate of change that is directly comparable to BBS trends. Trends in HY Proportion were calculated as the slope of the regression on year of the arcsine of the square root of the original indices. Detrended indices (residuals from regression of indices on year) were derived from regression of indices transformed as described above. All other statistics involving HY Proportion were also performed on transformed indices, which normalized their distribution. Results were considered significant if P < 0.05.

RESULTS

In all species, annual Raw HY Proportion indices were significantly correlated with annual indices of Adjusted HY Proportion from the same station (r ranged from 0.71 to 0.96, P < 0.01 in all cases). However, Adjusted HY Proportion was higher than Raw HY Proportion, and usually had lower variance (Table 1). All remaining analyses were run with both indices, and each produced similar results. In the remainder of this paper, unless noted otherwise, results and discussion are limited to Adjusted HY Proportion (hereafter referred to simply as HY Proportion).

The HY Proportion at both stations averaged about 0.73 (Table 1), which is typical of other inland banding stations in North America (Dunn and Nol 1980). Values were always slightly higher at the River station (Table 1), significantly (or nearly) so

TABLE 1. MEAN RAW AND ADJUSTED HY PROPORTION FOR TWO STATIONS, 1979–1991

Species	River			Marsh		
	Raw	Adjusted	N	Raw	Adjusted	N
Gray-cheeked Thrush	0.64 ± 0.09	0.68 ± 0.08	469	0.54 ± 0.10	0.57+0.11	321
Hermit Thrush	0.82 ± 0.08	0.85 ± 0.07	803	0.75 ± 0.06	0.80+0.07	1,260
Swainson's Thrush	0.82 ± 0.06	0.86 ± 0.08	2,638	0.72 ± 0.13	0.74+0.11	654
Magnolia Warbler	0.71 ± 0.13	0.73 ± 0.07	1,506	0.69 ± 0.13	0.69+0.10	1,101
Yellow-rumped Warbler	0.76 ± 0.09	0.83 ± 0.05	6,862	0.74 ± 0.11	0.79+0.07	754
Dark-eyed Junco	–	–	116	0.65 ± 0.11	0.68+0.09	1,057
White-throated Sparrow	0.64 ± 0.12	0.66 ± 0.08	1,243	0.60 ± 0.10	0.65+0.08	1,348

Notes: Values shown are mean ± SD of indices averaged across years.

for all species except White-throated Sparrow. HY Abundance also differed between stations in six of seven species (Table 2), but there was no consistency in which station had higher mean numbers.

There were no significant long-term trends in HY Proportion, but a few in HY Abundance (Table 3). Direction of trend in HY Abundance at the Marsh station matched direction of BBS trends from Michigan (four species only, all increasing), but not those from Ontario. Trends in HY Abundance at the River station did not agree with BBS trend directions from either region. White-throated Sparrow was notable in showing significant trends in HY Abundance at both banding stations, but in opposite directions.

To determine whether productivity indices fluctuated in parallel, we examined correlation of detrended indices. (Detrending prevents correlation resulting solely from trends in the two sets of indices.) HY Abundance indices were positively correlated between stations, sometimes significantly so, as were HY Proportion indices (Table 4). HY Abundance and HY Proportion tended to fluctuate in parallel with each other within stations.

DISCUSSION

Date, weather, and moon phase had significant effects on HY Proportion in most species (as also found by Hussell *this volume*). Raw HY Proportion is therefore a more biased index than Adjusted HY Proportion, although all analyses gave similar results regardless of which HY Proportion index was used. This suggests that Raw HY Proportion may be a minimally acceptable index of productivity, despite the added variance caused by date and weather effects. More importantly, the similarity of results using both HY proportion indices strengthens our confidence that migration season productivity indices actually reflect proportion of young birds

present in the population, and are not artifacts of weather effects.

Results indicated that young birds of all species were relatively more prevalent than adults at the River station, regardless of which station hosted the higher abundance (Tables 1 and 2). Not only were there differences between stations in absolute values of productivity indices, but occasionally in long-term trends as well (Table 3). HY proportions in migrants are also known to differ markedly between coastal and inland banding stations, and between samples of birds banded and those killed at lighted structures during nocturnal migration (Dunn and Nol 1980, Ralph 1981). These results show that productivity indices derived from migration banding are not reliable indicators of the absolute number of young produced per adult. Similar conclusions have been drawn for productivity indices derived from summer banding, in which there can be higher proportions of HY birds in particular habitats, and in samples of birds captured with particular trapping devices (Peach et al. 1996, Bart et al. 1999, Green 1999, Senar et al. 1999).

Nonetheless, even when summer productivity indices differ in absolute magnitude, they may fluctuate in parallel (Peach et al. 1996, Green 1999), showing that annual changes in the relative proportion of age groups can still be a good indicator of annual shifts in productivity. The same appears to be true of migration season productivity indices (Table 4). In this study, HY Abundance and HY Proportion fluctuated in parallel within and between stations in most species, although many of the correlations fell short of statistical significance. Parallel fluctuation occurred even when trends in these indices did not agree. For example, long-term trends in HY Abundance for White-throated Sparrow were significant at both stations but opposite in sign (Table 3), yet detrended annual indices fluctuated in parallel (Table 4). These results indicate that

TABLE 2. MEAN HY ABUNDANCE FOR TWO STATIONS, 1979–1991

Species	River		Marsh
Gray-cheeked Thrush	0.30 ± 0.11		0.29 ± 0.10
Hermit Thrush	0.72 ± 0.27	***	1.35 ± 0.49
Swainson's Thrush	1.02 ± 0.24	**	0.75 ± 0.29
Magnolia Warbler	0.65 ± 0.18	**	1.16 ± 0.42
Yellow-rumped Warbler	2.14 ± 0.57	***	0.55 ± 0.23
Dark-eyed Junco	0.11 ± 0.05	***	0.89 ± 0.48
White-throated Sparrow	0.64 ± 0.36	+	0.94 ± 0.33

Notes: Values shown are mean ± SD of values averaged across years. Symbols indicate significant difference (see text) between stations (paired t-tests between annual indices): *** = P < 0.001, ** = P < 0.01, * = P < 0.05, + = 0.5 < P < 0.1.

TABLE 3. TRENDS IN POPULATION SIZE AND PRODUCTIVITY INDICES, 1979–1991

	BBS		HY Abundance		HY Proportion	
Species	Ontario	Michigan	River	Marsh	River	Marsh
Hermit Thrush	2.6+	8.1**	2.3	8.5***	0.25	0.84
Swainson's Thrush	-1.2+	–	2.0	5.9	-0.24	-0.85
Magnolia Warbler	3.0+	9.2	-1.8	4.3+	-0.40	-0.63
Yellow-rumped Warbler	-2.8+	4.6*	1.2	3.8	-0.30	0.49
Dark-eyed Junco	-2.8	–	0.1	-4.9	-0.64	-0.18
White-throated Sparrow	-1.3*	1.2	-9.3**	7.4**	0.22	-0.52

Notes: BBS and HY Abundance trends are rates of change (%/yr). Trend in HY Proportion is average annual change (see Methods). Symbols indicate significance of trend (see text): *** = P < 0.001, ** = P < 0.01, * = P < 0.05, + = 0.5 < P < 0.1.

annual fluctuation in HY Abundance is quite strongly affected by reproductive success. Because it is also affected by annual change in population size, however, it is not as useful an indicator of reproductive success as is HY Proportion.

Several factors may have introduced bias into the productivity indices in this study, which could have reduced the strength of evidence for parallel fluctuation. Vegetation increased in height throughout the study period, and while nets at the Marsh station were moved to keep them in shrub habitat, at the River station they were not. Also, in some years there was a large berry crop at one station but not the other, and thrushes were noted to concentrate where berry crops were high, perhaps reducing correlation of HY Abundance between the stations.

In addition, there were methodological differences between the stations that may have affected results. Net numbers were not wholly standardized, with some nets added and others discontinued during the study period, and not all nets were opened on every day that netting took place. Such factors could alter the abundance, proportion, or both of HY birds

at one station relative to the other, particularly if certain nets were more likely to capture birds of a particular age class, or if nets were opened at only one station when there was an influx of birds with unusual age distribution.

The MAPS and CES programs pool productivity data from many stations to calculate regional values, such that anomalies at individual stations are evened out. The same approach with fall migration indices may strengthen results. One difficulty with this approach, however, is defining the region within which all monitoring stations are capturing individuals from the same breeding population. There is evidence, for example, that migrants moving through southern Michigan come from both Michigan and Ontario (Dunn et al. 1997). When BBS trends differ in different parts of the breeding range from which migrants are drawn (as in Yellow-rumped Warbler; Table 3), we do not know which trend is most important for comparison to fall migration productivity indices from southern Michigan. Similarly, we do not know to what extent a more distant station—for example, in central or northern Michigan—would be sampling

TABLE 4. CORRELATIONS OF ANNUAL PRODUCTIVITY INDICES BETWEEN STATIONS, AND WITH EACH OTHER WITHIN STATIONS

	Correlation between River and Marsh		Correlation between HY Abundance and HY Proportion	
Species	HY Abundance	HY Proportion	River	Marsh
Gray-cheeked Thrush	0.33	0.32	0.45	0.61*
Hermit Thrush	0.60*	0.49+	0.08	0.75**
Swainson's Thrush	0.14	0.27	0.44	0.73**
Magnolia Warbler	0.50+	0.93***	-0.19	0.33
Yellow-rumped Warbler	0.46	0.42	0.68*	0.54+
Dark-eyed Junco	0.38	–	0.72*	0.43
White-throated Sparrow	0.51+	0.70**	0.72**	0.04

Notes: Values shown are correlation coefficients between detrended indices (residuals from regression of appropriately-transformed indices on year) and significance levels (see text): *** = P < 0.001, ** = P < 0.01, * = P < 0.05, + = 0.5 < P < 0.1.

the same population as the stations in Kalamazoo. Fall migration productivity data from very nearby stations can certainly be pooled for analysis (assuming stations all follow the same protocol), but it may not be justifiable to pool data from very distant stations.

It will be hard to associate fall migration productivity indices with specific breeding populations because of uncertainty as to breeding origin, so migration season productivity indices will have limited value in assessing impact on productivity of locally varying factors such as predation levels. Nonetheless, accurate information on annual shifts in productivity of migrants should be useful for other purposes. For example, there are known cases of reproductive success varying with weather, either routinely or in response to unusual conditions (e.g., DeSante and Geupel 1987, Bradley et al. 1997). Because weather often affects large geographic areas, data from migrating birds might be especially well suited to the study of such weather effects.

This paper is one of the first to critically examine fall migration productivity indices (see also Hussell *this volume*). Although we found evidence that different stations detected similar annual changes in productivity, our primary conclusion is that a good deal more basic research is in order. A recent study of Pink-footed Geese (*Anser brachyrynchos*) showed the importance of cross-validation and study of biases in data sources, including productivity indices, even for well-studied populations with excellent data (Gantner and Madsen 2001). Similar kinds of work are needed on fall migration productivity indices, including effects of habitat and net location on ages of birds captured, and degree of parallel fluctuation in productivity indices among nearby stations. For example, Harrison et al. (2000) found that habitat change at his late summer banding station altered the relative proportions of age groups in some species but not in others. Similar kinds of research are needed to determine the circumstances and species for which fall productivity indices are meaningful. Even more important is the need to validate migration season productivity indices through comparison with independently collected data on reproductive success. The most suitable comparison would be with MAPS results from probable breeding areas.

In the meantime, we offer several recommendations for the study of productivity through capture of fall migrants. Banders should routinely record the technique they use for ageing each bird handled, and keep careful records of daily effort, net number, and location, so that users of age data can analyze and interpret them correctly. Recording the net number where each bird is captured should permit analysis of net-site effects on age proportion. Capture effort should be as standardized as much as possible (Ralph et al. *this volume a*), to avoid bias in the numbers of each age group captured. Finally, many species have differential timing of fall migration, so it is especially important for avoiding bias to collect evenly spaced (preferably daily) samples throughout the entire migration period of the species.

ACKNOWLEDGMENTS

OPTIMIZING THE ALLOCATION OF COUNT DAYS IN A MIGRATION MONITORING PROGRAM

LEN THOMAS, GEOFFREY R. GEUPEL, NADAV NUR, AND GRANT BALLARD

Abstract. Many migration monitoring stations cannot operate on every day of the migration period. In this paper, we used migration count data from two stations (Point Reyes Bird Observatory fall migration and Long Point Bird Observatory spring migration) to examine the relationship between the proportion of count days (frequency of sampling) and the statistical power to detect long-term population trends. We found that power to detect trends at a single station declined at an accelerating rate as the frequency of sampling decreased. Stations that operate on one or two days per week are unlikely to detect changes in abundance for most species that would be well monitored at higher sampling frequencies. The effect of missing counts can be mitigated to some extent by the choice of sampling design (method of allocating count days over the migration period). We compared a number of candidate designs and found that systematic sampling was the most accurate, although stratified random sampling may be preferred in situations where little is known about the pattern of migration. Designs that clump count days together, such as sampling only at weekends, should be avoided because adjacent count days tend to duplicate the same information.

Key Words: avian migration monitoring, population trends, power analysis, sampling frequency, survey design.

One of the principal objectives of songbird migration monitoring is to determine whether the abundance of birds arriving at a monitoring station has changed over time. To achieve this, birds are surveyed at the station on as many days as possible during the migration season. A number of survey techniques may be employed (including mist netting), depending on the characteristics of the location. Regardless of the survey method used, the daily counts are converted into annual indices of abundance, and population trends are estimated from the annual indices by regression. The use of count data to monitor migration in this way has been reviewed elsewhere (Dunn and Hussell 1995, Dunn in press).

At most of the larger migration monitoring stations in North America (e.g., Long Point, Point Reyes, and Manomet bird observatories), counts take place on essentially every day of the migration season. However, many smaller stations are constrained by funding or by the availability of volunteers and cannot operate every day. Gaps in the daily counts result, which introduce additional variability into the annual index estimates, and in turn reduce the ability of the station to detect population trends. This additional variability is called "sampling variance" and is a function of the "frequency of sampling" (the proportion of days on which counts take place) and the "sampling design" (method of allocating count days over the season). In this paper, we address two questions: (1) To what degree can gaps in the count data reduce our ability to detect long-term trends?

and (2) Can the effect of gaps be minimized by the choice of an appropriate sampling design?

To answer these questions, we used data from two stations where monitoring occurs continually throughout the migration season: Point Reyes and Long Point bird observatories. By analyzing the pattern of counts in these "complete" data sets, we could estimate the sampling variance that would arise from different frequencies of sampling and sampling designs. The two observatories differ in the environment of the stations, the methods of data collection, and the species seen. In addition, we used fall data from Point Reyes and spring data from Long Point. We reasoned that, by using very different datasets, any similarities in the results between stations would be of more general applicability to other migration monitoring stations in North America. This does mean, however, that we cannot interpret any differences in results between the data sets as being due to differences between stations, since they could also be due to differences between season.

The ability of a monitoring program to detect a given trend can be measured using the concept of "statistical power." Statistical power is the probability of getting a significant result in a statistical hypothesis test, given that there is an effect (i.e., trend) of specified magnitude (Cohen 1988, Gerrodette 1987, Nemac 1991, Steidl and Thomas 2001). In the context of avian population monitoring, the Monitoring Working Group of Partners in Flight have proposed that a successful monitoring

program is one that has a 90% chance of detecting a 50% decline in a species' population over a 25 year period (Butcher et al. 1993:199). Thus, the importance of gaps in the count data can be assessed by quantifying the impact of sampling frequency on the statistical power to detect a population change of this magnitude. To do this, we estimated statistical power to detect a 50% population change over 25 year for sampling frequencies ranging from one to seven days per week. We assumed that count days are selected at random, that the statistical test used to detect trends is a linear regression of annual indices against time, where annual indices are the mean of the log-transformed daily counts, and that the test was statistically significant when $P \leq 0.05$.

At monitoring stations that do not operate every day, there is often some flexibility in the way that the count days can be allocated through the season. A number of familiar sampling designs are discussed in standard textbooks on sampling (e.g., Cochran 1977), such as simple random, stratified random, and systematic. These designs vary in the ease with which they can be implemented, and in the sampling variance of the resulting annual indices. To quantify the differences in sampling variance that could be expected for migration monitoring, we computed the "design effect" of a number of candidate sampling designs, over a range of sampling frequencies. The design effect is the ratio of (1) the sampling variance obtained from the candidate sampling design divided by (2) the sampling variance obtained from simple random sampling at the same sampling frequency (reference in Cochran 1977:85). Design effects of <1 indicate an improvement in precision over random sampling, and the design with the lowest design effect should be preferred by those designing monitoring programs, all other factors being equal.

This paper is aimed at those designing a migration monitoring program at a single station. We do not consider the trade-off between the frequency of sampling at multiple stations versus the number of stations that can be sampled. The solution to this problem will depend in part upon the variability between stations, which is not well known for migration monitoring. A treatment of the topic in the context of extensive surveys such as the Breeding Bird Survey is given by Link et al. (1994).

METHODS

DATA USED

Data from Point Reyes were collected at the Palomarin Field Station in coastal California using constant-effort

mist netting (see DeSante and Geupel 1987 for details of the field methods). We used fall migration data collected between August 18 and November 26 (101 days) in the years 1980 to 1992 (13 years). The field protocol calls for 20 fixed nets to be placed for 6 h each day during the migration period, making 120 net-h in total. However, inclement weather and other eventualities sometimes prevented the protocol from being followed: the mean percent of days when no nets were set was 4.7/year, and the mean net-hours for the remaining days was 110.7. Our analyses required a complete data set, so we substituted for the missing value on days when no nets were run the mean of the counts from the previous two days and the next two days. We standardized all daily counts to the total new birds banded per 120 net-h.

Long Point is a peninsula on Lake Erie, Ontario. There, a combination of standardized transect counts, unstandardized netting, and casual observations were used to produce a daily estimated total of each bird species present at the monitoring station (see Dunn et al. *this volume a*, for a full description of the method of data collection). In this paper, we used spring migration data from the station at the tip of the Point, collected between 1961 and 1993. Of these data, we excluded years with any missing counts, leaving 16 years: 1963, 1967, 1971–1972, 1975, 1978, 1980, 1982–1985, 1987, 1989–1990, 1992–1993. The seasonal timing of data collection varied between years, with a mean starting date of April 10 and ending date of June 13, giving a mean of 65 daily counts per year.

Having derived daily counts from each station, we treated both datasets identically. At each station, we chose migration periods separately for each species using a procedure similar to that of Hussell et al. (1992), as follows. First, we excluded counts in the first or last two weeks of the data collection periods that were separated by more than four days from any other count in any other year. Then we selected start and end dates so as to encompass the middle 98% of days on which the species was recorded. To simplify the comparison of sample designs (see below), we further truncated the data such that all migration periods at the same station began on the same day of the week, and were an integer number of weeks in length.

After the selection of migration periods, species with a mean daily count over all years of less than 1.0 were excluded from the analysis. This criterion was necessary because the methods we used to estimate trends are known to be unreliable for species with low abundances (see DISCUSSION). The counts were then log-transformed (after the adding 0.5 to all counts, to prevent taking the log of a zero count). Annual indices were calculated as the mean of the logged counts.

To better understand our results, we performed a number of descriptive analyses of the log-transformed count data for each species at each station. We studied the frequency distribution of the counts, and the distribution of counts through the migration period. We also plotted "correlograms," which show the correlation between counts taken in the same season against the number of days the counts are spaced apart. Correlation between counts taken

in the same season is known as "autocorrelation." The degree of autocorrelation has important implications for the optimal choice of sampling design, as is discussed below.

All of the analyses described here were performed using SAS for OS/2 version 6.10 (SAS Institute 1993).

IMPORTANCE OF SAMPLING FREQUENCY

For each species, we calculated the power to detect a 50% change in annual indices over 25 year under the null hypothesis of no change as

$$\text{power} = 1 - \Phi\left(t_{1-\alpha/2, df, \delta}\right) + \Phi\left(-t_{1-\alpha/2, df, \delta}\right)$$

where $t_{x, df, \delta}$ denotes the xth quantile of the noncentral t-distribution, given df degrees of freedom and the noncentrality parameter δ, and $\Phi(t_{x, df, \delta})$ is the cumulative distribution function of the appropriate noncentral t-distribution, evaluated at x (Nemac 1991). The degrees of freedom were 23 throughout (number of years over which we hope to detect the 50% change - 2), and α was set at 0.05. The noncentrality parameter was calculated as:

$$\delta = \frac{(\beta_1 - 0)}{\sqrt{\left(S_{reg}^2 + S_{samp}^2\right)/SS_{yr}}}$$

where β_1 is the slope of the log-linear regression line under the alternate hypothesis (50% change in population size over 25 years is a slope of 0.277 on the log scale), SS_{yr} is the sum of squares of the year variable (which, with 25 years of continuous data, is 1,300), S_{reg}^2 is the variance of the annual indices that is unexplained by the regression line when there are no gaps in the daily count data, and S_{samp}^2 is the additional variance due to missing counts (i.e., the sampling variance). We estimated S_{reg}^2 for each species at each station as the mean square error from the linear regression of annual indices against year. We calculated statistical power at seven levels of S_{reg}^2, assuming sampling frequencies of one to seven days per week and a simple

random sampling design. At a sampling frequency of seven days per week (i.e., no gaps in the count data), $S_{samp}^2 = 0$. At the other six levels of sample frequency, was calculated for each species as the mean sampling variance over all years of count data. The calculation of the sampling variance for each year, assuming a simple random design, is outlined in the next section.

COMPARISON OF SAMPLING DESIGNS

At each station, we calculated the sampling variance for the sampling designs for each species and year and over a range of sampling frequencies. We then calculated the design effect of each sampling design for each species at each sampling frequency as (1) the mean sampling variance over years for the sampling design divided by (2) the mean sampling variance for random sampling at the same sampling frequency. The sampling designs are explained below, and are shown diagrammatically in Figure 1. The sampling frequencies used and the method of comparing sampling designs are described at the end of this section.

In "simple random sampling," our baseline design, the number of count days is fixed, but their location in the season is chosen at random. The variance of the annual index in a single year was calculated as

$$V_{rand} = \frac{S^2}{n} \frac{(N - n)}{N}$$

(Cochran 1977, formula 2.8), where S^2 is the variance of the log-transformed daily counts, N is the number of days in the migration period and n is the number of days on which counts took place. Here, n = frequency of sampling per week × N/7.

In "stratified random sampling," the season is divided into strata and the sample consists of count days randomly selected from within each stratum. For simplicity, we chose to use equal-size strata of one week. We used two different methods to allocate sample days within strata: "proportional
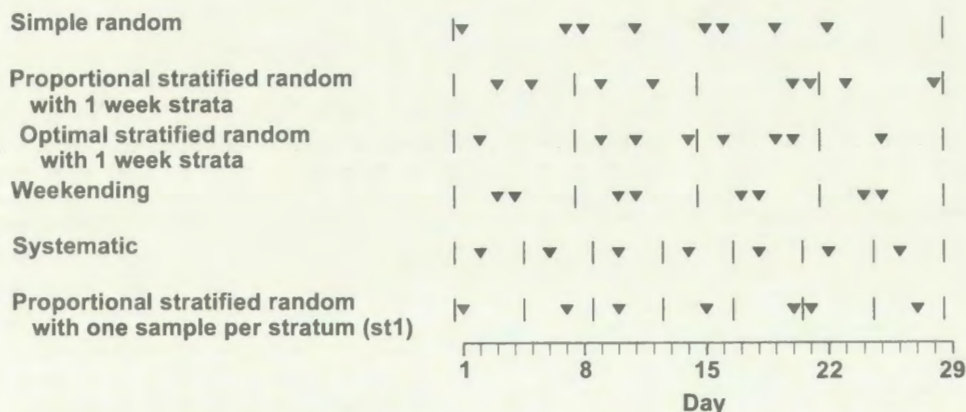

FIGURE 1. A schematic representation of possible sample allocations under six different sampling designs. Vertical lines represent the sampling strata for each design.

allocation" and "optimal allocation." In proportional allocation, the proportion of count days is the same in all strata. Because strata were all the same size, the number of sample days was the same in each stratum. In optimal allocation, strata in which the daily counts are more variable are given a greater frequency of sampling than strata with relatively homogeneous counts. With equal-size strata, the overall sample variance is minimized when the count days are allocated as follows:

$$n_h = n \frac{S_h}{\sum_{h=1}^{H} S_h}$$

(from Cochran 1977, formula 5.26), where n is the total number of count days, n_h is the number of count days in stratum h, and S_h is the standard deviation of the log-transformed daily counts in stratum h ($h = 1 .. H$). Because the timing of migration varied between species and years, no one allocation of count days could be optimal for all species in all years. We thus constructed a compromise allocation scheme at each monitoring station such that the number of count days in each stratum was proportional to the mean of the optimal allocation for that stratum over all years and species. When the compromise allocation of count days was non-integer, we used the nearest integer value, unless that value was zero, in which case we used one, or was greater than seven, in which case we used seven. For both stratified random designs, the sampling variance was calculated as

$$V_{strat} = \frac{1}{H^2} \sum_{h=1}^{H} \frac{S_h^2}{n_h} \left(\frac{N_h - n_h}{N_h} \right)$$

(Cochran 1977, formula 5.6), where H is the number of strata (number of weeks), N_h is the stratum size, n_h is the number of count days, and S_h^2 is the variance of the log-transformed daily counts in stratum h.

In the "weekending" design, counting is concentrated on two consecutive days each week. No formulae are available to determine the sampling variance of the annual indices under such a design, so we calculated the variance empirically. At a sample frequency of two days per week, there were seven possible sample allocations for each year, each subsample consisting of day i, $i+1$, $i+7$, $i+8$, $i+14$, $i+15$,... where $i = 1 .. 7$. We calculated the annual index from each subsample, and used the variance of these seven indices as an estimate of the sampling variance. At higher sample frequencies, additional counts were randomly located during each week, and we used 50 subsamples at each level of i to calculate the sampling variance.

In "systematic sampling," the sample consists of a fixed number of count days spaced at regular intervals throughout the season. The sampling variance is

$$V_{sys} = \frac{N-1}{N} S^2 - \frac{k(n-1)}{N} S_{wsy}^2$$

(Cochran 1977, formula 8.1), where S^2 and N are defined as with random sampling, k is the interval of the count days (e.g., when sampling on alternate days, $k = 2$), n is the

number of count days and S_{wsy}^2 is the mean within-sample variance. Here,

$$n = N / k$$

and

$$S_{wsy}^2 = \frac{1}{k(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i\bullet})^2$$

where y_{ij} is the jth log-transformed count in sample i, and $\bar{y}_{i\bullet}$ is the mean of the log-transformed counts in sample i. To simplify the calculations, we removed the last few counts when the migration period was not an integer multiple of the count interval (k). In these cases, we used the same data to calculate the sampling variance for random sampling when determining the design effect.

Systematic sampling is often compared with "proportional stratified random sampling with one sample per stratum" ($st1$) because the two designs differ only in the allocation of samples within strata (Fig. 1). In $st1$, the stratum size is equal to k and one sample is drawn from each of n strata. We calculated the variance for the $st1$ design using the same data as for systematic sampling and the formula given above for stratified random sampling.

We compared the sampling designs in two groups: (1) proportional stratified, optimal stratified, and weekending, and (2) systematic and $st1$. In the first group, we calculated design effects at all integer sampling frequencies from two to six days per week (two days per week being the minimum for the weekending design). In the second group, we calculated design effects at all integer count intervals from $k = 2$ (sampling every other day) to 7 (once per week), which gave sampling frequencies of 3.5, 2.33, 1.75, 1.4, 1.17, and 1 day per week. For both groups, we performed the comparison using a repeated-measures analysis of variance, with species as the subject and sampling design and sampling frequency as the within-subject factors. Because design effect is a ratio measurement, all comparisons took place on the log scale, and were then back-transformed to the arithmetic scale for the presentation of results.

RESULTS

A total of 38 species was observed at the Point Reyes station in fall, and 81 at Long Point in spring. Of these, six species occurred with a mean count of 1.0 bird/120 net-h or greater during the fall migration period at Point Reyes and 46 species had daily counts of 1.0 or greater during spring migration at Long Point. For these more common species, there was considerable variation in the distribution of the log-transformed daily counts (Fig. 2). Most of the species with low mean counts had very skewed distributions, with a majority of zero counts and a few days when many birds were seen. Species with greater mean counts tended to have less skewed distributions. Some differences were also evident between stations (Fig. 3). Species at Long Point tended

FIGURE 2. The mean (black dot), median (X), and range (horizontal line) of log-transformed daily counts for 52 species at two migration stations (see key to species codes in Appendix A).

to show greater variability in log-transformed counts (larger standard deviation), and have more skewed distributions than at Point Reyes.

The timing of migration varied between species, but the pattern of counts was similar for most species at both stations, showing a rise from low counts to a period of high counts and then a drop-off (Fig. 4). There were, however, distinct differences in the pattern of autocorrelation between the stations (Fig. 5). Species at Point Reyes tended to show an
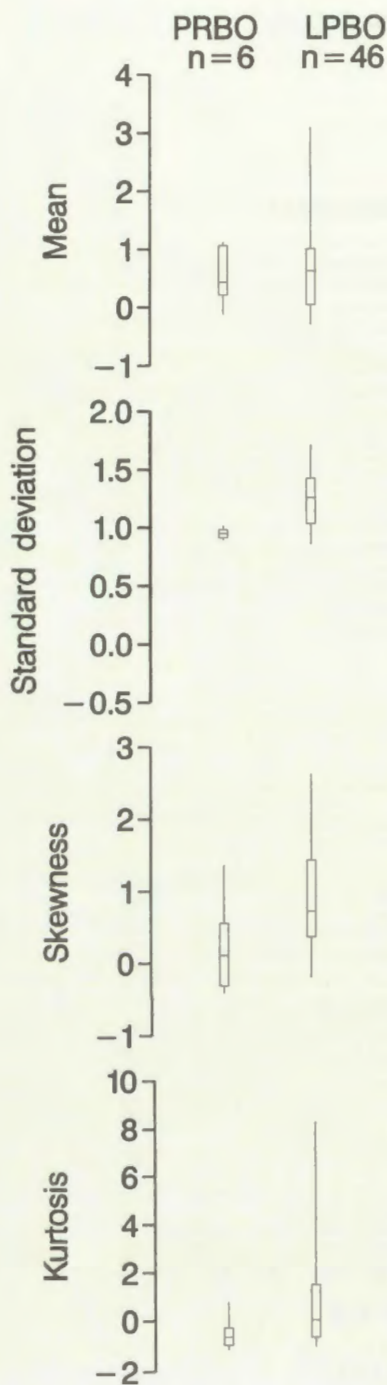
FIGURE 3. Boxplots showing the distribution of four statistics that describe the log-transformed daily counts for six species at Point Reyes Birds Observatory (PRBO) and 46 species at Long Point Bird Observatory (LPBO). The vertical lines show the median, the ends of the box show the inter-quartile range, and the horizontal lines show the range.

approximately linear decline in correlation between counts with increasing separation between count days. A few species at Long Point showed the same pattern, but the majority showed a sharp drop in autocorrelation so that the median correlation between counts spaced three days apart was close to zero. As noted previously, these differences could be due either to differences between stations or between seasons.

IMPORTANCE OF SAMPLING FREQUENCY

Statistical power declined at an accelerating rate with decreasing frequency of sampling for all species (Fig. 6). At Point Reyes, median power declined from 0.78 at a sample frequency of seven days per week to 0.59 at a sample frequency of one day per week. At Long Point, median power declined from 0.60 to 0.28 over the same range of sample frequencies.

The overall rate of decline differed between species, with some showing little effect of decreasing sampling frequency (e.g., Golden-crowned Kinglet [scientific names and key to four-letter codes in the Appendix] at Point Reyes; Slate-colored Junco at Long Point) and others being strongly affected (e.g., Hermit Thrush at Point Reyes; Chestnut-sided Warbler at Long Point). Species little affected were generally those with low power even at high sampling frequencies, while those strongly affected tended to have good power at high sampling frequencies (Spearman's rank correlation between rate of decline in power and power at a sampling frequency of seven days per week: $r = 0.89$, $N = 6$, $P = 0.02$ at Point Reyes Bird Observatory; $r = 0.26$, $N = 46$, $P < 0.001$ at Long Point Bird Observatory).

COMPARISON OF SAMPLING DESIGNS

There were consistent differences in design effect (and therefore in sampling variance) among sampling designs. These differences were statistically significant for both groups of sampling designs at Long Point, but not at Point Reyes (Table 1), although the sample size was low at the latter station (six species). Averaged over all sampling frequencies, optimal stratified sampling had the lowest design effect in group 1, and systematic was the lowest in group 2 at Long Point (Table 2). Although direct comparisons between all sampling designs are not strictly appropriate, since the two groups were measured at different sampling frequencies, it is clear that weekending performed very poorly (i.e., had the highest design effect), and the systematic design the best (i.e., had the lowest design effect; Table 2). Average design effects were
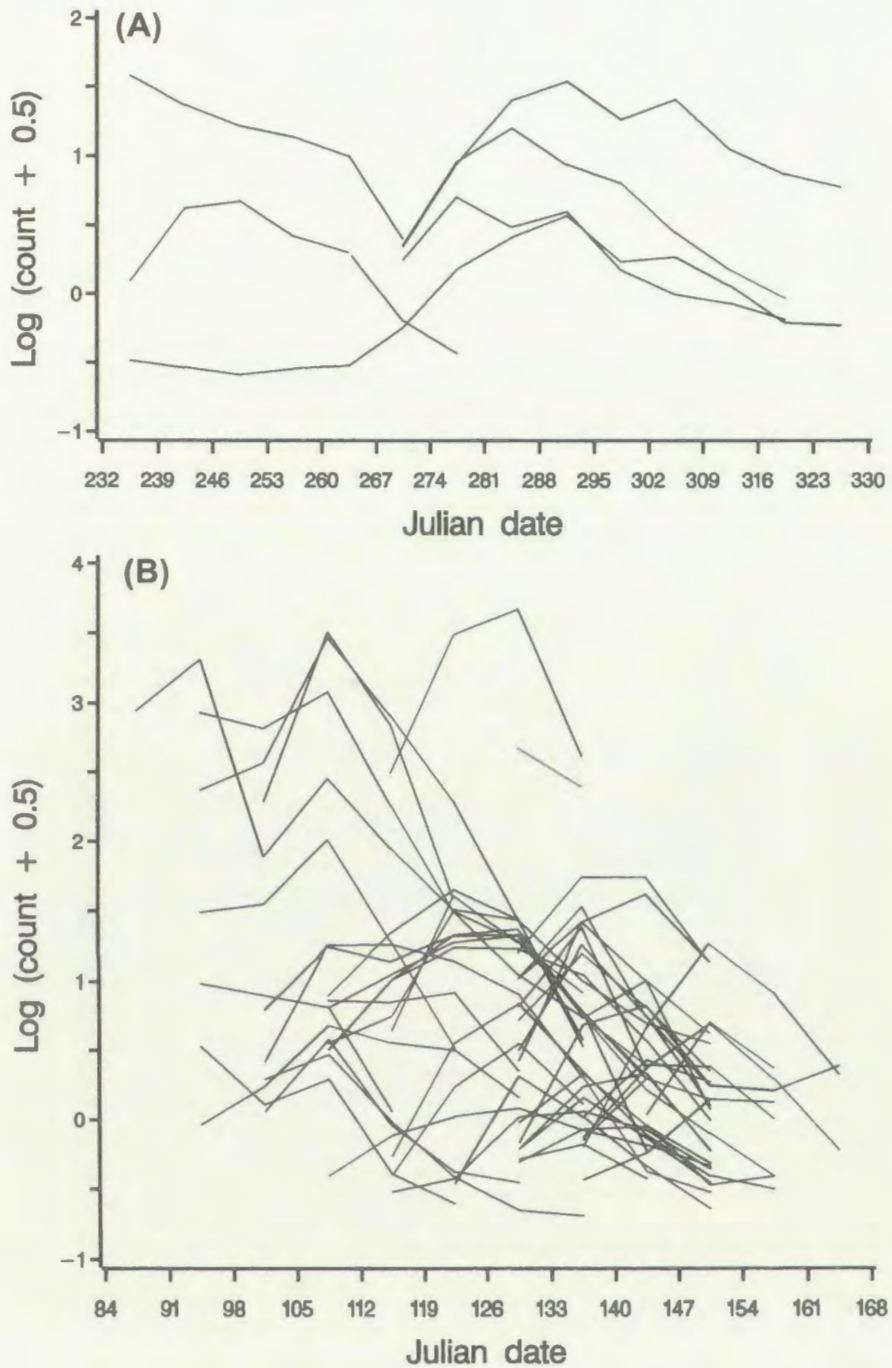
FIGURE 4. Change in daily counts over time at two migration monitoring stations. Lines show weekly mean (pooled across years) of the log-transformed counts during the migration period for each of six species at Point Reyes Bird Observatory (fall migration) and 46 species at Long Point Bird Observatory (spring migration).
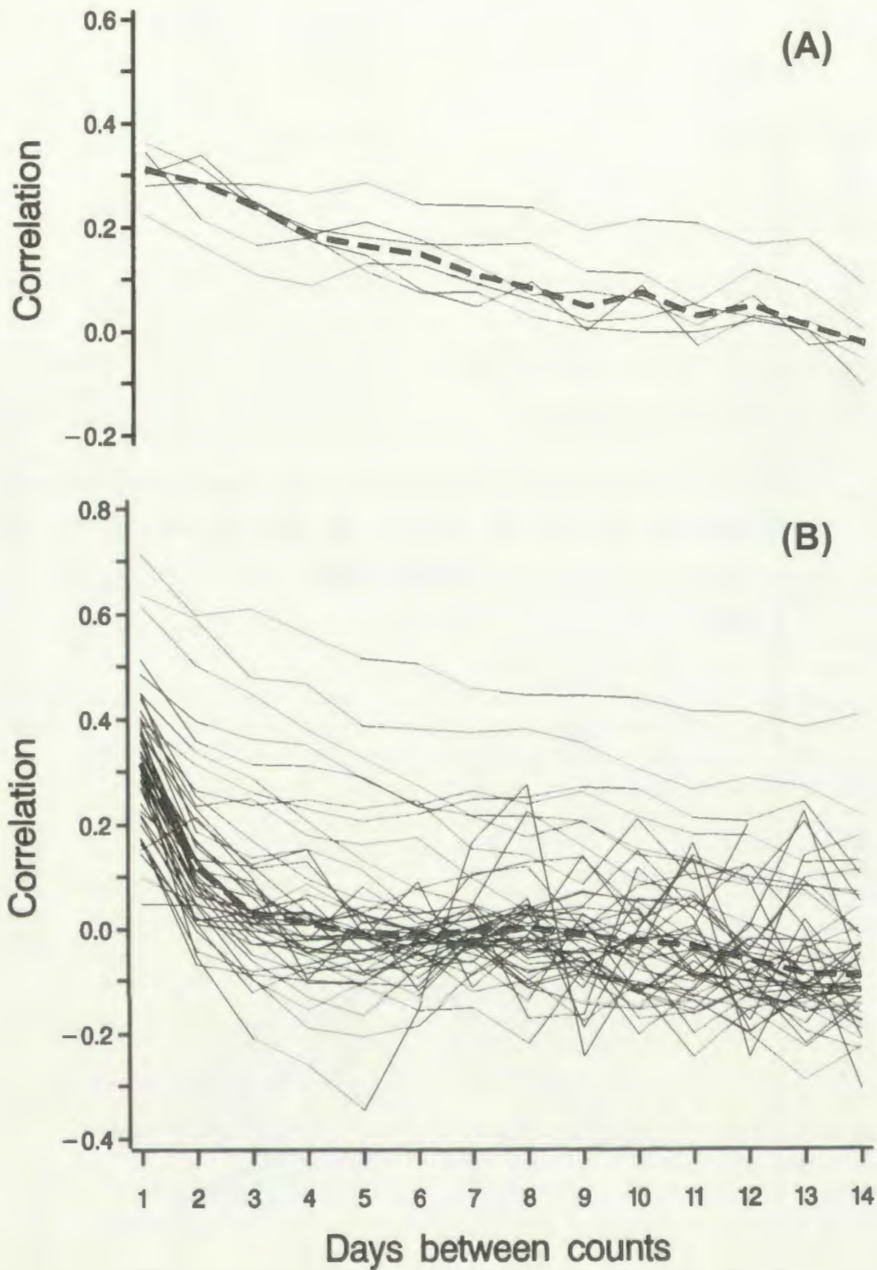
FIGURE 5. Autocorrelation between daily counts at two migration monitoring stations. Thin solid lines show the mean correlation over years between log-transformed counts for each of six species at Point Reyes Bird Observatory and 46 species at Long Point Bird Observatory. The thick dashed line shows the median of the species correlations.

similar among sampling designs at Point Reyes, and were not significantly different (Table 2).

There were also statistically significant effects of sampling frequency on design effect, and interactions between sampling frequency and sampling design

for both groups of sampling designs at Long Point (Table 1). The interactions are shown in Figure 7 (lower panel). For the group 1 designs, weekending performed very poorly at a sampling frequency of two days per week, and was the only sampling design to

TABLE 1. RESULTS OF REPEATED-MEASURES ANOVA TEST OF THE NULL HYPOTHESIS THAT DESIGN EFFECT WAS INDEPENDENT OF SAMPLING DESIGN AND SAMPLING FREQUENCY AT TWO MIGRATION MONITORING STATIONS.

Factor	Point Reyes			Long Point		
	df	F	P	df	F	P
Group 1:						
Sampling design	2, 10	1.37	0.30	2, 90	112.61	<0.001
Sampling frequency	4, 20	1.15	0.36	4, 180	112.04	<0.001
Sampling design × sampling frequency	8, 40	1.52	0.18	8, 360	72.95	<0.001
Group 2:						
Sampling design	1, 5	2.24	0.19	1, 45	44.92	<0.001
Sampling frequency	5, 25	0.70	0.63	5, 225	58.25	<0.001
Sampling design × sampling frequency	5, 25	0.43	0.83	5, 225	4.76	<0.001

Notes: Design effect measures the sampling variance of a sampling design relative to that of random sampling. Response variable was the log-transformed design effect for 6 species at Point Reyes Bird Observatory and 46 species at Long Point Bird Observatory. Group 1 designs (proportional stratified random with 1 week strata, optimal stratified random with 1 week strata, weekending) were tested at sampling frequencies of 2, 3, 4, 5 and 6 days per week. Group 2 designs (systematic, proportional stratified random with 1 sample per stratum) were tested at sampling frequencies of 1, 1.17, 1.4, 1.75, 2.33, 3.5 days per week. The sampling designs are defined in the text.

have a higher sampling variance than simple random sampling (i.e., design effect >1.0). At higher sampling frequencies weekending improved, becoming similar to the proportional stratified random design. Optimal stratified random sampling was similar to the proportional design at low sample frequencies, but improved as the frequency of sampling increased. For the group 2 designs, both systematic sampling and *st1* improved with increasing frequency of sampling, but in an erratic manner such that they were quite similar at a sampling frequency of 1.75 days per week ($k = 4$), but not similar at other frequencies.

At Point Reyes (Fig. 7, upper panel), the weekending design showed the same pattern of convergence upon the proportional stratified design with increasing sampling frequency, but, unlike Long Point, the optimal design and proportional designs were very similar at all frequencies of sampling. The *st1* and systematic designs showed no consistent patterns. For clarity, error bars were not shown on Figure 7, but they were very wide for all designs at

Point Reyes, due to the low number of species in the analysis.

DISCUSSION

Some readers will be disappointed to see that, even when there are no gaps in the count data, the median power to detect a population change of 50% over 25 years was 0.78 for the fall migration data from Point Reyes and 0.60 for the spring migration data from Long Point. This falls short of the goal suggested by the Monitoring Working Group of Partners in Flight (90% chance of detecting a decline of this magnitude; Butcher et al. 1993). However, a number of caveats should be made regarding our power analysis, and we begin the discussion by pointing these out. These limitations affect the level of power assigned to each species when there are no gaps in the count data; they do not greatly alter our main results regarding the relationship between statistical power and sampling frequency, which we

TABLE 2. GEOMETRIC MEAN DESIGN EFFECTS FOR FIVE SAMPLING DESIGNS AT TWO STATIONS. SMALLER DESIGN EFFECTS INDICATE SMALLER SAMPLING VARIANCE (RELATIVE TO RANDOM SAMPLING).

Design	PRBO		LPBO	
Group 1:				
Proportional stratified random with 1 week strata	0.69	A	0.77	B
Optimal stratified random with 1 week strata	0.69	A	0.75	C
Weekending	0.72	A	0.87	A
Group 2:				
Systematic	0.67	B	0.62	E
Proportional stratified random with 1 sample per stratum (st1)	0.66	B	0.70	D

Notes: Means were calculated over six species at Point Reyes Bird Observatory (PRBO) and 46 species at Long Point Bird Observatory (LPBO), and over five sampling frequencies for Group 1 and six sampling frequencies for Group 2. Values in a column with the same letters were not significantly different in paired comparison of means from a repeated-measures ANOVA for each group with log-transformed design effect as the response variable and design and sampling frequency as factors (Bonferroni *t*-tests with experimentwise $\alpha = 0.05$).
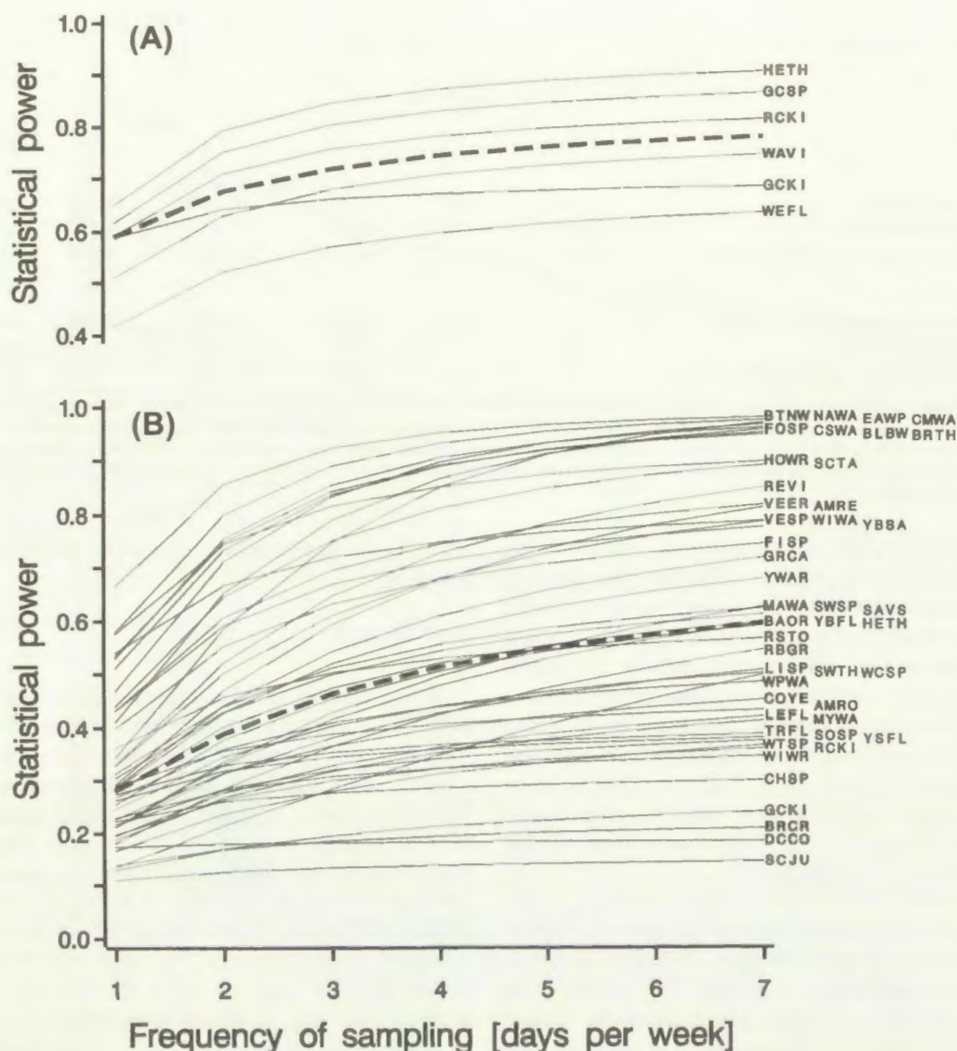
FIGURE 6. Statistical power to detect a 50% change in counts over 25 year and over a range of sampling frequencies at two migration monitoring stations, assuming random allocation of count days. Thin solid lines show the power for each species (species abbreviations at right of plot; see Appendix A for key). The thick solid line shows the median power at each sampling frequency.

discuss next. We finish by discussing the comparison of sampling designs.

LIMITATIONS OF THE POWER ANALYSIS

Our study was not designed as a rigorous evaluation of the ability of the two monitoring stations to detect population trends. To treat all species at both stations in the same manner, and to make the analysis tractable, we made a number of simplifying assumptions. Thus, a number of caveats should be made regarding our results.

First, we assumed that population trends could be estimated using a linear regression of annual indices against time on the log scale. However, for many species the indices do not conform very closely to the assumptions of a linear regression model. A number of other statistical models of trend could be used (such as non-parametric models or empirical smoothing, with different error models; Gerrodette 1987, Thomas and Martin 1996, Thomas 1997), which would almost certainly produce different estimates of power. In addition, we calculated annual indices as the mean of the logged daily counts. However, much of the
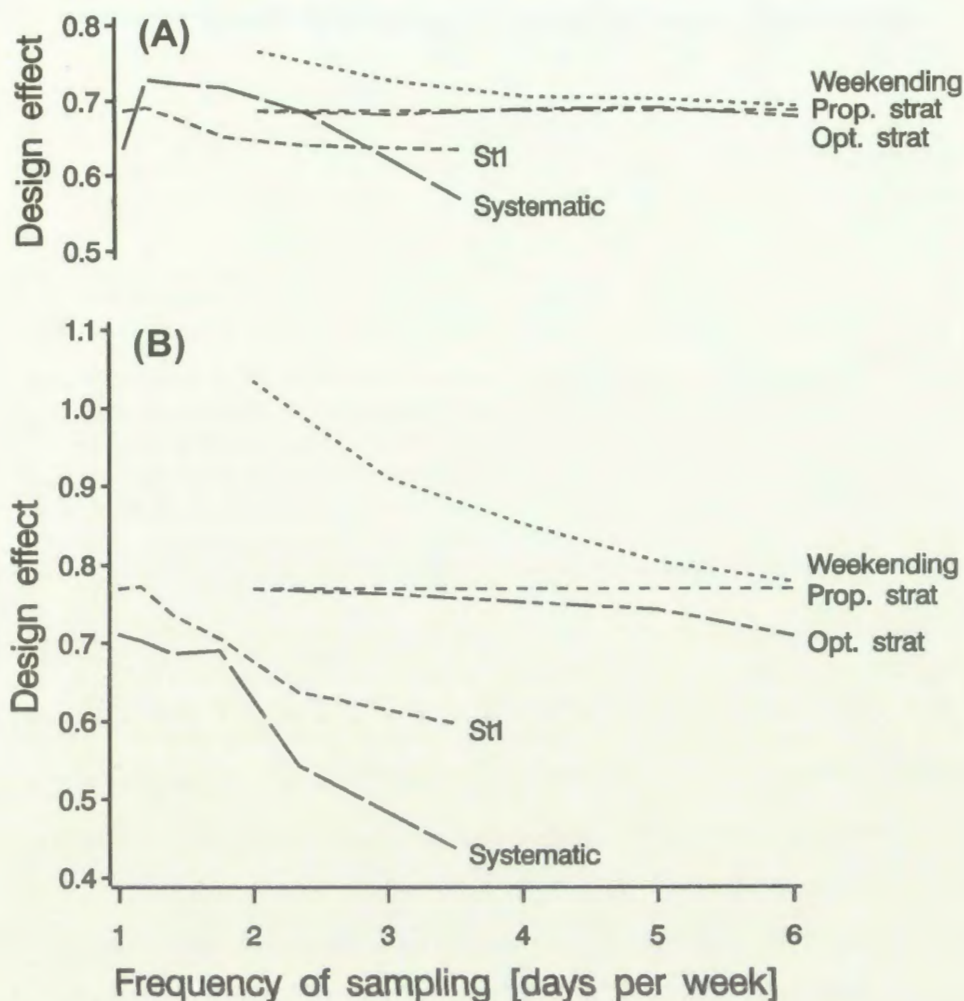
FIGURE 7. Design effect of five sampling designs over a range of sampling frequencies at two migration monitoring stations. Each line is the geometric mean design effect for the sampling design over six species at Point Reyes Bird Observatory and 46 species at Long Point Bird Observatory. Smaller design effects indicate smaller sampling variance (relative to random sampling). See Appendix A for key to species codes.

day-to-day variation in counts may be attributed to environmental factors such as wind direction and phase of the moon. Multiple regression approaches can be used to correct for these factors (e.g., Hussell et al. 1992, Pyle et al. 1994), which can lead to less variability in the trend estimates (Pyle et al. 1994) and increased power to detect trends.

The second caveat regards the selection of data. Our method of selecting a migration period for each species was designed to exclude data collected during the period when the species was not migrating, and also to produce migration periods that were an integer number of weeks in length. Our criterion of

excluding species with mean daily counts of less than 1.0 was designed to minimize the bias associated with adding a constant to the daily counts before log transformation (see below). All such criteria are arbitrary in nature, and differences among analysts will undoubtedly lead to different data being selected and thus to different estimates of power.

Third, we estimated power at only one magnitude of trend (50% over 25 years). Statistical power is dependent upon the size of trend, number of years of monitoring and significance criterion ("α-level"). Thus, even species with low power using our criterion will show a statistically significant trend if the
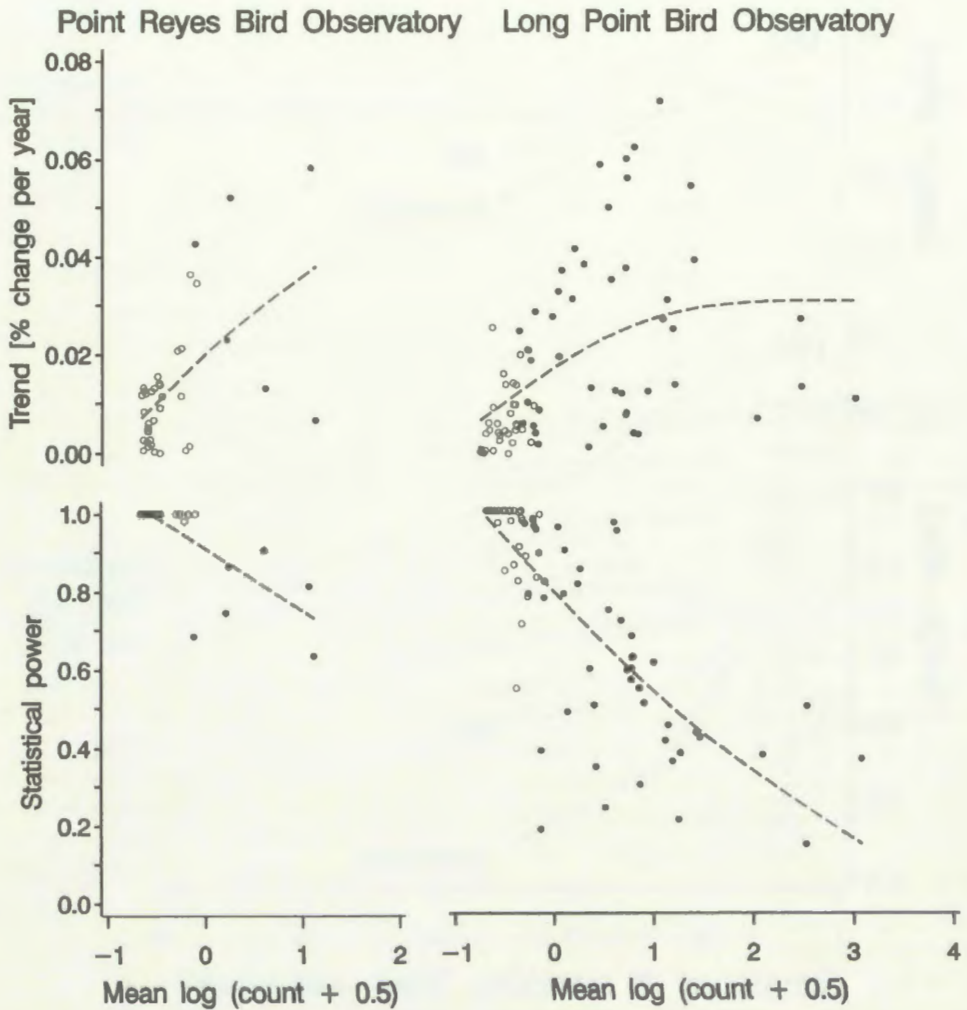
FIGURE 8. Relationship between absolute estimate of trend and statistical power (y-axis) and mean log-transformed count (x-axis) at two migration monitoring stations. Species that were excluded from our analysis (i.e., with a mean count of less than 1.0) are plotted with an open circle; those that were included are plotted with closed circles. A cubic spline (tension parameter 0.75) has been added to illustrate the pattern.

rate of change is steep enough, if there are enough years of data, or if the significance criterion is raised to greater than the customary level of 0.05. In addition, our analysis used two-tailed tests for population change, while the Partners in Flight criteria only refer to population declines. If we were not interested in detecting population increases then one-tailed tests could be used, which would result in higher statistical power (Gerrodette 1987, Cohen 1988, Steidl and Thomas 2001).

Lastly, we should emphasize that statistical power is a measure of the precision of an estimate,

assuming that the estimator is unbiased. In other words, we assumed that the annual indices, on average, reflect the true patterns of population change in the species they measure. There are two reasons why this may not be the case. Firstly, the estimate of trend will be biased if the proportion of the population that is counted varies with population size (see Sauer and Link *this volume*). The proportion of birds counted is often called the detectability, and cannot be measured directly from count data. Dunn and Hussell (1995) review the factors that may lead to differences in detectability in migration monitoring, and Sauer and

Link (*this volume*) show the statistical consequences. The second possible source of bias in the trend is the statistical procedure used to derive the estimate. In this case we added a constant before log transforming the counts. This is known to bias trend estimates (Link and Sauer 1994, Thomas 1997), causing underestimates of the size of the trend and also decreasing the variance of the estimate. The bias is greater at low abundances, where the constant is large relative to the counts, and so tends to swamp out any natural variability. Thus low abundance species will tend to have artificially high estimates of statistical power (because these estimates are based on the variance of the trend estimate), when in fact the bias towards no trend means that there is little chance of detecting a change in population size if one occurs. In our data, including all species observed at the two stations, there was a clear correlation between the size of species' trend estimates and their mean count, and between statistical power and mean count (Fig. 8). To control this bias, species with low counts are usually excluded from trend analyses that involve log-transformed data (e.g., linear route regression analysis of the Breeding Bird Survey; Geissler and Sauer 1990), as we did in our analyses by excluding species with a mean daily count of less than 1.0. This reduced the correlation between trend estimate and mean count, but the relationship between statistical power and mean count was still quite strong at Long Point (closed circles in Fig. 8). It is thus possible that our estimates of power for low-abundance species were overly optimistic.

In conclusion, due to the limitations outlined above, we caution readers not to place too much emphasis on the actual levels of power assigned to individual species, especially those with low mean counts. Further research is required before we can evaluate the importance of these limitations for assessing the ability of the two stations to detect trends. We do feel, however, that our results regarding the variation in statistical power with sampling frequency are qualitatively robust to these limitations. Because sampling variance increases at an increasing rate with decreasing frequency of sampling, missing count days will always be more important when the overall frequency of sampling is low. We discuss this further in the next section.

IMPORTANCE OF SAMPLING FREQUENCY

Gaps in the daily counts introduce sampling variance into the annual indices, which increases the unexplained variability about the trend line, and thus decreases the statistical power. Our results show that power declines at an accelerating rate with decreasing sampling frequency (Fig. 6), due to the accelerating rate of increase in sampling variance. Hence, a few missing count days have little effect on the power to detect trends for these species, but the effect of additional missing days becomes greater the more there are. Species most affected tended to be those with highest power, that is, those that are the best monitored.

We are reluctant to make general recommendations about a minimum frequency of sampling that should be used, because much depends upon the individual circumstances of each station. However it is plain from our results that frequencies of one or two days per week will likely lead to annual indices for most species that are too imprecise to be able to detect large population trends if they occur. This is supported by the recommendation in Hussell and Ralph (1998) that sampling take place on at least 75% of days within a species' migration window. In addition, if analysis methods that incorporate weather variables are to be used then a large sample of days is required to detect consistent effects (Francis and Hussell 1998).

COMPARISON OF SAMPLING DESIGNS

Many of the differences between sampling designs are explained by two features of the data: the strong seasonality in counts for each species (Fig. 4), and the autocorrelation between counts taken on days that are close to one another (Fig. 5). Seasonality favors sampling designs that lead to counts being taken throughout the migration period; these will tend to consistently capture seasonal variation in counts and, because of this consistency, will have low sampling variance. All of the designs we compared limited the number of count days per week (or per *k* in the case of systematic sampling and *st1*), and thus had lower sampling variances on average than simple random sampling (i.e., average design effects <1.0). The autocorrelation between counts taken on adjacent days was relatively high (median $\cong 0.3$), and in most species decreased with increasing distance apart of count days (Fig. 5). Thus in the weekending design, with two count days per week, the count data collected on the second day of each "weekend" contained similar information to that already collected on the first day, making it less efficient than the other designs. At higher sample frequencies, additional random days were sampled during the week, and allocation of count days became similar to that of proportional stratified random sampling.

Comparing the systematic and *st1* designs,

Cochran (1977:219–221) has shown that systematic sampling will necessarily have a lower design effect than *stl* if the shape of the correlogram is concave upwards. Many species in the Long Point spring migration data set exhibited correlograms that approximated this pattern, especially at four days between counts and less (Fig. 5). Species in the Point Reyes fall data set did not tend to show the same pattern, and systematic sampling did not appear to be better than *stl*, although the small sample size prevents us from making any strong inferences about differences between the designs at this station.

Overall, our results indicate that systematic sampling should be preferred over the other designs if sampling variance is the sole criterion. Systematic sampling also has the advantage of being easy to implement. However, three drawbacks of the design should be mentioned (see Cochran 1977 and references therein for details). The first is that there is no reliable way of calculating the sampling variance from the sample. Treating the data as if it came from a random sample will almost always result in an overestimate of the true variance. This is not a problem in the current application because the variance of the annual indices does not need to be calculated to estimate the variance of the trend estimate. Secondly, systematic sampling is very imprecise if the counts show a linear trend within the season. This could be a problem for migratory species with breeding populations at the station, because the abundance of birds will tend to show a monotonic increase in the spring and decrease in the fall. However, many migration stations use only the number of new birds captured in mist nets as the daily count (e.g., Point Reyes), or attempt to exclude the resident population from analysis using other techniques (see Dunn and Hussell 1995). We saw little evidence of linear trends in counts at either station in this study (Fig. 4). Where linear trends are suspected, the problem of imprecision may be avoided by making simple adjustments to the formula for calculating annual indices ("end corrections," Cochran 1977). Thirdly, systematic sampling can be very imprecise if the counts show regular periodic variation within the season. This is a potential problem at some stations, such as Long Point, where bird abundance at the station is thought to be related to the regular passage of weather fronts. However, even without adjusting the counts for environmental variables, we found systematic sampling to have a lower design effect than stratified random at Long Point. We thus suspect that imprecision due to periodic variation is unlikely to be a major concern.

Despite the higher expected design effects, stratified random designs have the advantage that it is always possible to derive an unbiased estimate of the sampling variance from the sample. Because of this, stratified random designs are often preferred when little or nothing is known about the distribution of the data being sampled (such as the possibility of periodic variation or linear trends in abundance within the migration period). In this study, optimal allocation performed similarly to proportional allocation at low frequencies of sampling, where the constraint of at least one sample per stratum made the two designs very similar. Even at higher frequencies, the optimal design was only slightly better. It appears that the optimal allocation for individual species were different enough to prevent the compromise optimum allocation from providing much overall benefit. In addition, it should also be noted that a real-life implementation of the optimal allocation formula would not have the true within-stratum variances to work with, but only estimates based on previous years' sampling. We conclude that optimal allocation schemes are probably not worth the extra effort involved in their implementation. If a stratified random design is chosen, perhaps because little is known about the region being sampled, then we recommend a proportional scheme over an attempt at optimization.

Sampling only on two adjacent days (weekending) produced the highest sampling variance of all the designs we compared. We recommend that if constraints are such that sampling can only occur on two days per week, then sampling one day in the weekend and once in the middle of the week should be encouraged.

CONCLUSIONS

1. The frequency of daily sampling will likely have an important effect on the ability of a migration monitoring station to detect trends for some species, regardless of the statistical method used to calculate annual indices and trends.

2. The effect of missing count days is not great when the overall frequency of sampling is high, but increases with decreasing sampling frequency. Species that are well monitored (i.e., high statistical power) are more strongly affected than species that are not well monitored. Single stations that operate on 1–2 days per week are unlikely to be able to detect large changes in the abundance of species that would be well monitored at higher frequencies of sampling.

3. Of the designs we compared, systematic sampling (i.e., counting at regular intervals) performed the best for the Long Point spring migration data,

yielding the lowest sampling variance over a range of sampling frequencies. We had too few species to tell whether the systematic design was any better than the others for the Point Reyes fall data, but it did not appear to be very much different. There are a number of problems associated with systematic sampling, but these are unlikely to be important in the context of migration monitoring. Also, a major advantage of this design is that it is easy to implement, because the sampling days are regular and predictable (e.g., every second day).

4. Stratified random sampling (i.e., dividing the season into one week strata and counting on random days within each week) yielded the next lowest sampling variance for the Long Point spring migration data. Proportional stratified random sampling (i.e., the same number of counts in each week) may be preferred over systematic sampling under conditions where an unbiased estimate of the variance of the annual indices is required. We also evaluated an optimal allocation scheme, which allocated more sampling effort to weeks in which the abundance of birds was more variable. This performed slightly better than proportional allocation, but would be difficult to implement in practice and so is not recommended.

5. Sampling only at weekends produced the largest estimates of variance of the designs we compared. If the station can only be open two days a week, we recommend counting once at the weekend and once during the middle of the week.

ACKNOWLEDGMENTS

APPENDIX. KEY TO SPECIES CODES IN FIGURES

Code	Species)
DCCO	Double-crested Cormorant (*Phalacrocorax auritus*)
YBSA	Yellow-bellied Sapsucker (*Sphyrapicus varius*)
YSFL	Northern Flicker (*Colaptes auratus*)
EAWP	Eastern Wood-Pewee (*Contopus virens*)
YBFL	Yellow-bellied Flycatcher (*Empidonax flaviventris*)
TRFL	Alder/Willow Flycatcher (*E. alnorum/E. traillii*)
LEFL	Least Flycatcher (*E. minimus*)
WEFL	Pacific-slope Flycatcher (*E. difficilis*)
WAVI	Warbling Vireo (*Vireo gilvus*)
REVI	Red-eyed Vireo (*V. olivaceus*)
BRCR	Brown Creeper (*Certhia americana*)
HOWR	House Wren (*Troglodytes aedon*)
WIWR	Winter Wren (*T. troglodytes*)
GCKI	Golden-crowned Kinglet (*Regulus satrapa*)
RCKI	Ruby-crowned Kinglet (*R. calendula*)
VEER	Veery (*Catharus fuscescens*)
SWTH	Swainson's Thrush (*C. ustulatus*)
HETH	Hermit Thrush (*C. guttatus*)
AMRO	American Robin (*Turdus migratorius*)
GRCA	Gray Catbird (*Dumetella carolinensis*)
BRTH	Brown Thrasher (*Toxostoma rufum*)
NAWA	Nashville Warbler (*Vermivora ruficapilla*)
YWAR	Yellow Warbler (*Dendroica petechia*)
CSWA	Chestnut-sided Warbler (*D. pensylvanica*)
MAWA	Magnolia Warbler (*D. magnolia*)
CMWA	Cape May Warbler (*D. tigrina*)
MYWA	Yellow-rumped Warbler (*D. coronata*)
BTNW	Black-throated Green Warbler (*D. virens*)
BLBW	Blackburnian Warbler (*D. fusca*)

APPENDIX. CONTINUED

Code	Species
WPWA	Palm Warbler (*D. palmarum*)
AMRE	American Redstart (*Setophaga ruticilla*)
COYE	Common Yellowthroat (*Geothlypis trichas*)
WIWA	Wilson's Warbler (*Wilsonia pusilla*)
SCTA	Scarlet Tanager (*Piranga olivacea*)
RSTO	Eastern Towhee (*Pipilo erythrophthalmus*)
CHSP	Chipping Sparrow (*Spizella passerina*)
FISP	Field Sparrow (*S. pusilla*)
VESP	Vesper Sparrow (*Pooecetes gramineus*)
SAVS	Savannah Sparrow (*Passerculus sandwichensis*)
FOSP	Fox Sparrow (*Passerella iliaca*)
SOSP	Song Sparrow (*Melospiza melodia*)
LISP	Lincoln's Sparrow (*M. lincolnii*)
SWSP	Swamp Sparrow (*M. georgiana*)
WTSP	White-throated Sparrow (*Zonotrichia albicollis*)
WCSP	White-crowned Sparrow (*Z. leucophrys*)
GCSP	Golden-crowned Sparrow (*Z. atricapilla*)
SCJU	Dark-eyed Junco (*Junco hyemalis*)
RBGR	Rose-breasted Grosbeak (*Pheucticus ludovicianus*)
BAOR	Baltimore Oriole (*Icterus galbula*)

USE OF MIST NETS FOR MONITORING LANDBIRD FALL POPULATION TRENDS, AND COMPARISON WITH OTHER METHODS

PETER BERTHOLD

Abstract. In Central Europe, a long-term trapping program based on mist netting has been carried out since 1972. In this "MRI-program," about 40 migratory landbird species are studied annually throughout the fall migratory period. Netting figures from this strictly standardized program are used to monitor trends of populations. Comparisons with other data show that the method can detect trends similar to those from breeding-season studies. Some illustrative examples are presented.

Key Words: migration, mist net, MRI-program, population monitoring.

In 1972, the Max Planck Research Centre for Ornithology, Vogelwarte Radolfzell, initiated a long-term bird trapping program that focuses on a variety of research fields, including migration studies. One of its main purposes is to monitor trends in population size. The program is based on mist netting of about 40 migratory landbird species during the complete fall migratory period. The program was named "Mettnau-Reit-Illmitz-Program" ("MRI-program"), after the large trapping stations in Germany and Austria where it was initially launched. It was extended to five stations in 1992 in eastern Germany and Russia. Up to 1993, about 400,000 individuals (first traps and retraps) were captured. One of the essential characteristics of the MRI-program is to keep the basic conditions for trapping birds as constant as possible, and it is the most standardized long-term trapping program in the world.

Several studies have used annual MRI trapping totals to detect long-term population trends. Trends from the MRI-program for the 10-year period 1974–1983 (Berthold et al. 1986), and from the Mettnau station for the 20-year period 1972–1991 (Berthold et al. 1993) and the 25-year period 1972–1996 (Berthold et al. 1998) were validated through comparison with trends from other studies. Here we show examples for four species.

METHODS

Nets are set up every year on June 30 and used continuously until November 7. Only in 1987 was there no trapping activity, due to flood conditions. All operations of the nets, handling of birds, and data collection have been described in detail elsewhere (Berthold et al. 1991, Kaiser and Berthold *this volume*). All aspects of operations were strictly standardized, even to the extent of nets being set at the same height above the ground each year, and with the same distances between shelf strings.

The study areas are large, mostly with climax vegetation. In addition, vegetation is trimmed to a constant height around the area of the nets.

Because netting is so standardized, there is no need to present results as effort-corrected capture totals, and we use total birds captured within the species-specific fall migration period as the annual index of abundance. Long-term trends in annual indices were calculated as the slope of the regression of annual total number of birds captured on year (Berthold et al. 1993).

RESULTS

Here we compare population data for four species from the MRI-program with independent data from other sources.

The Robin (*Erithacus rubecula*) is one of the few passerine species with no reported recent decline in a central, western, or northern European population. In fact, its populations are considered exceptionally stable (Bezzel et al. 1992, Bauer and Berthold 1997). This is reflected in extremely constant indices according to the Common Breeding Birds Census (CBC) in Britain since the middle 1960s (Marchant et al. 1990; Fig. 1). A strikingly similar pattern was found in the annual netting totals of the MRI-program, with one of the lowest variations from year to year (coefficient of variation 18.96%; Fig. 1).

The Redstart (*Phoenicurus phoenicurus*) is known to be a species with decades-long and essentially continuous decline in large parts of Europe (Hildén and Sharrock 1982, Bauer and Berthold 1997). Only recently have some European populations appeared to stabilize or even increase slightly (e.g., Marchant et al. 1990). Such a long-term decline with a tendency to a possible recent stabilization is also shown in the netting figures from the MRI-program (Fig. 2).

In the Whitethroat (*Sylvia communis*) various investigations have found a population crash of
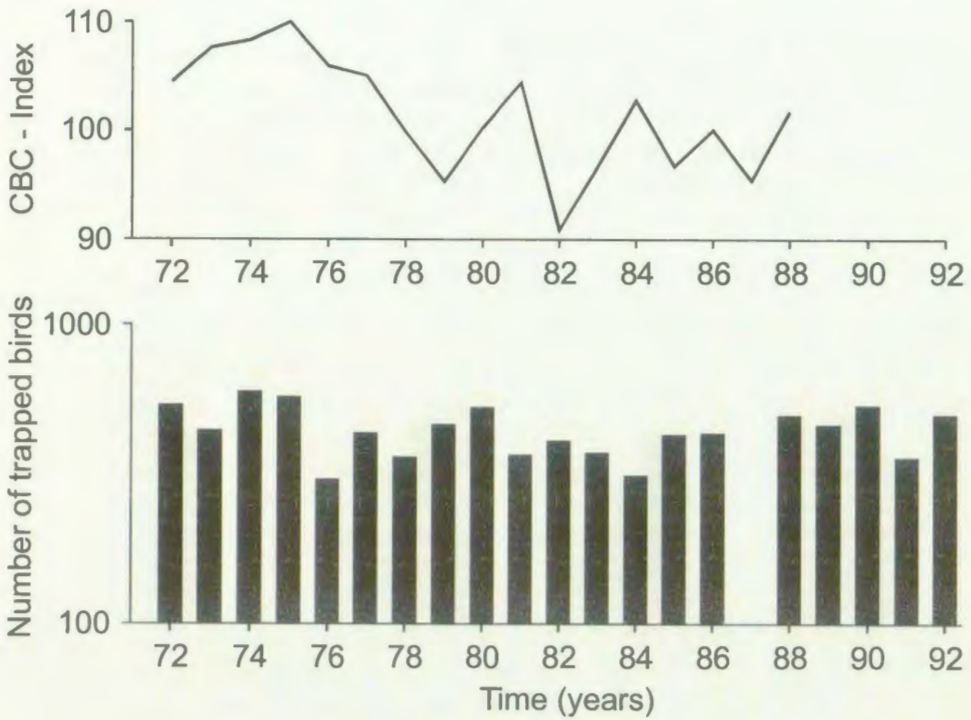
FIGURE 1. Population changes in the Robin. Above: CBC (Common Birds Census) indices from the British Isles (after Marchant et al. 1990). Below: fall netting totals from the MRI-program, Mettnau station, southern Germany (after Berthold et al. 1993).
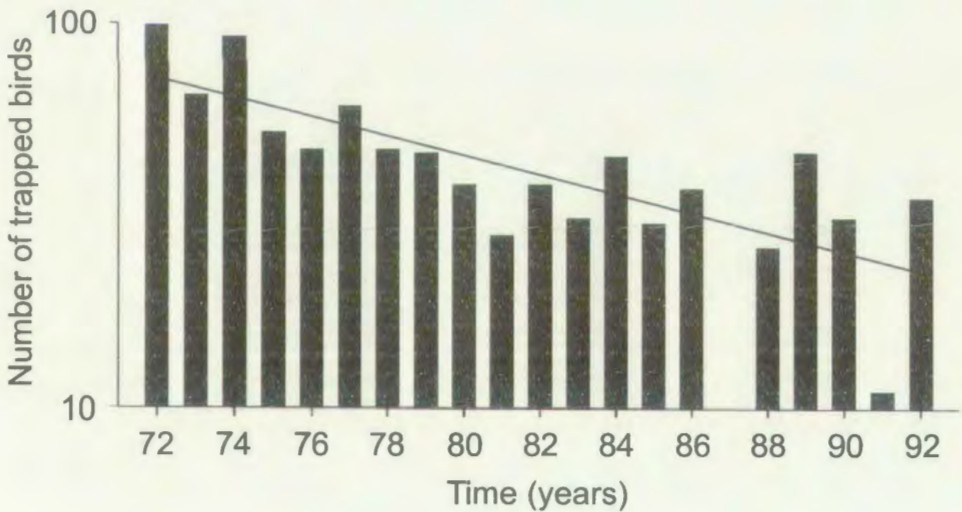

FIGURE 2. Decline in the Redstart, indicated by netting figures (fall totals) and the regression line from the MRI-program, Mettnau station, southern Germany. Slope of the regression analysis = -0.80, N = 909, P < 0.001 (after Berthold et al. 1993).

about 50 to 75% in large parts of Europe since 1968–1969, and a number of local populations have completely disappeared since 1969 (e.g., Berthold 1974, Bauer and Berthold 1997). An exceptionally severe drought in the Sahel zone, south of the Sahara, was recognized as the main cause of this sudden decline (Winstanley et al. 1974). We had just started standardized mist netting in a special "warbler program" on the Mettnau Peninsula in southern Germany in 1968, that is, one year before the population crash of the Whitethroat. These netting activities then merged directly into the MRI-program. They provided the unique opportunity to compare the observed population crash as assessed by our netting figures with the one deduced from the CBC by the BTO in Britain. The patterns of the crash and of the subsequent low population level obtained by the two methods are largely identical (Fig. 3).

The Willow Warbler (*Phylloscopus trochilus*) showed somewhat stable CBC indices in Britain from about 1965 to 1980 (Marchant et al. 1990). Then, it underwent a severe and almost continuous decline of about 60% over the following decade, with only very slight short-term recovery thereafter (Peach and Baillie 1993; Fig. 4). At the Mettnau station, netting rate was fairly constant until 1980, but since 1981 has been gradually dropping. The total decline amounted to 70% between 1981 and 1993. Again, trend in capture rates closely matched the trend in CBC figures for Great Britain (Fig. 4).

DISCUSSION

Factors affecting numbers of migrant birds at a particular stopover site were thoroughly reviewed by Dunn and Hussell (1995). Standardization of effort is important in ensuring that as constant a proportion as possible of the birds that are actually present will be captured on each day and in each year (Ralph et al. *this volume a*). The MRI methodology ensures that this will be the case, such that variation in numbers of birds captured will not simply reflect variation in effort or capture technique.

A crucial aspect of standardization that is often ignored by migration monitoring stations is the need to maintain habitat in the same condition, and vegetation at the same height, from year to year. Even if the same species and number of individuals were present from day to day, growth in vegetation alone could cause changes in the numbers of birds captured. For example, after vegetation grows higher than nets, a higher proportion of birds may fly over nets and avoid capture. Moreover, birds have habitat preferences that will cause them to move elsewhere if there are changes in preferred habitat type and structure (Bairlein 1981; see also Mallory et al. *this volume* regarding capture bias related to habitat structure). MRI stations control vegetation to prevent trends in capture rates over time that could be caused by change in vegetation rather than by change in bird numbers.
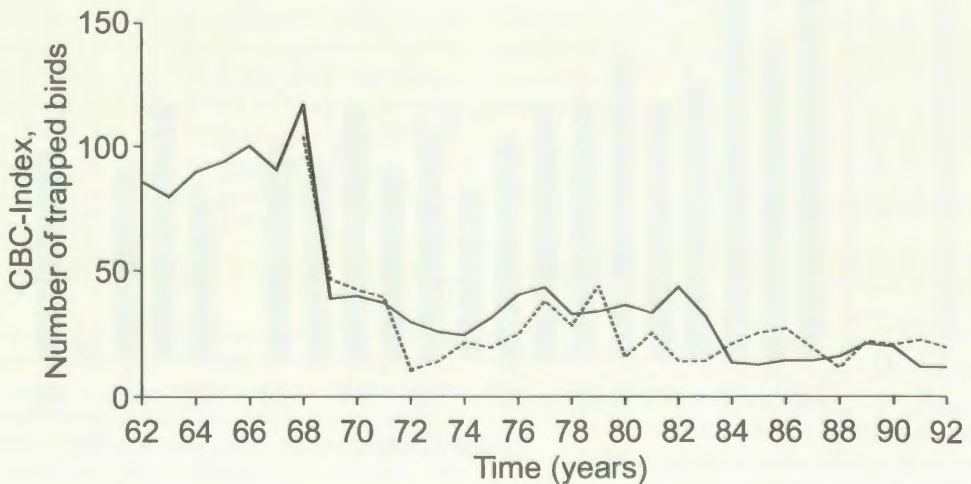

FIGURE 3. Population changes in the Whitethroat. Solid line: CBC indices from the British Isles (after Marchant et al. 1990). Broken line: fall netting totals from the MRI-program, Mettnau station, southern Germany (after Berthold et al. 1993).
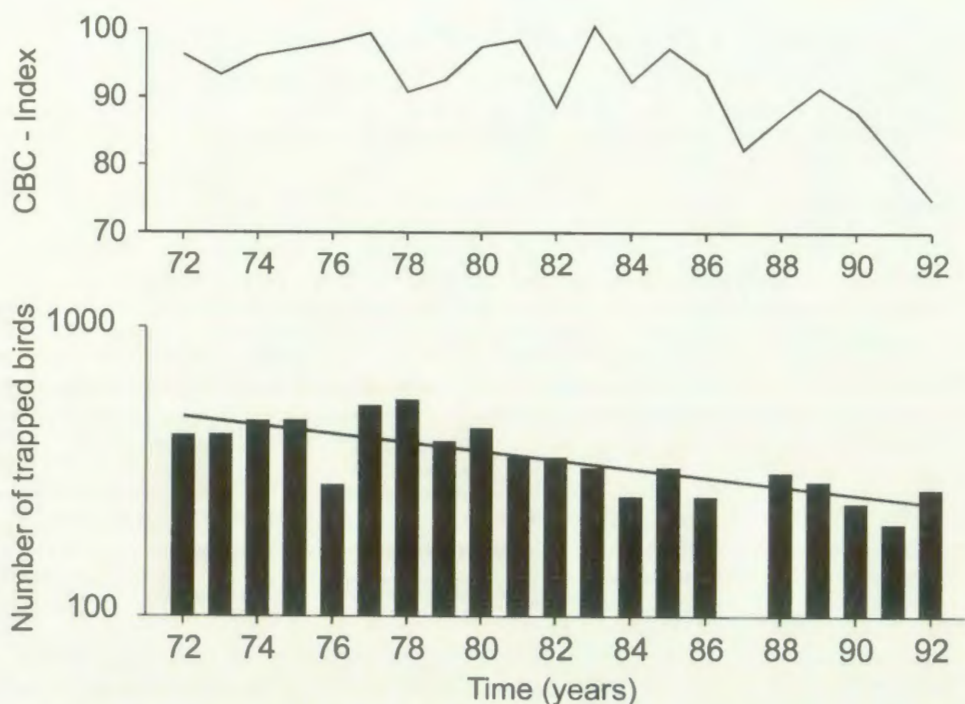
FIGURE 4. Population changes in the Willow Warbler, indicated by netting figures (fall totals) and the regression line from the MRI-program, Mettnau station, southern Germany. Slope of the regression analysis = -0.78, N = 7,240, P < 0.001.

Numbers of migrants captured are also affected by factors other than effort and habitat change, particularly weather (Dunn and Hussell 1995). Although daily sampling reduces the chances that a few days of large migratory flights will not dominate results, log-transformation of daily captures should be routinely used as a minimum treatment for migration counts (Dunn and Hussell 1995). More sophisticated analyses can be used to take into account data on season, weather, and other factors (Dunn et al. 1997, Francis and Hussell 1998), which further reduces variance in the data set and increases precision of population trend estimates (Dunn and Hussell 1995). However, even without any of these treatments, it is clear from the examples in this paper that migration capture data can mirror trends in breeding-population size as determined from independent data sources.

The examples shown compare migration capture data from Germany to breeding-population trends in Great Britain and Germany. Such agreement in the trends as demonstrated would only be expected for species that are changing in the same way over large areas. This will not be true of all species, and one of the unanswered questions for most migration monitoring stations is knowledge of the origin of migrants coming through their sites. In some cases birds from different breeding populations can be distinguished by plumage and measurement differences, and it is important that these data be collected to help identify the breeding populations that are being sampled (Berthold et al. 1991).

ACKNOWLEDGMENTS

A COMPARISON OF THREE COUNT METHODS FOR MONITORING SONGBIRD ABUNDANCE DURING SPRING MIGRATION: CAPTURE, CENSUS, AND ESTIMATED TOTALS

ERICA H. DUNN, DAVID J. T. HUSSELL, CHARLES M. FRANCIS, AND JON D. MCCRACKEN

Abstract. We compared long-term trends (1984–2001) based on three types of spring migration count data, from the three migration monitoring stations at Long Point Bird Observatory (southern Ontario), for 64 species. The three count methods consisted of daily capture totals from banding, sightings from a daily 1-h count on a fixed route ("census"), and "estimated totals" (ETs). The latter were estimates of birds detected in each study area each day, based on results from banding, census, and unstandardized "other observations." In the majority of species, ET annual indices were significantly positively correlated with both banding and census indices. Banding was not standardized, and variance of annual banding indices was higher than for other count methods, but trends based on banding alone were similar in magnitude to trends from census alone. Relative to trends based on banding or census alone, ET trends were positively biased, possibly as a result of change in estimation methods over time. Nonetheless, because ETs combine data from a variety of count methods, more species can be monitored, with greater precision, than by using one count method alone. Comparison of trends among stations suggested an influence of habitat change at one location. Biases should be minimized with strict standardization of all component count methods, adherence to a clear protocol for ETs, and management of vegetation to prevent systematic habitat change.

Key Words: banding, Breeding Bird Survey, census, estimated totals, habitat bias, migration monitoring, population trend, trend analysis.

Standardized counts of migrating birds can be used to calculate population trends, which have been shown to correlate with trends from the Breeding Bird Survey (BBS; Hussell et al. 1992, Dunn and Hussell 1995, Dunn et al. 1997, Francis and Hussell 1998). Recommended guidelines for migration counting (Hussell and Ralph 1998) state that each monitoring station should select the count method that is most suitable for the location, which may include daily banding, route surveys, counts of birds moving past a fixed point, or some combination of count methods. Different counting techniques may be more suitable for certain types of migratory species, and magnitude of counts will differ among methods, but as long as count protocol at any station is followed consistently, trends should be the same regardless of the type of migration count. However, this assertion has not previously been tested.

Here we present results of separate trend analyses for different count methods from the Long Point Bird Observatory (LPBO), in southern Ontario. At each of three stations (all within 30 km of one another), there was daily banding and a daily "census" (approximately 1-h survey of birds along a fixed route). In addition, records were kept of all birds detected during these and other activities in the day ("other observations"). At the end of the day, all personnel gathered to agree on "estimated totals"

(ETs). These were estimates of the total number of individuals detected in the defined study area that day, based on all available data. We estimated trends based on banding totals, census counts, and ETs separately, then compared them with each other and with trends from BBS.

Whatever methods are selected for migration counts, it is important to use them in a standardized and consistent manner from day to day and year to year, so that variation in counts will not reflect changes in methods (Ralph et al. *this volume a*). At Areas 1 and 2 (the two stations on the Long Point peninsula), early successional dune habitat consists of constantly shifting shorelines and vegetation patches, which has required periodic change in net locations. Moreover, the number of nets that can be operated safely, and the effectiveness of the nets, varies with wind strength at these exposed locations. Areas 1 and 2 each had a Heligoland trap (Woodford and Hussell 1961) that was often used in addition to nets, or in place of nets when weather precluded netting. Banding at Area 3 (the third station, at the mainland end of Long Point) was more standardized in net placement, but not necessarily in number of nets operated or daily operating hours. The census, on the other hand, has always been conducted in a consistent manner at all stations. A comparison of trends based on census or banding

alone should therefore allow us to examine the effect of standardization in banding on population trends. Comparison with ET trends should indicate the relative importance of each survey method for particular species, and show whether combining data from different count methods adds to the effectiveness of monitoring.

METHODS

Data were collected from mid-April to early June, 1984–2001, at LPBO's three stations on Long Point, on the north shore of Lake Erie. For each of 64 species of common migrants (Table 1), we calculated annual indices for three data sets (daily banding totals, census, and ETs) for each station separately, and in a composite analysis that produced indices for all stations together.

Banding data were the raw daily banding totals (new captures only), unadjusted for effort. Ideally, capture totals should be corrected for effort either through calculating birds per unit effort (e.g., net-h, trap-h; or, for Heligoland traps, trap-drives), or through including effort as a covariate. However the effort data have not been computerized, and extraction was ruled out for this analysis because time and cost were prohibitive. Even if the data were available, there is no simple way of combining effort-corrected results from each type of capture method.

The Long Point "census" was not a true total count, but rather a daily survey that recorded all birds identified by sight or sound along a fixed route that wound throughout the study area. The census was usually (but not always) done by one observer. Personnel often changed from day to day, and nearly always from year to year, so long-term trends should not be affected by systematic observer bias. Each walk lasted about 1 h and was conducted in early or mid-morning. The route at Area 1 was altered in 1986 and the route at Area 2 in 1988 to accommodate loss of area due to erosion, but otherwise the routes remained fixed.

"Other observations" consisted of sightings within the defined study area additional to census, but there was no standardization of the amount of time expended or number of observers contributing. As noted above, the "defined study area" was altered somewhat at Area 1 in 1986 and at Area 2 in 1988.

ETs were derived jointly at each day's end by all participants. The ETs were intended to be carefully considered estimates of numbers detected in the study area that day, based on banding, census, and other observations. Double-counting was avoided where possible by taking into account numbers retrapped and likelihood that independent sightings were actually of the same birds. The ET procedure was devised in part to overcome the problems posed by a banding program that could not be fully standardized, and the census was intended to provide consistent daily input. ETs were the best estimate by personnel at the station of birds detected each day, regardless of variation in effort put into the various component counts.

Data were included in analyses only for dates within a

species-specific time period judged to constitute the spring migration season of each species at Long Point (Hussell et al. 1992). Annual indices were calculated from a regression procedure designed to assign variability in daily counts to date, weather, moon phase, and year (Francis and Hussell 1998). Composite analyses (designed to produce indices combining data from all three stations) also included dummy variables for station, and for interactions between station and all other variables except those for year. Analysis methods are described in detail elsewhere (Hussell et al. 1992, Francis and Hussell 1998), and the following gives only a brief overview.

The dependent variable in the regression analyses was log (daily count + 1), in which the "daily count" was either the daily number of newly-banded birds, the number recorded on the daily census, or the daily estimated total (i.e., the analyses were run three times for each species). The constant was added to allow transformation of zeros, and 1 was chosen because it is the minimum change that can occur in daily counts. The log-transformed daily count was the dependent variable in a regression that included independent variables for year (dummy variables for each year except for one reference year; e.g., Y79 = 1 if the year was 1979, otherwise Y79 = 0), date (first through fifth order day terms), first and second order moon phase variables (days from nearest new moon and its square), and 12 weather variables. Weather variables were constructed using data from Erie, Pennsylvania (40 km south of the study locations), as detailed in Francis and Hussell (1998), and included daily values for horizontal visibility, cloud cover, first and second order terms for temperature difference from normal, and first and second order terms of four wind variables. Annual abundance indices were calculated from the coefficients of the dummy variables for year that were estimated in the regression. The annual abundance index was the adjusted mean for year plus one-half of the error variance of the regression (so the corrected index in the original scale represented an estimate of the mean rather than of the median; see references in Hussell et al. 1992), back-transformed to the original scale. The adjusted mean for year represented the mean of the transformed daily counts under standardized conditions of day, weather, and moon. The annual abundance indices therefore represented the estimated numbers of birds that would be counted each year on the same average date in the season, under average weather and moon conditions.

Trends were calculated as the slope from the weighted linear regression of log-transformed annual indices on year. Weights were proportional to the number of daily counts in the year represented by the index.

We performed bivariate correlations between annual banding and census indices to determine level of correspondence. To determine whether banding and census had independent effects on ET, we performed multiple regressions for each species, with log-transformed ET annual index as the dependent variable, and log-transformed banding and census indices as independent variables.

To detect difference in trends according to count method, we conducted species-specific analyses of

TABLE 1. RELATIONSHIPS AMONG ANNUAL INDICES (1984–2001) FROM BANDING AND CENSUS (DATA FROM THREE STATIONS COMBINED) AT LONG POINT, ONTARIO

Species	Banding-census r[a]	Contribution to ET[b]		R²[c]
		Census	Banding	
Black-billed Cuckoo (*Coccyzus erythropthalmus*)	0.66**	***		0.63
Red-headed Woodpecker (*Melanerpes erythrocephalus*)	0.92***	***		0.93
Yellow-bellied Sapsucker (*Sphyrapicus varius*)	0.74***	***		0.75
Northern Flicker (*Colaptes auratus*)	0.75***	***		0.83
Eastern Wood-Pewee (*Contopus virens*)	0.35	**	+	0.53
Yellow-bellied Flycatcher (*Empidonax flaviventris*)	0.41+	*	***	0.76
Least Flycatcher (*E. minimus*)	0.35		***	0.71
Eastern Phoebe (*Sayornis phoebe*)	0.77***	***	*	0.89
Great Crested Flycatcher (*Myiarchus crinitus*)	-0.04	**		0.28
Blue-headed Vireo (*Vireo solitarius*)	0.90***	*	**	0.90
Warbling Vireo (*V. gilvus*)	0.29	***		0.85
Philadelphia Vireo (*V. philadelphicus*)	0.67**		**	0.72
Red-eyed Vireo (*V. olivaceus*)	0.62**	+	***	0.73
Brown Creeper (*Certhia americana*)	0.85***		**	0.85
House Wren (*Troglodytes aedon*)	0.44+	***	***	0.86
Winter Wren (*T. troglodytes*)	0.76***	***	***	0.94
Golden-crowned Kinglet (*Regulus satrapa*)	-0.28	***	*	0.85
Ruby-crowned Kinglet (*R. calendula*)	0.74***	**	*	0.80
Blue-gray Gnatcatcher (*Polioptila caerulea*)	0.35	***	+	0.73
Veery (*Catharus fuscescens*)	0.59**	**	***	0.89
Gray-cheeked Thrush (*C. minimus*)	0.16	*	***	0.67
Swainson's Thrush (*C. ustulatus*)	0.56*	**	***	0.85
Hermit Thrush (*C. guttatus*)	0.67**		**	0.62
Wood Thrush (*Hylocichla mustelina*)	0.48*	*	***	0.71
American Robin (*Turdus migratorius*)	0.08	***		0.69
Gray Catbird (*Dumetella carolinensis*)	0.88***	**	*	0.91
Brown Thrasher (*Toxostoma rufum*)	0.79***	**		0.75
Tennessee Warbler (*Vermivora peregrina*)	0.81***	*	***	0.88
Nashville Warbler (*V. ruficapilla*)	0.78***	***		0.81
Yellow Warbler (*Dendroica petechia*)	0.73***	***	***	0.97
Chestnut-sided Warbler (*D. pensylvanica*)	0.70**		***	0.69
Magnolia Warbler (*D. magnolia*)	0.47*		***	0.85
Cape May Warbler (*D. tigrina*)	0.82***	**	*	0.86
Black-throated Blue Warbler (*D. caerulescens*)	0.67**		***	0.78
Yellow-rumped Warbler (*D. coronata*)	0.82***	***		0.78
Black-throated Green Warbler (*D. virens*)	0.67**	**		0.60
Blackburnian Warbler (*D. fusca*)	0.43+	*	*	0.55
Palm Warbler (*D. palmarum*)	0.36	***		0.75
Bay-breasted Warbler (*D. castanea*)	0.80***	*	*	0.80
Blackpoll Warbler (*D. striata*)	0.79***	***	**	0.91
Black-and-white Warbler (*Mniotilta varia*)	0.81***	*	+	0.73
American Redstart (*Setophaga ruticilla*)	0.59*		***	0.80
Ovenbird (*Seiurus aurocapilla*)	0.85***	**	***	0.93
Northern Waterthrush (*S. noveboracensis*)	0.79***	**	***	0.93
Mourning Warbler (*Oporornis philadelphia*)	0.33	**	***	0.72
Common Yellowthroat (*Geothlypis trichas*)	0.71**	**	**	0.80
Wilson's Warbler (*Wilsonia pusilla*)	0.20	**	***	0.78
Canada Warbler (*W. canadensis*)	0.34	+	**	0.61
Scarlet Tanager (*Piranga olivacea*)	0.60**	***		0.86
Eastern Towhee (*Pipilo erythropthalmus*)	0.75***	***	+	0.90
Chipping Sparrow (*Spizella passerina*)	0.79***	***	*	0.92
Field Sparrow (*S. pusilla*)	0.55*	**	**	0.76
Vesper Sparrow (*Pooecetes gramineus*)	0.54*	**	+	0.60

TABLE 1. CONTINUED

Species	Banding-census r[a]	Contribution to ET[b]		R[c]
		Census	Banding	
Savannah Sparrow (*Passerculus sandwichensis*)	0.83***	**		0.76
Fox Sparrow (*Passerella iliaca*)	0.63**	***	+	0.85
Song Sparrow (*Melospiza melodia*)	0.77***	*		0.59
Lincoln's Sparrow (*M. lincolnii*)	0.46+	***	***	0.86
Swamp Sparrow (*M. georgiana*)	0.81***	***		0.90
White-throated Sparrow (*Zonotrichia albicollis*)	0.85***		*	0.75
White-crowned Sparrow (*Z. leucophrys*)	0.73***	***	*	0.87
Dark-eyed Junco (*Junco hyemalis*)	0.79***	*		0.55
Rose-breasted Grosbeak (*Pheucticus ludovicianus*)	0.68**	***		0.85
Indigo Bunting (*Passerina cyanea*)	0.88***	*		0.76
Baltimore Oriole (*Icterus galbula*)	0.38	***		0.72

[a] Correlation coefficient between annual indices from banding and census.
[b] Significance of partial correlation coefficient in regression of ET indices on indices for banding and census, indicating whether the count method significantly influenced ET independently of the other count method (* = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$).
[c] Proportion of annual variation in ET indices explained by census and banding indices (R^2 of regression described in footnote b). All R^2 were significant (symbols not shown).

covariance with count method as the factor and year as covariate. We examined interactions between count method and year. Significant interactions indicated trends that differed in slope.

We compared variability in indices among count methods by calculating variance in the residuals from linear regressions of log-transformed indices on year (thereby removing variability related to long-term trends in the data).

To determine whether trends from different stations or those based on different count methods produced the same magnitude of trend (e.g., comparing the 64 species, trends based on census from Area 1 to those from Area 2), we conducted reduced major axis regression on pairs of trends (Bohonak 2002). If trends from the two sources correspond in magnitude, then the regression results would show an intercept of 0 and a slope of 1.

RESULTS

Analysis of annual indices based on data pooled from all three stations showed that banding and census indices were usually correlated with each other (73% of 64 species). In 35 species, banding and census each had independent influences on annual

ET indices, even though banding and census indices were usually correlated with each other (Table 1). In 20 additional species, banding did not add anything to ETs after census had been taken into account, and in 9 species the reverse was true. For these 29 species, the non-contributing count method usually had much lower mean counts than the other, and thus had little influence on the ET indices whether or not the banding and census indices were correlated with each other. A few species had very low R^2 values (most notably Great Crested Flycatcher [scientific names in Table 1]), indicating that ETs were heavily influenced by observations other than those from banding and census. Results were similar when analysed for each station separately.

Variance of detrended annual indices based on banding was highest at Area 1, lower at Area 2, and lowest at Area 3 (Table 2), but there were no significant differences. Variability of indices based on census was more similar among stations, and ET indices were the least variable, but for all three count methods, variability was lowest at

TABLE 2. COMPARISON OF VARIANCE IN DETRENDED ANNUAL INDICES OVER 17 YEARS FOR DIFFERENT COUNT METHODS AND STATIONS AT LONG POINT, ONTARIO

Station	Mean variance ± SD[a] of indices based on		
	Banding	Census	ET
Area 1	0.47 ± 0.26	0.31 ± 0.21	0.21 ± 0.16
Area 2	0.33 ± 0.25	0.29 ± 0.23	0.22 ± 0.19
Area 3	0.17 ± 0.16	0.21 ± 0.16	0.13 ± 0.07
All stations combined	0.12 ± 0.13	0.11 ± 0.08	0.09 ± 0.05

[a] Mean and SD across species of individual species' variance of detrended annual indices.

TABLE 3. COMPARISON OF TRENDS FROM 1984–2001 BASED ON INDICES FROM DIFFERENT COUNT METHODS AT LONG POINT, ONTARIO

Area	Count methods compared	Slope	Intercept	R^2
1	Census vs. band	0.85*	-0.57	0.56
	ET vs. band	0.70**	1.10**	0.73
	ET vs. census	0.83**	1.58**	0.83
2	Census vs. band	1.10	-0.81	0.29
	ET vs. band	0.90	1.40**	0.53
	ET vs. census	0.82**	2.07**	0.70
3	Census vs. band	0.78*	-0.78	0.09
	ET vs. band	0.76**	0.54	0.35
	ET vs. census	0.95	1.36**	0.63
All	Census vs. band	1.02	-0.34	0.51
	ET vs. band	0.93	1.16**	0.64
	ET vs. census	0.91*	1.46**	0.86

Notes: Slope, intercept, and R^2 from reduced major axis regression of the trends from the two count methods being compared (Bohonak 2002). Significance levels are for test of null hypothesis that slope is 1.0, and intercept is 0 (* = $P < 0.05$, ** = $P < 0.01$).

Area 3. Regardless of count method, variability was considerably reduced when indices were based on data from all three stations combined.

Trends from pairs of count methods were compared within stations, using reduced major axis regression. In Table 3, an intercept >0 indicates a tendency to a positive bias in the first count method relative to the second method in each pair. In seven of eight comparisons, ET trends were positively biased relative to banding and census. These eight comparisons also showed slopes <1 (significant in five cases), indicating that the positive bias was less in species with increasing trends than in those with decreasing trends (Table 3, Fig. 1). By contrast, census showed little bias relative to banding, although at two stations the slopes of the relationships were significantly <1, indicating a tendency to a negative bias in census relative to banding in increasing species and the opposite effect in decreasing species (Table 3).

A similar analysis compared trends within count methods between pairs of stations (Table 4). Trends at Area 3 were strongly more negative, for all count methods, relative to trends at Areas 1 and 2 (as shown by the negative intercepts). However, slopes tended not to differ between stations (seven of nine comparisons).

DISCUSSION

Lack of standardization in banding added variability to annual indices. Variability was highest at the station with least standardization (Area 1), and lowest where netting effort was most uniform

(Area 3; Table 2). Increased variability reduces trend precision, such that it will take longer to detect a significant population change. However, increased variance of banding indices did not have a detectable effect on magnitude of estimated trends, which showed the same relationship to census trends at all three stations (Table 3).

The ET procedure incorporates data from census as well as from banding (Table 1), and ET indices were less variable than banding or census indices alone (Table 2). ETs therefore performed their intended function of removing some of the variability from unstandardized banding effort and adding information from other count methods.

Compared to banding and census, ETs tended to be positively biased (Fig. 1). Although we cannot be sure which method best represents actual population trends, there are several reasons to suspect that ETs might be positively biased. First, there appears to have been a change in the way ETs were estimated, starting in about 1993, with observers becoming less conservative in their estimates (E. Dunn et al., unpubl. data). In addition, there may have been an increase over time in the number of personnel, and longer hours spent in the field. We were unable to correct for variable effort in our analyses, and effort-correction is in any case an imperfect and time-consuming solution, particularly when many types of effort are combined. However, additional work could be done to determine the relative importance of these sources of bias. Regardless of the source of bias in historical data at Long Point, bias in trends from other stations or from Long Point in future can be minimized by ensuring that every aspect of data
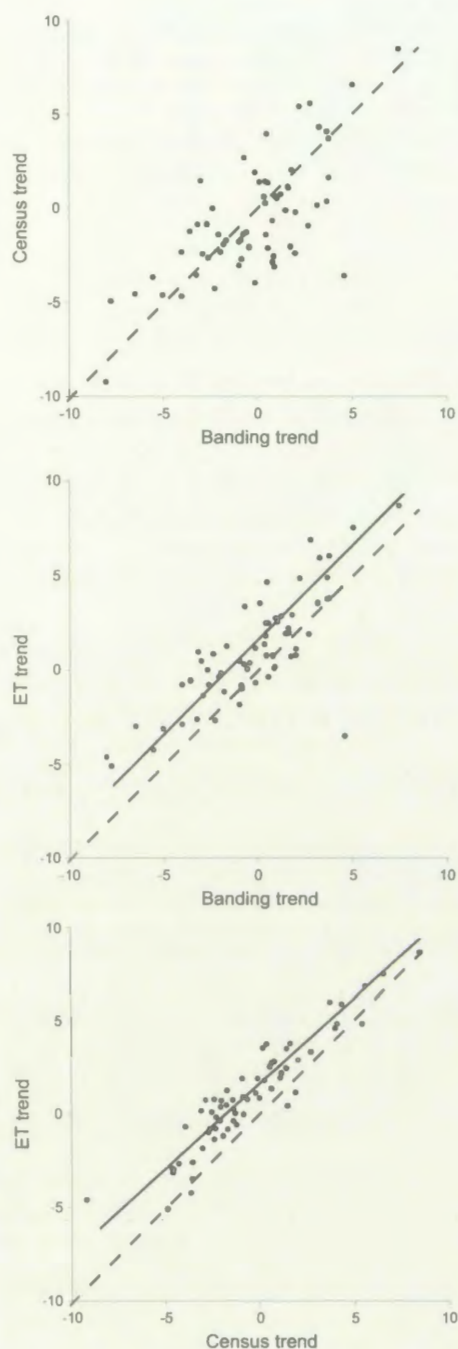
FIGURE 1. Comparison of population trends at Long Point, Ontario, based on different data sources (data pooled from all stations). ET trends were positively biased relative to trends based on banding or census alone. Dashed line indicates one-to-one correspondence between trends; solid line shows fit according to reduced major axis regression (shown only if different from the dashed line).

TABLE 4. COMPARISON OF TRENDS FROM 1984–2001 BASED ON INDICES FROM THREE DIFFERENT COUNT AREAS AT LONG POINT, ONTARIO

Count method	Areas compared	Slope	Intercept	R^2
Banding	2 vs. 1	0.75**	-0.41	0.22
	3 vs. 2	1.08	-3.19**	0.12
	3 vs. 1	0.80	-3.79**	0.03
Census	2 vs. 1	0.97	-0.73	0.24
	3 vs. 2	0.84	-2.72**	0.24
	3 vs. 1	0.74**	-3.42**	0.24
ET	2 vs. 1	0.95	0.01	0.30
	3 vs. 2	0.96	-3.35**	0.25
	3 vs. 1	0.91	-3.58**	0.43

Notes: Slope, intercept, and R^2 are from reduced major axis regression of the trends from the two areas being compared (Bohonak 2002). Significance levels are for test of null hypothesis that slope is 1.0, and intercept is 0 (* = $P < 0.05$, ** = $P < 0.01$).

collection is strictly standardized, as recommended by Ralph et al. (*this volume* a).

We found clear evidence of station differences in population trends. We have no reason to suspect that the strongly more negative trends at Area 3, relative to trends at the other two stations, were related to station differences in data collection. One possible explanation is differential habitat change among the three stations. Area 3 is a small woodlot surrounded by marsh and cottage. The vegetation at this station, especially the trees, grew steadily taller throughout the study period and understory was reduced. Many of the species for which the trend at Area 3 was the lowest (most negative) of the three stations, both for banding and census, are large and conspicuous. These species would probably have been detected if present, so we suspect they do not use the location now as often as in the past (e.g., Northern Flicker, Great Crested Flycatcher, nearly all thrushes, Brown Thrasher, Gray Catbird, Rose-breasted Grosbeak, Scarlet Tanager, Baltimore Oriole). However, another 23 species with their lowest trends at Area 3, made up mostly of warblers and vireos, could have been present but detected and captured in mist nets with lower probability as the canopy grew higher and more dense. In contrast to Area 3, Areas 1 and 2 are maintained at relatively early successional stages by storms and shifting of dunes. Although habitat at these two areas is certainly not constant, change appears to be less directional over time.

It is often stated in the migration monitoring literature that habitat change could bias population trends, but this is often ignored when study locations are selected and results are being interpreted. The difference between trends at Area 3 and the other

two sites at Long Point suggest that habitat effects could be substantial, and emphasizes the importance of having an effective habitat management protocol for long-term studies.

ACKNOWLEDGMENTS

A COMPARISON OF CONSTANT-EFFORT MIST NETTING RESULTS AT A COASTAL AND INLAND NEW ENGLAND SITE DURING MIGRATION

Christopher C. Rimmer, Steven D. Faccio, Trevor L. Lloyd-Evans, and John M. Hagan, III

Abstract. We compared population trends from spring and fall migration capture data from two constant-effort banding stations in New England: one coastal (Manomet Center for Conservation Sciences, hereafter "Manomet") and one inland (Vermont Institute of Natural Science, "VINS"). Data were examined for two time periods, 1981–1992 and 1986–1992. Twelve-year population trends were compared to regional Breeding Bird Survey (BBS) data for the same period. The two migration data sets showed little congruence. Of 22 species examined, Manomet data showed significant declines in 11 during one or both seasons, whereas seven species increased significantly at VINS. The number of significant trends at both sites increased between a 7-year and a 12-year sample. Among six species that were strictly transient at the two sites, five showed the same 12-year trend in fall. In general, Manomet tracked BBS data from the Northern Spruce-Hardwood region reasonably well, while VINS more closely tracked BBS trends from Northern New England. Neither site correlated well with BBS trends from Quebec. VINS captured significantly higher proportions of adult birds than did Manomet in 81% of species examined. However, the two sites tracked trends in age ratios largely independently. Several factors appeared to account for the weak congruence between sites, and we discuss the limitations in comparing these two data sets.

Key Words: age ratios, banding station, capture data, migration, New England, population trends.

Despite an extensive network of migration banding operations in North America and Europe, there have been relatively few studies to establish the validity of migration capture data to monitor bird population changes. Hagan et al. (1992) showed that a 19-year migration data set from the Manomet Center for Conservation Sciences (Manomet) in coastal Massachusetts correlated well with documented population changes in several passerine species that breed in northeastern North America. These included resident species (Tufted Titmouse and Northern Cardinal [scientific names listed in Table 1]), short-distance migrants (Golden-crowned Kinglet and Ruby-crowned Kinglet), and neotropical migrants (Tennessee Warbler, Cape May Warbler, and Bay-breasted Warbler). The Manomet data also corresponded with regional Breeding Bird Survey (BBS) data, as 24 of 38 species examined (63%) showed significant positive correlation of annual indices with those from BBS from at least one northeastern physiographic BBS stratum. Positive correlations between the Manomet and BBS data were more common for physiographic strata close to Manomet, suggesting geographic limitations to the usefulness of migration capture data.

However, Manomet trend data correlated poorly with those from another long-term migration banding station in eastern North America, the Powdermill Nature Reserve, located 800 km west–southwest of Manomet (Hagan et al. 1992). Of 40 species ana-lyzed in both data sets, only one showed a significant positive correlation between the two sites. This suggested that different source populations undergoing independent changes were likely sampled at each site, and that local habitat changes might have biased samples of migrants through time, particularly at the Powdermill site.

In a study comparing 1979–1991 banding totals among 13 transient species at two Michigan sites 0.75 km apart, Dunn et al. (1997) found significant positive correlations between trends from capture data and those from Michigan and Ontario BBS data. Moreover, the trends between banding and BBS data were of similar magnitude. Although the two sites showed little overall correlation in trends, this was due to discrepancies in one species, and trend differences were small in most cases. These results suggested that standardized mist netting can serve as a useful and accurate population monitoring tool.

Other comparisons of banding capture data with regionally appropriate BBS data have also shown good concordance between the two (Hussell et al. 1992, Dunn and Hussell 1995, Francis and Hussell 1998), but relatively few comparisons between or among migration banding stations have been made. Under the assumption that migration capture data can accurately track population changes, such comparisons might provide valuable information on population trends within a given region. In this study we compare data from two northeastern U.S. migration

TABLE 1. BREEDING STATUS, SAMPLE SIZE, AND POPULATION TRENDS FOR SELECTED SPECIES FROM MIGRATION CAPTURE DATA AT MANOMET AND VINS, AND BBS DATA FROM NORTHERN NEW ENGLAND (NNE) AND NORTHERN SPRUCE-HARDWOOD (NS-H) STRATA, 1981–1992

Species	Code	Status[a]		N captured				Trend (percent/yr)				BBS trend (percent/yr)	
		Manomet	VINS	Manomet		VINS		Manomet		VINS		NNE	NS-H
				Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall		
Eastern Phoebe (*Sayornis phoebe*)	EAPH	B	B	63	123	61	373	nd	1.9*	nd	4.9	7.4*	3.9*
Red-eyed Vireo (*Vireo olivaceus*)	REVI	B	B	106	973	33	378	-1.1	-6.8*	nd	7.8	-1.2*	1.7*
Black-capped Chickadee (*Poecile atricapilla*)	BCCH	B	B	760	9,680	158	603	-11.1*	-8.6	8.1*	-3.9	2.8*	1.2
Ruby-crowned Kinglet (*Regulus calendula*)	RCKI	T	T	399	524	134	334	0.8	1.2	6.7	0.7	nd	2.0*
Veery (*Catharus fuscescens*)	VEER	T	B	213	209	25	151	-3.7*	-3.7	nd	7.9*	-0.8	-3.1*
Swainson's Thrush (*C. ustulatus*)	SWTH	T	T	510	358	3	265	-3.4	-2.8	nd	-1.6	nd	-2.9*
Hermit Thrush (*C. guttatus*)	HETH	B	B	537	403	17	573	4.7	2.0	nd	10.6*	10.3*	3.2*
American Robin (*Turdus migratorius*)	AMRO	B	B	370	1,977	70	1,378	-1.8	-6.0	nd	23.4*	0.9	0.7
Gray Catbird (*Dumetella carolinensis*)	GRCA	B	B	471	5,922	177	722	-2.5	-5.9*	3.4	1.4	-1.0	-3.9*
Nashville Warbler (*Vermivora ruficapilla*)	NAWA	T	T	23	116	17	236	nd	-1.8*	nd	-11.0*	-4.9*	-1.6*
Yellow Warbler (*Dendroica petechia*)	YWAR	B	B	268	114	141	30	-0.1	-2.4	-5.0	nd	-0.1	-0.1
Magnolia Warbler (*D. magnolia*)	MAWA	T	T	772	232	27	231	2.1	-2.2*	nd	-3.3	-6.4*	0.7
Yellow-rumped Warbler (*D. coronata*)	MYWA	T	B	183	4,446	377	935	1.4	-9.8*	10.8	-2.7	5.9*	-0.03
American Redstart (*Setophaga ruticilla*)	AMRE	B	B	921	1,047	70	112	-2.6	-5.5*	nd	-5.3	-0.9	-2.0*
Ovenbird (*Seiurus aurocapillus*)	OVEN	B	B	340	228	10	198	0.3	-1.7	nd	4.7	1.4*	-1.0*
Common Yellowthroat (*Geothlypis trichas*)	COYE	B	B	1,266	497	286	1,159	-1.3	-2.8*	-20.0*	-24.0*	-1.8*	-1.3*
Canada Warbler (*Wilsonia canadensis*)	CAWA	T	T	524	150	48	112	-3.5	-2.3	nd	-5.3	-1.2	-2.9*
Song Sparrow (*Melospiza melodia*)	SOSP	B	B	291	869	122	881	-1.9	-0.1	6.7*	-0.9	0.7	0.1
Swamp Sparrow (*M. georgiana*)	SWSP	T	B	314	257	31	206	0.7	3.0*	nd	5.9*	4.3	0.5
White-throated Sparrow (*Zonotrichia albicollis*)	WTSP	T	B	1,853	1,697	209	1,792	-3.6	-5.4	11.8*	19.4*	-2.2*	-1.7*
Dark-eyed Junco (*Junco hyemalis*)	DEJU	T	B	106	415	90	288	-3.3	-4.0*	nd	6.9*	-0.2	-3.0*
Purple Finch (*Carpodacus purpureus*)	PUFI	T	T	10	195	27	283	nd	-3.4	nd	6.3	-3.6	-5.3*

Note. "nd" denotes insufficient data for analysis.

[a] Status: T = strictly transient, B = regularly breeds within 25 km of banding site.

* denotes P = 0.05

banding stations, Manomet and the Vermont Institute of Natural Science (VINS). We use population indices based on migration captures to examine correlations between the two data sets, and we compare trends from banding data at each site with regional BBS trends to further assess congruence. We briefly examine age ratios and their correlation through time of fall migrants at both sites. Finally, we discuss the validity of comparing these two data sets in light of between-site differences.

METHODS

STUDY AREAS

Manomet, located on the western shore of Cape Cod Bay, Plymouth Co., Massachusetts (41°50'N, 70°30'W), lies about 250 km southeast of VINS, located in Woodstock, Windsor County, Vermont (43°36'N, 72°32'W). Both sites are characterized by heterogeneous second-growth deciduous shrub-woodland, consisting of brushy thickets interspersed with groves of largely mature trees. Hagan et al. (1992) described dominant vegetation on the 7-ha Manomet study plot. On the 3-ha study plot used for the VINS banding operation, dominant trees include sugar maple (*Acer saccharum*), bigtooth aspen (*Populus grandidentata*), black cherry (*Prunus serotina*), and white ash (*Fraxinus americana*); dominant shrubs include willow (*Salix* spp.), autumn olive (*Elaeagnus umbellata*), steeplebush (*Spiraea tomentosa*), hawthorne (*Crataegus* spp.), dwarf juniper (*Juniperus communis*), and common buckthorn (*Rhamnus cathartica*).

The rate of successional habitat change differed between Manomet and VINS, although vegetation data were not systematically collected at either site. Because of its coastal exposure, Manomet underwent little successional change during the 1981–1992 study period. Because of VINS's more sheltered, inland location, and the gradual maturation of its habitats from open farmland prior to 1970, relatively more rapid plant succession occurred at that site. Limited vegetation management at VINS during the study slowed the rate of habitat change, but vegetation height around some VINS nets likely increased by 50% or more over the 12-year period. Any effects of successional change on netting totals were probably more pronounced at VINS than at Manomet.

The two study sites occupy contrasting landscapes. Manomet is a 7-ha "oasis" in a coastal belt that is highly fragmented by suburban development, with an increasingly dense human population. VINS lies on a largely forested 32-ha preserve in a predominantly forested and unfragmented rural landscape, with little human population growth.

DATA COLLECTION

From 1969–1992, inclusive, Manomet annually operated 45–50 nylon mist nets (12 × 2.6 m, 4-panel, 36-mm extended mesh) at fixed locations. During the spring (15 April through 15 June) and fall (15 August through 15 November) migrations, nets remained open at least five days a week from 0.5 h before sunrise to 0.5 h after sunset. From 1981–1992, inclusive, VINS operated 15–20 mist nets (12 × 2.6 m, 4-panel, 36-mm extended mesh) each year, generally from 15 April to 15 June in spring, and 1 August to 15 November in fall. Standardization of the VINS operation was less uniform than at Manomet, and differed between 1981–1985 and 1986–1992. During the earlier period, nets were opened on an average of two to three days a week for three to five morning hours. Between 1986–1992, nets were opened five days a week for 6 h, beginning 0.5 h before sunrise. Although some net site locations at VINS shifted during the study period, nets were maintained at fixed locations after 1987. At both Manomet and VINS, nets were closed under adverse weather conditions, and records were kept of opening and closing times of nets.

Data Analysis

We analyzed data only for the period 1981–1992, when both banding stations were in operation. We compared both spring and fall data. We compared only those species for which ≥ 100 captures were obtained at each site, combined over both seasons and all years. To restrict our analyses to migrant birds, we eliminated all known or presumed breeding individuals, that is, those with enlarged cloacal protuberances or well-developed brood patches. For each species we calculated a site-specific temporal migration window, defined as those dates after the 1st percentile and before the 99th percentile of captures within each migration season, all years combined.

For each species we calculated a daily population index for each date within its migration window. This was derived by dividing the number of captured individuals of a species by the number of net hours for that date, multiplying that number by 1,000 and adding a constant of 1, then taking the natural log. We calculated an annual population index for each species at each site by computing the mean of the logged daily indices. This procedure smoothed out variation due to days with unusually large numbers of captures (Dunn and Hussell 1995). Population trends were then calculated as the slope of the annual indices regressed on year, producing an estimated annual percent rate of change. Because three species each had annual indices of zero in one year, we did not back-transform indices and remove the constant of one prior to calculating population trends, as log transformation of zero would have resulted in a negative index.

We examined population trends at Manomet and VINS by dividing the data into two time periods: the entire 12-year period from 1981–1992, and a 7-year subset from 1986–1992, during which time the VINS operation was relatively standardized. We suspected that the lack of uniform standards at VINS during 1981–1985 might have obscured or biased actual population trends at that site over the longer 1981–1992 study period. We thus compared trends over both 7 and 12 years.

We obtained BBS population trends for 1981–1992 from the U.S. Geological Survey's Patuxent Wildlife Research Center webpage (http://www.mbr-pwrc.usgs.gov/bbs/bbs.html). We compared trends from spring and fall migration capture data at both sites with trends from BBS data for three regions: Northern New England (physiographic strata 27), Northern Spruce-Hardwood (physiographic strata 28), and the province of Quebec. We believe that these three areas represent the most likely geographic source of migrants sampled at Manomet and VINS. We calculated Spearman rank correlations (one-tailed significance tests) between trends from VINS and Manomet banding totals and those from the BBS (SYSTAT 1998).

We compared age ratios in the migration window in those species with adequate sample sizes at both sites (see criteria above). We used only those species for which the proportion of unknown-age birds was less than 5% and less than the proportion of adults at each site. We used Manomet capture data from 1969 to 1992 and VINS data from 1981 to 1992. We excluded spring migrants from our analysis of age ratios due to generally small samples of known-age (second-year and after second-year) individuals at each site. We examined differences of age ratio differences using a t-test, and we compared annual changes in age ratios at both sites using a Pearson product-moment correlation (SAS Institute 1985).

RESULTS

CORRELATION BETWEEN MANOMET AND VINS POPULATION TRENDS

The combined Manomet and VINS migration capture data set consisted of 22 species with sufficient sample sizes for between-site comparison (Table 1). Migration trends from 1981–1992 for all species combined for which there were data from both sites were uncorrelated both for fall ($r = -0.031$, $N = 21$ species, $P > 0.10$) and spring ($r = -0.238$, $N = 8$, $P > 0.10$; Fig. 1). For three species (American Robin, Common Yellowthroat, and White-throated Sparrow), VINS trends were biologically unrealistic (>15%/year increase or decrease), but excluding them did not improve the correlation among the remaining species in fall. Comparison of fall trends over the 7-year subset of data (1986–1992) revealed similarly poor correlation between the two sites ($r = 0.008$, $P > 0.10$).

DIRECTION OF MANOMET AND VINS POPULATION TRENDS

Over the 12-year period, 13 species showed significant ($P \leq 0.05$) population trends in one or both seasons at Manomet, whereas populations of 10 species changed significantly at VINS (Table 1). During the

fall season, there was moderate agreement in the direction of trends between the two sites, with 13 (62%) species agreeing and eight (38%) disagreeing (Table 2). At Manomet, two species significantly increased in fall, whereas eight species experienced significant declines. VINS data showed significant increases in six species and significant declines in two. Swamp Sparrow was the only species to increase significantly at both sites in fall, whereas Nashville Warbler and Common Yellowthroat declined significantly at both. Only Dark-eyed Junco showed an opposite significant trend at the two sites, declining at Manomet and increasing at VINS (Table 1).

During the spring, there was little correspondence in trend direction between the two sites, with four species agreeing and four disagreeing (Table 2). At Manomet, three species showed significant population trends, all declines. Despite small sample sizes at VINS, which reduced the number of species included in spring analyses to eight, three species showed significant increases and one a significant decline (Table 1). No species showed the same significant trend at both sites, but Black-capped Chickadee declined significantly at Manomet while increasing significantly at VINS.

A comparison of population trends over both 12 and 7 years indicated that although most trends became non-significant over the shorter time period, three species showed significant population changes only during this period. Three species (Swamp Sparrow at Manomet, Common Yellowthroat and White-throated Sparrow at VINS) showed the same significant trend in the same season over both 12 and 7 years. Two species (American Redstart and Canada Warbler) at Manomet showed significant declines during both periods, but in different seasons. No species showed opposite significant trends at Manomet and VINS during 1986–1992.

COMPARISON OF BANDING DATA WITH BBS DATA

For neither site did trends correlate well with BBS trends from Quebec. Trends for all species from Manomet were significantly correlated with those from the Northern Spruce-Hardwood region when Black-capped Chickadee (an irruptive species) was excluded ($r = 0.424$, $N = 20$, $P < 0.05$) but not otherwise ($r = 0.312$, $N = 21$, $P > 0.05$; Fig. 2). The relationship between Manomet trends and BBS trends from Northern New England was weaker ($r = 0.112$, $P > 0.10$ for all species; and $r = 0.205$, $P > 0.10$ excluding Black-capped Chickadee).

TABLE 2. MEAN AGE RATIOS OF 21 SPECIES WITH ADEQUATE SAMPLE SIZES FROM FALL CAPTURES AT MANOMET (1969–1992) AND VINS (1981–1992)

| | Mean age ratio (percent) | | | |
| | Manomet | | VINS | |
Species	After hatch year	Hatch year	After hatch year	Hatch year
Eastern Phoebe	10.99	89.01	7.43	92.57
Black-capped Chickadee	6.41	93.59	5.65	94.35
Ruby-crowned Kinglet	19.00	81.00	7.64	92.36
Veery	16.10	83.90	23.53	76.47
Swainson's Thrush *	5.57	94.43	36.47	63.53
Hermit Thrush	9.18	90.82	6.76	93.24
American Robin	4.86	95.14	9.51	90.49
Gray Catbird *	2.49	97.51	8.23	91.77
Red-eyed Vireo *	1.85	98.15	16.55	83.45
Nashville Warbler *	3.07	96.93	21.34	78.66
Magnolia Warbler *	5.08	94.92	33.57	66.43
Yellow-rumped Warbler	9.43	90.57	11.58	88.42
American Redstart *	2.86	97.14	18.32	81.68
Ovenbird *	2.69	97.31	17.49	82.51
Common Yellowthroat *	6.00	93.00	17.97	82.03
Canada Warbler *	3.18	96.82	21.45	78.55
Song Sparrow *	4.28	95.72	13.38	86.62
Swamp Sparrow *	1.87	98.13	28.50	71.50
White-throated Sparrow *	3.21	96.79	29.66	70.34
Dark-eyed Junco *	6.60	93.40	40.15	59.85
Purple Finch *	10.08	89.92	30.80	69.20

* denotes species with significantly different age ratios between sites (t-test, P < 0.05).

Similarly, trends from VINS were significantly correlated with those from Northern New England when White-throated Sparrow was excluded (r = 0.425, P < 0.05, n = 18), but not overall (r = 0.291, N = 19, P > 0.10; Fig. 3). Correlation with BBS from the Northern Spruce-Hardwood region was less strong (r = 0.083, P > 0.10).

Four specific examples, using fall data only from 1981–1992, illustrate the range of comparisons in population trends between the two sites and their congruence to regional BBS data:

Common Yellowthroat.—This species showed a highly significant decline in capture rate at both Manomet (r^2 = 0.410, P = 0.025) and VINS (r^2 = 0.851, P < 0.001; Fig. 4A). Both sites tracked a steady decline that was reflected in BBS data from both physiographic strata 27 and 28 (Table 1). However, the trend at VINS was so steep (24%/year) as to be biologically unrealistic.

Nashville Warbler.—Although population indices showed more variance over time for this species than for Common Yellowthroat, significant declines occurred at both Manomet (r^2 = 0.341, P = 0.046) and VINS (r^2 = 0.515, P = 0.009; Fig. 4B). These were also reflected in regional BBS data, as both strata 27 and 28 showed significant declines (Table 1).

Veery.—This species significantly increased at VINS (r^2 = 0.450, P = 0.017) and significantly decreased at Manomet (r^2 = 0.323, P = 0.054) during the study period (Fig. 4C). The population increase at VINS was due primarily to a pulse of migrants between 1989–1992. Regional BBS data indicated significant and nonsignificant declines in both strata. Veerys breed in the vicinity of the VINS banding station, and the increase at VINS may have reflected an increase in local breeding populations that masked a more widespread decline.

Red-eyed Vireo.—This species showed a significant decline at Manomet (r^2 = 0.482, P = 0.012), and a nonsignificant positive trend at VINS (r^2 = 0.287, P = 0.073; Fig. 4D). BBS data from the Northern Spruce-Hardwood region showed a significant increase in Red-eyed Vireos, whereas Northern New England BBS data showed a significant decline.

COMPARISON OF AGE RATIOS

Of the 21 fall migrant species for which we examined age ratios at the two sites, VINS captured a higher ratio of AHY (after hatching year) to HY (hatching year) birds for 17 (81%), whereas Manomet's ratio of AHY birds was higher for only
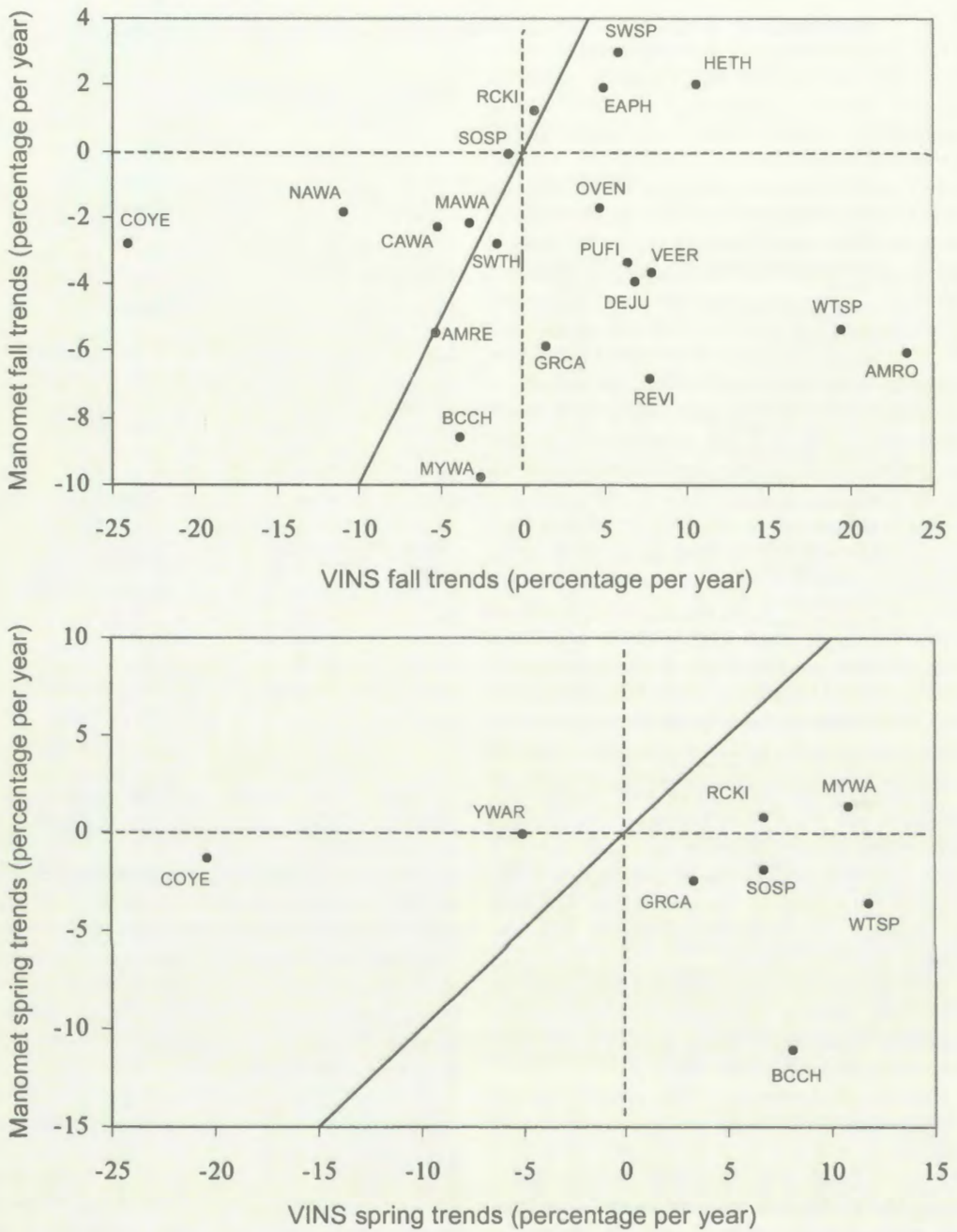
FIGURE 1. Rates of change (%/year) in fall (top) and spring (bottom) migration count indices at Manomet and VINS, 1981–1992. Solid line indicates one-to-one correspondence. See Table 1 for species codes.

four species (19%; Table 3). Fourteen of the 21 age ratio differences were significant (P < 0.05, t-test).

To assess the degree to which Manomet and VINS tracked changes in age ratios, we examined correlations among species over the 12-year period. We found no significant correlations; thus there appeared to be little year-to-year synchrony in age ratios at the two sites.
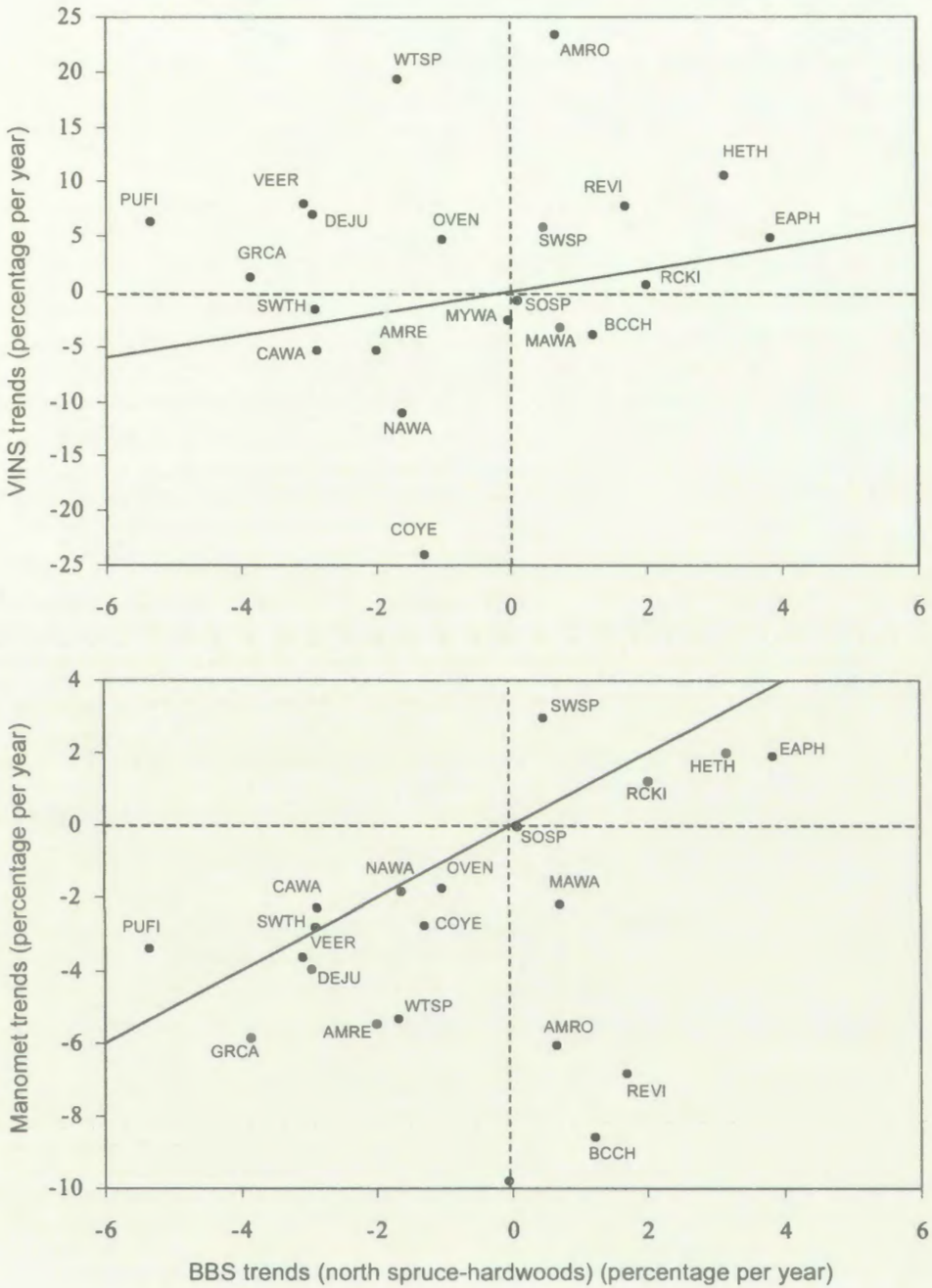
FIGURE 2. Rates of change (%/year) in fall migration capture indices at VINS (top) and Manomet (bottom) and BBS trends for Northern Spruce-Hardwoods physiographic stratum, 1981–1992. Solid line indicates one-to-one correspondence. See Table 1 for species codes.
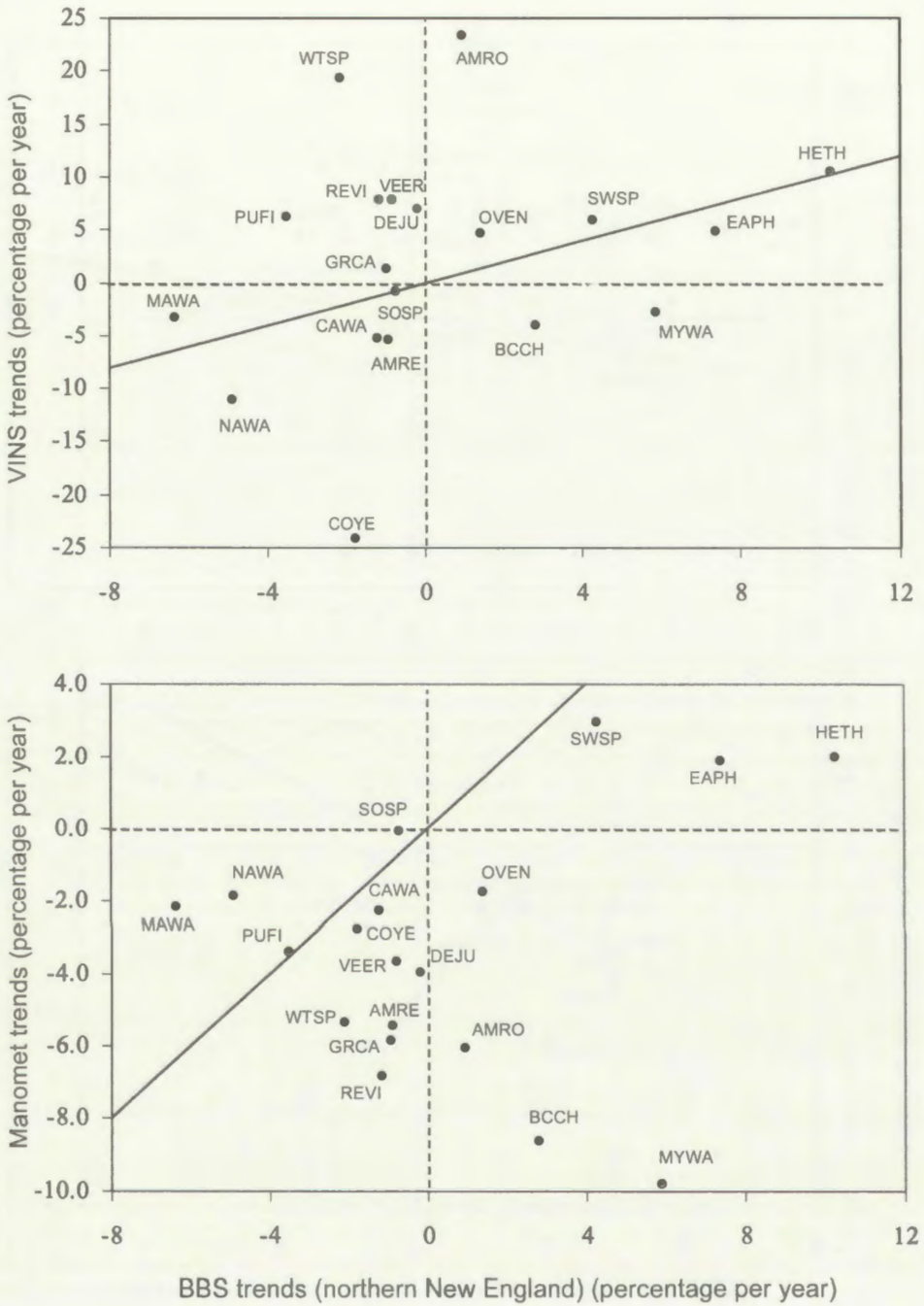
FIGURE 3. Rates of change (%/year) in fall migration capture indices at VINS (top) and Manomet (bottom) and BBS trends for Northern New England physiographic stratum, 1981–1992. Solid line indicates one-to-one correspondence. See Table 1 for species codes.
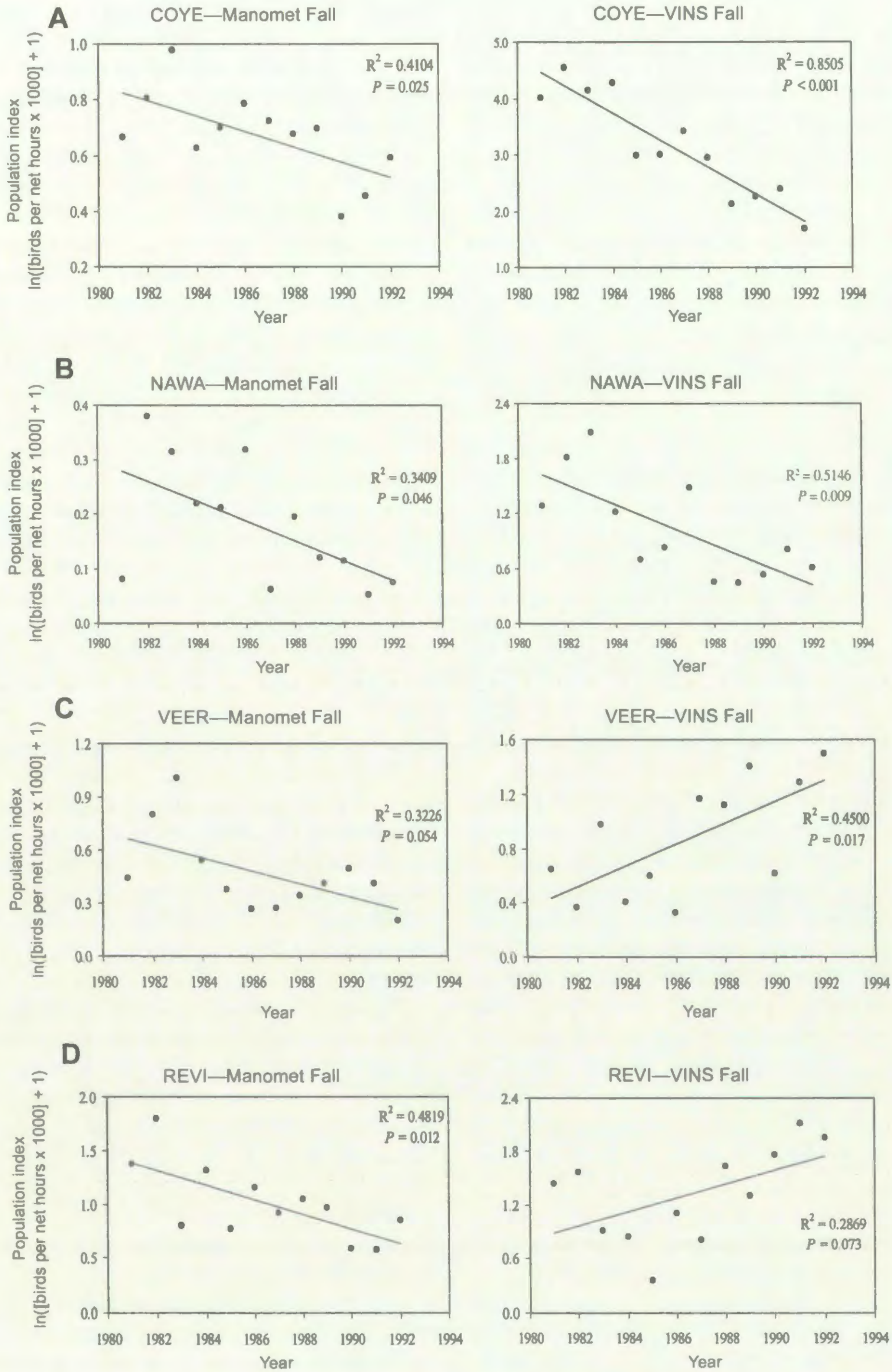
FIGURE 4. Linear regression of annual population indices from fall migration capture data for: (A) Common Yellowthroat; (B) Nashville Warbler; (C) Veery; and (D) Red-eyed Vireo during 1981–1992 at VINS and Manomet.

DISCUSSION

For three species, fall trends at VINS were biologically unrealistic (23.4% annual increase in American Robin, 24% annual decrease in Common Yellowthroat, and 19.4% annual increase in White-throated Sparrow). Even excluding these outliers, however, there was generally poor correlation between population trends calculated from Manomet and VINS migration capture data. Possible reasons include the following: (1) 7 or 12 years may be too short a period for detection of trends; (2) sample sizes at VINS, particularly during the 1981–1985 period, may have been inadequate for many species; (3) inconsistent standardization of methods at VINS during the study period may have obscured actual trends and reduced comparability of the two data sets; (4) changes in local breeding populations may have unduly biased VINS' data; (5) local or landscape-level habitat change may have biased population indices at either or both study sites, by differentially affecting the composition or abundance of species captured through time; and (6) different source populations may have been sampled by each station, such that population differences were real.

We suspect that the combination of small sample sizes (Table 1) and relative inconsistency of operating standards at VINS may have affected validity of many of the trend comparisons between VINS and Manomet. The minimum sample sizes we arbitrarily selected for analysis may have been too small, despite resulting in a number of significant trends for species captured in low numbers. For example, among species for which we obtained only 100–250 fall captures during 1981–1992, three of eight at Manomet and three of seven at VINS showed statistically significant population changes (Table 1). We can not be confident that trends based on such small samples are biologically meaningful. Inconsistent standardization of the VINS station, especially in 1981–1985 when fewer numbers of nets were used for shorter and more variable periods on fewer days each week than in 1986–1992, undoubtedly increased variance of the capture data in those early years. This unequal variance may in part explain the poor congruence of Manomet and VINS data. More rigorously standardized data collection at VINS would likely have resulted in more directly comparable data sets.

We suspect that the proportion of locally breeding and dispersing birds in the overall VINS sample was substantially higher than at Manomet. Among the VINS sample of 22 species, only six can be classified as true transients (regularly breeding >25 km from banding site), whereas 11 species captured at Manomet were wholly transient, or very nearly so (Table 1). Fluctuations in local breeding populations of migrants at and near VINS, as well as differing annual rates of dispersal onto and away from the site, may have obscured trends of transient populations at VINS, particularly in fall. However, of the six species that we judged to be strictly transient at both Manomet and VINS, five showed the same direction of trend in fall, whereas one species (Ruby-crowned Kinglet) with adequate capture data showed the same trend direction at both sites in spring (Table 1). Only Purple Finch showed opposite trends among fall transients, and this species is more of an irruptive than regular migrant at both sites (C. Rimmer and S. Faccio, unpubl. data; T. Lloyd-Evans, unpubl. data). Of the five transient species with similar fall trends at Manomet and VINS, 1982–1991 regional BBS data showed a corresponding trend for each (Table 1). This suggests that the two sites corresponded more closely in tracking population changes of fall transient species than of species with local breeding populations. Although we believe that the great majority of captures among all species at both Manomet and VINS were of migrant individuals rather than dispersing local breeders or fledglings, because of our migration window criteria, comparisons between sites could be strengthened if locally breeding species were excluded from trend analyses. Among the six transient species at both Manomet and VINS, the correspondence in trend directions (with the exception of Purple Finch) was not reflected in trend magnitudes, which correlated poorly (r = -0.371, P > 0.10).

Whereas the VINS site experienced greater vegetation succession than the Manomet site during the 12 years of study, Manomet may have been subject to greater landscape level habitat change, through increased suburbanization of coastal southeastern Massachusetts. Either type of change may have influenced the diversity and abundance of migrants using the two sites. The very large declines of some early to mid-successional species at VINS (e.g., Nashville Warbler, Common Yellowthroat) may have resulted in part from decreased habitat suitability of the maturing old field communities in the vicinity of the VINS banding station. Increased vegetation height around nets may also have reduced capture rates. At Manomet, increasing isolation of the 7-ha site as a habitat "fragment" in a predominantly suburban coastal landscape may have variably altered its use by stopover migrants over time. Local changes in vegetation at Manomet, while less pronounced than

those at VINS, may also have contributed to changes in migrant bird populations using the site. Because no quantitative habitat assessment was conducted at either site, we were unable to evaluate the extent of such changes. We believe that regular, standardized measurements of habitat features on both local and landscape levels are needed to evaluate the context of changes in migrant bird populations within and between sites.

Trends from each of the two stations were most congruent with regional BBS data from different strata. Although the presumed southeasterly direction of many fall migrants in the northeastern United States (e.g., Ralph 1978) might well have carried some Vermont birds to coastal Massachusetts, this can not be assumed. BBS data suggest that Manomet migrants may have originated largely in northeastern areas of New England and maritime Canada, whereas VINS's migrant sample may have been composed largely of birds from northwestern New England, southern Quebec and southeastern Ontario. These results suggest that each station may have tracked largely independent population changes, as suggested by Hagan et al. (1992) in their comparison of data from Manomet and a station at Powdermill, Pennsylvania. Without knowledge of the source populations being sampled, and of possible annual variation in the geographic composition of migrant captures, population trends at different sites must be compared cautiously. We believe that careful, species-specific analyses of BBS data from appropriate physiographic strata or specific geographic regions may be a good means of inferring the extent to which different banding stations sample similar source populations.

POPULATION TRENDS

The preponderance of declining species at Manomet and of increasing species at VINS is difficult to explain, even in light of potential within- and between-site biases. The possibility that one or both sites failed to track population changes accurately can not be discounted. However, Hagan et al. (1992) demonstrated that Manomet migration capture data collected over a 19-year period accurately measured known population changes in several species in northeastern North America. The VINS data are less clear in this regard. Although several species (e.g., Nashville Warbler, Common Yellowthroat) showed corresponding trend directions at both sites and in regional BBS data, others (e.g., Veery, White-throated Sparrow, Dark-eyed Junco) showed

poor congruence between Manomet and VINS. That Manomet trend data more closely matched those of the BBS stratum directly to its north than did VINS, which correlated with BBS data within its own stratum, leads us to believe that Manomet more accurately measured actual population changes among migrants. Although a more detailed, species-by-species analysis of the two migration capture data sets and data from appropriate BBS strata might have enabled us to more fully evaluate this, such an analysis was beyond the scope of this paper.

PRODUCTIVITY INDICES

The significantly higher proportion of HY birds at Manomet and of AHY birds at VINS conforms to the coastal–inland ratio typical of most autumn passerine migrants (e.g., Drury and Keith 1962; Ralph 1971, 1978, 1981). The "coastal effect" results from most adults following overland routes in fall while immatures travel both inland and on the coast, or from differential behavior of the age classes upon reaching coastlines (Dunn and Nol 1980). Manomet migration capture data, which consisted largely of HY birds, may have been more strongly influenced, and thus potentially biased, by weather-related phenomena affecting their abundance and behavior at the coast (see Dunn et al. *this volume b* and Hussell *this volume* for evidence that weather affects age ratios). Further, age ratios at Manomet and VINS may have differed in part due to sampling different source populations, as discussed above. Finally, different trends in age ratios at the two sites may have masked agreement in the annual directions of change (Dunn et al. *this volume b*).

CONCLUSION

We recognize that our comparison of these two data sets is an imperfect one. We believe, however, that it reflects the realities of comparing migration capture data from geographically distant sites subject to different sources of variability. We further believe that migration capture data collected under standardized conditions (Ralph et al. *this volume a*) can provide a valid means of assessing avian population trends, and we encourage more comparisons of data among migration banding stations. Careful analyses of migration capture data from a network of long-term banding stations might yield valuable information on regional population trends and demographics of migrant birds. Comparisons among multiple sites could provide needed independent tests of results obtained from breeding season studies.

MIST NETTING TRANS-GULF MIGRANTS AT COASTAL STOPOVER SITES: THE INFLUENCE OF SPATIAL AND TEMPORAL VARIABILITY ON CAPTURE DATA

THEODORE R. SIMONS, FRANK R. MOORE, AND SIDNEY A. GAUTHREAUX

Abstract. We used constant effort mist netting during spring migration to sample populations of trans-Gulf migrants at two coastal study sites from 1987 to 1992. Approximately 2,500 individuals of 70 species were netted each season with approximately 5,000 net-hours of effort. Although captures per net hour and total species captured were fairly consistent each year, the seasonal patterns of capture, arrival condition, stopover duration, diversity of species, and number of individuals showed considerable variation from year to year. Differences in seasonal and annual weather patterns, the arrival condition of migrants, and habitat quality at stopover sites all influenced the probability of capturing birds with mist nets at our coastal stopover sites. Mist-net capture rates from coastal stopover sites, migratory activity indicated by radar echoes, and counts of migrants from censuses at mainland sites were correlated within a geographic radius of 100–150 km.

Key Words: capture variability, migration, mist netting, stopover, trans-Gulf migrants.

Over 80% of North American birds are migratory to some extent, and about half of those species cross the Gulf of Mexico during migration (Lowery 1946, Rappole and Warner 1976, Moore and Kerlinger 1987). The trans-Gulf flight is a dangerous, energetically expensive phase of the annual cycle. A typical migrant like an Ovenbird (scientific names in Table 1) deposits 40–50% of its body weight in fat each spring before departing on a 15–20 h non-stop flight en route from its tropical wintering grounds to the breeding grounds in North America. Crossing a large ecological barrier like the Gulf of Mexico is a risky endeavor for migrants, exposing them to the unpredictable forces of spring cold fronts and thunderstorms (Buskirk 1980). For migrants, this unpredictability often means that they have little control over their precise migratory trajectories (Gauthreaux 1971, Rappole et al. 1979, Moore and Kerlinger 1991). The inherently unpredictable nature of migration may make it a limiting factor for some populations. The variability in migratory patterns that emerge each year have important implications for the interpretation of mist-netting data from migratory stopover sites along the northern Gulf coast.

The objectives of this paper are to examine how variability in seasonal patterns of capture, arrival condition, and stopover duration at stopover sites may confound estimates of larger scale population trends, and to compare mist-net capture data with indices of activity derived simultaneously from mainland censuses and weather surveillance radar.

METHODS

We worked at two study sites along the northern Gulf Coast from 1987 to 1992 (Fig. 1). Peveto Beach is a coastal woodland in southwestern Louisiana. East Ship and Horn islands are barrier islands in Mississippi Sound. The two stations are approximately 400 km apart. The vegetation and field methods have been described in detail elsewhere (Loria and Moore 1990, Moore and Kerlinger 1987, Moore et al. 1990, Kuenzi et al. 1991). Approximately 20, 12-m nets were run daily at each station from dawn to 1100 hours and from 1400 to 1800 hours. The field season ran from late March to early May each spring. Standard measurements were taken on all birds captured before they were banded and released. Levels of body fat were estimated according to the ordinal scale developed by Helms and Drury (1960). In 1992 we conducted 1-km strip transect censuses (Emlen 1977) in pine (N = 63) and deciduous forest (N = 63) habitats in coastal Mississippi (Simons et al. 2000). In that same year, we also analyzed the archived film record of the WSR-57 radar at Slidell, Louisiana from 23 March to 27 May (Gauthreaux 1971, 1992). To quantify the radar images we used a calibration curve that related the spatial extent of the migration echoes on the radar image (measured as the maximum radius in nautical miles) to the mean number of birds in the volume defined by the $1.75°$ conical radar beam (elevated 2.5°) sweeping $20°$ of azimuth at a range of 46.3 km (Gauthreaux 1994).

RESULTS

Trans-Gulf migration occurs in spring from mid-March to late May, although the peak of activity is concentrated in April. Approximately 70 species

TABLE 1. MEAN ANNUAL CAPTURES AT EAST SHIP ISLAND, 1987–1991

Species	Captures/1,000 net-h	CV
Yellow-billed Cuckoo (*Coccyzus americanus*)	2.33	1.42
EasternWood-Pewee (*Contopus virens*)	5.22	0.62
Yellow-bellied Flycatcher (*Empidonax flaviventris*)	1.05	1.16
Acadian Flycatcher (*E. virescens*)	4.76	0.23
Least Flycatcher (*E. minimus*)	0.57	0.93
Eastern Phoebe (*Sayornis phoebe*)	0.25	1.47
Great Crested Flycatcher (*Myiarchus crinitus*)	2.42	0.49
Eastern Kingbird (*Tyrannus tyrannus*)	1.26	0.76
White-eyed Vireo (*Vireo griseus*)	62.67	0.74
Yellow-throated Vireo (*V. flavifrons*)	10.48	0.35
Blue-headed Vireo (*V. solitarius*)	0.40	0.97
Warbling Vireo (*V. gilvus*)	0.07	2.24
Philadelphia Vireo (*V. philadelphicus*)	1.47	0.70
Red-eyed Vireo (*V. olivaceus*)	127.39	0.44
Black-whiskered Vireo (*V. altiloquus*)	0.12	1.38
Barn Swallow (*Hirundo rustica*)	0.13	2.24
Red-breasted Nuthatch (*Sitta canadensis*)	0.27	2.24
House Wren (*Troglodytes aedon*)	1.55	0.94
Ruby-crowned Kinglet (*Regulus calendula*)	1.44	1.40
Blue-gray Gnatcatcher (*Polioptila caerulea*)	0.69	1.43
Veery (*Catharus fuscescens*)	13.07	0.70
Gray-cheeked Thrush (*C. minimus*)	8.02	0.76
Swainson's Thrush (*C. ustulatus*)	12.24	0.93
Hermit Thrush (*C. guttatus*)	0.11	2.24
Wood Thrush (*Hylocichla mustelina*)	13.10	0.67
Cedar Waxwing (*Bombycilla cedrorum*)	0.07	2.24
Blue-winged Warbler (*Vermivora pinus*)	3.61	0.62
Golden-winged Warbler (*V. chrysoptera*)	0.42	0.90
Tennessee Warbler (*V. peregrina*)	9.31	0.55
Orange-crowned Warbler (*V. celata*)	0.07	2.24
Northern Parula (*Parula americana*)	4.39	0.39
Yellow Warbler (*Dendroica petechia*)	15.03	0.59
Magnolia Warbler (*D. magnolia*)	9.35	0.54
Cape May Warbler (*D. tigrina*)	4.79	1.40
Black-throated Blue Warbler (*D. caerulescens*)	1.16	0.68
Yellow-rumped Warbler (*D. coronata*)	2.25	1.60
Black-throated Green Warbler (*D. virens*)	2.95	0.79
Blackburnian Warbler (*D. fusca*)	1.31	0.92
Yellow-throated Warbler (*D. dominica*)	0.90	0.95
Prairie Warbler (*D. discolor*)	2.00	1.17
Palm Warbler (*D. palmarum*)	0.84	0.86
Bay-breasted Warbler (*D. castanea*)	4.84	0.67
Blackpoll Warbler (*D. striata*)	12.34	0.86
Cerulean Warbler (*D. cerulea*)	0.98	0.36
Black-and-white Warbler (*Mniotilta varia*)	16.17	0.27
American Redstart (*Setophaga ruticilla*)	8.50	0.34
Prothonotary Warbler (*Protonotaria citrea*)	9.41	0.67
Worm-eating Warbler (*Helmitheros vermivorus*)	9.64	0.53
Swainson's Warbler (*Limnothlypis swainsonii*)	1.13	0.92
Ovenbird (*Seiurus aurocapilla*)	13.53	0.66
Northern Waterthrush (*S. noveboracensis*)	6.75	0.56
Louisiana Waterthrush (*S. motacilla*)	0.56	0.77
Kentucky Warbler (*Oporornis formosus*)	8.15	0.74
Connecticut Warbler (*O. agilis*)	0.11	1.42
Common Yellowthroat (*Geothlypis trichas*)	14.29	0.66