Spatial variation in shorebird nest success: Implications for inference

BRIAN J. MCCAFFERY[1] & DANIEL R. RUTHRAUFF[2]

U.S. Fish and Wildlife Service, Yukon Delta National Wildlife Refuge, PO Box 346, Bethel, Alaska 99559, USA. brian_mccaffery@fws.gov; [2]Department of Wildlife, Humboldt State University, Arcata, California 95521, USA. Present address: U.S. Geological Survey, Alaska Science Center, 1011 E. Tudor Rd., Anchorage, Alaska 99503, USA

McCaffery, B.J. & Ruthrauff, D.R. Spatial variation in shorebird nest success: Implications for inference. *Wader Study Group Bull.* 103: 67–70.

Estimates of nest success are widely applied in order to evaluate a multitude of theoretical and practical issues. Frequently, however, researchers fail to limit their inferences to the appropriate spatial scale. We evaluated small-scale variation in nest success of Western Sandpipers *Calidris mauri* during a four-year study on the Yukon-Kuskokwim Delta in western Alaska. We use these data to demonstrate that small-scale variation in nest success can significantly alter a researcher's interpretation of the factors affecting that reproductive parameter. In the absence of a statistically valid sampling design, researchers must be very careful about making inferences for areas beyond their actual study site. Properly designed studies allow for broader inferential power, but the logistical and financial hurdles involved in designing and implementing such a study are daunting. Metareplication can enhance one's confidence in the interpretation of local results, but should not be seen as a substitute for well-designed sampling schemes implemented across broad geographic scales.

INTRODUCTION

Studies of nesting success across a broad spectrum of avian taxa have multiplied dramatically over the last decade. Estimates of nest success have been used to evaluate a wide range of theoretical and practical issues, including the effects of habitat fragmentation, brood parasitism, and predation on nest success (e.g., Fauth 2000, Davison & Bollinger 2000, Mermoz & Reboreda 1998, Zanette & Jenkins 2000, Burke & Nol 1998, Hersek *et al.* 2002); the effects of habitat management (e.g., Duguay *et al.* 2000, Yahner 2000, Popotnik & Guiliano 2000, Reynolds *et al.* 2001, Garrettson & Rohwer 2001); landscape level components of habitat quality (e.g., Thogmartin 1999, Rodenwald & Yahner 2001, Clark *et al.* 1999); and the effects of varying reproductive strategies (e.g., Caffrey 1999, Dewey & Kennedy 2001). Manolis *et al.* (2000) provided a constructive review of Mayfield nest success estimates in particular, and proposed a more standardized approach to the interpretation and analysis of nest success data.

To date, however, there has been fairly little attention paid to the implications of spatial variation in nest success (but see Chase 2002, Tarvin & Garvin 2002). In many studies, researchers select a plot (or plots), and infer that plot-specific nest success values are representative of adjacent and/or nearby areas. The spatial scale over which such inferences are made may range from the surrounding woodlot to the species' continental distribution. In lieu of the application of formal sampling theory (e.g., random plot selection, replication, formal iterative determination of required sample sizes), however, such inferences may be invalid. This issue can be particularly troublesome when large-scale inferences are drawn from a handful of small, arbitrarily selected plots. A survey of 50 articles reporting nest success in recent (2000–2001) peer-reviewed ornithological journals found that only one-third of studies defined their sampling universe, only 12% of studies limited their inferences to the correct spatial scale, and only 4% used a true sampling design for plot selection (McCaffery, unpubl.).

In the shorebird literature, there are frequent attempts to correlate productivity in arctic-breeding shorebirds with variation in environmental factors such as weather or the cyclic abundance of arvicoline rodents (e.g., Underhill *et al.* 1993, Blomqvist *et al.* 2002, Tomkovich & Soloviev 2002). Nest success is apparently often higher in years with lemming and/or vole population peaks because a) predation risk is spread more broadly across a higher density of potential prey items, and/or b) predators may preferentially target rodents when the latter are abundant. Conversely, in years following highs (i.e., when rodent numbers have crashed), high predator populations (including offspring produced during the cyclic rodent peak) can have dramatically negative effects on nesting shorebirds (Roselaar 1979, Summers & Underhill 1987, Underhill *et al.* 1989). Much of the empirical support for this hypothesis consists of inferences based upon estimates of annual production in both waders and geese at a number of widely-separated staging and wintering areas. Empirical support from the breeding grounds, however, is derived from spatially limited studies that may not represent a true sample from the geographic areas over which inferences are made. The issues of spatial and temporal variation in arctic shorebird nesting success have not been adequately addressed from a quantitative perspective.

In this paper, we compare estimates of nest success between adjacent study plots on the Yukon-Kuskokwim Delta, and demonstrate the potential differences in interpretation that would have arisen if only one of the plots had been sampled. This example will illustrate the difficulties in making correlations between locally-derived estimates of shorebird nest success and regional phenomena.

STUDY AREA AND METHODS

We studied Western Sandpipers *Calidris mauri* from May to July 1999–2002 at the Kanaryarmiut Field Station, a heath tundra site near the Aphrewn River on the Yukon-Kuskokwim Delta, Alaska (61°21.80'N, 165°07.53'W). The study site was not randomly selected from within the distribution of the Western Sandpiper. The study site was also quite small, divided into a 16 ha core study plot and an additional 30 ha buffer plot immediately adjacent to the core plot (for details, see Ruthrauff 2002). For the purposes of this paper, the core and buffer plots are referred to as the primary and secondary plots, respectively. The term "combined" plot refers to the primary and secondary plots considered as a single unit.

One to 3 persons searched the primary plot for nests daily from late May to late June 1999–2002. The secondary plot was searched less frequently. We checked the status of nests approximately every 5 days during incubation, and daily near the time of hatch. We calculated nest daily survival rates using the Mayfield method (Mayfield 1975) and applied a period length of 25 days (4 days for laying period and 21 for incubation; Ruthrauff & McCaffery, unpubl.) to calculate nest success. We defined a successful nest as one at which at least one egg hatched. We compared estimates of daily survival rate between plots and across years using the logistic regression technique outlined in Aebischer (1999). All analyses were conducted using SAS statistical software (SAS Institute 1999) and results were considered statistically significant at $\alpha \le 0.05$.

RESULTS

Nest daily survival rate varied significantly by plot and by year; additionally, there was a significant year by plot interaction (Table 1). Within years, estimates of nest success varied dramatically between the primary and secondary plots (Fig. 1). Secondary plot values varied from 6% to 214% of the primary plot values. Primary and secondary plot estimates were closest in 2001, but the secondary plot point estimate of nest success was still 38% higher in that year. Furthermore, there was no agreement between the two plots in identifying the years of highest and lowest nest success, respectively. On the primary plot, the lowest success was in 2002, and the highest was in 2000. On the secondary plot, the lowest nest success was in 1999, and the highest was in 2001.

DISCUSSION

Shorebird researchers often attempt to correlate annual estimates of nest success with year-specific phenomena (e.g., severe storms, rodent population highs or lows). The validity of such correlations, and the inferences based upon them, depend upon both the scale and design of the sampling used to derive those nest success estimates. Inferences based on small and/or non-randomly selected plots may be misleading, regardless of how "obvious" such inferences may seem. The differences in nest success between our primary and secondary plots reveal the types of problems encountered when such an approach is used.

Within-year variation in nest success estimates between plots does not necessarily preclude the possibility of correctly identifying the effects of annual environmental variation. For example, across a series of years, unusually heavy predation in year *x* might result in the lowest estimate of nest success on both plots in year *x*, even if the estimates for that year differed significantly between the two. The data from our study, however, do not show this pattern. Instead, there was a significant interaction between plot and year (Table 1); the plots tell different stories in different years (Fig. 1). If we had been attempting to track temporal trends in productivity, data from the two plots might have led to contradictory conclusions. On the primary plot, nest success estimates were highest in the first two years, and lowest in the next two years; the opposite pattern was even more dramatically exhibited on the secondary plot.

The implications for interpretation are dramatic. Consider hypotheses about the effects of rodent population highs on sandpiper nest success. On the Yukon-Kuskokwim Delta, rodent populations have exhibited cyclic peaks every four years since at least 1984 (McCaffery, unpubl. data). During the course of our study, rodent populations peaked in 2000. If we had tried to assess the impacts of such a peak on the primary plot, we would have seen that sandpiper nest success was highest in 2000, and thus we might have concluded that nest predators focused their foraging effort on rodents rather than bird nests in that year (of course, we would have been ignoring the fact that we should draw no such conclusions based on a sample size of 1 [i.e., one rodent population cycle]). If we had only sampled the secondary plot, however, we would have seen that nest success was extremely low and perhaps concluded that high rodent numbers led to high predator numbers, which in turn led to higher-than-average predation rates on sandpiper nests. We would have been even more likely to accept this hypothesis if our study had not been initiated until 2000. In summary, both data sets would have suggested that rodent population highs have marked effects on sandpiper nest success. The dramatically different estimates of nest success on adjacent plots, however, would have led to diametrically opposed interpretations about what those effects were.

One might argue that the obvious solution is to base inferences on data from the combined plot. These data show no significant differences in nest success between years ($\chi^2_3 = 4.71$, $P > 0.05$), so we might have concluded that rodent population highs have little effect on sandpiper nest success, a conclusion different from both of the interpretations generated by looking at data from either the primary or secondary plot alone. The problem with this approach, however, is that there is still no basis for concluding that the data from the combined plot are any more representative than those from the individual smaller plots. In lieu of a formal sampling scheme, the combined plot is simply another small, non-randomly selected plot that happens to be just incrementally larger than either the primary or secondary plots alone. The spatial scope of our inference should be limited to the combined plot; we can draw no statistical inferences about the effects of rodent population highs on sandpiper nest success at any spatial scale beyond that. Thus, despite the temptation to do so, we would not be justified in generalizing our findings to larger geographic areas such as the Yukon-Kuskokwim Delta, western Alaska, or the American arctic.

We readily concede that the small size of our plots may have made our results susceptible to very local effects, and thus perhaps inappropriate for evaluating large-scale phenomena such as regional rodent cycles. For example, because the home range of certain predators at our study site (e.g.,

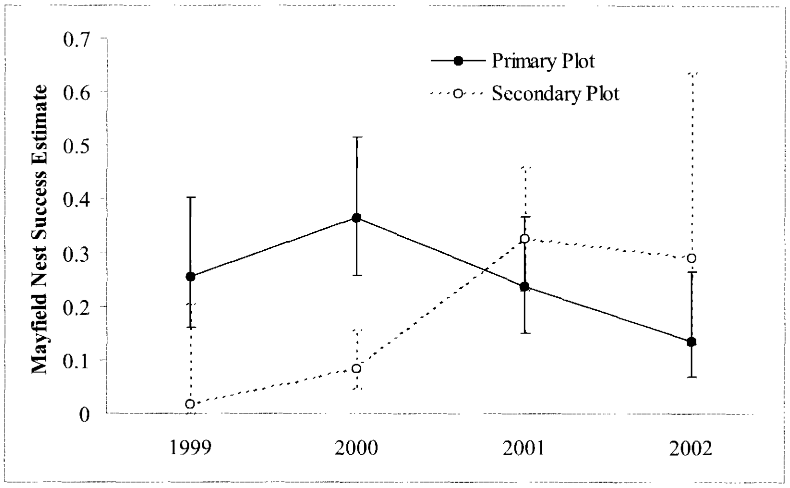

Fig. 1. Mayfield nest success estimate (±95% confidence interval) for Western Sandpiper nests on primary and secondary plot at Kanaryarmiut Field Station, Alaska, 1999–2002. Number of nests used in calculations (primary, secondary plot respectively): 51, 15 in 1999; 53, 72 in 2000; 54, 77 in 2001; 42, 16 in 2002.

Arctic Fox *Alopex lagopus*, Red Fox *Vulpes vulpes*, Mink *Mustela vison*, Long-tailed Jaeger *Stercorarius longicaudus*) may have entirely or nearly overlapped our study plots, differences in predation intensity between the two plots, both between and within years, may simply have been a function of slight variation in home range use patterns of individual predators. As a result, the differences we found between adjacent plots may be more extreme than those found in other situations where larger plots are used. We contend, however, that our findings remain germane for two reasons. First, at least in the North American arctic, the sizes of study plots used by shorebird researchers are often between 10 and 100 ha; our plots (at 16, 30, and 46 ha, respectively) are well within this range. Thus, many research efforts may indeed be vulnerable to the types of problems our data illustrate. Secondly, even at sites where much larger study areas are selected, the potential sampling universe is still often huge by comparison, and the same principle applies: when plots are either non-randomly selected and/or not replicated, there is little statistical basis for valid inference.

An appropriate solution to this problem of limited inference is easy to suggest, but difficult to implement. Just as schemes for estimating population size require statistically valid sampling designs (e.g., with randomization, replication, and the potential for samples to be drawn from anywhere within the area of inference), so, too, do schemes for estimating demographic parameters. Given the magnitude of effort needed to derive estimates of parameters such as productivity or survival at single sites, however, conducting enough replicates to statistically distinguish differences across space and/or time may be prohibitively expensive.

Thus, we are confronted with a dilemma. Current practice is frequently inappropriate, but often appropriate practice is difficult to achieve. Although a detailed solution is beyond the scope of this paper, four steps are essential to moving forward. First, when using inferential statistics, we must be conscientious about limiting our inferences to the area from which the samples were drawn, and we must be explicit that inferences beyond that area are really just hypotheses yet to be tested. Second, if we wish to draw inferences about demographic parameters over large geographic areas, we need to determine study design requirements that will generate statistically valid estimates rigorously. Third, when limited funding or logistics preclude getting the precision of estimates required by managers, or when potentially important information (e.g., low survival of adults) is available only from small study sites not selected via formal sampling procedures, we need to explore options for decision-making in the face of uncertainty. Such analyses might involve elements of formal decision theory and Bayesian analyses.

Finally, we should keep in mind the value of metareplication, i.e., the replication of studies themselves by different researchers, at different sites, and in different years (Johnson 2002). "Metareplication provides us greater confidence that certain relationships are general. Obtaining consistent inferences from studies conducted under a wide variety of conditions will assure us that the conclusions are not unique to the particular set of circumstances that prevailed during the study" (Johnson 2002, p. 930). Such an approach is implicitly being used by the International Arctic Birds Breeding Conditions Survey (IABBCS) coordinated by M.Y. Soloviev and P.S. Tomkovich (e.g., Soloviev & Tomkovich 2002), and has also been proposed by E. Pierce and H. Meltofte for the Pan-Arctic Shorebird Research Network (PASRN). In both programs, results from shorebird researchers from across the Arctic are pooled to evaluate the effects of rodent cycles, weather, and/or climate change on shorebird population dynamics. Thus, even when we are unable to apply the tools of statistical inference to these questions, there would be some reason for confidence in our conclusions if and when there is concordance among dozens of studies and sites from across the Arctic.

Table 1. Logistic regression output comparing variation in daily survival rate for Western Sandpiper nests on primary and secondary plot at Kanaryarmiut Field Station, Alaska, 1999–2002.

Variable	Degrees of Freedom	Chi-square Value	*P*
Year	3	8.45	0.0376
Plot	1	4.60	0.0319
Year × Plot	3	24.90	<0.0001

Despite its attractiveness, however, metareplication should not be seen as a panacea. The conclusions drawn from such efforts are only as valid and powerful as the individual studies that were replicated. Those of us who do, or are planning to, contribute to the IABBCS and/or the PASRN must ensure that our individual studies are conducted as rigorously as possible, and that alternatives to the prevailing causal hypotheses are not ignored. Similarly, although metareplication can allow for a certain degree of "common sense" inference, we should not ignore opportunities to design and implement standardized, statistically valid sampling protocols for assessing changes in shorebird population size and demography, even at spatial scales as large as ecoregions and continents.

ACKNOWLEDGMENTS

We appreciate the support provided by the entire staff of Yukon Delta National Wildlife Refuge. We would also like to thank our field assistants for their competence and camaraderie: Zach Fairbanks, Chadd Fitzpatrick, Laird Hamblin, Amanda Niehaus, Ron Opsahl, John Pearce, Daniel Rizzolo, and Melanie Spies. Chadd Fitzpatrick and Melanie Spies helped compile and analyze references for nest success studies. We appreciate Hans Schekkerman's thoughtful comments on an earlier draft, as well as Colleen Handel's patient statistical assistance. John Pearce provided insightful editorial comments. Finally, we thank Bob Gill for constantly sharing his wealth of shorebird knowledge and for encouraging us to formally develop and share the ideas presented in this paper.

REFERENCES

Aebischer, N.J. 1999. Multi-way comparisons and generalized linear models of nest success: extensions of the Mayfield method. *Bird Study* 46: S22–31.

Blomqvist, S., N. Holmgren, S. Åkesson, A. Hedenström & J. Pettersson. 2002. Indirect effects of lemming cycles on sandpiper dynamics: 50 years of counts from southern Sweden. *Oecologia* 133: 146–158.

Burke, D.M. & E. Nol. 1998. Influence of food abundance, nest-site habitat, and forest fragmentation on breeding ovenbirds. *Auk* 115: 96–104.

Caffrey, C. 1999. Feeding rates and individual contributions to feeding at nests in cooperatively breeding western American crows. *Auk* 116: 836–841.

Chase, M.K. 2002. Nest site selection and nest success in a Song Sparrow population: The significance of spatial variation. *Condor* 104: 103–116.

Clark, W.R., R.A. Schmitz & T.R. Bogenschultz. 1999. Site selection and nest success of ring-necked pheasants as a function of location in Iowa landscapes. *J. Wildlife Management* 63: 976–989.

Davison, W.B. & E. Bollinger. 2000. Predation rates on real and artificial nests of grasslands birds. *Auk* 117: 147–153.

Dewey, S.R. & P.L. Kennedy. 2001. Effects of supplemental food on parental-care strategies and juvenile survival of Northern Goshawks. *Auk* 118: 352–365.

Duguay, J.P., P.B. Wood & G.W. Miller. 2000. Effects of timber harvests on invertebrate biomass and avian nest success. *Wildlife Soc. Bull.* 28: 1123–1131.

Fauth, P.T. 2000. Reproductive success of wood thrushes in forest fragments in northern Indiana. *Auk* 177: 194–204.

Garrettson, P.R. & F.C. Rohwer. 2001. Effect of mammalian predator removal on production of upland-nesting ducks in North Dakota. *J. Wildlife Management* 65: 398–405.

Hersek, M.J., M.A. Frankel, J.A. Cigliano & F.E. Wasserman. 2002. Brown-headed cowbird parasitism of ovenbirds in suburban forest fragments. *Auk* 119: 240–243.

Johnson, D.H. 2002. The importance of replication in wildlife research. *J. Wildlife Management* 66: 919–932.

Manolis, J.C., D.E. Andersen & F.J. Cuthbert. 2000. Uncertain nest fates in songbird studies and variation in Mayfield estimation. *Auk* 117: 615–626.

Mayfield, H.F. 1975. Suggestions for calculating nest success. *Wilson Bull.* 87: 456–466.

Mermoz, M.E. & J.C. Reboreda. 1998. Nesting success in brown-and-yellow marshbirds: effects of timing, nest site, and brood parasitism. *Auk* 115: 871–878.

Popotnik, G.J. & W.M. Guiliano. 2000. Response of birds to grazing of riparian zones. *J. Wildlife Management* 63: 976–989.

Reynolds, R.E., T.L. Shaffer, R.W. Renner, W.E. Newton & B.D.J. Batt. 2001. Impact of the conservation reserve program on duck recruitment in the U.S. prairie pothole region. *J. Wildlife Management* 65: 765–780.

Rodenwald, A.D. & R.H. Yahner. 2001. Avian nesting success in forested landscapes: influence of landscape composition, stand and nest-patch microhabitat, and biotic interactions. *Auk* 118: 1018–1028.

Roselaar, C.S. 1979. Variation in numbers of Curlew Sandpipers (Calidris ferruginea) [in Dutch with English summary]. Watervogels *4: 202–210.*

Ruthrauff, D.R. 2002. Seasonal and age-related trends in the reproductive output of Western Sandpipers (*Calidris mauri*) at Kanaryaraq, Alaska. M.Sc. thesis, Humboldt State University, Arcata, California.

SAS Institute. 1999. Version 8 Edition. SAS Institute, Cary, North Carolina.

Soloviev, M.Y. & P.S. Tomkovich. 2002. A word from the compilers. *Arctic Birds* 4: 1.

Summers, R.W. & L.G. Underhill. 1987. Factors related to breeding production of Brent Geese *Branta b. bernicla* and waders (Charadrii) on the Taimyr Peninsula. *Bird Study* 34: 161–171.

Tarvin, K.A. & M.C. Garvin. 2002. Habitat and nest success of Blue Jays (*Cyanocitta cristata*): Importance of scale. *Auk* 119: 971–983.

Thogmartin, W.E. 1999. Landscape attributes and nest-site selection in wild turkeys. *Auk* 116: 912–923.

Tomkovich, P.S. & M.Y. Soloviev. 2002. Bird breeding conditions in the Arctic in 2001. *Arctic Birds* 4: 23–27.

Underhill, L.G., M. Waltner & R.W. Summers. 1989. Three-year cycles in breeding productivity of Knots *Calidris canutus* wintering in southern Africa suggest Taimyr Peninsula provenance. *Bird Study* 36: 83–87.

Underhill, L.G., R.P. Prys-Jones, E.E. Syroechkovski, Jr., N.M. Groen, V. Karpov, H.G. Lappo, M.W.J. van Roomen, A. Rybkin, H. Schekkerman, H. Spiekman & R.W. Summers. 1993. Breeding of waders (Charadrii) and Brent Geese *Branta bernicla bernicla* at Pronchishcheva Lake, northeastern Taimyr, Russia, in a peak and a decreasing lemming year. *Ibis* 135: 277–292.

Yahner, R.H. 2000. Long-term effects of even-aged management on bird communities in central Pennsylvania. *Wildlife Soc. Bull.* 28: 1102–1110.

Zanette, L. & B. Jenkins. 2000. Nesting success and nest predators in forest fragments: a study using real and artificial nests. *Auk* 117: 445–454.