ESTIMATING POPULATION TRENDS WITH A LINEAR MODEL: TECHNICAL COMMENTS

Authors: Sauer, John R., Link, William A., and Royle, J. Andrew

Source: The Condor, 106(2) : 435-440

Published By: American Ornithological Society

URL: https://doi.org/10.1650/7431

COMMENTARY

The Condor 106:435–440

ESTIMATING POPULATION TRENDS WITH A LINEAR MODEL: TECHNICAL COMMENTS

JOHN R. SAUER,[1] WILLIAM A. LINK AND J. ANDREW ROYLE

USGS Patuxent Wildlife Research Center and U.S. Fish and Wildlife Service, 12100 Beech Forest Road, Laurel, MD 20708-4039

Abstract. Controversy has sometimes arisen over whether there is a need to accommodate the limitations of survey design in estimating population change from the count data collected in bird surveys. Analyses of surveys such as the North American Breeding Bird Survey (BBS) can be quite complex; it is natural to ask if the complexity is necessary, or whether the statisticians have run amok. Bart et al. (2003) propose a very simple analysis involving nothing more complicated than simple linear regression, and contrast their approach with model-based procedures. We review the assumptions implicit to their proposed method, and document that these assumptions are unlikely to be valid for surveys such as the BBS. One fundamental limitation of a purely design-based approach is the absence of controls for factors that influence detection of birds at survey sites. We show that failure to model observer effects in survey data leads to substantial bias in estimation of population trends from BBS data for the 20 species that Bart et al. (2003) used as the basis of their simulations. Finally, we note that the simulations presented in Bart et al. (2003) do not provide a useful evaluation of their proposed method, nor do they provide a valid comparison to the estimating-equations alternative they consider.

Key words: BBS, estimation, North American Breeding Bird Survey, population change, surveys, trends.

Estimando Tendencias Poblacionales con un Modelo Lineal: Comentarios Técnicos

Resumen. A veces ha surgido controversia sobre la necesidad de considerar las limitantes del diseño de muestreo al estimar cambios poblacionales a partir de datos de conteos de aves. Los análisis de muestreos como el Muestreo de Aves Reproductivas de América del Norte (North American Breeding Bird Survey [BBS]) pueden ser bastante complejos; es natural preguntarse si esta complejidad es necesaria, o si los análisis estadísticos son desmedidos. Bart et al. (2003) proponen un análisis muy simple que sólo involucra regresión lineal simple, y contrastan su enfoque con los procedimientos basados en modelos. Nosotros revisamos los supuestos implícitos en el método que ellos proponen y documentamos que estos supuestos no son probablemente válidos para muestreos tales como el BBS. Una limitante fundamental de un enfoque basado exclusivamente en el diseño es la ausencia de controles para factores que influencian la detección de aves en los sitios de muestreo. Mostramos que el hecho de no modelar los efectos del observador en los datos de muestreo lleva a sesgos substanciales en las estimaciones de las tendencias poblacionales de las 20 especies que Bart et al. (2003) usaron como la base de sus simulaciones a partir de datos del BBS. Finalmente, notamos que las simulaciones presentadas en Bart et al. (2003) no brindan una evaluación útil del método que proponen ni tampoco ofrecen una comparación válida para la alternativa de estimación de ecuaciones que ellos consideran.

Bart et al. (2003) describe a design-based approach for estimation of population change from count survey data. They promote their method as ''simple, self-weighting, and versatile'' (p. 371), and conduct simulations based on data from the North American Breeding Bird Survey (BBS) and the International Shorebird Survey. They contrast their estimator with an estimating-equations estimator, and suggest that their procedure is superior to the alternative approach. The estimation technique advocated by Bart et al. (2003) has serious conceptual and practical limitations; its use is likely to lead investigators to invalid conclusions about population change. In this note, we identify a few of these deficiencies, and suggest that users be cautious in implementing methods that stress simplicity while ignoring critical design issues of the survey and important features of the data. We show that the proposed analysis is based on unreasonable assumptions about the nature of count data, that the evaluation of the method is conducted under conditions that predispose it to be favored, and that the Bart et al. (2003) analyses provide biased estimates of population trends from BBS data.

DESIGN-BASED ANALYSES

Bart et al. (2003) place great importance on the notion that their estimator is design-based, and briefly dismiss model-based analysis with suggestions that it is biased, complicated, and that ''different methods perform best

---

Manuscript received 15 August 2003; accepted 20 January 2004.

[1] E-mail: john_r_sauer@usgs.gov

in different situations'' (p. 367). They say that design-based estimators ''assume only that the sampling plan was followed and that the sample size is large enough to make inferences based on the central limit theorem'' (p. 367). The second half of this assertion is false: no normality assumption is required for design-based analysis. The first half is crucial, and virtually never satisfied by count surveys such as the BBS upon which we focus our discussion. Design-based procedures are based on assumptions that values of quantities measured at sample sites are used to estimate a population parameter, and that a random sampling scheme is used as the basis of assessing the precision of the estimator. To meet these assumptions, the procedure must have the following features: (1) There is a well-defined finite population of sites. (2) The data are collected subject to a clearly defined scheme for random sampling of the sites. (3) Associated with each site there is a quantity that can be measured without error. (4) The population parameter to be estimated is defined as a function of the site-specific quantities. We address each of these points below.

Finite population of sites with clearly defined scheme for random sampling. Although there is an element of randomness in selection of starting points of BBS survey routes, the routes cannot be viewed as random selections from a sample frame. The most obvious deficiency is that BBS routes do not survey habitats >0.4 km from secondary roads. Any BBS sampling frame would be restricted to roadsides along a subset of roads. Even along roadsides, route selection procedures are not based on a random selection from a predetermined sampling frame of possible routes, and routes often cross other routes. Sampling intensity varies temporally and regionally, without regard to a preestablished design (Sauer, Fallon, and Johnson 2003). Finally, the definition of the BBS sample unit is vague, since counts cover an unknown area and area covered by a route undoubtedly differs due to quality of observers (Link and Sauer 1998a). BBS sampling methods cannot guarantee either a census or a known fixed area of sampling, facts well known to the originators of the survey (Robbins et al. 1986). Consequently, one cannot conceive a ''population of sites'' from which BBS routes are sampled, except as an abstraction (i.e., as a model, Link and Sauer 1998a).

From their discussion and simulations, it is clear that Bart et al. (2003) regard BBS routes as a simple random sample from a finite population, sampled as replicates at the continental scale. A great deal of evidence based on both the limitations of defining a sample frame at local scales and the need to impose regional strata indicate that this is an oversimplification (e.g., James et al. 1996, Peterjohn et al. 1995, Link and Sauer 1998a, Sauer, Fallon, and Johnson 2003). As Bart et al. (2003) acknowledge, samples from many other surveys, such as the International Shorebird Survey, are even less appropriately viewed as random selections from sampling frames.

Measurable quantities at sites. Definition of counts at sites has been the source of considerable controversy in the analysis of count data, as the distinction between counts, indexes, and actual population sizes is often obscured. Bart et al. (2003:368) suggest that analysis of count survey data should begin with

> ''an estimate of, or an index to, population size Y_j in year j; and that any adjustments to account for change in detection rates (e.g., due to change in average observer ability) have been made. Methods for incorporating such adjustments into the trend analysis, rather than making them beforehand, will be discussed in a future paper. The Y_j may thus be regarded as the true population sizes, if survey methods permit an unbiased estimate, or more generally, as the expected values of the survey means in each year.''

This statement sets the stage for an inappropriate analysis. In the first sentence Y_j is the population size; an unnamed ''estimate of, or index to'' Y_j is to be used as the basis of trend analysis. In the third sentence, Y_j is used for the index itself, but with the assurance that it may ''be regarded as the true population size.'' The qualification that the index may be regarded ''more generally, as the expected value of the survey means'' is tautological. Are we interested in doing an analysis of trends in the expected values of the survey means, or trends in the animal populations?

The difference between indices, estimates, and true population sizes is neither superficial nor insignificant. It is simply incorrect to treat indices or estimates as true population sizes in trend analysis, without taking into account the manner in which (necessarily model-based) ''adjustments'' were made. Nor need the scientific community wait for further research to show how to simultaneously adjust data and perform trend analysis; there is already a substantial literature (James et al. 1996, Fewster et al. 2000, Link and Sauer 2002).

Site-specific variation. Bart et al. (2003) choose to overlook all variability except that which occurs among sites. Given the amount of within-site variability associated with BBS data, the availability of site-specific covariables associated with observer characteristics known to contribute to this variability, and the documented confounding of such variability with population trend, application of the method to BBS data would be irresponsible. A first step to a more realistic description of site-specific variation would be to address measurement error (e.g., Fuller 1987); in our view the capacity for model-based analysis of site-specific variation is a necessary component of analysis of BBS data. Count surveys such as the BBS do not have detectability estimation as a component of the design, and the only way to accommodate factors that influence counts is through site- and time-specific covariables (Link and Sauer 1998a, Bennetts et al. 1999). We discuss this in more detail in a later section.

Defining a parameter of interest. Bart et al. (2003:368) define trend as follows:

> ''We assume that a scatterplot of the true population sizes, plotted against time, would reveal a pattern that is well described using an exponential curve. We do not assume the true population sizes fall on this exponential curve (this would be an assumption typical of model-based approaches), only that the exponential curve would describe the pattern in a useful way.''

This statement is misleading and vague. First, one can hardly say that assuming the population sizes fall precisely on an exponential curve is typical of model-based approaches. Second, the assumption as stated is vague: does it mean that

$$Y_j = \exp(\alpha + \beta t) + \varepsilon_j,$$

where ε_j are independent and identically distributed mean-zero variables? Or are the errors assumed to be additive on the log scale, viz.,

$$\ln(Y_j) = \alpha + \beta t + \varepsilon_j.$$

In either case, are we to assume homoscedasticity? The implicit answer seems to be that it does not matter. There seems to be no way to judge, either: the exponential curve is merely a ''useful'' description. A more precise definition of usefulness is needed.

At any rate, none of this ambiguity is necessary. From their simulation study, it is apparent that Bart et al. (2003) define R_{exp} by β^*, where $\{\alpha^*, \beta^*\}$ is the minimizer of

$$Q_1(\alpha, \beta) = \sum_{j=1}^{J} [Y_j - \exp(\alpha + \beta j)]^2.$$

Here, and subsequently, we use Y_j to denote a true population total in year j.

It is important to note that Bart et al. (2003) choose R_{exp} as their parameter, but then estimate a different quantity. The quantity they estimate is R_{lin}, defined by b^*, where $\{a^*, b^*\}$ is the minimizer of

$$Q_2(a, b) = \sum_{j=1}^{J} [Y_j - (a + bj)]^2.$$

In our view, the choice of estimator should be governed by the choice of parameter; Bart et al. (2003) instead choose an estimator for convenience, then attempt to rationalize its use. On the one hand, the parameter b may often closely approximate the parameter β, as demonstrated in Bart et al.'s (2003) Table 1. (We note that Table 1 was computed assuming population sizes falling precisely on the linear curve, a condition which favors the approximation.) On the other hand, it is possible that the approximation will not be so close: for example, the collection of population sizes {1207, 1009, 251, 512, 655, 356, 377, 469, 556, 389, 659, 266,673,477,871,609,743,1074,1064,1150} has $R_{exp} = 1.0218$ and $R_{lin} = 1.0321$. It might be suggested for these data that the exponential curve is not a ''useful'' description of the pattern. However, usefulness cannot be defined, and discrepancy from the posited pattern cannot be measured, without resort to a model-based analysis.

Our experience has been that much pointless discussion on the topic of trend analysis arises due to a failure to begin by defining just what trend is. The methods described by Bart et al. (2003) have the virtue of being based on a precise (though unclearly articulated) notion of trend, but we argue that analytical methods should be based on the definition, rather than an approximation. We suggest that a better definition of trend is the geometric mean rate of change over a particular interval: this definition does not require qualitative assessments as to the usefulness of exponential or linear patterns, and seems to closely approximate the informal usage of the term ''trend.'' See Link and Sauer (1998a) for a discussion of this definition of trend.

HISTORICAL PERSPECTIVE

We view the Bart et al. (2003) approach as motivated by the same general design considerations that guided earlier investigations of count survey data, but as failing to adopt the lessons learned by applications of those methods. The innovation claimed by Bart et al. (2003) is that site-specific trends can be aggregated in a design-based framework to estimate an overall trend. Geissler and Noon (1981), and Geissler (1984) used similar ratio estimators of site-specific population change in estimation of a composite change as an average of site-specific change. These early efforts and a number of subsequent developments (Geissler and Sauer 1990, James et al. 1990, Link and Sauer 1994) shared the notion of averaging site-specific estimates, but were generally complicated by two technical issues: There was controversy about (1) the appropriate way to estimate change for each route in the face of needed covariates and model-fitting problems (Link and Sauer 1997a), and (2) the need to weight each route to accommodate detectability and consistency of survey. Bart et al. (2003) dismiss these issues as either model-based complications or as items requiring future innovations. These issues are critical components of the analysis; for example, observer covariates are an absolute necessity to avoid bias in trend estimates (e.g., Sauer et al. 1994, Link and Sauer 1998b). In the next section, we specify the consequences of omitting observer covariates.

SITE-SPECIFIC COVARIATES

The need for site-specific covariates for factors that influence counting efficiency is well established for the BBS and most count-based surveys. In the BBS, Sauer et al. (1994) clearly documented the bias in estimation associated with the failure to include observer information as covariates in BBS analysis. Kendall et al. (1996) documented the presence of further observer effects associated with the first year of counting on routes. Link and Sauer (1998b) explicitly modeled a ''new observer'' effect that expresses the increase in counts associated with improvement in observer quality over time in the BBS. James et al. (1996) included observer covariates in their semiparametric analysis of BBS data. A purely design-based estimation procedure cannot accommodate observer covariates. Here, we illustrate the consequences of their omission. Bart et al. (2003) analyze data from 20 species in the BBS. Although they label the analysis ''true trend,'' it is actually a complete analysis of a subset of the larger BBS dataset, which even in its entirety, would not provide a true trend at the population level. We analyzed these same species for the same time period using all available data, but following the estimating-equations approach described in Link and Sauer (1994). Note that this is not the estimating-equations procedure followed by Bart et al. (2003; see below). We performed our analyses with and without controls for observer effects, and predicted that omission would overall lead to more positive trend estimates as documented in the literature (Sauer et al. 1994, Link and Sauer 1998a). Letting *EEW* denote the estimating effects estimator *with* observer effects, and *EEWO* denote the estimating effects estimator *without* observer effects, we conducted paired *t*-tests and found results consistent with our prediction (mean difference $EEWO - EEW =$ 0.20% per year, one-sided paired *t*-test, $t_{19} = -2.05$, $P = 0.03$). We believe that the Bart et al. (2003) analysis, as implemented in this paper for BBS data, will lead to biased estimates because of its failure to incorporate observer effects.

We also note that avoiding the bias due to ignoring observer effects comes at an unavoidable cost. Adding controls for observers diminishes precision. In the present case variances from estimates with observer effects were almost twice as large as variances from estimates without observer effects. It is however, a false economy to suppose that one may use an unrealistically small estimate of precision in making inference about population trends, or in planning future studies. It is sometimes true that biased estimates have smaller mean squared error than corresponding unbiased estimates: there is a trade-off between accuracy and precision. However, it is also sometimes true that biases lead to incorrect inferences. Conclusions drawn from monitoring programs and associated analyses must withstand scrutiny. Systematic bias is a fatal flaw, unless demonstrably of inconsequential magnitude.

SIMULATIONS

Many practitioners, seeing the simulation results presented in Bart et al. (2003), would be convinced that the method they present is at least as good as other published works, and in some cases much better than existing methods. We believe that deficiencies of the simulations limit their usefulness.

First, the ‘‘data sets’’ considered (and treated as hypothetical populations) correspond to no real populations, despite Bart et al.'s desire to ‘‘make maximum use of real data’’ (p. 369). As noted previously, BBS data include a large component of variability due to observers. This, and other sources of variation which could well be confounded with population change, have been ignored. Bart et al. (2003) have simply defined away many of the possible problems which necessitate a model-based analysis. In our view, this narrow view of what is to be estimated is not relevant.

The satisfactory performance of Bart et al.'s estimator is simply a consequence of the design-based sampling in their simulation, and the satisfactory approximation of R_{exp} by R_{lin} in the hypothetical populations. The relevant questions to be addressed are whether the design-based estimator will work when sampling is not design based and when nonlinear patterns of change exist, but the Bart et al. simulations provide no information on these questions. In addition, to apply Bart et al.'s proposal to the BBS, the question must be addressed as to whether observer effects can be overlooked. These questions have not been addressed.

Using our *EEWO* (estimating effects estimator *without* observer effects) and *EEW* (estimating effects estimator *with* observer effects) results described above, we compared our results to the Bart et al. (2003) results. We predicted that in the absence of controls for observer effects our results would correspond to those of Bart et al. (2003), but that with observer effects controlled for, the trend estimates would be lower. Letting *D* denote Bart et al.'s design-based estimator, we conducted paired *t*-tests, finding *P*-values of 0.48 for comparison of *D* and *EEWO*, and 0.03 for comparison of *D* and *EEW*, consistent with our predictions.

The simulations presented by Bart et al. (2003) present an alternative estimating-equations estimator in a rather poor light. The comparison is unfair: the estimating-equations estimator does not estimate R_{exp}, but rather, a precision- and abundance-weighted average of site-specific trends. The performance of this estimator of population trend must depend on the definition of population trend, and on the appropriateness of the weights applied. We cannot comment on the weights chosen by Bart et al. (2003), except to note that they are not the same as those described in Link and Sauer (1994, Geissler and Sauer 1990), appearances notwithstanding. Although it is not our intent to defend the estimating-equation procedures discussed in Bart et al. (2003) or those presented by Link and Sauer (1994), we direct readers to other simulations (Thomas 1996) and comparisons (Link and Sauer 1996, Sauer, Hines, and Fallon 2003) of analyses based on estimating equations and other route regression approaches, which we believe more satisfactorily evaluate the performance of alternative methodologies.

DISCUSSION

Analyses of count data can be controversial, and a great deal of caution is needed to avoid weakening our credibility as scientists and managers by making unwarranted statements based on flawed analyses. Simplicity is a great virtue in analysis of survey data, but, as the comment attributed to Einstein says, ‘‘Things should be made as simple as possible—but no simpler.’’ The risk is that excessive simplicity may compromise the credibility of results obtained from count survey data. In our view, the notion that BBS data can be effectively treated as a random sample of population sizes is wrong, and the primary failure of the Bart et al. (2003) approach is that it perpetuates this view by ignoring important features of the analysis and by structuring simulations of BBS data as though the counts were actual populations. Their approach also ignores the lessons from the history of the survey. The original conception of the BBS was that of a design-based survey, but this view was abandoned when it became apparent that model-based adjustments were needed to accommodate uneven survey coverage and covariates that influence counts. While it is useful to occasionally evaluate assumptions and to refine analyses, new analyses should not ignore features that have been shown to be important in past analyses. Although we have focused our discussion on the BBS, we note that even greater constraints exist on analysis of other continental-scale surveys such as the Christmas Bird Count (Link and Sauer 1999) and the International Shorebird Survey (Howe et al. 1989).

Modern approaches to the analysis of BBS data reflect the necessity of accommodating the large changes in number and consistency of routes surveyed over time, and attempts to analyze or simulate the survey must appropriately incorporate the variation induced by these logistical constraints. Sauer, Fallon, and Johnson (2003, their table 4) document that the amount of missing data in the BBS varies regionally, and often exceeds that considered by Bart et al. (2003). The BBS database is characterized by constant addition of new survey routes that add an additional challenge for analysts. Pattern in mean counts can be induced by adding new survey routes, and this observation has been used as evidence of the failure of simple design-based approaches to analysis of BBS data (Geissler and Noon 1981, James et al. 1990). Rather than develop methods that ignore these complications, a need exists to inform users about appropriate analyses and to identify situations when simple approaches are inadequate. With increased availability of BBS data via the Internet, there is increased risk of misuse of the information due to simplistic analyses. Clearly, a need exists for more extensive metadata associated with the survey to guide users to appropriate analyses, and metadata provided by Sauer, Hines, and Fallon (2003) initiate an attempt to define some of these issues.

Although this commentary has focused on the technical aspects of trend estimation, we also note that implementation of a procedure such as that proposed by Bart et al. (2003) has strategic implications for bird conservation. Increasing information needs for management and increased sophistication of analysis methods provide an opportunity to better integrate monitoring data with scientific and management activities such as adaptive resource management (Ruth et al. 2003). Analyses such as that described by Bart et al. (2003) step away from these opportunities by rejecting model-based approaches that allow direct modeling of the influence of environmental variables on counts and by focusing on the very limited goal of trend analysis. We agree with James et al. (1996) on the limits of use of trend information in science and management, and suggest that any modern analysis of BBS or International Shorebird Survey data should have the capability of directly modeling both more-general aspects of population change and covariates that influence detectability and population size.

We suggest that investigators interested in estimating population change from the BBS or other count-based bird surveys use one of the many approaches that accommodate the constraints of the surveys. Examples of these methods include hierarchical models (Link and Sauer 2002), overdispersed Poisson models (Link and Sauer 1997b), generalized additive models (James et al. 1996, Fewster et al. 2000), as well as more traditional approaches (e.g., Sauer and Droege 1990). Most of these analyses are readily available using computer programs or Internet-based programs; estimating-equation and general-additive-model estimation approaches are presently available on the BBS Analysis and Summary Internet site (Sauer, Hines, and Fallon 2003).

We thank J. D. Nichols, G. W. Pendleton, and an anonymous reviewer for comments on the manuscript.

LITERATURE CITED

BART, J., B. COLLINS, AND R. I. G. MORRISON. 2003. Estimating population trends with a linear model. Condor 105:367–372.

BENNETTS, R. E., W. A. LINK, J. R. SAUER, AND P. W. SIKES JR. 1999. Factors influencing counts in an annual survey of Snail Kites in Florida. Auk 116: 316–323.

FEWSTER, R. M., S. T. BUCKLAND, G. M. SIRIWARDENA, S. R. BAILLIE, AND J. D. WILSON. 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology 81:1970–1984.

FULLER, W. A. 1987. Measurement error models. 9th ed. Wiley, New York.

GEISSLER, P. H. 1984. Estimation of animal population trends and annual indices from a survey of call-counts or other indications. Proceedings of the American Statistical Association, Section on Survey Research Methods 1984:472–477.

GEISSLER, P. H., AND B. R. NOON. 1981. Estimates of avian population trends from the North American Breeding Bird Survey. Studies in Avian Biology 6:42–51.

GEISSLER, P. H., AND J. R. SAUER. 1990. Topics in route regression analysis, p. 54–57. *In* J. R. Sauer and S. Droege [eds.], Survey designs and statistical methods for the estimation of avian population trends. USDI Fish and Wildlife Service Biological Report 90(1), Washington, DC.

HOWE, M. A., P. H. GEISSLER, AND B. A. HARRINGTON. 1989. Population trends of North American shorebirds based on the International Shorebird Survey. Biological Conservation 49:185–199.

JAMES, F. C., C. E. MCCULLOCH, AND D. A. WIEDENFELD. 1996. New approaches to the analysis of population trends in land birds. Ecology 77:13–27.

JAMES, F. C., C. E. MCCULLOCH, AND L. E. WOLFE. 1990. Methodological issues in the estimation of trends in bird populations with an example: the Pine Warbler, p. 84–97. *In* J. R. Sauer and S. Droege [eds.], Survey designs and statistical methods for the estimation of avian population trends. USDI Fish and Wildlife Service Biological Report 90(1), Washington, DC.

KENDALL, W. L., B. G. PETERJOHN, AND J. R. SAUER. 1996. First-time observer effects in the North American Breeding Bird Survey. Auk 113:823–829.

LINK, W. A., AND J. R. SAUER. 1994. Estimating equations estimates of trends. Bird Populations 2:23–32.

LINK, W. A., AND J. R. SAUER. 1997a. New approaches to the analysis of population trends in land birds: a comment on statistical methods. Ecology 78: 2632–2634.

LINK, W. A., AND J. R. SAUER. 1997b. Estimation of population trajectories from count data. Biometrics 53:63–72.

LINK, W. A., AND J. R. SAUER. 1998a. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological Applications 8:258–268.

LINK, W. A., AND J. R. SAUER. 1998b. Estimating relative abundance from count data. Austrian Journal of Statistics 27:83–97.

LINK, W. A., AND J. R. SAUER. 1999. Controlling for varying effort in count surveys—an analysis of

Christmas Bird Count data. Journal of Agricultural, Biological and Environmental Statistics 4:116–125.
LINK, W. A., AND J. R. SAUER. 2002. A hierarchical model of population change with application to Cerulean Warblers. Ecology 83:2832–2840.
PETERJOHN, B. G., J. R. SAUER, AND C. S. ROBBINS. 1995. The North American Breeding Bird Survey and population trends of Neotropical migrant birds, p. 3–39. *In* T. E. Martin and D. Finch [eds.], Ecology and management of Neotropical migrant birds. Cambridge University Press, New York.
ROBBINS, C. S., D. BYSTRAK, AND P. H. GEISSLER. 1986. The Breeding Bird Survey: its first fifteen years, 1965–1979. USDI Fish and Wildlife Service Resource Publication 157, Washington, DC.
RUTH, J. M., D. R. PETIT, J. R. SAUER, M. D. SAMUEL, F. A. JOHNSON, M. D. FORNWALL, C. E. KORSCHGEN, AND J. P. BENNETT. 2003. Science for avian conservation: priorities for the new millennium. Auk 120:204–211.
SAUER, J. R., AND S. DROEGE [EDS.]. 1990. Survey designs and statistical methods for the estimation of avian population trends. USDI Fish and Wildlife Service Biological Report 90(1), Washington, DC.
SAUER, J. R., J. E. FALLON, AND R. JOHNSON. 2003. Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. Journal of Wildlife Management 67:372–389.
SAUER, J. R., J. E. HINES, AND J. FALLON [ONLINE]. 2003. The North American Breeding Bird Survey, results and analysis 1966–2002. Version 2003.1 <http://www.mbr-pwrc.usgs.gov/bbs/bbs.html> (11 August 2003).
SAUER, J. R., B. G. PETERJOHN, AND W. A. LINK. 1994. Observer differences in the North American Breeding Bird Survey. Auk 111:50–62.
THOMAS, L. 1996. Monitoring long-term population change: why are there so many analysis methods? Ecology 77:49–58.

The Condor 106:440–443

ESTIMATING TRENDS WITH A LINEAR MODEL: REPLY TO SAUER ET AL.

JONATHAN BART[1,3], BRIAN COLLINS[2] AND R. I. G. MORRISON[2]

[1]*USGS Forest and Rangeland Ecosystem Science Center, 970 Lusk Street, Boise, ID 83706*
[2]*Canadian Wildlife Service, National Wildlife Research Centre, Raven Road, Carleton University, Ottawa, ON K1A 0H3, Canada*

Abstract. Sauer et al. (2004) advocate the use of trend estimation models that adjust counts for differences among observers. We agree that such adjustments are sometimes needed, and we noted (Bart et al. 2003) that they may readily be carried out prior to using the estimation method we described. Including observer covariates, however, is not always necessary and substantially reduces precision, as Sauer et al. (2004) acknowledge. Furthermore, under plausible conditions, including observer covariables introduces bias rather than removing it. In addition, the weighting scheme used in the estimating-equations approach may introduce bias. Our method avoids these sources of bias, is simpler and more flexible than the estimating-equations approach (e.g., carrying out power and sample-size calculations is much easier with our approach), and has smaller standard errors than the estimating-equations approach, especially when counts fluctuate widely. Model-based methods, including the estimating-equations approach, also have advantages, particularly in assessing complex influences on the counts. We recommend that analysts consider both approaches; comparing results obtained with the different methods may be especially informative.

Key words: bias, Breeding Bird survey, indices, trends.

Estimación de Tendencias con un Modelo Lineal: Respuesta a Sauer et al.

Resumen. Sauer et al. (2004) recomiendan el uso de modelos de estimación de tendencias que ajusten los conteos a las diferencias existentes entre observadores. Nosotros estamos de acuerdo en que dichos modelos podrían ser útiles, y sugerimos que estos ajustes pueden incorporarse fácilmente antes de usar el método de estimación que describimos. Nosotros introdujimos nuestro método porque es más sencillo y más flexible que el método que requiere estimar ecuaciones (e.g., realizar cálculos de poder estadístico y de tamaños de muestra es mucho más fácil con nuestro método), y porque el nuestro se desempeñó mejor que el de estimación de ecuaciones cuando los conteos fluctuaron ampliamente. Adicionalmente, el procedimiento de pesaje usado en el método de estimación de ecuaciones podría introducir sesgos, mientras que el procedimiento lineal que nosotros describimos se pesa a sí mismo y no es susceptible a este error. Sin embargo, el método de estimación de ecuaciones también ofrece ventajas, particularmente en su habilidad para manejar modelos complejos. Recomendamos que los análisis consideren ambos procedimientos; comparar los resultados obtenidos mediante ambos métodos podría ser particularmente informativo.

In their commentary, Sauer et al. (2004) acknowledge that our method of trend estimation (Bart et al. 2003) performed ‘‘in some cases much better than existing methods’’ but they are concerned that analysts using our approach might not adjust counts for observer differences, a step they view as essential for unbiased estimates. They also raise questions about the North American Breeding Bird Survey (BBS) and about design-based analytic methods in general. Our main responses, presented in detail below, are (1) we noted

Manuscript received 2 February 2004; accepted 4 February 2004.
[3] E-mail: jon_bart@usgs.gov

that analysts can make adjustments for observer effects, or other spurious influences, prior to employing our method; (2) such adjustments are sometimes useful, but they reduce precision and in some cases introduce, rather than remove, bias and should thus be undertaken cautiously rather than automatically; (3) we disagree with some of Sauer et al.'s (2004) comments about the BBS but note, more importantly, that our purpose was not to evaluate the BBS but rather to describe a general trend estimation method; and (4) their criticisms of design-based methods apply with equal force to the model-based approaches they favor.

Sauer et al. (2004) assert that analysts should always use methods that adjust counts for observer effects. We believe that these adjustments are valuable in some cases, and we noted (Bart et al. 2003) that they can be made prior to using the trend estimation method we described. For example, to adjust for observer effects one might carry out multiple regression, with observer covariates, and record the slope and mean from this regression rather than from the simple regression we used to illustrate our method. Other methods might be used to adjust the counts for weather influences, change in extraneous noise, or other spurious influences. Thus, our method can readily be combined with initial steps to adjust the counts in any way that the analyst deems suitable.

We do not agree that counts should always be adjusted, as they are in the estimating-equations approach (Link and Sauer 1994). The approach described by Link and Sauer (1994) includes the assumption that detection rates within observers show no long-term trends. But observers' abilities improve during the first several years they conduct surveys, and later in life the proportion of birds they detect declines as hearing and visual acuity decline. By calculating observer-specific trends, the Link and Sauer method confounds change in detection rate with change in population size. For example, if detection rates are increasing for a majority of the surveyors (due either to increasing skill or familiarity with the route), then the estimated trend will be positive even if population size is stable. If detection rates are decreasing, the trend will be negative even if the population is stable. These within-surveyor trends do not cause bias in the method we introduced. Our method, like any method based on an index, requires that there be no long-term trend in the ''index ratio,'' the ratio of the expected survey result to population size (Bart et al. 1997). That assumption, however, might readily be met. For example, even if most surveyors were early in their career, the proportion of observers with k years experience, $k = 1, 2, 3, \ldots,$ might be about the same each year due to the annual arrival of new surveyors and disappearance of previous participants. Thus, while long-term trend in average detection rates is a serious problem, the Link and Sauer (1994) method may be significantly biased even if no such trend occurs. In contrast, our method is essentially unbiased under that condition.

A separate reason for concern about the method Link and Sauer advocate is worth mentioning. In their method, data from long-term observers are weighted more heavily than data from short-term observers. Long-term surveyors, however, probably have declining detection rates as noted above. Estimates for regions with a small number of observers, and a few long-term observers, thus may have negative bias. In our method, results are not weighted by number of years the surveyor participates (or any other measure of within-route precision).

It is also worth noting, as Sauer et al. (2004) acknowledge, that including observer covariates reduces precision. In simulations Sauer et al. (2004) describe, variances were almost twice as large when adjustments for observer covariates were included. Thus, including such adjustments, if they are not necessary, is costly.

In summary, adjusting counts with observer or other covariates reduces precision, is not always necessary, may introduce bias rather than remove it, and may need to be carried out in different ways with different data sets (e.g., correcting for observer differences and for weather effects may require quite different approaches). That is why we separated these two efforts in our description. We did not intend, however, to imply that counts should never be adjusted to remove spurious influences.

We now briefly respond to the other comments made by Sauer et al. (2004). Headings refer to their paper. Space limits preclude our addressing all of their comments.

DESIGN-BASED ANALYSES

We noted (Bart et al. 2003:367) that design-based methods ''assume only that the sampling plan was followed and that the sample size is large enough to make inferences based on the central limit theorem.'' Sauer et al. (2004) say that the first assumption is ''crucial, and virtually never satisfied,'' but this is incorrect. For example, in the BBS, the sampled population is the roadsides (not the entire region) and a well-defined sampling plan is used to select locations from this population. Sauer et al. (2004) also claim that the second part of our sentence is incorrect, noting that ''no normality assumption is required for design-based analysis.'' We did not say that normality is required for design-based methods. Our point, in fact was just the opposite: under the central limit theorem the t-distribution may be used, if the sample size is large enough, regardless of the underlying distribution.

Finite population of sites. Sauer et al. (2004) argue that BBS ''routes cannot be viewed as selections from a sample frame'' because they ''do not survey habitats >0.4 km from secondary roads'' and they say that ''Bart et al. (2003) regard BBS routes as a simple random sample'', apparently because we treated a data set, derived from the BBS and used to evaluate our trend estimation method, as a simple random sample. As noted above, the BBS is a (stratified) random sample from a well-defined sampling frame (roadsides). More to the point, however, we were not attempting to evaluate the BBS or make claims about how BBS data should be analyzed, we simply used this data set to construct a hypothetical population. We agree that trend estimates based on BBS data, for specified areas (as opposed to our hypothetical population), should be analyzed with methods for stratified sampling. The extension from simple random sampling to stratified sampling is straightforward in general and for our trend estimation method: calculate point and interval estimates for each stratum and combine them using stratum sizes as weights. This feature is incorporated into a comprehensive trend analysis program we are preparing for general distribution.

Measurable quantities at sites. Sauer et al. object to our definition of Y_j as either (a) population size in year *j*, or (b) as the expected value of the survey result in year *j*, but we do not understand their concern. Our first case arises when, for example, subjects in a random sample of plots are enumerated so that density per plot yields an unbiased estimate of population density (and thus population size can be estimated). The second case arises with index methods in which, by definition, the detection rate is not estimated. We simply provide a definition and notation which accommodates either case.

Site-specific variation. Sauer et al. (2004) object to our focus on among-site variability as opposed to within-site variability, and they refer to the ''documented confounding of such (within-site) variability with population trend.'' Actually, such documentation is rare in the bird survey literature, and we believe they overlook the fact that within-site variability does not, by itself, cause any bias; only a long-term trend in the index ratio does. We agree, however, that such trends in the index ratio must be detected and removed or trend estimates will be biased. In some cases, Sauer et al.'s (2004) focus on observer detection rates and their method (Link and Sauer 1994) for removing trend due to this factor will be appropriate. In other cases, other factors or methods may be more appropriate. By separating adjustment of counts from estimation of the trend, we let analysts tailor this effort to the specific features of their data set.

Defining a parameter of interest. Sauer et al. (2004) question the accuracy of the approximation (R_{lin}) we used for our parameter of interest (R_{exp}), but they present an example in which the pattern of counts is distinctly *U*-shaped. Such data should not be analyzed using either our method or the estimating-equations approach, a point we emphasized in our report. We agree that deciding whether to use an exponential curve to describe trend is sometimes difficult and that better guidelines for making this decision might be useful. The lack of such guidelines, however, hinders decisions about whether to use the estimating-equations approach just as much it affects whether to use our method.

SITE-SPECIFIC COVARIATES

Sauer et al. (2004) carried out analyses using the estimating-equations approach with and without observer covariates and found consistent differences, and they argue that because we did not correct for observer effects, our approach is biased. But as noted above (a) analysts who want to include these adjustments can readily do so and then employ our method, and (b) it is not entirely clear which approach produces smaller bias because the ''with observer covariates'' approach confounds within-observer trends with trend in population size. We thus regard this analysis as unnecessary and inconclusive. We agree with their view that factors which might cause a long-term trend in the detection ratio should not be ignored.

SIMULATIONS

Sauer et al. (2004) acknowledge that our method performed well but they are concerned that our population did not correspond perfectly to a real population. Our data set, however, was collected on real BBS routes and is far more realistic than the data sets others (e.g., Link and Sauer 1994) have used to evaluate model performance. They also point out again that we excluded observer effects. As we have stressed above, we regard adjustment of counts as an important, but separate, issue. Sauer et al. (2004) assert that the real question is how well the method will work when sampling is not design-based and trends are not linear. We disagree with this view. We did not (and would not) recommend our method except when a random sampling plan has been followed and trends are approximately linear. Sauer et al. (2004) then state that the estimating-equations approach does not estimate the parameter (R_{exp}) that we defined but rather ''a precision- and abundance-weighted average of site-specific trends''. This definition, however, is too vague to be useful. Managers and researchers want to know how population size is changing. The parameter we estimate, R_{exp} (the annual rate of change of an exponential curve, fit to the population sizes), provides a clear description of how population size is changing. Sauer and colleagues need to explain more clearly how their parameter relates to change in population size.

DISCUSSION

Sauer et al. (2004) say again that we believe ''BBS data can be effectively treated as a random sample of population sizes.'' BBS locations are randomly selected from roadsides and thus can be treated as a random sample from this population. Questions do arise about spatial limits of the sampled population (i.e., how far from the road the surveyed area extends), and about the magnitude of selection bias (i.e., a difference between the roadside and regionwide trend), but these issues were not the focus of our report and are just as problematic if the estimating-equations approach is used. Furthermore, we do not view BBS counts as estimates of population sizes; they are an index. Sauer et al. (2004) also state that model-based adjustments are needed to compensate for uneven coverage through time, but we disagree. Such adjustments can also be applied using design-based methods, though this is usually not done due to concern that the adjustment may be correlated with the response variable. Sauer et al. (2004) note that the fraction of data missing in BBS data sets sometimes exceeds the fraction we used in the simulations. This is true, but they give no reason for believing results would have been different had a higher fraction been used. They note that addition of new routes can in some cases lead to bias. We agree, but this issue is distinct from estimating trend in the existing data set; both our method and the estimating-equations approach would need modification to remove bias due to this cause. Finally, they urge that reliable methods be used and note that many other trend-estimation methods have been developed. We recommend that analysts consider using these methods. All of them, however, are complex, make assumptions that are difficult to evaluate, and are difficult for most users to implement, whereas our method is simple to understand and use, makes only limited assumptions, and can easily be used in combination with adjustments for observer differences or other influences. In addition, power and sample-size analyses are much easier using our method than using the other methods.

In conclusion, we reiterate that models which adjust counts for observer differences or other spurious influences are useful tools, especially for exploring complex interactions between observer effects, annual effects, and effects of environmental variables. These approaches, however, inevitably entail more assumptions than our approach, and when these assumptions are incorrect may result in larger bias than analysis of the uncorrected counts. We prefer an analytic strategy in which several approaches, making different assumptions, are available so that the most plausible assumptions for the particular data set may be identified and employed in the analysis. Our method facilitates this approach and, in addition, is simpler and outperforms the estimating-equations approach when counts fluctuate widely, as was true in the shorebird data set we investigated.

LITERATURE CITED

BART, J., B. COLLINS, AND R. I. G. MORRISON. 2003. Estimating population trends with a linear model. Condor 105:367–372.

BART, J., M. J. FLIGNER, AND W. NOTZ. 1998. Sampling and statistical methods for behavioral-ecologists. Cambridge University Press, Cambridge, UK.

LINK, W. A., AND J. R. SAUER. 1994. Estimating equations estimates of trends. Bird Populations 2:23–32.

SAUER, J. R., W. A. LINK, AND J. A. ROYLE. 2004. Estimating population trends with a linear model: technical comments. Condor 106:435–440.