

Comments

Ecology, 81(11), 2000, pp. 3256–3259
© 2000 by the Ecological Society of America

DEMOGRAPHY OF THE YELLOWSTONE GRIZZLY BEARS: COMMENT

L. Lee Eberhardt¹ and Steve Cherry²

This analysis of demography of the Yellowstone grizzlies by Pease and Mattson (1999) (hereafter referred to as “P&M”) evidently resulted from a dedicated and extensive effort by the authors. An interesting model was developed that may provide the basis for useful further analyses. The present results should not, however, be taken as a definitive accounting of the status of the Yellowstone grizzly population for the reasons given in the following.

Lack of information about statistical methods

It is repeatedly stated (e.g., P&M 1999: 961) that the maximum-likelihood method was used for parameter estimation. However the likelihood equations described at the bottom of page 973 are not mathematically valid likelihoods as written. Likelihoods are probability density functions viewed as functions of the parameters with the data held fixed. Viewed as functions of the data (time) the likelihood equations in P&M are not valid densities (i.e., they do not integrate to 1). Further, it is not obvious how these equations were used to build the joint likelihood that would have served as the basis for parameter estimation. Applying the theory of maximum-likelihood estimation to functions that are not mathematically valid likelihoods is not guaranteed to produce valid results. This is particularly true with respect to computation of standard errors and, in this case, Akaike Information Criterion (AIC) values. The AIC (Burnham and Anderson 1996, 1998) depends on the likelihood model. It is thus not possible to determine whether it is appropriate for the approach to estimation used by P&M until the actual likelihood statements are available.

An explicit representation of the likelihood is also

Manuscript received 14 May 1999; revised 20 March 2000; accepted 21 March 2000; final version received 10 April 2000.

¹ 2528 West Klamath Avenue, Kennewick, Washington 99336 USA.

² Corresponding author. Department of Mathematical Sciences, Montana State University, Bozeman, Montana 59717 USA. E-mail: cherry@math.montana.edu

needed to shed light on independence assumptions. Individual grizzly bear histories are complex. For instance, it is fairly common to have an interrupted sequence of data on a given bear because a collar was dropped and the bear was later recaptured and recollared. P&M recognized this prospect (pp. 959, 972), but it is not clear that it is taken into account in their model (Tables 1 and 2) and the statement (p. 958) that “A monitoring sequence ended either when a bear died . . . or when a live bear dropped its radio collar”). Possibly repeated sequences of observations on the same bear were used independently, but this further complicates any interpretation of the model as a stochastic process.

There are no clear descriptions of how AIC was applied to generate the so-called “best” model of Table 3. There is thus little or no prospect of any verification or repetition of the process by interested scientists. In most recent applications of AIC some sort of diagrammatic representation of the trial models is exhibited, along with resulting AIC values (Gaillard et al. 1993, Jorgenson et al. 1997, Toigo et al. 1997). P&M do provide useful information. What is lacking is a way to compare competing models that differ with respect to more than one explanatory variable, and a way to compare much simpler models with their “best” but quite complex one.

Model fitting and estimation of variances

The method used to estimate standard errors is based on large-sample theory. Although a sample size of 202 would appear to be large, P&M are attempting to estimate a large number of parameters. Their final model has 43 parameters. Thus, the sample sizes available to estimate some of the relevant parameters are too small. Even if the likelihood-based approach taken by P&M is valid their estimated standard errors, calculated according to large-sample theory, are invalid. A logical approach would be to use resampling methods (Efron and Tibshirani 1993, Manly 1997) to compute standard errors.

The small sample size (relative to the number of parameters being fit) also complicates the model-building process. The potential for overfitting and spurious results is high in such cases. A small-sample version of AIC is available (AICc) as described in Burnham et al. (1995) and Burnham and Anderson (1998). P&M could have used AICc. It is quite possible that a different “best” model would have resulted.

Leslie matrix calculations

A “generalized Leslie matrix” (P&M: p. 962 and Table 4) is used to estimate survival and reproductive

rates in years in which whitebark pine mast is or is not believed to be available to bears, and leads to the statement that "The population declines 5% in whitebark pine nonmast years . . . and increases 7% in mast years" (p. 964). Evidently this statement results from calculations of eigenvalues for the mast and nonmast matrices. This procedure is not legitimate inasmuch as the mast and nonmast years (listed below Table 2) are intermingled in time. Presumably a statement of this kind might be made conditional on a long series of mast or nonmast years, but with the presumed favorable and unfavorable conditions intermingled in time, the statement is not appropriate. It appears that estimates of λ for mast and nonmast years were combined by simply computing a geometric mean to obtain an overall estimate (p. 964). This incorrect estimate apparently leads to the main conclusion of the paper—that the point estimate of λ (1.01) is lower than previously published values. It is important to note that there were just five deaths of adult females and three of subadult females in the data used by Eberhardt (1995) to calculate λ for the 1983–1995 period. There is thus little prospect of reliably estimating separate survival rates for mast and nonmast years.

"Management" trapping

P&M acknowledged (p. 967) that their definition of "management trapping" differs from that commonly used. Their subpopulation of "management-trapped bears" included not only bears captured as a result of management activities (removal of a nuisance bear) but also bears caught for research purposes in 1 of 17 areas listed on page 960. The difference makes comparison of their results with other research findings problematic.

Bears caught in the 17 areas listed on page 960 are assumed by P&M to be unwary and thus prone to higher mortality. This notion is carried to what we believe is the unrealistic extreme of being defined as a heritable trait (p. 960). We do not dispute the prospect that cubs who are with their mothers when they are trapped due to some management action may be more likely to get into trouble themselves later. Meagher and Fowler (1989) provided empirical evidence of this. But using that evidence to extend this assumption to cubs whose mothers are caught in certain areas in traps used for research purposes is speculative. So is extending that result to cubs born to mothers who were trapped in management actions as sub-adults and moved but who subsequently avoided trouble. P&M carried out an analysis assuming no heritability and reported an estimate of the growth rate of 1.04 (p. 964). It would have been interesting to see the results of the fit of the model from that analysis. Clearly additional research into this issue is needed.

It is later stated (p. 967) that "Our analysis predicted that 73% of the population is management-trapped." P&M acknowledged that this value is too high. P&M provided four reasons for why the value of 73% is too high but dismissed the first two (their assumptions that "management-trapped bears never become non-management-trapped" and that "management-trapped females produce only management-trapped cubs"). But their fourth reason ("overestimation of the rate at which bears move from non-management-trapped status to management-trapped") could arise in part from violations of the first two assumptions.

P&M believed that the research sample used by Eberhardt et al. (1994), Boyce (1995),⁴ and Eberhardt (1995) was biased because it did not include bears originally caught in management actions. But the sample of research bears used by both Eberhardt and Boyce does contain bears caught at the 17 sites listed on page 960. P&M claim that Eberhardt et al.'s (1994) and Eberhardt's (1995) analyses did not allow for movement from the "non-management-trapped" to "management-trapped" subpopulations but then acknowledge that some bears (actually 23%) in the research sample were later captured in management situations. Thus the research sample also contains bears that later got into trouble as a result of contacts with humans. Assuming that the research sample is representative of the population, then it is not at all clear why the results are claimed to be biased. Adding bears initially caught as a result of management actions to the research sample would tend to bias the resulting mortality estimates high. Even if the research sample is biased it is not clear that adding what amounts to a self-selected sample of bears originally caught as a result of management activities to the sample of research bears corrects the bias, at least in a statistically acceptable manner.

The probability of a population decline

P&M estimate the finite rate of population growth to be $\hat{\lambda} = 1.01$ with a standard error of 0.04. On page 964 they wrote: "The standard error of λ [sic] . . . quantifies uncertainty arising from sampling error. Because a Gaussian distribution with mean 1.01 and standard deviation of 0.04 has ~40% of its probability mass below 1.00, there is an approximate 40% probability that the number of Yellowstone grizzlies has declined from 1975 to 1995, and, conversely a 60% chance that the population increased over this time." This is not correct. Standard errors are properties of random samples and estimators, not populations and parameters. In

⁴ M. Boyce, 1995, *unpublished report* to the Interagency Grizzly Bear Committee. Available in somewhat updated version as Boyce, M. S., B. Blanchard, R. R. Knight, and C. Servheen. 2000. Population viability for grizzly bears: a critical review. *Ursus monograph* 5, *in press*.

the frequentist statistical approach used by P&M, there is no variability associated with the parameter and P&M cannot compute the probability that λ is less than 1.00. It either is or is not, although they do not know which of these is true.

Comparing their estimate of λ with others

On occasion P&M appear to ignore the uncertainty in their estimation procedure. On page 966 they note that other authors (Eberhardt et al. 1994, Boyce 1995,⁴ and Eberhardt 1995) have provided point estimates of λ equal to 1.05. They state, "The differences between their results and ours are real, even though their estimate falls within our 95% CI (0.93, 1.09). There is no issue of statistical significance." The authors are not correct. Assessing the statistical distance between their estimate of 1.01 and the disputed value of 1.05 is a reasonable thing to do and their confidence interval is adequate for that task. As P&M note, the disputed value of 1.05 falls within their confidence interval of (0.93, 1.09). Given the uncertainty in their estimator they cannot statistically rule out 1.05 as a plausible value for the growth rate even if they feel that previous authors used inappropriate methods to compute that value.

Of course, the point estimate of 1.05 is itself uncertain. In fact confidence intervals from all the studies would include both the point estimates of 1.01 and 1.05. P&M are justified in comparing their analysis to the analyses conducted by other researchers, but not justified in basing the comparison on the point estimates while ignoring the uncertainties associated with the point estimates. The "real differences" postulated by P&M are not nearly as profound as claimed once the statistical uncertainties are acknowledged. For instance, starting with an assumed population of 300 bears in 1975 they estimate that there were only 370 bears in the population in 1995 (p. 966). They argue that this value is more reasonable than estimates obtained using results from Boyce (1995),⁴ Eberhardt et al. (1994), and Eberhardt (1995). But the argument loses much of its force when the corresponding 95% confidence interval of 65 to 1833 bears is computed.

The large standard error indicates that P&M's estimate of λ has low precision, i.e., the data were not adequate to do a good job of estimating the finite rate of population change. P&M were not even able to provide convincing evidence for a direction of change. Their results are simply inconclusive, being consistent both with the views of those who believe the population has done well (at least over the short term) and those who think that the population has done poorly.

Further issues

There are various other questionable features of the paper.

- P&M invent a "generalized jackknife" procedure (p. 962) without citing a reference or providing more than a heuristic justification. The underlying idea is just that of the various "information" criteria (Burnham and Anderson 1998), and the term has been previously used to describe Tukey's (1958) original jackknife (Gray and Schucany 1972).

- Senescence is assumed to occur at age 13 (p. 959). Field data suggest that bears do not become senescent until much later in life (Eberhardt et al. 1994, Eberhardt 1995). Assuming early senescence will, of course, result in lowering the estimate of λ .

- A possibly less important point is that nothing is said about the importance of age structure in calculating λ from the Leslie matrices. References on demography (Keyfitz 1968, Caswell 1989) point out the need to take into account the effect of age structure in estimating λ .

Although their analyses are questionable in various respects, we do agree with the authors final doubts (p. 969) about an "optimistic long-term prognosis for the Yellowstone grizzly bear population". Maintaining the population will require continual vigilance and improvements in data collection. About 70% of the variance in Eberhardt's (1995) estimate of λ is associated with the data on subadult female survival. It may thus be quite feasible to narrow the confidence interval on λ by radiocollaring more subadult females, but rapid improvements do not seem to be in reach, and a reduction in research efforts can readily result in failing to recognize a downward trend when it occurs. From what we know at present, however, there is room for short-term optimism.

Literature Cited

- Burnham, K. P., and D. R. Anderson. 1996. Data-based selection of an appropriate biological model: the key to modern data analysis. Pages 16–30 in *Wildlife 2001*. D. R. McCullough and R. H. Barrett, editors. Elsevier Applied Science, New York, New York, USA.
- Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.
- Burnham, K. P., G. C. White, and D. R. Anderson. 1995. Model selection strategy in the analysis of capture-recapture data. *Biometrics* 51:888–898.
- Caswell, H. 1989. Matrix population models: construction, analysis, and interpretation. Sinauer Associates, Sunderland, Massachusetts, USA.
- Eberhardt, L. L. 1995. Population trend estimates from reproductive and survival data. Pages 13–19 in R. R. Knight and B. M. Blanchard, editors. *Yellowstone grizzly bear investigations; annual report of the Interagency Study Team 1994*. National Biological Service, Bozeman, Montana, USA.
- Eberhardt, L. L., B. M. Blanchard, and R. R. Knight. 1994. Population trend of the Yellowstone grizzly bear as estimated from reproductive and survival rates. *Canadian Journal of Zoology* 72:360–363.

- Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, New York, New York, USA.
- Gaillard, J.-M., D. Delorme, J.-M. Boutin, G. Van Laere, B. Boisaubert, and R. Pradel. 1993. Roe deer survival patterns: a comparative analysis of contrasting populations. *Journal of Animal Ecology* **62**:778–791.
- Gray, H. L., and W. R. Schucany. 1972. The generalized jackknife statistic. Marcel Dekker, New York, New York, USA.
- Jorgenson, J. T., M. Festa-Bianchet, J.-M., Gaillard, and W. D. Wishart. 1997. Effects of age, sex, disease, and density on survival of bighorn sheep. *Ecology* **78**:1019–1032.
- Keyfitz, N. 1968. Introduction to the mathematics of population. Addison-Wesley, Reading, Massachusetts, USA.
- Manly, B. F. J. 1997. Randomization, bootstrap, and monte carlo methods in biology. Second edition. Chapman & Hall, London, UK.
- Meagher, M., and S. Fowler. 1989. The consequences of protecting problem grizzly bears. Pages 141–144 in *Bear-people conflicts: proceedings of a symposium on management strategies*. Northwest Territories Department of Renewable Resources, Yellowknife, Northwest Territories, Canada.
- Pease, C. M., and D. J. Mattson. 1999. Demography of the Yellowstone grizzly bears. *Ecology* **80**:957–975.
- Toigo, C., J.-M. Gaillard, and J. Michallet. 1997. Adult survival of the sexually dimorphic Alpine ibex (*Capra ibex ibex*). *Canadian Journal of Zoology* **75**:75–79.
- Tukey, J. W. 1958. Bias and confidence in not quite large samples. *Annals of Mathematical Statistics* **29**:614.

ERRATA

In the article by Ragan Callaway, Thomas DeLuca, and Wendy Belliveau (1999) entitled “Biological-control herbivores may increase competitive ability of the noxious weed *Centaurea maculosa*,” published in *Ecology* **80**(4):1196–1201, there is an error in the *Methods* section (p. 1197) for the date that *Agapeta zoegana* were applied in the common garden experiment. In the sixth sentence of the first paragraph in the *Methods*, the words “on 12 August” should be substituted for “in June.” After correction the sentence should read “Mesh cages were placed over all 20 plots, and four *Agapeta zoegana* moths were introduced into each of a randomly chosen subset of 10 plots on 12 August 1994 (see Müller-Schärer 1991).” Consequently, *Agapeta zoegana* were on target *Centaurea* plants during parts of two growing seasons rather than for “two growing seasons” as stated later in the *Methods*. This does not affect the fundamental conclusion of the paper, “that some biocontrols may have indirect negative effects on native species that are not currently recognized.”

An error has been discovered in the electronic supplement to the article by Valerio DePatta Pillar (1999), entitled “How sharp are classifications?” [*Ecology* **80**(8):2508–2516]. The program originally published in *Ecological Archives* as E080-014 contained a bug that may cause wrong results when the user selects minimum variance as the clustering method. The error did not affect the results published in the paper. A corrected version of the program has now been posted as *Ecological Archives* E081-028.