On bias, precision and accuracy in wildlife aerial surveys

Article *in* Wildlife Research · January 2008

DOI: 10.1071/WR07144

Jim Hone
University of Canberra

CSIRO PUBLISHING

www.publish.csiro.au/journals/wr *Wildlife Research*, 2008, **35**, 253–257

On bias, precision and accuracy in wildlife aerial surveys

Jim Hone

Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601, Australia.
Email: jim.hone@canberra.edu.au

Abstract. Bias, precision and accuracy have been studied extensively in wildlife population estimation including aerial surveys. A review of the literature shows that the concepts of bias and precision are used broadly consistently. Aerial survey data from known populations of feral pig carcases and white-tailed deer show that few density estimates are unbiased and precise. Research is needed, however, to clarify how much bias and how much precision are enough for the various types of wildlife management activities. Accuracy is used in two closely related but different ways. One set of definitions of accuracy relates to deviations from the true value (bias) and the second set relates to squared deviations from the true value (bias and precision). The implications are that authors are encouraged to clearly state which definition of accuracy they use, or focus solely on bias and precision.

Introduction

The topics of bias, precision and accuracy have been examined in aerial surveys over many years. The use of aerial surveys to estimate wildlife abundance or density has progressed from early efforts of unknown bias, precision and accuracy to intensive efforts to obtain estimates closer to true density (unbiased) and with small variances and standard errors (high precision). The reviews of Caughley (1974) and Pollock and Kendall (1987) demonstrated the need for closer study of bias as previous studies showed that many animals were missed even by good observers during aerial surveys. Such studies examined the topic of how biased estimates of density could be (that is, how close to the true value). Other studies in population estimation generally have debated the topic of how unbiased estimates should be. For example, one study considered that if estimates could get within 10% of true density then many biologists might feel this was adequate (Anderson and Southwell 1995). Another study considered that if an estimate was underestimating by 10% then it was not a good estimate (Lancia *et al*. 2005). Other studies of sampling, for example Caughley (1977a) and Sinclair *et al*. (2006), described sampling procedures to achieve high precision, such as by sampling without replacement and using a high sampling intensity.

This note reviews the topics of bias, precision and accuracy in wildlife aerial survey and their application. The implications of the results are then discussed.

Definitions of bias, precision and accuracy

A review of the relevant scientific literature shows that bias is defined as the difference between a sample mean (m) and the true population value (μ) (Cochran 1977, p.13). Other wildlife literature, for example, Williams *et al*. (2002, p. 43), Bart *et al*. (2004, p. 1244), Lancia *et al*. (2005, p. 108) and Garton *et al*. (2005, p. 51) use a broadly similar definition. Precision is defined as the size of deviations from the mean obtained by repeated application of the sampling procedure (Cochran 1977, p. 16). Other statistical and ecological literature, for example Sokal and Rohlf (1981, p. 13), Caughley and Sinclair (1994, p. 192), Krebs (1999, p. 7), Ford (2000, p. 134), Williams *et al*. (2002, p. 44), Lancia *et al*. (2005, p. 108), Garton *et al*. (2005, p. 51), Sinclair *et al*. (2006, p. 221), and Greenwood and Robinson (2006, p. 14) are broadly consistent with this definition. The fine details of wording may differ between the many publications but the essence of the definitions are consistent for bias and for precision.

In aerial surveys data are often used in line-transect or mark–recapture analyses to obtain a density estimate and its associated variance. Bias is then the difference between the estimate and the true density. Precision is related to the size of the variance (high variance = not precise, low variance = precise).

The topics of bias and precision are combined in the principle of parsimony that describes the conceptual trade-off between squared bias and variance versus the number of estimable parameters in a model (estimator) used to estimate wildlife density (Burnham and Anderson 2001, 2002). As the number of parameters used to estimate density is increased then bias squared decreases but the variance of the estimate increases so precision decreases. Analyses such as Akaike's information criterion involve a trade-off between bias and precision. Hence the use of such analyses is likely to produce some bias in density estimates. Statisticians often caution against 'overfitting', which corresponds to obtaining density estimates of low precision (high variance). As a result, and since bias is often unknown, Cochran (1977, p. 16) suggested that the focus is often on precision, not accuracy.

Cochran (1977, pp. 12–16) examined an aspect of combining bias and precision. He showed that the effect of bias on the accuracy of an estimate is negligible if the bias is less than one-tenth of the standard error of the estimate. He stated that if the ratio were less than 0.2 the effect was modest. Fewster and Buckland (2004) described the result as a useful rule of thumb.

The definition of accuracy differs between publications, and the difference may not be fully acknowledged and appreciated. The first set of definitions (Definition 1) is illustrated by the use

 10.1071/WR07144 1035-3712/08/040253

in an ecological textbook: accuracy is the closeness of a measured value to its true value (Krebs 1999, p. 7). Similar definitions are given, in general discussions of estimation, not specifically wildlife aerial survey, by Sokal and Rohlf (1981, p. 13), Bibby *et al*. (1992, p. 34), Caughley and Sinclair (1994, p. 192), Underwood (1997, p. 39), Zar (1999, p. 5), Ford (2000, pp. 133–134), Engeman (2005, p. 204), Lancia *et al*. (2005, p. 108), Sinclair *et al*. (2006, p. 221) and Greenwood and Robinson (2006, p. 14). The reviews of bias in aerial surveys by Caughley (1974) and Pollock and Kendall (1987) implicitly use this definition of accuracy. This definition focuses on bias, specifically deviations from the true value.

Another set of definitions of accuracy (Definition 2) uses different wording, and they represent something very similar but distinctly different. Cochran (1963, p. 15) and Cochran (1977, p. 15) discuss the use of the mean square error (MSE) as a criterion of accuracy of an estimator. MSE is the sum of the variance and the squared bias (Eqn 1) and is derived from the expected value of squared deviations from the true value (Cochran 1977, p. 15). Cochran (1953) did not discuss the topic. In drawing an analogy between population estimation and rifle bullets hitting a target, Overton and Davis (1969, p. 406–407) state 'The accuracy of the estimator is analogous to the spread of the group about the bull's-eye. MSE is used to measure this property, and the greater the value of MSE, the less the accuracy of the estimator'. Note that the literature appears to use mean square error and mean squared error as synonymous. Note also that mean square error, as discussed above, is not to be confused with error mean square, which is the denominator of the *F* ratio in a fixed-factor analysis of variance. The definition of accuracy using the MSE is used elsewhere, for example, Ratti and Garton (1994, p. 9), Levy and Lemeshow (1999, p. 38), Williams *et al*. (2002, p. 45), Bart *et al*. (2004, p. 1242), Lancia *et al*. (2005, p. 108) and Garton *et al*. (2005, p. 51). The latter publication states that 'if an estimate is both unbiased and precise, we say it is accurate (defined as an estimator with small mean-squared error, Cochran 1963)'. Similarly, McCallum (2000, p. 23) considered that lack of bias, together with precision, is often termed accuracy, and Ramsey *et al*. (2005, pp. 232, 235) viewed accuracy as bias and precision. Hence, Definition 2 of accuracy combines the concepts of bias and precision. In contrast, Definition 1 of accuracy relates only to bias. The differences follow from focusing on the deviations from the true value (first set) or squared deviations from the true value (second set). A related, but different, perspective is given by Seber (1982, p. 5), who describes the coefficient of variation (standard deviation divided by the mean) as a 'useful measure of the 'accuracy' of' an estimate.

$$\text{MSE} = \text{variance} + \text{bias}^2 \quad (1)$$

The analogy of the rifle shot hits of the target is used by several authors, with agreement on many aspects but one important difference that highlights the conceptual difference between definitions. White *et al*. (1982, fig. 2.4, p. 22) shows four targets and classifies the shot patterns as (a) unbiased and precise, (b) unbiased but not precise, (c) biased but precise, and (d) biased and not precise. Ratti and Garton (1994, fig. 2, p. 9) and Garton *et al*. (2005, fig. 3, p. 51) have the same diagram and use the same wording, except they each add an extra component to the wording with the figures. The figure named 'unbiased but not precise' (Part b of the figures in all three publications) has the extra words 'not accurate' included. White *et al*. (1982) did not use those extra words within their fig. 2.4. In contrast, Bibby *et al*. (1992, p. 28) in box 2.5 used the number line to graphically illustrate the same four combinations of bias and precision. The 'unbiased but not precise' combination of White *et al*. (1982) was labelled 'imprecise and accurate' by Bibby *et al*. (1992). Hence Bibby *et al*. (1992) equate unbiased with accurate, and separate precision and accuracy.

The concepts of bias and precision can also be illustrated in a graph, as shown in Fig. 1. The solid diagonal line in Fig. 1 illustrates points with equal numerical value of the MSE, associated with different combinations of bias squared and variance. Definition 2 of accuracy describes all such points of equal MSE as being of equal accuracy.

A survey of a wildlife population that provides a total count has no bias and a variance of 0 (exceptionally high precision from sampling without replacement of all possible samples), so corresponds to the origin in Fig. 1. This is unlikely to occur in aerial surveys of wildlife because of cost and practical issues such as the need for very accurate navigation and animals moving in response to aircraft, so the MSE is expected to be greater than zero. A survey that is biased and not precise (high variance, top right of Fig. 1) may be useless for wildlife management. High variance (low precision = not precise) could occur from sampling with replacement and using a low sampling intensity.

A density index is a relative value of true density. A precise index may have substantial bias and low variance (biased but precise, bottom right of Fig. 1) and an index with low bias and a higher variance occurs in the top left of Fig. 1. In the literature on wildlife population estimation the use of an index is described (Caughley 1977b, pp. 12–25; Seber 1982) and discouraged (Anderson 2001). A whole issue of *Wildlife Research* (2005, volume 32, issue 3) was devoted to density indices and the issues were examined and debated therein. So where do

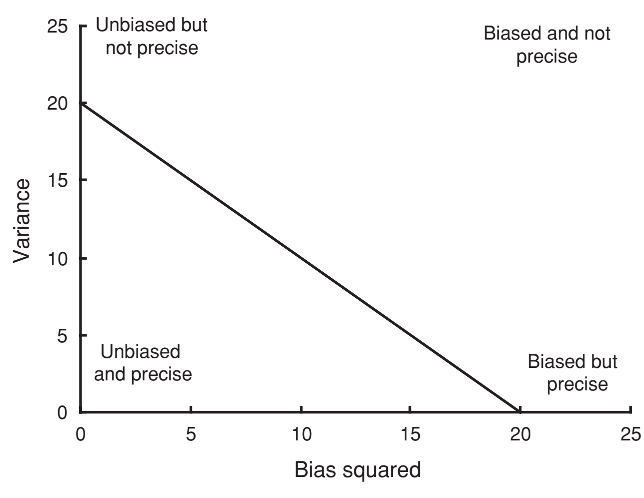

Fig. 1. A classification of wildlife density estimates based on the variance of a density estimate (as a measure of precision) and the bias squared. The solid diagonal line represents estimates of equal numerical value of the mean square error. The numerical values on the axes are hypothetical.

density estimates obtained by aerial survey occur in the parameter space of Fig. 1?

Methods

Two studies are examined in which aerial survey was used to estimate density and true density was also known. Hence bias and precision can be assessed. The bias assessment uses a simple version of that described by Cox (2006) based on knowledge of true density. The first study was of carcases of feral pigs (*Sus scrofa*) (Hone 1988) and the second of white-tailed deer (*Odocoileus virginianus*) (Potvin and Breton 2005). The pig carcase study used eight line-transect estimators (Cox's method using two different strip widths, Fourier series, exponential power series, half-normal, negative exponential, exponential polynomial, hermite polynomial and hazard rate), and four strip-transect estimates during one aerial survey. The latter used strip widths of 0–25 m, 0–50 m, 0–75 m and 0–100 m (Hone 1988) and did not attempt to correct for visibility bias. The white-tailed deer study used two estimators, but for this note only one, the visual double-count (mark–recapture) estimator, was examined. Five surveys were conducted in August and one in January. True density was estimated from population reconstruction using harvest data and estimates of winter mortality. Ideally, true density should be estimated from complete removal counts or by placing known numbers of animals in enclosures and surveying them soon thereafter. Further details of surveys are described in the original publications. Related studies of mule deer (*Odocoileus hemionus*) (White *et al*. 1989) and black rhinoceros (*Diceros bicornis*) (Brockett 2002) had known densities but reported results as relative bias (%), not absolute bias, so the results were not used in the present analysis. The results are graphed using the same axes as shown in Fig. 1 and used to illustrate the concepts of bias, precision and accuracy.

Results

Estimates of carcases of feral pigs were obtained during one aerial survey. A plot of the variance of each estimate and bias squared shows a scatter of points with several estimates being unbiased (on, or very close to, the vertical axis in Fig. 2). The most precise estimate had high bias (Fig. 2). The population of white-tailed deer was surveyed six times and density estimates obtained from a visual double-count estimator. The graph of variances of estimates and bias squared shows a scatter of points with no estimate being unbiased and precise (Fig. 3).

The numerical scales on the x (bias squared) and y (variance) axes in Figs 2 and 3 show that estimates of MSE (Eqn 1) would be dominated by bias squared, as the numerical scale for bias squared shows a greater range than that for the variances. This is particularly so in Fig. 3 for deer.

Discussion

Aerial surveys for estimating wildlife density and abundance have undergone tremendous change over the last four decades. The early reviews and results of Caughley (1974, 1977*a*) and Pollock and Kendall (1987) were very useful. However, this note demonstrates that there is still scope for further research on the combined topics of bias, precision and accuracy. The empirical results from the two aerial survey studies show that few density estimates are unbiased and precise (Figs 2, 3), and suggest that the estimates of variances and biases have large sources of variation. Such variation could include observers misidentifying animals, double counting, errors in estimating distances that were used in analyses, use of single versus multiple observers, different estimators and differing numbers of parameters used in various estimators. Rigorous standardisation of procedures is encouraged to reduce such variation. Obviously, the results here come from only two studies and discussion of the generality of the results requires results from many more such studies.

The literature on bias and precision is broadly consistent with respect to definitions. Clearly, differences exist in how the concept of accuracy is defined and used. Authors are encouraged to clearly state their definition of accuracy when using the concept, in particular whether their use relates solely to bias, as described by Krebs (1999) and others, or to bias and variance combined, as in mean square error, as described by Cochran

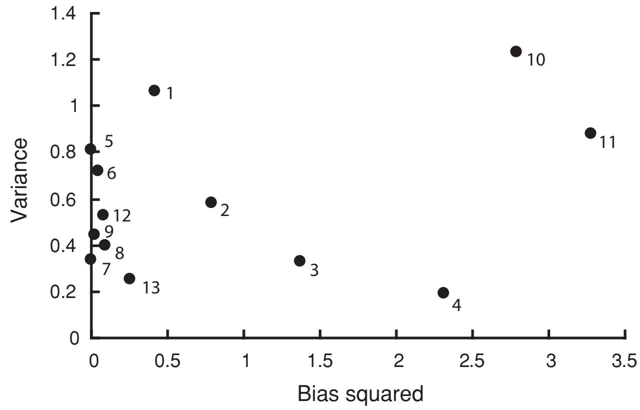

Fig. 2. The variances of density estimates and the bias squared of estimates of carcase density of feral pigs. The numbers in the figure correspond to the various density estimators: 1, ratio 0–25 m; 2, ratio 0–50 m; 3, ratio 0–75 m; 4, ratio 0–100 m; 5, Cox 25 m; 6, Cox 50 m; 7, Fourier series; 8, exponential power series; 9, half-normal; 10, negative exponential; 11, exponential polynomial; 12, hermite polynomial; and 13, hazard rate.

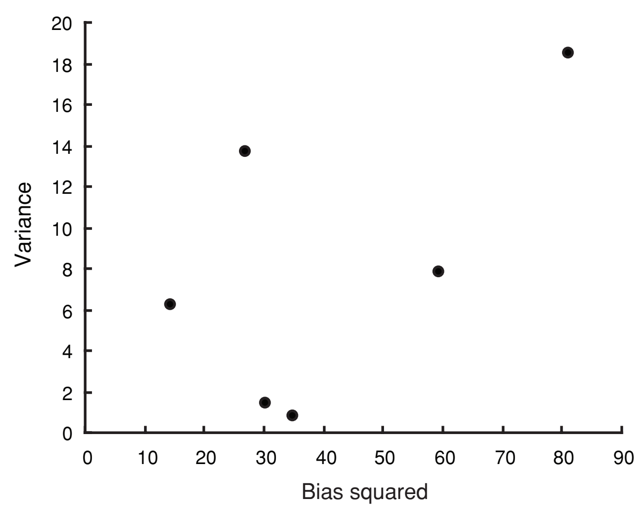

Fig. 3. The variances of density estimates and the bias squared of estimates of the density of white-tailed deer.

(1963) and others. Alternatively, authors should focus only on bias and precision, an approach consistent with the principle of parsimony (Burnham and Anderson 2002).

One aspect of bias not addressed here is availability bias. Some animals may be unobservable in an aerial survey as they are hidden, such as by vegetation or boulders, and no estimator can correct for their being missed by observers. Extra time to observe and extra observers will not decrease the bias. Hence the estimated density will be negatively biased (underestimation occurs).

The implications for aerial surveys are that there is a need to address the issue of how much bias is acceptable and how precise estimates need to be for a variety of wildlife management applications, namely conservation, harvest and pest control. In practice, this corresponds to asking: where in the parameter space of Fig. 1 should a density estimate occur to be useful for a specific wildlife management application? The literature discussing bias, for example Anderson and Southwell (1995) and Lancia *et al*. (2005), demonstrate differences of opinion. The required levels of relative accuracy, discussed implicitly as bias, can differ between preliminary studies, management uses and research studies (Seber 1982, p. 64; Krebs 1999, p. 29), with the lowest levels needed in preliminary studies and the highest levels in research studies. Such results were suggested for mark–recapture studies, so their relevance to wildlife aerial surveys needs further study. Ford (2000, p. 134) suggested that researchers must judge what level of accuracy is necessary to test a given postulate. The required level of precision is often investigated, to estimate sample size (Manly 1992) and using a power analysis (Stirrat *et al*. 2001) to estimate how many surveys are needed to be, say, 95% confident of detecting a 20% decrease in density. This does not provide guidance on the required levels of bias. Further research on the topics of bias, precision and accuracy is encouraged.

Acknowledgements

I thank Peter Fleming and John Tracey for the invitation to participate in the aerial survey symposium at the 2006 annual meeting of the Australasian Wildlife Management Society. I thank the University of Canberra for support, and D. Ramsey and D. Pederson for statistical advice. P. Caley, C. Krebs and D. Pederson provided useful comments on a draft manuscript.

References

Anderson, D. R. (2001). The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin* **29**, 1294–1297.

Anderson, D. R., and Southwell, C. (1995). Estimates of macropod density from line transect surveys relative to analyst expertise. *Journal of Wildlife Management* **59**, 852–857. doi:10.2307/3801966

Bart, J., Droege, S., Geissler, P., Peterjohn, B., and Ralph, C. J. (2004). Density estimation in wildlife surveys. *Wildlife Society Bulletin* **32**, 1242–1247. doi:10.2193/0091-7648(2004)032[1242:DEIWS]2.0.CO;2

Bibby, C. J., Burgess, N. D., and Hill, D. A. (1992). 'Bird Census Techniques.' (Academic Press: London.)

Brockett, B. H. (2002). Accuracy, bias and precision of helicopter-based counts of black rhinoceros in Pilanesberg National Park, South Africa. *South African Journal of Wildlife Research* **32**, 121–136.

Burnham, K. P., and Anderson, D. R. (2001). Kullback–Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* **28**, 111–119. doi:10.1071/WR99107

Burnham, K. P., and Anderson, D. R. (2002). 'Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach.' 2nd edn. (Springer: Berlin.)

Caughley, G. (1974). Bias in aerial survey. *Journal of Wildlife Management* **38**, 921–933. doi:10.2307/3800067

Caughley, G. (1977*a*). Sampling in aerial survey. *Journal of Wildlife Management* **41**, 605–615. doi:10.2307/3799980

Caughley, G. (1977*b*). 'Analysis of Vertebrate Populations.' (John Wiley & Sons: New York.)

Caughley, G., and Sinclair, A. R. E. (1994). 'Wildlife Ecology and Management.' (Blackwell: Oxford.)

Cochran, W. G. (1953). 'Sampling Techniques.' (John Wiley & Sons: New York.)

Cochran, W. G. (1963). 'Sampling Techniques.' 2nd edn. (John Wiley & Sons: New York.)

Cochran, W. G. (1977). 'Sampling Techniques.' 3rd edn. (John Wiley & Sons: New York.)

Cox, D. R. (2006). 'Principles of Statistical Inference.' (Cambridge University Press: Cambridge.)

Engeman, R. M. (2005). Indexing principles and a widely applicable paradigm for indexing animal populations. *Wildlife Research* **32**, 203–210. doi:10.1071/WR03120

Fewster, R. M., and Buckland, S. T. (2004). Assessment of distance sampling estimators. In 'Advanced Distance Sampling'. (Eds S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas.) pp. 281–306. (Oxford University Press: Oxford.)

Ford, E. D. (2000). 'Scientific Method for Ecological Research.' (Cambridge University Press: Cambridge.)

Garton, E. O., Ratti, J. T., and Giudice, J. H. (2005). Research and experimental design. In 'Techniques for Wildlife Investigations and Management'. 6th edn. (Ed. C. E. Braun.) pp. 43–71. (The Wildlife Society: Bethesda, MD.)

Greenwood, J. J. D., and Robinson, R. A. (2006). Principles of sampling. In 'Ecological Census Techniques: A Handbook'. 2nd edn. (Ed. W. J. Sutherland.) pp. 11–86. (Cambridge University Press: Cambridge.)

Hone, J. (1988). A test of the accuracy of line and strip transect estimators in aerial survey. *Australian Wildlife Research* **15**, 493–497. doi:10.1071/WR9880493

Krebs, C. J. (1999). 'Ecological Methodology.' 2nd edn. (Addison Wesley Longman: Menlo Park, CA.)

Lancia, R. A., Kendall, W. L., Pollock, K. H., and Nichols, J. D. (2005). Estimating the number of animals in wildlife populations. In 'Techniques for Wildlife Investigations and Management'. 6th edn. (Ed. C. E. Braun.) pp. 106–153. (The Wildlife Society: Bethesda, MD.)

Levy, P. S., and Lemeshow, S. (1999). 'Sampling of Populations. Methods and Applications.' 3rd edn. (John Wiley & Sons: New York.)

Manly, B. F. J. (1992). 'The Design and Analysis of Research Studies.' (Cambridge University Press: Cambridge.)

McCallum, H. (2000). 'Population Parameters. Estimation for Ecological Models.' (Blackwell Science: Oxford.)

Overton, W. S., and Davis, D. E. (1969). Estimating the numbers of animals in wildlife population. In 'Wildlife Management Techniques'. 3rd edn (rev.). (Ed. R. H. Giles.) pp. 403–455. (The Wildlife Society: Washington.)

Pollock, K. H., and Kendall, W. L. (1987). Visibility bias in aerial surveys: a review of estimation procedures. *Journal of Wildlife Management* **51**, 502–510. doi:10.2307/3801040

Potvin, F., and Breton, L. (2005). Testing 2 aerial survey techniques on deer in fenced enclosures – visual double-counts and thermal infrared sensing. *Wildlife Society Bulletin* **33**, 317–325. doi:10.2193/0091-7648(2005)33[317:FTFTAS]2.0.CO;2

Ramsey, D., Efford, M., Ball, S., and Nugent, G. (2005). The evaluation of indices of animal abundance using spatial simulation of animal trapping. *Wildlife Research* **32**, 229–237. doi:10.1071/WR03119

Ratti, J. T., and Garton, E. O. (1994). Research and experimental design. In 'Research and Management Techniques for Wildlife and Habitats'. 5th edn. (Ed. T. A. Bookhout.) pp. 1–23. (The Wildlife Society: Bethesda, MD.)

Seber, G. A. F. (1982). 'Estimation of Animal Abundance and Related Parameters.' 2nd edn. (Charles Griffin & Company: London.)

Sinclair, A. R. E., Fryxell, J. M., and Caughley, G. (2006). 'Wildlife Ecology, Conservation and Management.' 2nd edn. (Blackwell: Oxford.)

Sokal, R. R., and Rohlf, F. J. (1981). 'Biometry.' 2nd edn. (W.H. Freeman & Co.: New York.)

Stirrat, S. C., Lawson, D., Freeland, W. J., and Morton, R. (2001). Monitoring *Crocodylus porosus* populations in the Northern Territory of Australia: a retrospective power analysis. *Wildlife Research* **28**, 547–554. doi:10.1071/WR00079

Underwood, A. J. (1997). 'Experiments in Ecology. Their Logical Design and Interpretation using Analysis of Variance.' (Cambridge University Press: Cambridge.)

White, G. C., Anderson, D. R., Burnham, K. P., and Otis, D. L. (1982). 'Capture–Recapture and Removal Methods for Sampling Closed Populations.' (Los Alamos National Laboratory: Los Alamos, NM.)

White, G. C., Bartmann, R. M., Carpenter, L. H., and Garrott, R. A. (1989). Evaluation of aerial line transects for estimating mule deer densities. *Journal of Wildlife Management* **53**, 625–635. doi:10.2307/3809187

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). 'Analysis and Management of Animal Populations.' (Academic Press: New York.)

Zar, J. H. (1999). 'Biostatistical Analysis.' 4th edn. (Prentice Hall International Inc.: London.)

Manuscript received 14 September 2007, accepted 17 December 2007