Ardeola 57(Especial), 2010, 23-32

OBSERVER ERROR ASSOCIATED WITH BAND ALLOCATION IS NEGLIGIBLE IN LARGE SCALE BIRD MONITORING SCHEMES, BUT HOW PRECISE IS THE USE OF BANDS AT ALL?

EL ERROR ASOCIADO A LA ASIGNACIÓN DE CONTACTOS A BANDAS DEBIDO AL OBSERVADOR ES DESPRECIABLE EN LOS PROGRAMAS DE MONITOREO DE AVES A GRAN ESCALA, PERO ¿CUÁN PRECISO ES USAR BANDAS?

Javier QUESADA[1] *, Santi GUALLAR[1], Natàlia J. PÉREZ-RUIZ[1],
Joan ESTRADA[1] and Sergi HERRANDO[1]

SUMMARY.—*Observer error associated with band allocation is negligible in large scale bird monitoring schemes, but how precise is the use of bands at all?.*

This study assesses the observer error of the Catalan Common Bird Survey (SOCC), a bird monitoring programme based on line transects. Specifically, it studies the error associated with assigning every detected bird to the correct distance band, and its influence in density estimates. Thirty SOCC participants were asked to allocate 20 objects detected along their line transect to 3 distance bands, while an evaluator measured the exact distances to the objects with a laser range-finder. Density estimates were derived both from exact and distance band data, the latter corrected or not corrected by the assessed error. Observer accuracy was high and precision low, although these varied with band boundaries. Estimates based on exact distances were significantly smaller than those based on bands. Band correction did not affect density estimates significantly. We suggest guidelines for detecting and reducing bias in observer error and for obtaining more reliable density estimates.

Key words: Catalan Common Bird Survey, density estimates, detectability, Distance Sampling, line transects.

RESUMEN.—*El error asociado a la asignación de contactos a bandas debido al observador es despreciable en los programas de monitoreo de aves a gran escala, pero ¿cuán preciso es usar bandas?*

Este estudio evalúa el error del observador en el programa de monitoreo de aves SOCC basado en transectos lineales. En concreto, se estudia el error en la asignación de cada contacto en la banda correcta y su influencia en la estimación de la densidad. Treinta participantes asignaron 20 elementos contactados en las tres bandas de su transecto SOCC mientras que un evaluador medía las distancias

[1] Institut Català d'Ornitologia, Museu de Ciències Naturals. Passeig Picasso s/n, E-08003 Barcelona, Spain.

* Corresponding author: analisi@ornitologia.org

exactas a ellos con un medidor láser. Se calcularon estimaciones de densidad usando las distancias exactas y las de las bandas, esto último corrigiendo o no el error realizado. La exactitud de los observadores fue alta y la precisión fue baja, aunque esto varió con los límites de la banda considerados. Las densidades calculadas con las distancias exactas fueron significativamente menores que las basadas en bandas. La corrección de los errores realizados en las bandas no afectó las estimaciones de densidad. Se sugieren pautas para detectar y reducir el sesgo en el error del observador y para obtener estimaciones de densidad más fiables.

Palabras clave: detectabilidad, estimas de densidad, muestreo de distancias 'Distance Sampling', Seguimiento de Aves Comunes Catalanas (SOCC), transectos lineales.

INTRODUCTION

Monitoring schemes are used for determining population trends (Gregory *et al*., 2005), establishing conservation priorities (Herrando *et al*., 2010a), estimating population sizes (Robertson *et al*., 1995) and defining habitat requirements (Sutherland, 1998). Unfortunately, the results of such investigations can be compromised if the sampling design is not representative of the area of interest or the field methods are inadequate. Even if the design and methods of a study are adequate, measurement errors can compromise their results. A prominent source of measurement error is observer bias, caused by incorrect identification, poor hearing, etc. (Bibby *et al*., 2000).

At present, the most widely used analytical technique for deriving species densities, and from them population size estimates, is Distance Sampling. Distance Sampling consists of a set of methods in which distances from a line or point to detections are recorded, from which the density of objects is estimated and detectability is calculated (Thomas *et al*., 2010). Among the several count methods to which Distance Sampling can be applied, the line transect with parallel bands is the most frequently implemented in monitoring schemes (e.g. Newson *et al*., 2008). Here, allocating an observation to the correct band is fundamental to obtaining accurate estimates (Bibby *et al*., 2000; Gregory *et al*., 2004); however, this depends on the ability of the observer to carry out this work in field conditions, and it is likely that this will vary greatly among observers.

Despite being crucial for estimating population densities, observer bias in relation to distance assessment has been rarely studied in monitoring schemes (e.g. Frederick *et al*., 2003). In this study, we present a field test from which we evaluated the observer error in assigning birds to the correct distance band and its influence on resulting density estimates. We also present guidelines for detecting and reducing this error and for improving density estimates for monitoring schemes that use transects with parallel bands.

MATERIALS AND METHODS

Study design and field methods

The experiment was performed by Catalan Common Bird Survey (SOCC) volunteers, and an examiner who recorded the data (N. Pérez-Ruiz). The field work was conducted during 30 mornings between April and December 2008. The SOCC is a monitoring scheme based on a stratified network of 3-km line transects where the observer records the individuals of all the species he or she detects, and allocates them into one of three bands (0-25 m, 25-100 m, 100-1,000 m; see more details in Herrando *et al*., 2008).

We took a random sample of 30 of the 122 active SOCC participants in 2008. To obtain a representative sample of the landscapes found in Catalonia, we stratified the observers across the five main bioregions defined in Herrando *et al*., (2010b), and selected a number of SOCC proportional to the area occupied by each bioregion.

Observers were asked to walk their transects and allocate to a band every object the examiner pointed out to him or her. The examiner chose 20 objects of three types: five inanimate objects and 15 real birds of which at least three were aural contacts whose location could be established with certainty. We included inanimate objects for increasing our sample size (birds fly away very fast, and frequently the observer can not see them before they are far away or simply gone). We are aware of the large bias towards seen birds, but replicating the proportion of aural contacts in the SOCC would have demanded considerable effort by the observer.

For every object, the examiner took the perpendicular distance (PD hereafter) to the transect axis with a calibrated laser rangefinder (Leica rangemaster 900 CRF ™). The examiner also annotated the allocated band (observer's band) and the real band. We considered an observer to have made an error if they allocated the object to the wrong band. When an observer allocated the object to a band closer to the observer than the correct band, the value of the error was negative. Otherwise, the error was positive. Regardless of the error sign, the error value was the difference between the exact distance of the object (measured by the range-finder) and the nearest point of the band-boundary to which the observer allocated it. For instance, if a bird was at 115 m (100-1,000 m band) and the observer allocated the bird to the second band (25-100 m) the observer error was –15 m. When the observer allocated the bird correctly, the error was zero.

Statistical procedures

Evaluation of accuracy and precision in band allocation. We evaluated the observer error by describing precision and accuracy in band allocation and determined which elements drive them. To do this, we considered the 'absolute error' as an estimate of the precision. Absolute error is the absolute value of the error (regardless of the error sign). Balanced error is the observer error considering the sign of the error (positive or negative) and we considered it as an estimate of accuracy.

To study precision, we considered the absolute observer error (mean ± SE; range) and determined how it varied for each band-boundary (25 m, 100 m). We conducted a Generalized linear Model (GLZ) using the absolute error as a dependent variable (setting Poisson distribution and logarithmic link function), and band boundary, object (inanimate object, seen bird or heard bird), observer and the interaction band-boundary*object as predictors. Our data sample (inanimate objects, seen and heard birds) was assumed to have a similar distribution across bands. However, when we compared the perpendicular distances (PD) of all objects to the observer among the three types of objects, we observed that their mean values were statistically different (ANOVA: $F_{2,591} = 8.23$, $P < 0.01$), which implies that inanimate objects, seen and heard birds were not randomly chosen. To correct this, we included PD as a covariate in the model. Thus, our final model was:

Absolute error (~Poisson) =
PD + band-boundary + observer + object +
*band-boundary * object*

We proceeded in a similar way to study accuracy. We considered the balanced error as dependent variable and analysed the effects of the band-boundary, observer, object and PD. In this case we fitted the data to a normal distribution:

Balanced error (~Normal) =
PD + band-boundary + observer + object +
*band-boundary * object*

Balanced errors are expected to cancel out, rendering a high accuracy. Thus, if an observer makes one positive mistake and one negative they will cancel out, yielding a smaller balanced error, closer to zero. Thus, in an ideal monitoring program, observer errors would tend to cancel out, therefore we would expect a null balanced error in our experiment. For this reason, we evaluated whether the balanced errors associated with our two band-boundaries (25 m, 100 m) were 0 by using a single means t-test. In this test, the observed mean (balanced error) is compared to an expected or theoretical mean (in our case 0) (StatSoft, 2010).

Error estimates and error effects. We evaluated the consequences of the observer error by estimating the density of our objects using Distance 5.0 (Thomas *et al.*, 2010). To do this, we derived four types of density estimates. Firstly, from perpendicular distances ('Perpendicular distance' estimate); we took this estimate as a control because it does not break assumption 3 of Distance Sampling (Buckland *et al.*, 2001), and therefore it should yield *a priori* the most reliable estimate. Secondly, directly from raw observer data ('Observer band' estimate), i.e. without correcting the observer error. Thirdly, from the observer data once the misallocations were corrected ('Corrected band' estimate). Finally, from raw observer data applying the mean balanced error to correct the band boundaries (at 25 and 100 m) ('Corrected boundary' estimate).

We estimated density by fitting half-normal distributions to the data set with a Hermite polynomial series expansion in all cases. For the band data set, we generated two models: one truncating the outer band and one without truncating it. For the perpendicular distance data set, we generated three models: a first one without truncating, a second one truncating at 100 m (hence using the same data as the truncated model for the band data set), and a third one truncating at 300 m (which yields the best fit).

Results

Precision and accuracy in band allocation

Each observer made 20 observations but in five cases it was not possible to obtain a distance measure, and in one case the target was beyond the range of the range-finder (> 500 m).

The percentage of misallocated objects was rather high (24.0%), and all observers misallocated at least one object (range 1-10). The absolute error was relatively low (3.21 ± 0.33 m) since most objects (76.0%) were correctly allocated to their real band. However, the maximum error was high (63 m); this extreme case corresponded to an object at a perpendicular distance of 163 m but allocated to the 25-100 m band.

Regarding precision, we found significant differences between observers in absolute error (table 1), as well as between types of objects (inanimate object, seen bird or heard bird) so that the observer error was larger for inanimate objects than for real birds (either seen or heard). However, this effect was only significant on the furthest band-boundary (100 m; figure 1, table 1). Regarding band-boundaries, observer errors were smaller in the first band-boundary (25 m) than in the second band-boundary (100 m). These results suggest that the further the band-boundary is from the observer, the larger the probability of misallocation.

Observers showed significant differences in accuracy (figure 2, table 1), but we did not find differences in accuracy between band allocation of inanimate objects and of birds

TABLE 1

Generalized linear model (GLZ; Sequential Partitioning Variance Method) of the observer precision (absolute error) and accuracy (balanced error) in band allocation. Several predictors were included in the model: the observer, kind of object (inanimate object, seen bird or heard bird) and band-boundary influence precision and accuracy.
[Modelo lineal generalizado (GLZ; método de partición de la varianza secuencial) de la precisión (error absoluto) y exactitud (error balanceado) en la asignación de banda. Se incluyeron diversas variables predictoras: el observador, el tipo de objeto (objeto, ave vista u oída) y el límite de banda influencian significativamente la precisión y exactitud.]

Precision *Precisión*	**Parameter** *Parámetro*	**D.F.**	χ^2	**P**
Perpendicular distance (PD) *Distancia perpendicular*	–0.06	1	337.02	< 0.001
Band boundary *Límite de banda*	–0.93	1	916.85	< 0.001
Observer *Observador*		29	710.51	< 0.001
Object *Objeto*		2	75.34	< 0.001
Object x Band boundary *Objeto x Límite de banda*		2	8.95	< 0.02
Accuracy *Exactitud*	**Parameter** *Parámetro*	**D.F.**	χ^2	**P**
Perpendicular distance (PD) *Distancia perpendicular*	0.02	1	6.3	< 0.02
Band boundary *Límite de banda*	0.13	1	142.0	< 0.001
Observer *Observador*		29	152.5	< 0.001
Object *Objeto*		2	0.3	0.88
Object x Band boundary *Objeto x Límite de banda*		2	9.5	< 0.02

(either seen or heard), although observers tended to be less accurate at 100 m for seen birds than for heard birds and inanimate objects (figure 1, table 1). Additionally, observers were significantly more accurate at the 25 m band-boundary than at the 100 m band-boundary.

Balanced errors were significantly different from zero: at the 25 m band-boundary it was positive (mean ± 95% CI: 0.52 ± 0.44 m; one-sample test: $t = 2.29$, $P < 0.05$), and at the 100 m band-boundary it was negative (mean ± 95% CI: -2.26 ± 1.85; one-sample test: $t = -2.40$, $P < 0.05$). This means that, on average, the observers place the 25 m band-boundary at 25.52 m and the 100 m band-boundary at 97.74 m. Hence, we used these values for correcting band-boundaries (Corrected boundary) in density estimate.

Density estimates

In all approaches, estimates based on truncated models are greater than non-truncated models. We did not detect any signifi-

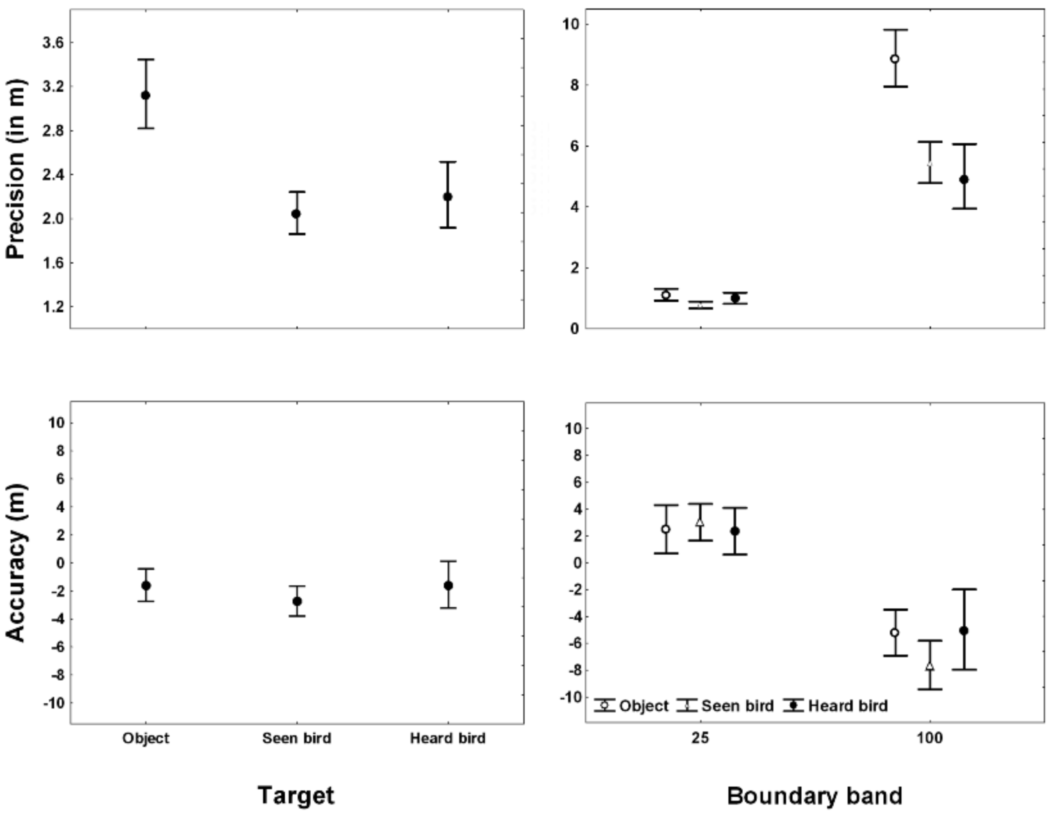

FIG. 1.—Differences in precision (absolute error) and accuracy (balanced error) in relation to the item and the band-boundary considered. Mean ± 95% CI are shown.
[Diferencias en precisión (error absoluto) y exactitud (error balanceado) en función del objeto y el límite de banda considerado. Se muestra la media ± IC.]

cant difference between the three density estimates based on bands (table 2). Interestingly, estimates based on exact distances ('Perpendicular distance' estimate) are significantly smaller than those based on bands ('Observer band', 'Corrected band' and 'Corrected boundary' estimates), even when truncating at 100 m (that is, using exactly the same data set used in truncated models based on bands), but especially when compared to the model with best fit truncation at 300 m (table 2).

TABLE 2

Different density estimates obtained using Distance 5.0. We deemed the density obtained from perpendicular distance data truncated at 300 m as the best estimate (it shows the best fit). For comparison purposes, we considered, as a control estimate, the density obtained from perpendicular distance data truncated at 100 m, and then compared it to densities obtained from band distance data (i.e. following the SOCC line transect design: 0-50, 50-100 and 100-1,000 m) by considering two possible approaches: firstly, without correcting the observer error (Observer band) and secondly, with two correction methods (Corrected band and Corrected boundary). We also show the density estimate resulting from line transect data grouped in four belts (4 bands: 0-25, 25-50, 50-100, 100-1,000 m).
*[Diferentes estimaciones de densidad obtenidas utilizando Distance 5.0. Se consideró la densidad obtenida de datos de distancia perpendicular truncada en 300 m, como la mejor estimación (muestra el mejor ajuste). A efectos comparativos, se consideró como una estimación control, la densidad obtenida de datos usando la distancia perpendicular truncada en 100 m, y luego se comparó con las densidades obtenidas a partir de datos de bandas (siguiendo el diseño del transecto SOCC: 0-50, 50-100 y 100-1.000 m), considerando dos enfoques posibles: el primero, sin corregir el error del observador (*Observer band*) y el segundo, con dos métodos de corrección (*Corrected band *y* Corrected boundary*). También se muestra la densidad de estimación a partir de los datos de transectos lineales agrupados en cuatro bandas (4 bandas: 0-25, 25-50, 50-100, 100-1.000 m).]*

	Estimate (Ind/km^2) *Estima (Ind./km^2)*	**-95% CI**	**+95% CI**	**Coefficient variation (%)** *Coeficiente de variación (%)*
Observer band *Banda del observador*	70.23	57.70	85.48	10.03
Corrected band *Bandas corregidas*	69.32	56.80	84.59	10.02
Boundary band *Límites corregidos*	68.79	56.51	83.73	10.03
Perpendicular distance (100) *Distancia perpendicular (100)*	34.98	28.77	42.53	9.96
Perpendicular distance (300) *Distancia perpendicular (300)*	31.05	26.29	36.82	8.68
Band redesign (4 bands) *Rediseño de bandas (4 bandas)*	45.45	36.36	56.81	11.40

Discussion

The minimisation of errors is an important issue in monitoring schemes. Effective identification and evaluation of the causes of errors are necessary to ensure good quality results (Sutherland, 1998; Bibby *et al.*, 2000; Gregory *et al.*, 2004). Skill, experience and perceptual capabilities of observers are important factors in collecting reliable data. Precise and accurate band allocation can influence the final results derived from monitoring scheme data when analysed using Distance Sampling. Our study confirms common sense: observers have different distance perceptions and make more errors the further the distance to the object, but our study shows that differences in band allocation may be found even in transects with wide bands. Landscape features (slope, vegetation, etc.) may influence observer perception of his or her distance to an object. Another factor that can increase error is the perpendicular distance to the object. The more distant an object, the higher the probability of error. Our results show that both the precision and accuracy were lower at greater distances, which agrees with previous evidence (e.g. Burnham *et al.*, 1985).

A surprising result of this study is that the observer error is similar for visual and aural contacts. However, this result could have changed if the observer had been asked to estimate the exact distance to the item rather than just allocating it to a coarse band.

Consequences of the observer error

An important assumption for reliable estimation to be made using Distance Sampling is that the perpendicular distance to the object is correctly measured, or that items are allocated to the correct band in line transects (Buckland *et al.*, 2001). Our results show that the observer error, although significantly different from zero, is negligible for density estimates based on strip line transects (comparison between table 2, figure 2). However, the most striking result is the large differences between the estimates obtained from grouped data (i.e. in bands) and ungrouped data (i.e. exact distances). If this effect is widespread in monitoring schemes that are based on the strip line transect and that are analysed by means of Distance Sampling, the population sizes of many species would tend to be overestimated.

Marques (2004) describes four types of error in Distance Sampling: species identification, rounding effect, biased random error (i.e. distance assessment) and unbiased random error. Our results show that assigning all contacts to the mid point of a given distance interval (band) has an influence larger than observer band misallocation.

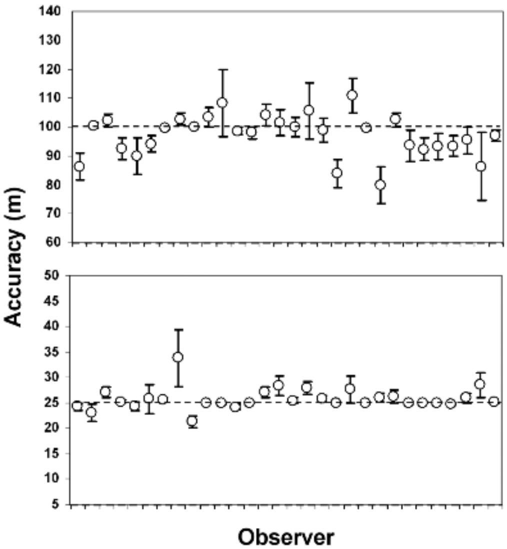

Fig. 2.—Differences in accuracy among observers assessed at the 25 m and 100 m band-boundaries. Mean ± 95% CI are shown.

[Diferencias en exactitud entre observadores en los límites de banda a 25 m y 100 m. Se muestra la media ± IC.]

Ways to reduce errors in band allocation and obtain more reliable estimates

Our results show that the average SOCC observer is highly accurate. However, since observer error can be very variable (figure 2), the effect of this bias may be more evident when using a low number of transects or when working at a local scale. In addition, several studies have shown that even where accuracy is extremely good, a low precision may generate poor density estimates (Chen, 1998; Marques, 2004). This could be the case here. Our results show that absolute errors were important for both 25 m and 100 m band-boundaries.

To improve distance estimation we could use range-finders (Marques, 2004). Unfortunately, this is impractical in large monitoring schemes, and in many cases unaffordable. A possible solution would be to improve the observer precision through training and/or providing satellite images of their transects with the virtual bands superimposed. Also, one partial solution would be to apply a calibration equation to the analysis (Buckland *et al.*, 2001).

In an experiment like ours, the errors of every observer can be plotted by means of a Pareto chart to identify the least competent observers (Bibby *et al.*, 2000). Small-scale studies could select their observers using this system.

Finally, any method for obtaining reliable density estimates from large monitoring schemes such as the SOCC (which are based on distance band data) should take into account two drawbacks: it is unfeasible to collect perpendicular distances, and estimates tend to be inflated. Here, we propose an intermediate approach to this problem consisting of rearranging the line transect design by applying what we have learnt from our experiment:

1) The less grouped the data, the better.
2) The further the band-boundaries are from the observer, the larger the observer errors.

For instance, if we consider four bands by dividing our 25-100 m band of our experiment into two bands (25-50 m, 50-100 m) and we convert perpendicular distances in bands (for instance: 12 m = 0-25 m band, 31 m = 25-50 m band, 77 m = 50-100 m band, 120 m = 100-1,000 m band), the resulting density estimate is much closer to the 'Perpendicular distance' estimate (table 2, figure 2) than the 'Corrected band' estimate. Hence, implementing additional bands may help in obtaining more accurate estimates and, more importantly, it is technically quite feasible. This extra band would allow the analyst to:

1) Gain more precision by suppressing the furthest band in most cases, which is much too wide and whose extreme band boundary (1,000 m) is usually far beyond the observer's visual and aural horizon.
2) Carry out tests of best fit (currently, Distance users must select between truncated and non-truncated models just by the shape of the plot).

Our approach does not take into account the new observer error (i.e. the one associated to the new band-boundary at 50 m) and, therefore, the results derived from this reanalysis are slightly more optimistic than they would have been. Future research should test if the observer error increases with an increasing number of bands.

ACKNOWLEDGEMENTS.—The authors are grateful to all the SOCC observers, and especially those who volunteered in this study. They also thank two anonymous reviewers and Javier Seoane for their helpful comments to the manuscript. The SOCC is a project developed by the Catalan Ornithological Institute and the Catalan Government. Research project to CGL 2009-08 798 to JQ from Science and Innovation Ministry partially funded this study.

BIBLIOGRAPHY

BIBBY, C., BURGUESS, N. D. and HILL, D. A. 2000. *Bird census techniques*. Cambridge University Press. Cambridge.

BUCKLAND, S. T., ANDERSON, D. R., BURNHAM, K. P., LAAKE, J. L., BORCHERS, D. L. and THOMAS, L. 2001. *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press. Oxford.

BURNHAM, K. P., ANDERSON, D. R. and LAAKE, J. L. 1985. Efficiency and bias in strip and line transect sampling. *Journal of Wildlife Management*, 49: 1012-1018.

CHEN, S. X. 1998. Measurement errors in line transect surveys. *Biometrics*, 57: 732-742.

FREDERICK, P. C., HILTON, B., HEATH, J. A. and RUANE, M. 2003. Accuracy and variation in estimates of large numbers of birds by individual observers using an aerial survey simulator. *Journal of Field Ornithology*, 74: 281-287.

GREGORY, R. D., GIBBONS, D. W. and DONALDSON, R. 2004. Bird census and survey techniques. In, W. J. Sutherland, I. Newton and R. E. Green (Eds.): *Bird ecology & conservation; a handbook of techniques*, pp. 17-56. Oxford University Press. Oxford.

GREGORY, R. D., VAN STRIEN, A., VORISEK, P., MEYLING, A. W. G., NOBLE, D. G., FOPPEN, R. P. B. and GIBBONS, D. W. 2005. Developing indicators for European birds. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 360: 269-288.

HERRANDO, S., BROTONS, L., ESTRADA, J. and PEDROCCHI, V. 2008. The Catalan common bird survey (SOCC): a tool to estimate species population numbers. *Revista Catalana d'Ornitologia*, 24: 138-146.

HERRANDO, S., BROTONS, L., ESTRADA, J., GUALLAR, S. and ANTON, M. 2010b. *Atles dels ocells de Catalunya a l'hivern 2006-2009*. ICO and Lynx Edicions. Barcelona.

HERRANDO, S., BROTONS, L., GUALLAR, S. and QUESADA, J. 2010a. Assessing regional variation in conservation value using fine-grained bird atlas. *Biodiversity and Conservation*, 19: 867-881.

MARQUES, T. A. 2004. Predicting and correcting bias caused by measurement error in line transect sampling using multiplicative error models. *Biometrics*, 60: 757-763.

NEWSON, S. E., EVANS, K. L., NOBLE, D. G., GREENWOOD, J. J. D. and GASTON, K. J. 2008. Use of distance sampling to improve estimates of national population sizes for common and widespread breeding birds in the UK. *Journal of Applied Ecology*, 45: 1330-1338.

ROBERTSON, A., SIMMONS, R. E., JARVIS, A. M. and BROWN, C. J. 1995. Can bird atlas data be used to estimate population-size? A case-study using Namibian endemics. *Biological Conservation*, 71: 87-95.

STATSOFT, Inc. 2010. Electronic statistics textbook. *http*://www.statsoft.com/textbook/. StatSoft. Tulsa.

SUTHERLAND, W. J. 1998. *Ecological census techniques: A handbook*. Cambridge University Press. Cambridge.

THOMAS, L., BUCKLAND, S. T., REXSTAD, E. A., LAAKE, J. L., STRINDBERG, S., HEDLEY, S. L., BISHOP, J. R., MARQUES, T. A. and BURNHAM, K. P. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, 47: 5-14.

[Recibido: 17-05-2010]
[Aceptado: 28-09-2010]