

# DENSITY: software for analysing capture–recapture data from passive detector arrays

M. G. Efford, D. K. Dawson & C. S. Robbins

Efford, M. G., Dawson, D. K. & Robbins, C. S., 2004. DENSITY: software for analysing capture–recapture data from passive detector arrays. *Animal Biodiversity and Conservation*, 27.1: 217–228.

## Abstract

*DENSITY: software for analysing capture–recapture data from passive detector arrays.*— A general computer–intensive method is described for fitting spatial detection functions to capture–recapture data from arrays of passive detectors such as live traps and mist nets. The method is used to estimate the population density of 10 species of breeding birds sampled by mist–netting in deciduous forest at Patuxent Research Refuge, Laurel, Maryland, U.S.A., from 1961 to 1972. Total density ( $9.9 \pm 0.6 \text{ ha}^{-1}$  mean  $\pm$  SE) appeared to decline over time (slope  $-0.41 \pm 0.15 \text{ ha}^{-1}\text{y}^{-1}$ ). The mean precision of annual estimates for all 10 species pooled was acceptable ( $CV(\hat{D}) = 14\%$ ). Spatial analysis of closed–population capture–recapture data highlighted deficiencies in non–spatial methodologies. For example, effective trapping area cannot be assumed constant when detection probability is variable. Simulation may be used to evaluate alternative designs for mist net arrays where density estimation is a study goal.

Key words: Passive detector arrays, Density estimation, Capture–recapture, Mist–netting, Birds.

## Resumen

*DENSITY: programa empleado para el análisis de datos de captura–recaptura procedentes de matrices de detectores pasivos.*— En este estudio se describe un método general de cómputo intensivo que permite ajustar las funciones de detección espacial a datos de captura–recaptura procedentes de baterías de trampas pasivas, como las trampas de cebo y las redes japonesas. Este método es utilizado para estimar la densidad de población de 10 especies de aves reproductoras, muestreadas mediante la colocación de redes japonesas en un bosque de árboles de hoja caduca del Centro de Investigación Patuxent, en Laurel, Maryland, Estados Unidos, desde 1961 hasta 1972. La densidad total ( $9,9 \pm 0,6 \text{ ha}^{-1}$  promedio  $\pm$  EE) parecía disminuir con el tiempo (gradiente  $-0,41 \pm 0,15 \text{ ha}^{-1}\text{y}^{-1}$ ). La precisión media de las estimaciones anuales correspondientes a la totalidad de las 10 especies recogidas fue aceptable ( $CV(\hat{D}) = 14\%$ ). El análisis espacial de los datos de captura–recaptura de la población cerrada revelaron deficiencias en las metodologías no espaciales. Así, por ejemplo, no puede suponerse que el área efectiva de colocación de trampas sea constante cuando la probabilidad de detección es variable. En los casos en que la estimación de la densidad sea objeto de estudio, la simulación permite evaluar diseños alternativos para baterías de redes japonesas.

Palabras clave: Baterías de trampas pasivas, Estimación de la densidad, Captura–recaptura, Redes japonesas, Aves.

Murray G. Efford, Landcare Research, Private Bag 1930, Dunedin, New Zealand.— Deanna K. Dawson & Chandler S. Robbins, USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708, U.S.A.

Corresponding author: Dr M. G. Efford. E–mail: [effordm@landcareresearch.co.nz](mailto:effordm@landcareresearch.co.nz)

## Introduction

Rigorous sampling of animal populations to estimate or index density raises the problem of incomplete detection (e.g. Burnham, 1981; MacKenzie & Kendall, 2002; Pollock et al., 2002; Rosenstock et al., 2002; Thompson, 2002). Detection probability generally has been described by a single parameter  $p$ . An estimate of  $p$  may be used to obtain a population estimate  $N$  from a count  $C$ :

$$\hat{N} = \frac{C}{\hat{p}} \quad (1)$$

Estimation of detection probability protects the population estimate or index from the confounding effects of season, time of day, observer, weather, habitat, etc. If detections relate to a known area  $A$  then population density may be calculated as

$$\hat{D} = \frac{\hat{N}}{A} \quad (2)$$

This strategy applies to "limited-area" counts when a stationary observer makes instantaneous observations of bird locations and includes only those within a known area (e.g. the double-observer approach of Nichols et al., 2000 and the removal method of Farnsworth et al., 2002). These methods are called "active" because they require continuous attention and discrimination by the observer.

"Passive" counts are obtained when a detector (e.g. trap, mist net, or camera) records individuals at a point. Individuals are included in the count only when they encounter and interact with a detector. Passive detectors are commonly deployed in arrays of varying geometry and size. The general term "passive detector array" (PDA) is suggested to emphasize the common features of spatial capture data from diverse field studies (table 1). Passive detection combined with mark and release on a series of occasions spaced closely in time is a common source of data for closed-population experiments (e.g. Otis et al., 1978).

Equations (1) and (2) and their capture-recapture equivalents do not provide an adequate framework for estimating animal density from capture data from PDAs. The area  $A$  is unknown and difficult to define. It follows that both  $N$  and  $p$  are ill-defined for these data except in an operational and probably circular sense (e.g. " $N$  is the number of animals potentially exposed to the PDA"). Thus  $N$  varies both with animal behaviour and with the configuration of the PDA. The widespread use of the term "abundance" for  $N$  acknowledges its vagueness in this context. Further complications arise because the component detectors (traps, mist nets) of a PDA may interact. Interaction commonly occurs when an animal detained in one trap is not immediately available for capture in a different trap.

An alternative framework is advocated for estimating density from closed population data from

arrays of passive detectors. The framework is conceptually consistent with that of distance sampling (Buckland et al., 1993; Rosenstock et al., 2002), but it offers major advantages for passive count data. Here we introduce the spatial detection model for PDAs and a numerical method for model fitting (Efford, 2004), along with software designed to make the method generally accessible. We assess the potential of the method for estimating the population density of birds captured in mist nets, using a dataset collected in Maryland, U.S.A., by CSR.

## Spatial model for the detection process

Assume that animals occupy stationary home ranges whose centres are a realization of a homogeneous random spatial point process with intensity (density)  $D$ . Populations that have a natural boundary are explicitly excluded. Passive sampling uses detectors in a known spatial configuration to sample the unknown distribution of animals. An individual-based model is proposed for the detection process. The core of the model is a spatial detection function  $g(r)$  for the simplest possible case: one animal and one detector. The probability of detecting animal  $i$  is assumed to be a decreasing function of the distance  $r$  between its range centre and the detector. The simplest useful detection function has two parameters. In the formulation discussed here, these correspond to measures of home range size ( $\sigma$ ) and susceptibility to capture ( $g(0)$ ). This definition of  $g$  is more useful than a global one at the level of the entire array, as parameter estimates are "portable" to other detector configurations (i.e. different arrays).

Given some ancillary information, the three parameters  $D$ ,  $g(0)$  and  $\sigma$  define the detection process. The required ancillary information is: (i) the configuration of the detector array (i.e. x-y coordinates of detectors), (ii) the nature of the spatial point process (here assumed to be Poisson), (iii) a model for resolving conflicts between incompatible detection events (e.g. animal caught in two traps at once), and (iv) the shape of the detection function (assumed here to be half-normal). Writing a computer algorithm to simulate capture data from this model is straightforward except for (iii), which is addressed later.

## Fitting the spatial detection model

Our formulation of closed population sampling in terms of  $D$ ,  $g(0)$  and  $\sigma$  is useful only if there is a practical method of estimation. An expression for the likelihood is currently lacking, and therefore maximum likelihood estimators cannot be derived. Instead,  $D$ ,  $g(0)$  and  $\sigma$  are estimated by simulation and inverse prediction (Carothers, 1979; Pledger & Efford, 1998). Briefly, this method uses Monte Carlo sampling of populations with known  $D$ ,  $g(0)$

and  $\sigma$  to generate data that may be "matched" to the field data. "Matching" uses statistics from the data as surrogates for  $D$ ,  $g(0)$  and  $\sigma$ . Each statistic is chosen for its conditional monotonic relation to a parameter. The statistics used here are the closed population estimate  $\hat{N}$ , the corresponding estimate of mean detection probability  $\hat{p}$ , and the mean distance between successive detections of the same individual  $\bar{d}$ .

Inverse prediction (Brown, 1982; Pledger & Efford, 1998) provides a formal framework for estimating  $D$ ,  $g(0)$  and  $\sigma$  from  $\hat{N}$ ,  $\hat{p}$  and  $\bar{d}$ , complete with prediction standard errors. The parameter vector  $\mathbf{x}$  and simulated observations  $\mathbf{y}$  may be used to fit the multivariate multiple regression

$$\mathbf{y} = \boldsymbol{\lambda} + \mathbf{B}\mathbf{x} + \mathbf{E}$$

where  $\boldsymbol{\lambda}$  is a 3 x 1 vector of intercepts,  $\mathbf{B}$  is a 3 x 3 matrix of coefficients, and  $\mathbf{E}$  is a 3 x 1 vector of error terms with multivariate normal  $MVN(\mathbf{0}, \mathbf{V})$  distribution. With sufficient replications, the elements of  $\boldsymbol{\lambda}$ ,  $\mathbf{B}$  and  $\mathbf{E}$  are estimated virtually without sampling error.

Given a single observation  $\mathbf{y}_p$  the point estimates of  $D$ ,  $g(0)$ , and  $\sigma$  (together, the vector  $\mathbf{x}_p$ ) are given by

$$\mathbf{x}_p = \mathbf{B}^{-1}(\mathbf{y}_p - \boldsymbol{\lambda}) \quad (3)$$

The model equation rearranges to:

$$\mathbf{x} = \mathbf{B}^{-1}(\mathbf{y} - \boldsymbol{\lambda}) + \mathbf{B}^{-1}\mathbf{E}$$

where the random error  $\mathbf{B}^{-1}\mathbf{E}$  has distribution  $MVN(\mathbf{0}, \boldsymbol{\Gamma})$ , where  $\boldsymbol{\Gamma} = \mathbf{B}^{-1} \mathbf{V} \mathbf{B}^{-1T}$ .

Detection conflicts

As noted previously, simulations with a one–animal, one–detector detection function may lead to conflicts when multiple animals interact with multiple detectors. The problem is severe when either animals or detectors are at high density (e.g. when traps become "saturated"). This issue was addressed by a discrete–event simulation of the trapping process (see also Efford, 2004). Potential detection events (animal  $i$  at detector  $j$ ) were treated as competing Poisson processes. The time to first occurrence of each potential event then follows an exponential distribution with rate parameter  $\lambda_{ij} = -\ln(1-g(r_{ij}))$  where  $r_{ij}$  is the distance between animal  $i$  and detector  $j$  and  $g(r) = g(0)\exp(-r_{ij}^2/(2\sigma^2))$ . Independent pseudorandom exponential variates, one for each potential event, were sorted by magnitude to establish priority among events. Simulated events were discarded (i.e. did not occur) if they were inconsistent with previous events or they occurred after one unit of elapsed detection time. Mist nets allow multiple simultaneous detections; for this detector type the only consistency constraint was that an individual could appear at no more than one detector per sample.

Table 1. Examples of passive detectors used in animal population studies. Detectors in which animals are detained provide the further option of selective non–release: D. Detections; Eb. Effect on behaviour; \* Individual natural marks are required for capture–recapture analysis.

*Tabla 1. Ejemplos de métodos de trampeo utilizados en estudios de poblaciones animales. Los métodos de trampeo en los que los animales quedan retenidos proporcionan la opción más avanzada de no liberación selectiva: D. Detecciones; Eb. Efecto en el comportamiento; \* Para el análisis de captura–recaptura se precisan marcas naturales individuales.*

Type	D	Eb
Sherman trap	single	detained
Mist net	multiple	detained
Pitfall trap	multiple	detained
Crab pot	multiple	detained
Fixed camera*	multiple	not detained

Software

The Windows ® program DENSITY ([www.landcare-research.co.nz/services/software/density](http://www.landcare-research.co.nz/services/software/density)) analyses closed population capture–recapture data from arrays of passive detectors. Two input text files are required. One file contains the locations (x–y coordinates) of the detectors. The second file records detection events (individual ID, sample number and detector ID). Detection events may for convenience be stratified by "session". Each session is analysed separately. A graphic interface enables the visualization of spatial detection data (fig. 1). Program usage is described in an online help file.

DENSITY implements the proposed method for simulation and inverse prediction. Starting values of the parameters are determined automatically (see the Appendix for a description of how these are calculated). Model fitting proceeds by Monte Carlo simulation of detection samples from random populations with known parameter vectors. In each simulation a new set of animal locations is generated for a rectangular area that includes the PDA and a buffer zone. The width of the buffer zone should be at least  $3\sigma$  in order to include all individuals with a reasonable chance of detection. Parameter values follow a full factorial design, i.e. they lie at the vertices of a "box" in parameter space. The dimensions of the parameter "box" are fixed as a percentage (e.g.  $\pm 10\%$ ) of the current best esti-

mate of the parameter values. Statistics ( $\hat{N}$ ,  $\hat{\rho}$  and  $\hat{\sigma}$ ) from the multiple simulations conducted at each vertex are averaged to remove sampling variance. A sample size of 20–100 simulations appears adequate, but this requires further investigation. Estimation follows equation (3). If the estimated vector ( $D$ ,  $g(0)$ ,  $\sigma$ ) lies outside the initial "box" then the estimate becomes the starting point for another simulation cycle. This avoids extrapolation of the linear approximating function. Once a satisfactory prediction is obtained (i.e. the estimate lies inside the box), further simulations are conducted to estimate the variance–covariance matrix in statistic space  $\mathbf{V}$ , and its equivalent in parameter space  $\mathbf{\Gamma}$  (see above).

The automatic algorithm sometimes fails to find a parameter "box" that includes the fitted parameter values, in which case a degree of supervision may be required to fit the model. It is usually sufficient to provide better starting values, either manually or by applying a constant scalar adjustment (e.g.  $\times 0.5$ ) to the "automatic" initial value for  $g(0)$ . The adjustment may be stored for further use with the same detector configuration. It may also be necessary to increase the size of the "box" or the number of replicate simulations at each vertex.

Outputs of DENSITY include both the " $N$ - $p$ " analyses of conventional closed population models (Otis et al., 1978; Chao & Huggins, in press a, in press b) and numerical estimates of  $D$ ,  $g(0)$  and  $\sigma$  by inverse prediction. The user may also simulate sampling with novel detector arrays to identify efficient ways to allocate sampling effort and to predict precision and bias. This meta-functionality is described as "power analysis" to distinguish it from the simulations embedded in the estimation of  $D$ ,  $g(0)$  and  $\sigma$  by inverse prediction.

#### Test of assumptions

Goodness-of-fit tests for the present spatial detection model have yet to be developed. However, a test has been developed for one key assumption. This is the assumption that capture locations are sampled from a stationary distribution subject only to the modelled effects of competition for and among detectors and, in particular, that capture location is not affected by previous capture. A suitable test statistic is the  $t$ -value for a comparison of the mean distance between first and second capture when captures are in consecutive samples versus the mean distance between first and second capture when captures are separated by more than one sample. Values of this statistic may be compared to its bootstrap sampling distribution from simulated realizations of spatial sampling with appropriate  $D$ ,  $g(0)$  and  $\sigma$  (see DENSITY online help).

#### Methods

Birds were mist-netted on a forested site on the Patuxent Research Refuge, Maryland, U.S.A.

(39° 3' N, 76° 48' W) in early summer from 1959 to 1972. The initial study design used 21 12-m nets at 61-m (200-foot) intervals on the arms of a cross (Stamm et al., 1960). This was changed in 1960 to a 4 x 11 grid, with net spacing 100 m between rows and 61 m along rows. On six non-consecutive days in late May and early June, nets were open between about 0600 hours and 1800 hours Eastern Daylight Time and were checked every 2 hours. On initial capture birds were ringed with uniquely numbered aluminium rings, identified to species and, when possible, to age and sex, measured, and released. The ring number and the date and location of capture were also recorded for previously marked birds. Noon & Sauer (1990) presented some analyses of survival and recruitment from these data. Here we analyse only closed-population data from the grid layout (1961 to 1972). Thirty-seven bird species netted over 1961–1972 were likely to have been breeding in the forest. We focus on the 10 species with at least 200 captures each (*Baeolophus bicolor*, *Cardinalis cardinalis*, *Empidonax vireescens*, *Hylocichla mustelina*, *Oporornis formosus*, *Piranga olivacea*, *Seiurus aurocapillus*, *Setophaga ruticilla*, *Vireo olivaceus*, *Wilsonia citrina*).

#### Analysis of mist-netting data in DENSITY

Captures were pooled by day (i.e. within-day recaptures were ignored). Most species were captured and recaptured in only small numbers in each sampling session (year), and it was necessary to group data for analysis. Analyses were conducted by species for data pooled over three consecutive years (1961–1963, 1964–1966, etc.), considering only recaptures within a year. In addition both annual and 3-year-pooled analyses were conducted on the pooled sample from the 10 species most commonly caught. This enabled us to assess empirically the effect of pooling species with different detection functions.

Half-normal spatial detection functions were fitted by simulation and inverse prediction. Bird range centres were assumed to follow a Poisson distribution. For closed-population estimation ( $\hat{N}$ ,  $\hat{\rho}$ ), Chao's second coverage estimator for model  $M_{th}$  was used (Otis et al., 1978; Lee & Chao, 1994). This estimator is a sensible and conservative option given the likely presence of non-spatial heterogeneity. However, some study designs do not allow formal probabilistic estimation of  $N$  (e.g. when the locations of detectors only partly overlap between samples). In these situations it is still possible to estimate density by the method presented here, but it is necessary to use an ad hoc surrogate for  $\hat{N}$  such as the number of individuals caught ( $M_{t+1}$ ). To evaluate the effect of this substitution on the estimates, the density of all 10 species pooled was also estimated using  $M_{t+1}$  for  $\hat{N}$  and setting

$$\hat{\rho} = \frac{n_1 + n_2}{2M_{t+1}}$$

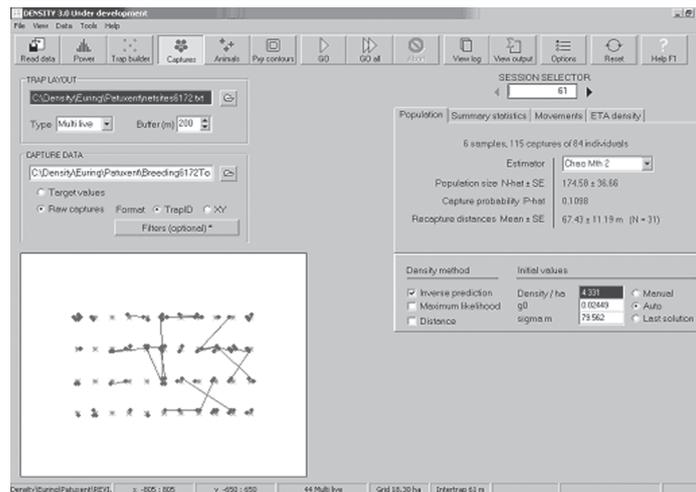


Fig. 1. Graphic interface to DENSITY.

Fig. 1. Interfaz gráfica del DENSITY.

Initial parameter values were determined as described in the Appendix. The factorial design in parameter space spanned  $\pm 10\%$  of the initial values; statistics were averaged from 100 simulations with each combination of parameter values. Nets were modelled as multi-catch detectors with marking and live release. The variance-covariance matrix was estimated by conducting 200 further simulations at the fitted values. The precision of density estimates was expressed as  $CV(\hat{D}) = SE(\hat{D})/\hat{D}$ .

Results

The detection model was fitted successfully to the 3-year grouped data for all species. *Vireo olivaceus* maintained the highest population density throughout the study (table 2). Precision depended strongly on the number of recaptures in the sample (fig. 2). Relative number of captures was a poor measure of relative density; for example, captures of *Seiurus aurocapillus* outnumbered those of *Empidonax virescens*, but the estimated density of *Empidonax virescens* was always more than twice that of *Seiurus aurocapillus* (table 2). This is consistent with large species differences in the fitted detection functions (table 2; fig. 3). Three-year estimates for the 10 species pooled (table 2) were close to the sums of the individual species estimates for each interval (1961–1963 12.9 ha<sup>-1</sup>; 1964–1966 11.0 ha<sup>-1</sup>; 1967–1969 10.3 ha<sup>-1</sup>; 1970–1972 7.6 ha<sup>-1</sup>). Thus the pooled data appear to provide usable estimates of total density despite species differences in detection.

The number of within-year recaptures ( $\Sigma m$ ) for all 10 species pooled ranged from 52 to 129 ( $97.4 \pm 7.9$

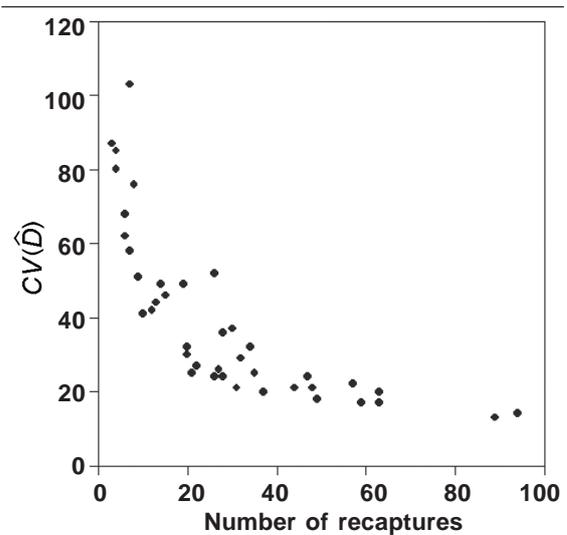


Fig. 2. Precision of density estimated by inverse prediction ( $CV(\hat{D})\%$ ) for 10 bird species mist-netted at Patuxent Research Refuge, Maryland, U.S.A., 1961–1972, as a function of the number of recaptures. Each point represents one 3-year pooled estimate for one species.

Fig. 2. Precisión de la densidad estimada por predicción inversa ( $CV(\hat{D})\%$ ) para 10 especies de aves capturadas con redes japonesas en el Centro de Investigación Patuxent, Maryland, Estados Unidos, 1961–1972, en función del número de recapturas. Cada punto representa una estimación combinada de tres años para una especie.

Table 2. Density ( $\hat{D}$  ha<sup>-1</sup>) and spatial detection parameters ( $\hat{g}(0)$ ,  $\hat{\sigma}$  m) of breeding bird species at Patuxent Research Refuge, Maryland, U.S.A., in 3-year intervals 1961–1972. Estimates by inverse prediction (SE). Also shown is the number of captures, including recaptures, over the entire study (NC).

Tabla 2. Densidad ( $\hat{D}$  ha<sup>-1</sup>) y parámetros de detección espacial ( $\hat{g}(0)$ ,  $\hat{\sigma}$  m) de especies de aves reproductoras en el Centro de Investigación Patuxent, Maryland, Estados Unidos, en intervalos de tres años: 1961–1972. Estimaciones por predicción inversa (EE). También se indica el número de capturas, incluyendo las recapturas, a lo largo de todo el estudio (NC).

Species	Year interval												
	1961–63			1964–66			1967–69			1970–72			
	NC	$\hat{D}$	$\hat{g}(0)$	$\hat{\sigma}$									
<i>Vireo olivaceus</i>													
	1,015	5.50	0.034	55	3.35	0.038	61	3.55	0.041	66	2.85	0.027	59
		(0.92)	(0.007)	(5)	(0.56)	(0.008)	(6)	(0.51)	(0.006)	(5)	(0.80)	(0.009)	(8)
<i>Hylocichla mustelina</i>													
	743	1.68	0.041	90	1.43	0.019	85	1.11	0.028	103	1.36	0.035	84
		(0.24)	(0.007)	(8)	(0.34)	(0.009)	(16)	(0.22)	(0.006)	(12)	(0.25)	(0.008)	(9)
<i>Seiurus aurocapillus</i>													
	436	0.42	0.076	89	0.33	0.070	111	0.37	0.050	114	0.30	0.035	118
		(0.09)	(0.019)	(10)	(0.08)	(0.018)	(13)	(0.09)	(0.015)	(16)	(0.11)	(0.014)	(22)
<i>Empidonax virescens</i>													
	385	1.28	0.036	75	1.06	0.039	68	0.88	0.043	64	0.97	0.020	78
		(0.27)	(0.011)	(10)	(0.29)	(0.012)	(10)	(0.22)	(0.016)	(10)	(0.45)	(0.011)	(20)
<i>Oporornis formosus</i>													
	366	0.53	0.072	78	0.42	0.023	124	0.66	0.038	87	0.40	0.055	81
		(0.11)	(0.019)	(8)	(0.16)	(0.009)	(24)	(0.20)	(0.012)	(12)	(0.10)	(0.020)	(13)
<i>Wilsonia citrina</i>													
	288	0.36	0.058	97	0.50	0.023	93	0.19	0.037	155	0.08	0.043	180
		(0.09)	(0.018)	(12)	(0.25)	(0.012)	(24)	(0.06)	(0.013)	(28)	(0.04)	(0.018)	(55)
<i>Baeolophus bicolor</i>													
	262	0.51	0.027	95	0.52	0.013	115	0.61	0.057	80	0.44	0.008	111
		(0.16)	(0.014)	(18)	(0.26)	(0.007)	(37)	(0.12)	(0.016)	(10)	(0.35)	(0.008)	(64)
<i>Setophaga ruticilla</i>													
	230	1.19	0.037	44	1.68	0.037	55	1.40	0.010	69	0.38	0.029	63
		(0.90)	(0.023)	(11)	(0.50)	(0.018)	(10)	(1.19)	(0.010)	(36)	(0.24)	(0.029)	(29)
<i>Cardinalis cardinalis</i>													
	214	0.35	0.027	96	0.60	0.010	114	0.87	0.015	92	0.30	0.017	132
		(0.14)	(0.015)	(27)	(0.62)	(0.008)	(73)	(0.36)	(0.008)	(21)	(0.13)	(0.010)	(43)
<i>Piranga olivacea</i>													
	200	1.04	0.012	85	1.09	0.015	71	0.64	0.008	145	0.52	0.004	142
		(0.60)	(0.009)	(43)	(0.74)	(0.009)	(22)	(0.35)	(0.005)	(46)	(0.45)	(0.005)	(87)
Pooled													
	4,139	11.39	0.035	79	11.05	0.024	84	10.06	0.029	89	7.11	0.022	96
		(0.81)	(0.003)	(3)	(0.94)	(0.003)	(4)	(0.77)	(0.003)	(4)	(0.78)	(0.003)	(6)

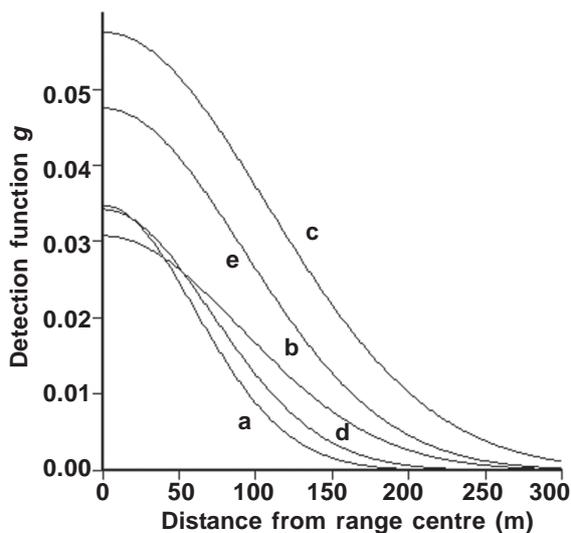


Fig. 3. Fitted detection functions for the five species most often caught in mist nets at Patuxent Research Refuge, Maryland, U.S.A., 1961–1972: a. *Vireo olivaceus*; b. *Hylocichla mustelina*; c. *Seiurus aurocapillus*; d. *Empidonax virescens*; e. *Oporornis formosus*. Parameters were obtained by averaging 3-year pooled estimates for each species (table 2).

Fig. 3. Funciones de detección ajustadas para las cinco especies más frecuentemente capturadas con redes japonesas en el Centro de Investigación Patuxent, Maryland, Estados Unidos, 1961–1972 (a. *Vireo olivaceus*; b. *Hylocichla mustelina*; c. *Seiurus aurocapillus*; d. *Empidonax virescens*; e. *Oporornis formosus*). Los parámetros se obtuvieron calculando el promedio de las estimaciones combinadas de tres años para cada especie (tabla 2).

mean  $\pm$  SE), sufficient to allow us to compute annual estimates of density. The estimated density of common breeding birds ranged between  $4.8 \text{ ha}^{-1}$  (1971) and  $12.2 \text{ ha}^{-1}$  (1966) ( $9.93 \pm 0.64 \text{ ha}^{-1}$  mean  $\pm$  SE) and appeared to decline over time (slope of linear trend  $-0.41 \pm 0.15 \text{ ha}^{-1} \text{ y}^{-1}$ ; fig. 4). Density estimates using  $M_{t+1}$  for  $\hat{N}$  ( $8.16 \pm 0.54 \text{ ha}^{-1}$ ) were lower by  $17 \pm 3\%$  than those using  $M_{th}$ , but this discrepancy was much less than that between the alternative estimates of  $N$  ( $51 \pm 2\%$ ). Estimates of  $g(0)$  and  $\sigma$  were inversely correlated between years ( $r = -0.83$ ,  $P < 0.001$ ), which may be due in part to their inverse sampling covariance. The precision of annual density estimates was high ( $CV(\hat{D}) = 0.144 \pm 0.010$ ) and only slightly worse than that of population size estimates ( $CV(\hat{N}) = 0.123 \pm 0.008$ ).

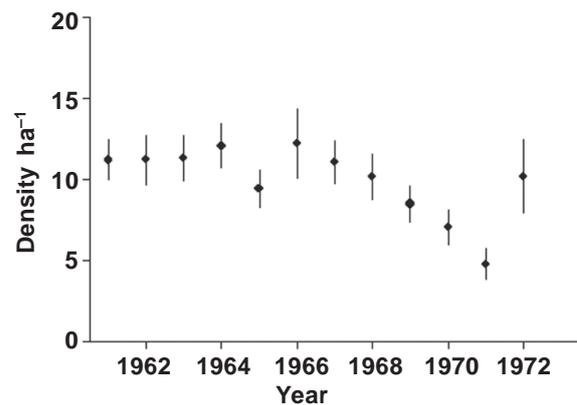


Fig. 4. Temporal variation in annual population density of common breeding bird species at Patuxent Research Refuge, Maryland, U.S.A., 1961–1972, estimated by inverse prediction  $\pm 1$  SE. (See text for details.)

Fig. 4. Variación temporal en la densidad anual de la población de especies de aves comunes reproductoras en el Centro de Investigación Patuxent, Maryland, Estados Unidos, 1961–1972, estimada mediante predicción inversa  $\pm 1$  EE (Para más detalles ver el texto.)

## Discussion

The conventional parameterization of closed-population models in terms of  $N$  and  $p$  is incomplete because it neglects space. Conversely, a spatial parameterization ( $D$ ,  $g(0)$ ,  $\sigma$ ) has major benefits where the underlying dispersion model (localized detection) fits the biology of the study animal. More often than not, ecologists want to measure population density  $D$  rather than  $N$ . Detection functions  $g(r)$  are fundamental to distance analysis, which uses a sample of detection distances (Buckland et al., 1993). Our approach does not model detection distances as such, but estimates parameters of the detection function from the pattern of recaptures. By our definition, the parameters are independent of a particular detector configuration. This means that simulations may be conducted to compare the efficiency of alternative, novel configurations using values of  $g(0)$  and  $\sigma$  estimated from the field.  $\sigma$  is also a convenient measure of home range size. There may be pathological detector configurations for which  $\hat{\sigma}$  is a biased estimate of  $\sigma$ , but this remains to be investigated.

### Closed-population estimator selection

Our method uses an empirical  $\hat{N}$  as an input to inverse prediction. Analyses were presented using

both (1) Chao's sample coverage estimators that are believed to be robust to temporal and individual heterogeneity ( $M_{th}$ ), and (2) the number of distinct individuals caught, which is almost certainly less than the number that would be caught over a longer period. This raises the issue of how to select an appropriate estimator of  $N$  from among the many available (e.g. Otis et al., 1978; Chao & Huggins, in press a, in press b). More work needs to be done on this topic, but the outcome is not critical for the adoption of the method.  $\hat{N}$  is used here only in the context of a particular PDA and detection function. Biases inherent in the estimator and context will also arise during simulation and be automatically down-weighted or removed. This applies specifically to spatial heterogeneity in access by animals to detectors and the general negative bias of incomplete counts from a small number of samples. In some situations (e.g. detection on a continuously shifting array),  $M_{tt+1}$  may be the only available measure of  $N$ , and an estimate of density using  $M_{tt+1}$  and inverse prediction may be acceptable. Nevertheless, density estimates of common breeding birds at Patuxent did depend on the estimator for  $N$ , and similar effects have been observed with field data for other species (e.g. Efford, 2004). Our provisional interpretation is that such field datasets include unmodelled heterogeneity that causes modest negative bias in  $\hat{D}$ , particularly with  $N$ -estimators that are not robust to individual heterogeneity. The available tests for heterogeneity in the closed-population capture histories (e.g. Otis et al., 1978) do not distinguish heterogeneity caused by spatial location (see below) from other heterogeneity. They are therefore inappropriate for model selection in this context. Until new methods are developed the use of robust estimators such as those for  $M_{th}$  and  $M_h$  is recommended.

Simulations reported elsewhere (Efford, 2004) suggest that  $\hat{D}$  is robust to the arbitrary choice of 2-D distribution (Poisson vs even) and detection function (half-normal vs uniform).

#### Conventional capture probability and the spatial detection function

Detection in conventional  $N$ - $p$  frameworks is described at the level of the detector array. In other words,  $p$  is the probability that an average individual of the target population is detected somewhere in the array. When density is low or detectors may register multiple individuals (e.g. mist nets) competition between animals for detectors may be ignored. Then the cumulative array-level probability of detection  $p_{xy}$  of an individual at  $x,y$  can be predicted from its pairwise interactions with detectors:

$$p_{xy} = 1 - \prod_j (1 - g_{xy}(j))$$

Here  $g_{xy}(j)$  represents the probability of detecting an animal with range centred at  $x,y$  in detector  $j$  over a

given time interval.  $p_{xy}$  is a spatial variable that may be contoured for given  $g$  and detector locations (example in fig. 5). This provides a useful perspective on the functioning of the entire PDA as a sampling device, discussed in the next section.

#### Constant-area assumption

The conventional  $N$ - $p$  parameterization relies on constant area  $A$ . Our fully spatial description of the detection process allows us to consider this assumption in more detail. Clearly the assumption does not hold when there is variation in the scale of movement by individuals, indexed by the parameter  $\sigma$  of the detection function. In our experience  $\sigma$  is often a decreasing function of population density. However, there is also reason to believe that  $A$  depends on the detection function parameter  $g(0)$ .  $A$  may be defined operationally as the area within which every animal is close enough to a detector that it is "counted" by the population ( $N$ ) estimator. For animals whose detection function declines gradually towards the edge of their range, this implies a fixed threshold  $p_T$  of individual detection probability for inclusion in  $N$ . Consider the effect of increasing the non-spatial parameter of the spatial detection function  $g(0)$  to say  $g(0)'$  while keeping  $\sigma$  constant. Animals outside  $A$  will now be counted (included in  $N$ ) because  $p'_{xy} > p_T$ .  $A$  will correspondingly increase to  $A'$  defined by the locus of points at which  $p'_{xy} = p_T$ . Our postulate that  $N$  estimators may be characterized by a threshold  $p_T$  is unproven. Nevertheless, the argument provides strong reason for doubting the common, if implicit, assumption that  $A$  is sensitive only to the scale of movement (home range size). "Abundance" as conventionally measured with passive detector arrays is confounded with variation in both the spatial and non-spatial components of detection, and with the size and configuration of detector arrays. The force of arguments against index methods (Anderson, 2003) should lead us also to reject  $\hat{N}$  as a surrogate for population density. The method described here allows researchers to overcome the problem by estimating density itself.

#### Pollock's robust design

It is conceptually simple to substitute  $\hat{D}_i$  for  $\hat{N}_i$  in the robust design of Pollock (1982). The appropriate unit for recruitment  $B_i$  is then animals per unit area per unit time. However, it is untidy to use a spatial parameterization ( $g(0)$ ,  $\sigma$ ) for detection probability in the closed population model and a non-spatial parameterization ( $p$ ) in the open population model. Further work is needed to determine whether it is beneficial and practical to incorporate a spatial detection function in the open population model used to estimate apparent survival ( $\phi_i$ ). We envisage modelling between-session home-range shifts in the open model, which would require at least one additional time-specific parameter.

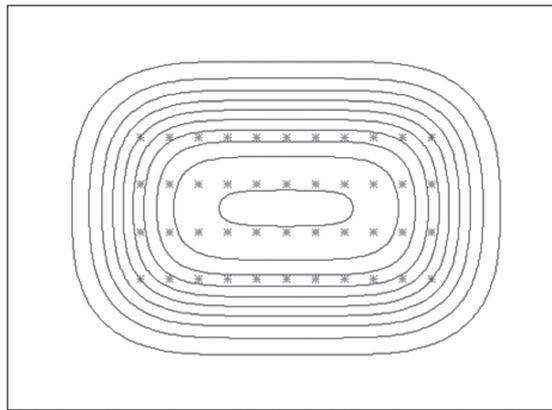


Fig. 5. Contours of daily array-level detection probability  $p_{xy}$  for common breeding birds at Patuxent Research Refuge, Maryland, U.S.A. Detection was assumed to follow a half-normal function with parameters equal to the estimated means for 1961–1972 ( $g(0) = 0.029$ ,  $\sigma = 87.3$  m). Contour levels range upwards in increments of 0.02 from 0.02 at the edge. Stars on the plot represent the locations of mist nets (vertical spacing 100 m).

*Fig. 5. Contornos de probabilidad de detección diaria  $p_{xy}$  de los distintos niveles de la batería de captura para aves comunes reproductoras en el Centro de Investigación Patuxent, Maryland, Estados Unidos. Se asumió que la detección seguía una función seminormal con parámetros iguales a los promedios estimados para 1961–1972 ( $g(0) = 0,029$ ,  $\sigma = 87,3$  m). Los niveles de contornos varían en sentido ascendente en incrementos de 0,02, desde 0,02 en el borde. Las estrellas que aparecen en la representación gráfica representan los lugares donde se colocaron las redes de niebla (distancia vertical de 100 m).*

#### Mist-netting to estimate the density of bird populations

The application of spatial detection functions in the analysis of capture–recapture data from mist nets is now briefly discussed. The main requirement for such analysis is that birds occupy equal-sized home ranges that are more or less stationary for the duration of a detection session. The effect of transients and of heterogeneous home range sizes on spatial detection estimates remains to be investigated. A related requirement is that recaptures within a detection session provide information on the spatial scale of movements. This can be met if the PDA is large enough to span several home ranges and either capture rates are high or netting continues for many occasions.

"Net shyness" is sometimes invoked to explain a declining trend in capture numbers during a mist-netting session (e.g. Swinebroad, 1964; MacArthur & MacArthur, 1974; Karr, 1981). Avoidance by birds of nets at which they have been captured previously has the potential to bias estimates of density by inverse prediction. The general occurrence of net shyness appears controversial, and alternatively may be explained as avoidance of areas of human activity (Murray, 1997). If the learned response is to a site-specific hazard rather than to the device itself, then a solution is to shift the location of nets part way through a detection session (e.g. Stamm

et al., 1960). Spatial data from such a design may be analysed in DENSITY by specifying the occasions on which detectors were operated at each location.

Finally it is noted that the fitted density and detection function for breeding birds at Patuxent (in round terms  $D = 10 \text{ ha}^{-1}$ ,  $g(0) = 0.03$ ,  $\sigma = 90$  m) is likely to be a good basis for simulations to optimize detector configurations in future field studies of similar species.

#### Acknowledgements

Our use of inverse prediction draws on previous work with Shirley Pledger, Andrew Tokeley and Dave Ramsey, for whose help we are most grateful. We thank Paul Lukacs, Gary White, and an anonymous referee for comments on a draft, David Fletcher for mathematical advice, and Christine Bezar for editing. Funding for MGE was provided by Landcare Research and the New Zealand Animal Health Board.

#### References

- Anderson, D. R., 2003. Index values rarely constitute reliable information. *Wildlife Society Bulletin*, 31: 288–291.

- Brown, P. J., 1982. Multivariate calibration. *Journal of the Royal Statistical Society, Series B*, 44: 287–321.
- Buckland, S. T., Burnham, K. P., Anderson, D. R. & Laake, J. L., 1993. *Density estimation using distance sampling*. Chapman and Hall, London.
- Burnham, K. P., 1981. Summarizing remarks: environmental influences. *Studies in Avian Biology*, 6: 324–325.
- Carothers, A. D., 1979. Quantifying unequal catchability and its effect on survival estimates in an actual population. *Journal of Animal Ecology*, 48: 863–869.
- Chao, A. & Huggins, R. M. (in press a). Classical closed population models. In: *The handbook of capture–recapture methods* (B. Manly, T. McDonald & S. Amstrup, Eds.). Princeton Univ. Press, Princeton.
- (in press b). Modern closed population models. In: *The handbook of capture–recapture methods* (B. Manly, T. McDonald & S. Amstrup, Eds.). Princeton Univ. Press, Princeton.
- Efford, M. G., 2004. Density estimation in live–trapping studies. *Oikos*, 106: 598–610.
- Farnsworth, G. L., Pollock, K. H., Nichols, J. D., Simons, T. R., Hines, J. D. & Sauer, J. R., 2002. A removal model for estimating detection probabilities from point–count surveys. *Auk*, 119: 414–425.
- Karr, J. R., 1981. Surveying birds with mist nets. *Studies in Avian Biology*, 6: 62–67.
- Lee, S.–M. & Chao, A., 1994. Estimating population size via sample coverage for closed capture–recapture models. *Biometrics*, 50: 88–97.
- MacArthur, R. H. & MacArthur, A. T., 1974. On the use of mist nets for population studies of birds. *Proceedings of the National Academy of Sciences*, 71: 3230–3233.
- MacKenzie, D. I. & Kendall, W. L., 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecology*, 83: 2387–2393.
- Murray, B. G. Jr., 1997. Net shyness in the Wood Thrush. *Journal of Field Ornithology*, 68: 348–357.
- Nichols, J. D., Hines, J. E., Sauer, J. R., Fallon, F. W., Fallon, J. E., & Heglund, P. J., 2000. A double observer approach for estimating detection probability and abundance from point counts. *Auk*, 117: 393–408.
- Noon, B. R. & Sauer, J. R., 1990. Population models for passerine birds: structure, parameterization, and analysis. In: *Wildlife 2001: populations*: 441–464 (D. R. McCullough & R. H. Barrett, Eds.). Elsevier Applied Science, London.
- Otis, D. L., Burnham, K. P., White, G. C. & Anderson, D. R., 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, No. 62.
- Pledger, S. & Efford, M., 1998. Correction of bias due to heterogeneous capture probability in capture–recapture studies of open populations. *Biometrics*, 54: 888–898.
- Pollock, K. H., 1982. A capture–recapture design robust to unequal probability of capture. *Journal of Wildlife Management*, 46: 752–757.
- Pollock, K. H., Nichols, J. D., Simons, T. R., Farnsworth, G. L., Bailey, L. L. & Sauer, J. R., 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics*, 13: 105–119.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T., 1989. *Numerical recipes in Pascal: the art of scientific computing*. Cambridge Univ. Press, Cambridge.
- Rosenstock, S. S., Anderson, D. R., Giesen, K. M., Leukering, T. & Carter, M. F., 2002. Landbird counting techniques: current practices and an alternative. *Auk*, 119: 46–53.
- Stamm, D. D., Davis, D. E. & Robbins, C. S., 1960. A method of studying wild bird populations by mist–netting and banding. *Bird–Banding*, 31: 115–130.
- Swinebroad, J., 1964. Net–shyness and Wood Thrush populations. *Bird–Banding*, 35: 196–202.
- Thompson, W. L., 2002. Towards reliable bird surveys: accounting for individuals present but not detected. *Auk*, 119: 18–25.

Appendix. Automatic calculation of initial values for inverse prediction search algorithm.

*Apéndice. Cálculo automático de los valores iniciales para buscar el algoritmo de predicción inversa.*

Initial values of the three parameters  $D$ ,  $g(0)$ ,  $\sigma$  required for inverse prediction are here denoted by subscript  $S$ . Good initial values often make the difference between a speedy, fruitful search and failure. Calculations in DENSITY use the simplifying assumption of negligible competition among animals for detectors. Calculation of  $\sigma_S$  and  $g(0)_S$  uses Monte Carlo integration; accuracy depends on the number of random points sampled within the nominal detection area  $A$ . The user may vary the sampling intensity as an option in DENSITY.

Initial detection scale  $\sigma_S$

The expected distance between recaptures may be inferred from  $\sigma$  and the detector configuration:

$$E(\bar{d}) = \frac{\int_A \sum_i \sum_j R_{ij} p_i p_j dA}{\int_A \sum_i \sum_j p_i p_j dA} \quad (1)$$

where the indices  $i$  and  $j$  refer to traps,  $p_i$  is the "naïve" probability of an animal located somewhere within area  $A$  being caught in trap  $i$ , and  $R_{ij}$  is the distance between traps  $i$  and  $j$ .

With a half-normal detection model

$$p_i = g(0) \exp(-r_i^2 / (2\sigma^2)) \quad (2)$$

where  $r_i$  is the distance between an animal's range centre at  $x,y$  and trap  $i$ .

The integrals are evaluated by sampling points  $x,y$  within an area  $A$ . The area  $A$  is limited to locations where animals are "detectable" by some criterion (e.g.  $P > 0.01$ , see below). A factor of  $g(0)^2$  appears in both numerator and denominator of equation 1, and cancels out.

Numerical minimization (the "golden" routine of Press et al., 1989) is used to find the value of  $\sigma$  for which  $E(\bar{d})$  matches the observed  $\bar{d}$  for the given detector configuration. Evaluation of equation (1) is time-consuming ( $O(T^2)$  where  $T$  is the number of traps). Only the lower triangle and diagonal of the symmetric  $R_{ij} p_i p_j$  matrix need be evaluated.

Initial core detection probability  $g(0)_S$

A similar but faster approach is used to obtain  $g(0)_S$ . Given a value for  $\sigma$  equation (2) may be used to estimate the "naïve" probability that an animal is caught somewhere in the PDA:

$$P = 1 - \prod_i (1 - p_i)$$

To what observable quantity should we relate  $g(0)$ ? We propose the mean number of captures within a detection session, conditional on an animal having been caught once:

$$\bar{C} = \frac{\sum_t n_t}{M_{t+1}}$$

This may be an unreliable indicator of  $g(0)$  when there is a large "learned trap response" (i.e.  $c \ll p$  or  $c \gg p$  in the notation of Otis et al., 1978). However, it has the advantage of not requiring an estimate of  $N$  or  $D$ .

$$E(C) = 1/A \int_A \frac{tP}{1 - (1-P)^t} dA \quad (3)$$

where  $t$  is the number of capture occasions and  $A$  is the area within which  $P > 0$ .

The procedure is again to sample  $P$  from  $A$  and use numerical minimization to find the value of  $g(0)$  for which  $E(\bar{C})$  equals the observed  $\bar{C}$ , given  $\sigma_S$ ,  $t$  and the detector configuration.

Appendix. (Cont.)

#### Initial DENSITY $D_s$

When  $D$  has been estimated by inverse prediction it is possible retrospectively to infer the boundary strip width  $W$  (e.g. Otis et al., 1978) that, if applied as a buffer around the PDA, would have yielded the "correct" density. Inferred values of  $W$  typically show a quadratic relationship to  $\sigma$  (unpubl. results). We base our initial estimate of density on this relationship as follows:

$$D_s = \frac{\hat{N}}{A_w}$$

where  $A_w$  is the effective area corresponding to  $W$ . A quadratic is used to predict  $W$ :

$$W = a\sigma_s^2 + b\sigma_s + c$$

The default polynomial coefficients in DENSITY are  $a = 0$ ,  $b = 2$ ,  $c = 0$  (i.e.  $W = 2\sigma_s$ ). Coefficients may be changed by the user if experience with a particular species and detector configuration provides more information.

The relationship between  $\sigma$  and  $W$  appears to vary with the properties of the estimator for  $\hat{N}$ , and also with the detector configuration and possibly the duration of the study. Estimators that are robust to individual heterogeneity ( $M_h$ ) tend to yield larger  $\hat{N}$ , and therefore require a smaller  $W$  to yield the same  $\hat{D}$ .