A shortcut to obtain reliable estimations of detectability in extensive multispecific census programs.

Luis M. Carrascal*[1], **Javier Seoane**** & **Vicente Polo*****
* Dept. Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain.
** Dep.. Ecología, Univ. Autónoma, Madrid, Spain.
*** Univ. Rey Juan Carlos, Getafe, Madrid
1 lmcarrascal@mncn.csic.es

The problem

Obtaining reliable estimates of animal abundance and density in field studies is a main concern for ecologists and conservation biologists, who often use distance-based sampling methods that rely on measuring distances to contacts accurately. These methods calculate the effective strip width (ESW), which is the distance within one can assume is doing a complete census (vs. a sampling).

However, when the study is aimed at an assemblage of numerous species and the field work is collaborative –thus involving a lot of people with different degree of expertise– measuring exact distances to each contact is unfeasible. Moreover, the estimation of the perpendicular distance of each bird to the observer's trajectory is not always exact, because devices such as laser range-finder cannot be applied precisely to birds heard but not seen, or in habitats with dense vegetation cover.

The goal

We propose a new method for estimation of reliable effective strip width to estimate absolute densities, based upon geometric and mathematical properties of distributions describing distances to bird contacts in line transects: the threshold method with a fixed belt **d**. We use a mathematic (integration calculus), simulation and empirical approach and compare the results provided by this method with those obtained by DISTANCE.

Material and methods

Census method

We censused 46 bird species in different environments of Madrid (pinewoods, 24 km), Gran Canaria (pinewoods, 60 km) and Fuerteventura (arid habitats, 592 km) islands. Breeding bird surveys were carried out in 2005, 2006, 2008 and 2009.

The survey method was the line transect, frequently used in extensive assessments of abundance, general distribution patterns and habitat preferences of birds (Bibby et al., 2000).

The transects were carried out on windless and rainless days, walking cross country or by little used dirt tracks at a low speed (1-3 km/h approximately), during the 4 hours after dawn and the 2.5 hours before dusk. For each detected bird, the perpendicular distance to the observer's trajectory was estimated (a few overflying birds sighted were disregarded). Training with a laser range-finder (Leica Rangemaster LRF 900) helped to improve distance estimates and to reduce inter-observer variability.

Detectability estimation by means of distance sampling

Detectabilities were estimated with distance sampling methods (Thomas *et al.*, 2002). For modeling the detectability, we fitted three canonical models (half-normal, negative exponential and hazard-rate, trying in each one to include a suitable series expansion –cosine or polynomial-) that are commonly used to explain the loss of detectability as a function of the distance from the transect line (the further the distance the lower the probability of detecting an individual).

These models were used to estimate the probability of detection and the effective census strip width (ESW). Models were evaluated according to AICc. We calculated a weighted average of the detection probabilities derived from the models according to weights obtained from AICc values (Burnham & Anderson, 2002). Detectability models were built with Distance 5.0 software (Thomas et al., 2004).

Detectability estimation by means of distance sampling

Detectabilities were estimated with distance sampling methods (Thomas *et al.*, 2002). For modeling the detectability, we fitted three canonical models (half-normal, negative exponential and hazard-rate, trying in each one to include a suitable series expansion –cosine or polynomial-) that are commonly used to explain the loss of detectability as a function of the distance from the transect line (the further the distance the lower the probability of detecting an individual).

These models were used to estimate the probability of detection and the effective strip width (ESW). Models were evaluated according to AICc. We calculated a weighted average of the detection probabilities derived from the models according to weights obtained from AICc values (Burnham & Anderson, 2002). Detectability models were built with Distance 5.0 software (Thomas et al., 2004).

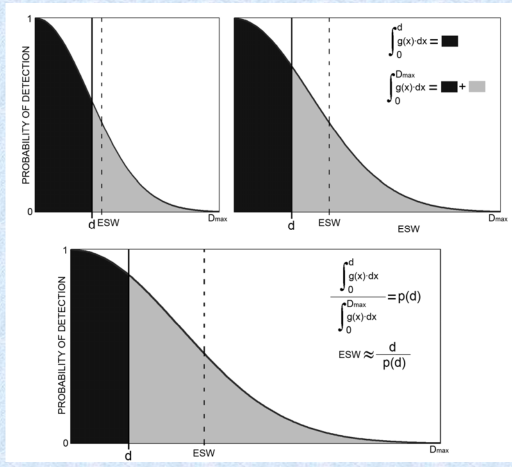
Simulations of distance data

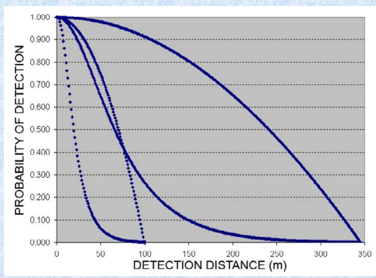
We produced simulated distance data using 550 different distributions. Several distribution functions have been used, including uniform with cosine adjustment, half-normal, negative exponential, hazard rate and double catenary equation. The distributions generated have very large sample sizes (2,000 to 25,000 distances) and cover a wide range of maximum detection distances (35 to 405 m perpendicular to the line transect). They retain the enormous variation in detectability functions ussually found under natural conditions while censusing birds (species, seasons, habitat structure, ...).

Some formulae ... graphically

For any detectability function ***g(x)*** there exists a relationship between the ratio **d/p(d)** and the effective strip width (**ESW),** being **d** the threshold distance and **p(d)** the proportion of contacts detected within that threshold [**ESW =d/p(d)**).

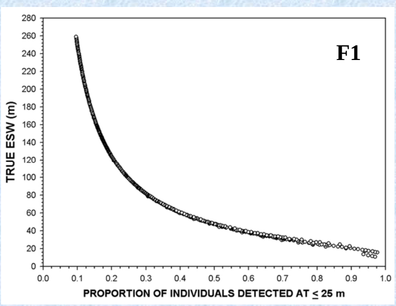
By integral calculus, it can be demonstrated that for any detectability function ***g(x)***, **ESW = d/p(d) + *t*)**, where ***t*** is a complex term depending on the parameters describing the detectability function ***g(x)***.

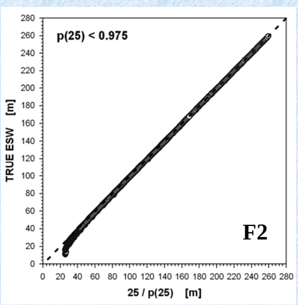


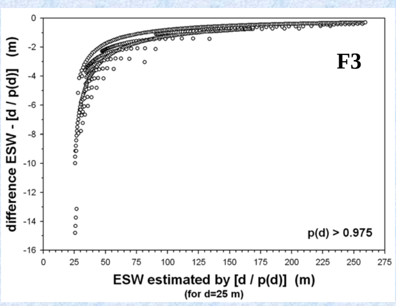


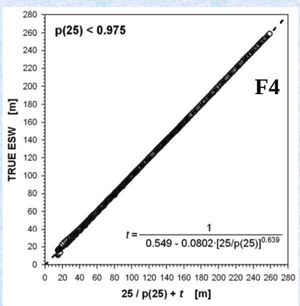
Example of extreme distribution functions used in this poster.

Simulated data









There exists a perfect linear relationship between the ratio **d/p(d)** and the effective strip width (ESW)**,** being **d** the threshold distance and **p(d)** the proportion of contacts detected within that threshold (**ESW =d/p(d)**; R^2>0.99 for **d** = 5, 10, 25, 50 and 100 m, and given that **p(d)**<0.975).

Nevertheless, the addition of a term ***t***, obtained after the integration of distribution functions, provide perfect predictions of the effective strip width (**ESW = d/p(d) + *t***), especially when using large threshold distances **d** and large **p(d)** proportions >0.5 (R^2>0.9999 in all models)

The equations for thresholds distances 5 to 100 m are shown below:

d = 5 m,	$ESW = [5/p(5)] + 1 / (9.459 - 1.2606\cdot[5/p(5)]^{0.870})$	(… butterflies)
d = 10 m,	$ESW = [10/p(10)] + 1 / (1.978 - 0.3296\cdot[10/p(10)]^{0.735})$	(… lizards)
d = 25 m,	$ESW = [25/p(25)] + 1 / (0.549 - 0.0802\cdot[25/p(25)]^{0.639})$	(… many birds)
d = 50 m,	$ESW = [50/p(50)] + 1 / (0.158 - 0.0110\cdot[50/p(50)]^{0.758})$	(… cuckoos, orioles)
d = 100 m,	$ESW = [100/p(100)] + 1 / (0.056 - 0.0010\cdot[100/p(100)]^{0.955})$	(… birds of prey)

F1- Relationships between the true ESWs and the proportion of contacts detected within belts of 25 m at both sides of the observer (i.e., p(25)) for 550 distributions of 2,000 to 25,000 distances.

F2 - Relationships between the true ESWs and the ratios **d/p(d)** between the threshold distance **d** and the proportions of contacts detected within those thresholds (**p(d)**) for the threshold of 25 m at both sides of the observer. The dashed line show the perfect fit between both estimations. The proportions **p(25)** are lower than 0.975.

F3 - Relationship between the estimation of ESW using the equation of the threshold method [d/p(d); dots in F2] and the difference between the this estimate and the true ESW (dashed line in F2).

F4 - Relationship between true (ESW) and predicted (ESW') effective strip widths in line transects with threshold distance **d** = 25 m. **p(25)**: proportion of contacts detected within belts of 25 m at both sides of the observer. ESW' = **d/p(d) + *t***. The differences between true and estimated effective strip width (ESW minus **d/p(d)**) are tightly related using a Harris model where:

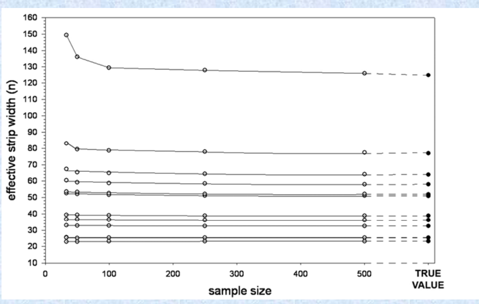
$$ESW - d/p(d) = 1 / (a + b\cdot[d/p(d)]^c) \approx t$$

Effects of sample size on effective strip width estimation

The Figure shows the variation of the estimated ESWs using the **threshold method with a fixed belt of 25 m** according to sample size. When true ESW are lower than 50 m, the average estimated ESWs are nearly identical to true ESWs even under low sample sizes of 33 or 50 distances. Estimated ESWs tend to be slightly larger than true ESWs when the true value is between 50 and 100 m under low sample sizes of less than 100 distances. This trend is accentuated when the true ESW is larger than 100 m at low sample sizes. The correlation (r) between true ESW and average estimated ESW is 0.975 when sample size (n) is 33, r=0.983 when n=50, r=0.991 when n=100, r=0.997 when n=250, and r=0.998 when n=500.

Similar results are observed when working with the **threshold method with a fixed belt of 50 m**, although less skewed estimations, and with narrower confidence intervals, are obtained for distributions with larger true ESW. For example, for the distribution with a true ESW=124.8 m, an average figure of 135.0 m and a relative spread of 86.1% around the true ESW are obtained with sample sizes of 33 distances, values more accurate than the previously observed of 149.4 m and 271% of relative variation of estimated ESWs around the true average using a fixed belt of 25 m.

Summarizing, the calculus of the effective strip width using the threshold method with a fixed belt of 25 m provides very accurate and stable results, with narrow confidence intervals, even under low sample sizes of census distances. Only large sample sizes (i.e., > 100 distances) are needed to estimate accurately the effective strip width in those more detectable species (ESW > 100 m).



Convergence between ESWs obtained with DISTANCE and with the threshold method with a fixed belt of 25 m

Simulated data

The Table shows the results of estimations of census strip widths for six virtual species using two different methodological approaches: threshold method with a fixed belt of 25 m (ESW d/p[d]) and the regression approach implemented by DISTANCE. The distributions of known parameters generated for the six species have very large sample sizes (2,779 to 14,954 distances) and cover a wide range of maximum detection distances (50 to 251 m perpendicular to the line transect).

Estimated and true strip widths are highly correlated in both methods (r=0.99999, n=6 virtual species).

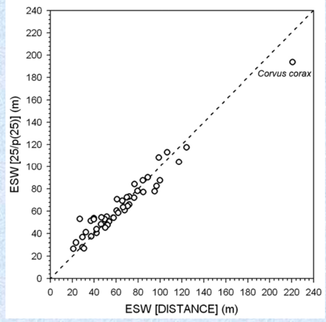
Both the threshold method with a fixed belt, and the DISTANCE regression approach, produce very accurate strip census width estimations, although DISTANCE figures are slightly more skewed towards larger figures, but have relatively narrower confidence intervals for more detectable species, than the threshold method with a fixed belt. Nevertheless, bear in mind that with DISTANCE the exact value of detection distances with all contacts are required, while with the threshold method with a fixed belt of 25 m only the proportion of those detected at ≤ 25 m are needed.

	SP1	SP2	SP3	SP4	SP5	SP6
N	2,779	4,794	6,421	9,541	12,430	14,954
Dmax	50	75	117	149	201	251
TRUE SW	28.1	48.3	64.8	96.1	125.3	150.9
p(25)	0.757	0.496	0.375	0.258	0.198	0.165
ESW d/p(d)						
estimate	27.7	48.0	65.1	96.0	125.2	150.6
LCI 95%	26.7	46.5	63.0	92.7	120.8	145.4
UCI 95%	28.7	49.6	67.4	99.4	129.8	156.3
DISTANCE						
estimate	29.5	49.8	66.7	97.9	127.5	153.5
LCI 95%	28.4	48.2	64.7	95.7	124.9	150.6
UCI 95%	30.7	51.5	68.8	100.1	130.1	156.5
	HNp	HNp	HNp	HNp	HNp	HNp
	HRp	HRp	HRp	HRp	HRp	HRp
	NEp	NEp	NEp	NEp	NEp	NEp
	HRc	NEc	NEc	HRc	HRc	NEc
	NEc	HRc	HRc	NEc	NEc	HRc
	HNc	HNc	HNc	HNc	HNc	HNc

Field data

The Figure shows the relationship between the estimates of ESWs for 46 bird species using the regression approach implemented by DISTANCE and the threshold method with a fixed belt of 25 m (see the table for more details on sample sizes and confidence intervals of ESW estimates). Both estimations of ESW distances are highly correlated (R^2=94.4%, p<<0.001) and very similar. One-way repeated measures ANOVA provides non-significant differences between estimations derived from both methods ($F_{1,45}$=0.013, p=0.911). Applying the more convenient fixed belt of 100 m in the threshold method for the most detectable species with the longest ESW (*Corvus corax*), a very similar ESW is obtained compared with that provided by DISTANCE: 216.4 m.

In summary, working with 'normal' sample sizes ranging from 30 to 3000 contacts, the threshold method with a fixed belt of 25 m provides nearly identical estimations of ESW to those obtained with DISTANCE regression approach. The bonus is that you have only to define if contacts with birds are shorter or longer than a previously defined threshold distance. You have only to train to that distance.



		THRESHOLD METHOD			DISTANCE		
	N	estimate	95% CI		estimate	95% CI	
Fuerteventura arid environments							
Alectoris barbara	238	87.6	72.3	110.6	99.9	89.9	111.2
Anthus berthelotii	3217	60.4	57.8	63.2	60.4	58.7	62.1
Bucanetes gitaginea	801	37.3	34.8	40.0	37.4	35.1	39.9
Burhinus oedicnemus	172	53.1	44.7	64.4	26.8	18.5	39.2
Calandrella rufescens	3545	55.0	52.9	57.4	51.3	48.7	54.2
Carduelis cannabina	285	40.5	35.9	45.8	42.0	37.8	46.7
Chlamydotis undulata	67	82.5	59.1	133.0	96.9	71.3	133.5
Columba livia	145	52.8	43.9	65.3	49.8	40.1	61.8
Corvus corax	101	193.6	127.5	399.0	220.9	176.9	276.9
Coturnix coturnix	191	51.5	43.9	61.6	37.0	30.8	44.5
Cursorius cursor	69	53.6	41.1	74.1	39.7	31.0	50.7
Cyanistes teneriffae	95	36.8	30.0	45.3	29.2	23.5	36.4
Falco tinunculus	146	113.1	86.0	164.3	106.5	88.4	128.3
Lanius meridionalis	541	66.0	59.1	74.6	71.9	66.4	77.8
Miliaria calandra	59	104.4	70.3	197.0	117.2	91.4	151.5
Passer hispaniolensis	661	41.3	38.2	44.8	32.3	28.5	36.5
Pterocles orientalis	165	90.5	71.8	121.5	88.8	73.9	106.9
Saxicola dacotiae	413	43.8	39.6	48.8	42.0	38.5	45.8
Streptopelia decaocto	225	28.0	24.5	31.7	29.4	26.0	33.4
Streptopelia turtur	115	32.2	26.7	38.3	23.5	18.8	29.4
Sylvia conspicillata	961	50.8	47.3	54.9	53.4	50.7	56.2
Sylvia melanocephala	175	26.8	22.8	30.7	21.1	17.9	24.9
Upupa epops	298	52.7	46.2	60.8	39.8	31.7	50.1
Gran Canaria pinewoods							
Anthus berthelotii	201	108.3	85.8	146.0	99.1	87.4	112.6
Columba livia	53	64.6	46.4	101.6	70.4	54.3	92.1
Cyanistes teneriffae	333	48.1	42.7	54.6	51.8	47.0	57.0
Dendrocopos major	89	78.1	58.8	114.4	79.4	63.5	100.0
Erithacus rubecula	75	60.7	46.3	85.8	67.8	55.9	82.5
Fringilla coelebs	47	72.0	49.9	123.3	76.2	54.7	107.0
Fringilla teydea	32	87.7	54.4	209.3	84.2	60.9	117.5
Phylloscopus canariensis	411	54.1	48.3	61.2	57.7	51.2	65.3
Serinus canaria	322	73.1	62.9	87.0	71.9	57.6	90.4
Streptopelia turtur	90	117.5	83.0	198.6	124.4	103.9	149.1
Turdus merula	62	72.4	52.4	113.5	70.0	53.2	92.3
Madrid pinewoods							
Certhia brachydactyla	111	48.2	39.4	60.5	46.2	37.8	56.9
Columba palumbus	38	77.9	51.2	152.7	95.1	68.5	133.4
Erithacus rubecula	68	69.3	51.2	104.4	65.8	51.2	85.4
Fringilla coelebs	183	58.4	49.1	71.4	62.0	50.8	76.6
Lullula arborea	45	54.3	39.1	83.5	46.6	32.3	68.7
Lophophanes cristatus	98	47.7	38.5	60.8	52.3	40.8	67.5
Parus major	52	70.8	50.0	116.1	60.7	46.5	79.6
Periparus ater	34	63.8	42.6	116.6	66.5	52.6	84.6
Serinus serinus	88	45.3	36.4	58.1	50.0	40.9	61.3
Sylvia cantillans	49	27.0	19.9	34.5	30.8	20.9	46.3
Turdus merula	65	84.3	60.0	138.4	76.8	59.8	99.0
Turdus viscivorus	44	77.3	52.3	140.8	84.8	66.8	108.0