

## Indirectly accounting for uneven sampling effort to estimate population sizes from atlas data

NICOLAS TITEUX<sup>1,2</sup>, ANTOINE DEROUAUX<sup>3</sup>, JEAN-PAUL JACOB<sup>4</sup>, THIERRY KINET<sup>4</sup>,  
JEAN-YVES PAQUET<sup>4</sup>

<sup>1</sup> Université catholique de Louvain, Biodiversity Research Centre - Croix du Sud 4-5, B-1348 Louvain-la-Neuve, Belgium

<sup>2</sup> Public Research Centre - Gabriel Lippmann, Department Environment and Agro-biotechnologies - Rue du Brill 41, L-4422 Belvaux, Luxembourg (titeux@lippmann.lu)

<sup>3</sup> Aves-Natagora, Centrale Ornithologique Aves - Rue Fusch, 3, B-4000 Liège, Belgium

<sup>4</sup> Aves-Natagora, Département Études - Rue du Wisconsin 3, B-5000 Namur, Belgium

**Abstract** – One important by-product of bird atlases is to provide population estimates to evaluate species conservation status. Unfortunately, large-scale atlas data are nearly always characterized by an uneven sampling effort among sampling units. This prevents from calculating straightforwardly reliable population estimates. We propose a procedure to deal with this unevenness when generating regional population estimates. The *Breeding Bird Atlas of Wallonia 2001-2007* (Southern Belgium) was used to illustrate our method. Data were collected according to a grid composed of 508 sampling units. Each unit was assigned to one or several observers who gathered information on species abundance and breeding status according to an unrestricted procedure. In addition, a series of sub-sampling units were surveyed within each sampling unit to collect information according to a standardized procedure. Based on the data collected according to the unrestricted surveys in the sampling units and the corresponding standardized surveys in the sub-sampling units, we calculated three indices representing different aspects of the sampling effort at the sampling unit level. Those indices are not sensitive to the natural disparity in the nature/diversity of the habitats among the sampling units. The indices were then combined into a single score that was used to weight the contribution of each sampling unit to the calculation of the regional population estimates. Depending on the species, the consideration of the unevenness in the sampling effort led to considerable discrepancies in the population estimates.

### INTRODUCTION

The mapping of bird distributions using grid-based atlases at different spatial scales is one of the most frequent form of ornithological survey and is based on the coordinated effort of dedicated volunteers in most cases (Hagemeijer and Blair 1997, Gibbons *et al.* 2007). These atlases collect data on species occurrence or abundance from a continuous grid composed of equal-sized sampling units to cover the whole region that is under examination (Bibby *et al.* 1992).

Atlas data offer many potential avenues of investigation (reviewed by Donald and Fuller 1998), including the estimation of regional population sizes (Robertson *et al.* 1995, Estrada *et al.* 2004). These estimates are relevant for descriptive purposes, for population ecology studies, but also for the evaluation of species conservation status and for the implementation of conservation strategies (Robertson *et al.* 1995).

This implies to rely on local information on species

abundance within each sampling unit. In modern atlases (Hustings and Vergeer 2002, Estrada *et al.* 2004), observers are asked to evaluate the abundance of each species in the sampling units (local estimates). Population estimates for the whole region that is covered by the atlas project can next be evaluated when the local estimates are combined. Nevertheless, important variation in the amount of effort expended by the manifold observers covering the different sampling units could prejudice the reliability of the regional population estimates (Vaisanen 1998, Link and Sauer 1999). Besides variations in the time spent in the field, significant disparity in identification skills and field experience among observers may determine important biases (Sauer *et al.* 1994). Accordingly, heterogeneous coverage of data is an inherent limitation that confounds large-scale atlas data (Donald and Fuller 1998, Link and Sauer 1998). Regional population estimates would therefore greatly benefit from making use of methodologies able to mitigate these limitations (Link and Sauer 1999), with

a view to differentiating true species abundance patterns from artefacts.

Most existing approaches use local estimates that have been scaled by sampling effort, ranging from simple division of local estimates by effort (Root and McDaniel 1995) to complex parametric adjustment of the relationship between local estimates and effort (Link and Sauer 1999). These methods rely either on directly known effective sampling time, when reported (Link and Sauer 1999), or on indirect estimation of sampling effort (Vaisanen 1998, Telfer *et al.* 2002, Estrada *et al.* 2004).

Here, we propose a procedure to compute regional population estimates by controlling for the variation in the effective sampling effort, when the latter is not reported directly by the observers. We evaluate and discuss to what extent the procedure is able to counterbalance the unevenness in the sampling effort. The data collected between 2001 and 2007 for the *Breeding Bird Atlas of Wallonia* (Southern Belgium) (Jacob 2000) were used to illustrate our method.

## METHODS

### Study area and data collection

The grid-based *Breeding Bird Atlas of Wallonia 2001-2007* covers the Southern part of Belgium and is composed of 508 fundamental sampling units of 40 km<sup>2</sup> (5x8 km). The fieldwork was conducted by more than 350 amateur ornithologists.

Each sampling unit was assigned to one or several volunteers who surveyed all habitats during the whole breeding season. They collected information on species abundance and reproduction with no standardized procedure (see ‘Unrestricted surveys of sampling units’). In addition, several 1x1 km sub-sampling units were selected within the sampling units, according to a regular systematic sampling procedure, to reach a 17% coverage of the region (see details in Jacob 2000). The sub-sampling units were surveyed according to a standardized procedure (see ‘Timed surveys of sub-sampling units’) by the same observers as the corresponding unrestricted survey in most cases.

### Unrestricted surveys of sampling units

The observers were asked to provide the highest possible breeding evidence for each species according to the EOAC classification, i.e. non-breeding, possible, probable and confirmed breeding (Timothy and Sharrock 1974, Hagemeyer and Blair 1997). Non-breeding species were excluded for the subsequent analyses. The local population estimates were supplied by observers in the form of 9

abundance classes derived from a geometric progression with a common ratio set to 2 (see Tab. 1). Three values were retained to describe the local population estimates: lower, central, and upper values of the abundance classes.

### Timed surveys of sub-sampling units

Within the different sub-sampling units, walked itineraries were delineated by the observers, so that the itineraries covered the diversity of broad habitats in representative proportions. The observers walked along the itineraries twice a year (March-April/May-June), at low speed (1-hour walk in total) and in the first 4 hours after sunrise. All encountered individuals of all species (detected either by sight or by sound) were recorded and counted (Schmid *et al.* 1998, Estrada *et al.* 2004, Schmid 2008 for similar approaches).

### Estimation of the effective sampling effort in the sampling units

A score was assigned to the different sampling units that were completed in early 2007 (N = 334) in order to represent the variations in the effective sampling effort put by the observers. This score was based on the computation of three indices based on three corresponding criteria that accounted for different aspects of the coverage quality. The species that are considered as strictly linked to lakes and ponds (i.e. *anatidae*, grebes, coots and *laridae*) were removed for the subsequent analysis, since variations in the presence or abundance of those habitats among adjacent sampling units or among sub-sampling units within sampling units would lead to significant discrepancies in the indices calculated below, irrespective of the sampling effort.

1. *Index of richness variation between unrestricted and timed surveys (RS)*. The number of species reported for a given sampling unit depended mostly on the nature/diversity of the habitats, but also on the sampling effort. The nature/diversity of habitats within the sampling units were assumed to be proportional to the number of species encountered during the timed surveys (Estrada *et al.* 2004). For each sampling unit, we have therefore computed the ratio (on a logarithmic scale) between the number of species reported from the unrestricted survey ( $N_{Unrestricted}$ ) and the total number of species reported from the corresponding timed surveys ( $N_{Timed}$ ), as follows:

$$RS = \exp\left(\frac{N_{Unrestricted}}{N_{Timed}}\right)$$

This allowed controlling for the inherent disparity in

**Table 1.** Abundance classes (number of breeding pairs) as counted at the sampling unit level and used by the *Breeding Bird Atlas of Wallonia 2001-2007*.

Class	Boundaries	Central value
A	1 - 5	3
B	6 - 10	8
C	11 - 20	15
D	21 - 40	30
E	41 - 80	60
F	81 - 160	120
G	161 - 320	240
H	321 - 640	480
I	> 640	960

the nature/diversity of the habitats among sampling units (Estrada *et al.* 2004) and informed us on the amount of effort put by the observers in the unrestricted surveys, assuming that their detection skills did not differ between unrestricted and timed surveys.

2. *Index of neighbourhood similarity (NS)*. For each sampling unit, the number of species reported from the unrestricted survey ( $N_{Unrestricted}$ ) was compared to the mean number of species reported from the unrestricted surveys in the  $k$  adjacent sampling units ( $N_{Unrestricted(k)}$ ), as follows:

$$NS = N_{Unrestricted} - \frac{\sum_1^k N_{Unrestricted(k)}}{k}$$

where  $k = 8$  except for border sampling units.

The *NS* index highlighted the fine-scale artefactual variations in the sampling effort, by pointing out the suspiciously poorly surveyed units in a complementary way as compared with the *RS* index.

3. *Index of confirmation of breeding (CB)*. The relative proportions ( $p_j$ ) of possible ( $j = 1$ ), probable ( $j = 2$ ) and confirmed ( $j = 3$ ) breeders were calculated for each sampling unit. Corresponding weighting coefficients ( $w_j$ ), ar-

bitrarily set at 1, 2 and 4 respectively, were attributed to these breeding evidences, so as to represent the average effort needed to collect them. For each sampling unit, the weighted sum of the  $p_j$  proportions was then computed as follows:

$$CB = \sum_{j=1}^3 w_j \times p_j$$

The *CB* index was used to evaluate the attention paid and the effort made to collect breeding evidence for all the species (Vaisanen 1998).

#### *The effective sampling effort score (S)*

The three indices were standardized between 0 and 1 (by subtracting the minimum value of the index among the whole sampling units and then by dividing this difference by the range between the minimum and maximum values of the index) to make their relative contributions to the final score comparable. For each sampling unit, the standardized indices were then combined into a single multi-criterial score (*S*) using their geometric mean. This resulting score represented indirectly the effective sampling effort for each sampling unit (including time/effort spent collecting data in the field and skills/experience of the observers).

#### *Regional population estimates*

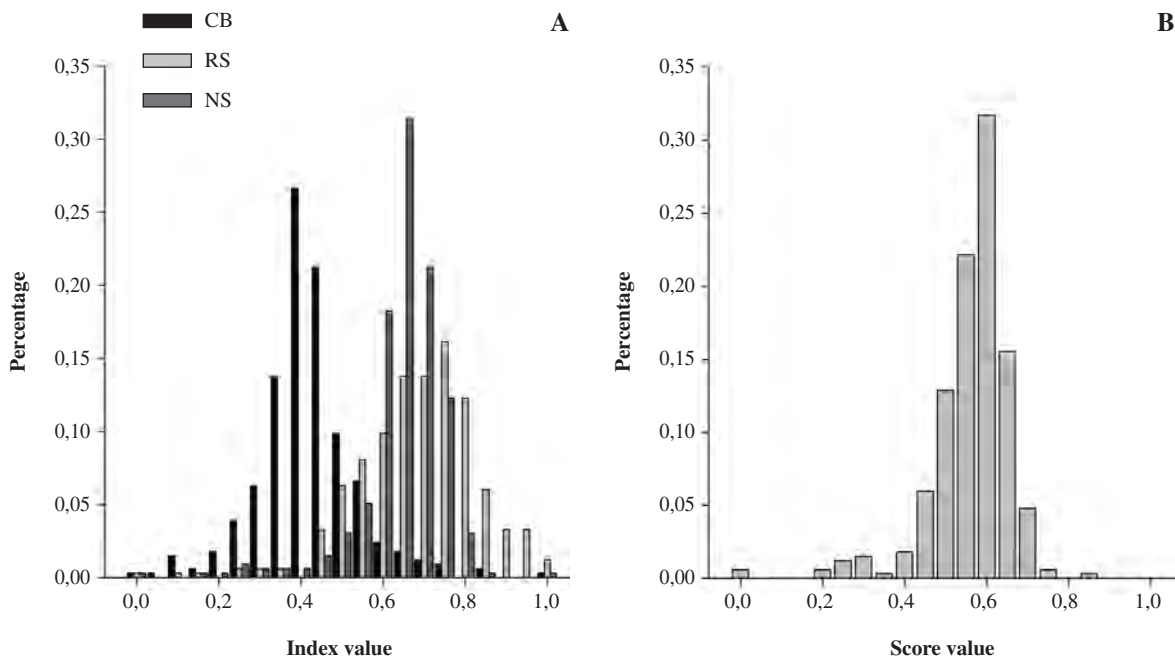
For each species, the mean abundance among the sampling units was calculated as the weighted mean of the different local estimates given by the observers (lower, central and upper values in turn), using the sampling effort score *S* as a weighting coefficient. Multiplying these mean local estimates by the number of sampling units provided the lower, central and upper weighted regional population estimates (*WPE*). For the comparison exercise, the non-weighted population estimates (*NWPE*) were also calculated as the sum of the different local estimates given by the observers in the sampling units, irrespective of the estimated sampling effort score *S*. The difference between the lower and upper regional estimates defined the regional estimation range for each species.

## RESULTS

The individual indices (*RS*, *NS* and *CB*) and the overall sampling effort score (*S*) exhibited a bell-shaped distribution (Fig. 1). A few sampling units were off-centred as compared with the tight distribution of the score, which means that a small number of sampling units were either very well or very poorly surveyed as compared with the mean coverage quality. Although the three indices were

**Table 2.** Pairwise Spearman correlation tests for each pair of indices (*RS*, *NS* and *CB*).

Pair	Spearman rho	p-value	N
NS-RS	0.24	< 0.001	334
CB-NS	0.29	< 0.001	334
CB-RS	0.12	0.024	334



**Figure 1.** Frequency distribution of the individual indices (*CB*, *RS* and *NS*) and the overall sampling effort score (*S*).

significantly correlated, fairly low correlation coefficients pointed out poor collinearity among them (Tab. 2 and Fig. 2) and indicated that the indices stressed different aspects of data-covering.

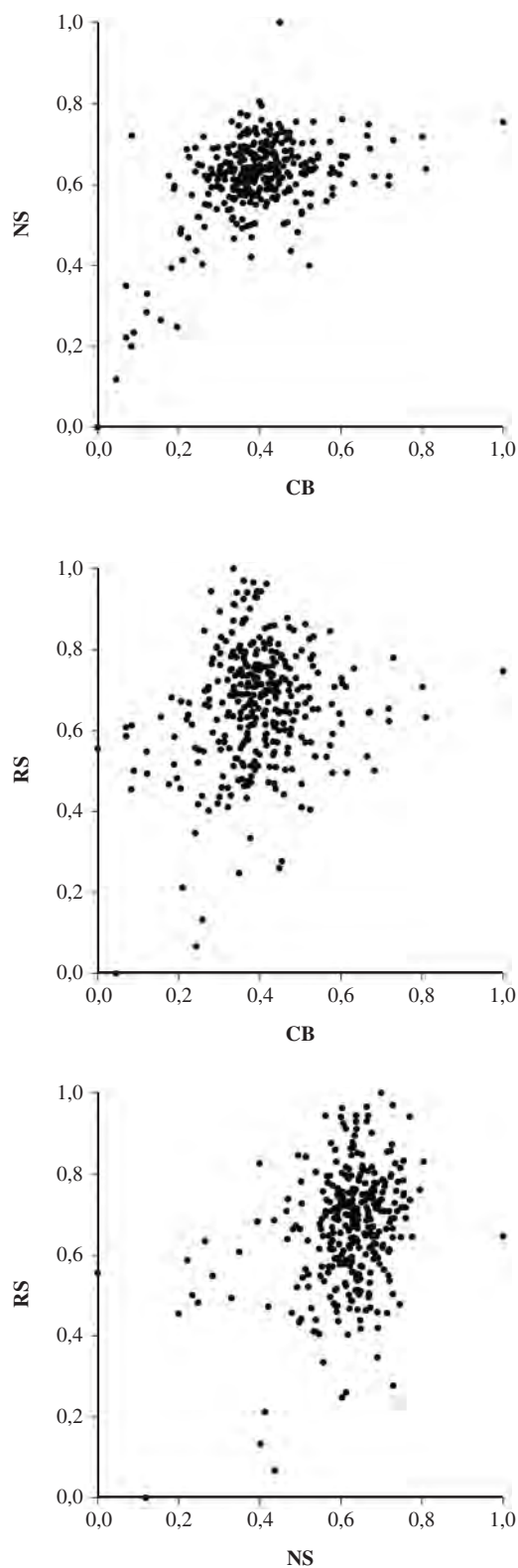
The lower, central and upper regional estimates were significantly higher (paired comparisons using Wilcoxon signed rank test,  $W = 257.5$  for the central value,  $N = 153$ ,  $p < 0.0001$ ) when rescaling the relative contribution of the sampling units proportionally to the sampling effort (*WPE*) than when ignoring the sampling effort (*NWPE*) (Fig. 3). This indicates that, on average, the regional estimation ranges shifted towards slightly higher values when considering the uneven sampling effort (mean relative difference = 5% [ $SD = 8.3\%$ ]).

On the other hand, the magnitude of the regional estimation range reflects the imprecision in the local estimates (due to the use of abundance classes) and is higher than 15% of the central regional estimate for most of the species, irrespective of the procedure (*WPE* or *NWPE*). When compared with this magnitude, the difference between *WPE* and *NWPE* can be considered as negligible in most cases. Nevertheless, this difference was quite relevant for a series of uncommon species, including the Whinchat (*Saxicola rubetra*, mean relative difference 16%) or the Eurasian Siskin (*Carduelis spinus*, mean relative difference 18%), most of them being of conservation-concern and/or quite elusive.

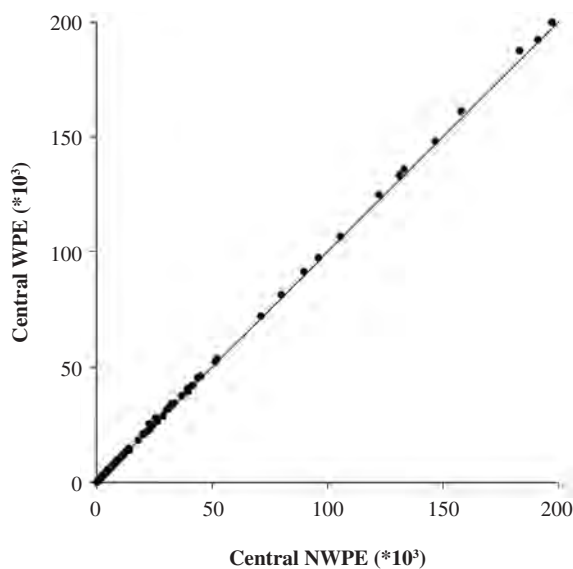
## DISCUSSION

This paper suggests a modus operandi to deal with the issue of heterogeneous data coverage in atlases, with a view to refining regional population estimates. All sampling units were evened out to represent equal sampling effort by using a weighting procedure. The method relied on a dual data collection and was implemented in the analytical framework of the *Breeding Bird Atlas of Wallonia 2001-2007*.

Three indices were computed on the basis of criteria reflecting different aspects of the attention paid and effort made to collect data in each sampling unit. The first index compared the data reported from the unrestricted surveys in the sampling units and from the timed surveys in the sub-sampling units to quantify the relative reliability of the unrestricted surveys. The second index evaluated to what extent the species richness reported from the unrestricted surveys in the sampling units departed from the species richness reported in their neighbouring units, which pointed out suspiciously poorly surveyed units. The third index was based on the relative proportions of species with possible, probable and confirmed breeding evidence reported from the unrestricted surveys and enabled us to evaluate the attention paid to collect the breeding bird data. These indices were next combined into a single score (see also Vaisanen 1998) defining the effective sampling effort in



**Figure 2.** Pairwise scatter plots for each pair of indices (*RS*, *NS* and *CB*).



**Figure 3.** Scatter plots comparing *WPE* and *NWPE* for all the species (except for species linked to lakes and ponds, i.e. Anatidae, grebes, coots and Laridae). The diagonal line represents the situation in which *WPE* is equal to *NWPE*.

the sampling units. This score was finally used to weight the relative contribution of each unit to the regional population estimates, by giving a lesser (respectively greater) importance to the units that were poorly (respectively well) covered.

As a general rule, weighted regional population estimates (*WPE*) were slightly higher than non-weighted regional population estimates (*NWPE*), therefore indicating that, on average, well-covered units reported higher local estimates than poorly covered units. In comparison with the magnitude of the regional estimation ranges (related to the inaccuracy of the local population estimates as a result of the use of abundance classes), the importance of the weighting procedure appeared, however, to be globally negligible (Fig. 3), except for some particular species. The field collection of more precise information than abundance classes (as adopted in this atlas) would undoubtedly contribute to reducing the estimation ranges for many species. In addition, the weak differences between *WPE* and *NWPE* can be explained by the existence of outlying sampling units along the range of score values (Fig. 1). A few sampling units exhibited either dramatically low or high score values due to their sensitivity to the different criteria used to evaluate the sampling effort, whereas most of the units were restricted to a narrow range of score values. As a result, when using the score values as a weighting coefficient for the generation of regional population estimates, most of the sampling units were set on equal terms, result-

ing in low differences between *WPE* and *NWPE*. Accordingly, further investigation should focus on the calculation of indices that would be able to discriminate between poorly- and well-surveyed units in a more gradual way so as to better reflect the whole gradient of disparity in the sampling effort.

The objective of the weighting coefficient proposed here is to derive the unevenness in the effective sampling effort from the field data provided by the observers. For this reason, this sampling effort score could be used in a variety of atlas data analyses in addition to population size estimations, such as comparison of sampling effort in space (e.g. detection of under-surveyed sub-regions) or in time (e.g. consideration of differences in local sampling effort for between-atlas comparisons).

**Acknowledgements** – We thank all the numerous participants in the *Breeding Bird Atlas of Wallonia 2001-2007* for fieldwork, Marc Fasol, Christophe Dehem and Marc De Sloover for data management, and the Walloon Region for financial support. N.T. was funded by the Belgian National Fund of Scientific Research during part of the study. A.D., J.P.J., T.K. and J.Y.P. were funded through projects with the Service Public de Wallonie (DGO3). This is contribution BRC172 of the Biodiversity Research Centre (UCL).

## REFERENCES

- Bibby CJ, Burgess ND, Hill DA 1992. *Bird Census Techniques*. Academic Press, London.
- Donald PF, Fuller RJ 1998. Ornithological atlas data: a review of uses and limitations. *Bird Study* 45: 129-145.
- Estrada J, Pedrocchi V, Brotons L, Herrando S 2004. *Atles dels Ocells Nidificants de Catalunya (1999-2002)*. Lynx Editions, Barcelona.
- Gibbons DW, Donald PF, Bauer HG, Fornasari L, Dawson IK 2007. Mapping avian distributions: the evolution of bird atlases. *Bird Study* 54: 324-334.
- Hagemeijer WJM, Blair MJ 1997. *The EBCC atlas of European breeding birds: their distribution and abundance*. Poyser, London/San Diego.
- Hustings F, Vergeer JW 2002. *Atlas van de Nederlandse broedvogels 1998-2000*. Nederlandse Fauna 5, Nederland.
- Jacob JP 2000. *Atlas des oiseaux nicheurs de Wallonie 2001 - 2005*. Notice à l'intention des collaborateurs. Document Aves Liège.
- Link WA, Sauer JR 1998. Estimating population change from count data: Application to the North American Breeding Bird Survey. *Ecological Applications* 8: 258-268.
- Link WA, Sauer JR 1999. Controlling for varying effort in count surveys - An analysis of Christmas bird count data. *Journal of Agricultural Biological and Environmental Statistics* 4: 116-125.
- Robertson A, Simmons RE, Jarvis AM, Brown CJ 1995. Can bird atlas data be used to estimate population-size - A case-study using Namibian endemics. *Biological Conservation* 71: 87-95.
- Root T, McDaniel L 1995. Winter population trends of selected songbirds. In: LaRoe ET, Farris GS, Puckett CE, Doran PD, Mac MJ (eds). *Our Living Resources*. Washington, DC: USDI, National Biological Service, pp. 21-23.
- Sauer JR, Peterjohn BG, Link WA 1994. Observer differences in the North-American breeding bird survey. *Auk* 111: 50-62.
- Schmid H 2008. How to count birds in the field. In: Vorisek P, Klvanová A, Wotton S, Gregory RD (eds). *A Best Practice Guide for Wild Bird Monitoring Schemes*. CSO/RSPB, Czech Republic, pp. 60-63.
- Schmid H, Luder R, Naef-Daenzer B, Graf R, Zbinden N 1998. *Atlas des oiseaux nicheurs de Suisse 1993-1996*. Station ornithologique suisse, Sempach, Suisse.
- Telfer MG, Preston CD, Rothery P 2002. A general method for measuring relative change in range size from biological atlas data. *Biological Conservation* 107: 99-109.
- Timothy J, Sharrock R 1974. Minutes of the second meeting of the European Ornithological Committee. *Acta Ornithologica* 14: 404-411.
- Vaisanen RA 1998. Changes in the distribution of species between two bird atlas surveys: the analysis of covariance after controlling for research activity. *Ornis Fennica* 75: 53-67.