



Hidden patterns of colony size variation in seabirds: a logarithmic point of view

Roger Jovani, Roddy Mavor and Daniel Oro

R. Jovani (roger.jovani@ufz.de), Dept of Conservation Biology, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas (C.S.I.C.). Pabellón del Perú, Avenida María Luisa s/n, ES-41013 Sevilla, Spain. Present address: UFZ, Helmholtz Centre for Environmental Research – UFZ, Dept of Ecological Modelling, Permoserstr. 15, DE-04318 Leipzig, Germany. – R. Mavor, Seabird Colony Team, J.N.C.C., Dunnet House, 7 Thistle Place, Aberdeen, AB10 1UZ, UK. – D. Oro, Inst. Mediterrani d'Estudis Avançats IMEDEA (CSIC-UIB), Miquel Marques 21, ES-07190 Esporles, Mallorca, Spain.

Explaining the huge variability present in bird colony sizes within and between species is intimately related to the understanding of the proximate and ultimate reasons for bird coloniality. However, natural patterns of colony size frequency distributions (CSFDs) remain poorly known. It is widely believed that colonial birds have similar long-tailed (highly right-skewed) CSFDs and that species mainly differ in their maximum colony sizes (in the length of the ‘tail’ of their CSFDs). We used data from the Seabird 2000 project (20 species; 19 978 colonies; 3 779 919 nests), the largest and most detailed dataset currently available, to analyse the CSFDs of seabird breeding in Britain and Ireland. Log-transformations of colony sizes revealed that the often reported long-tailed CSFDs in common histograms were hiding contrasting patterns, mainly log-normal but also power law CSFDs. The different statistical characteristics of CSFDs did not co-occur at random within species and were in fact highly correlated (e.g. a large geometric mean correlated with a large coefficient of variation). A PCA with these characteristics revealed a smoothed transition between species’ CSFD. Therefore, (a) a logarithmic analysis will allow different aspects of what is currently only referred to as ‘colony size variation’ to be quantified; (b) we challenge the current idea that all species show similar long-tailed CSFDs; (c) we offer a new (unified) view of colony size variation and discuss how these new patterns confirm, challenge and may advance theoretical and applied research into bird coloniality.

The study of the scaling properties of group size frequency distributions in animals is considered a key step towards a better understanding of grouping patterns and the underlying processes that govern them (Bonabeau et al. 1999, Sjöberg et al. 2000). A spectacular animal grouping phenomena that has received much attention from the public and researchers alike is the habit of around 13% of bird species – almost all seabirds – to breed in colonies (Wittenberger and Hunt 1985, Rolland et al. 1998).

In bird coloniality research, colony size is a recurrent theme (Kharitonov and Siegel-Causey 1988, Brown et al. 1990, Siegel-Causey and Kharitonov 1990) and it has been argued that any theory of bird coloniality should explain the huge variation in colony size that exists both within and between species (Brown et al. 1990, Brown and Brown 2001). This is because both ultimate (adaptive, e.g. lowering predation risk) and proximate (mechanisms; e.g. conspecific attraction) explanations of bird coloniality are intimately linked to the size of colonies. Although the study of patterns of colony size variation seems to be a justifiable research priority, natural patterns of colony sizes remain poorly studied, even though the rarely quantified concept of ‘colony size variation’ is widespread in the literature.

When reviewing existing data on bird colony size variation Brown et al. (1990) found that for all species studied, colony sizes range from just a few to many nests, and that species greatly differ in their maximum colony size. Moreover, colony size frequency distributions (hereafter CSFDs) are often reported to be long-tailed (highly right skewed) in common histograms (Götmark 1982, Garrison et al. 1987, Brown et al. 1990, Tella et al. 2001, Crawford 2003). Thus, the general consensus on colony size variation is that different species are alike in the shape of their CSFD, but that some species achieve larger colony sizes than others, and have histograms with longer ‘tails’.

However, here we show that the current methods used to investigate colony size variation are particularly insensitive to long-tailed distributions and can lead to mistaken ideas that do not allow for a quantitative approach to this central topic in bird coloniality research.

We analysed the CSFDs of seabirds breeding in Britain and Ireland with appropriate statistical tools for long-tailed frequency distributions (Pueyo 2003, 2006). Additionally, we studied the ways in which the CSFDs of different species resemble and differ from each other. Overall, this study offers the most precise and extensive discussion yet of

CSFDs for a long list of species and gives a new unified view to (sea)bird colony size variation.

Methods

Data

The data used in this study were obtained from the Seabird 2000 project (see Mitchell et al. 2004 for a detailed report on methods and results). In brief, specific survey methodologies for each seabird species were designed on the basis of the long-term experience of previous surveys in the UK (Cramp et al. 1974). The project was conducted from 1998 to 2002 by over 1000 surveyors who counted colonies of the 25 seabird species that breed inland and along the coastlines of Britain and Ireland. This is the most exhaustive and reliable dataset on bird colony sizes available for such a large area.

Bird colonies are not clear-cut entities (Coulson and Dixon 1979, Nelson 1980, Brown and Brown 2001, Coulson 2002, Gaston 2004, Mitchell et al. 2004, Jovani and Tella 2007). This is evident for northern fulmars where nests continuously spread along cliff stretches, making it difficult to say where one colony finishes and the next one starts. In addition, in colonies located on isolated islands, as for northern gannets *Morus bassanus*, there are cases where it is unclear whether or not two close ‘colonies’ are in fact two parts of a single ‘colony’ (Nelson 1980). In an extreme case, Jovani and Tella (2007) reported a fractal-like spatial nest distribution in a population of white storks *Ciconia ciconia*, where nests occurred in clusters (‘colonies’) within clusters at different scales, without a preferred ‘colony’ scale. However, Jovani and Tella (2007) found that irrespective of the scale used to define ‘colonies’, the shape of the resulting CSFD was alike, giving some hope that CSFDs are robust enough to allow comparisons between species. Thus, here, rather than relying on the always subjective definitions of what constitutes a colony for each species and in each particular case, we preferred to use as our colony scale the ‘sub-site’; unit (i.e. stretches of coastline of around 1 km in length) as used in the Seabird 2000 project. ‘Sub-sites’ are defined by Mitchell et al. 2004 as equating in most cases to what is thought to be a colony (p. 40 in Mitchell et al. 2004).

To explore to what extent the results shown here were much dependant on the spatial scale used to study colonies, we repeated the analyses considering the ‘site’ scale (Mitchell et al. 2004) as our colony scale. In contrast to ‘sub-sites’ that are areas of similar size, ‘sites’ differ widely from one another. ‘Sites’ are composed of a variable number of ‘sub-sites’ (median = 2, 5th percentile = 1, 95th percentile = 10; Supplementary material Appendix 1). Within ‘sites’, nearest distances between ‘sub-sites’ are a median of 1 km (5th percentile = 224 m, 95th percentile = 6.5 km; Supplementary material Appendix 1). As expected, less and larger colonies emerged by using ‘sites’ as colonies, but species were very robust to scaling on the shape of their CSFDs. Therefore, only results using ‘sub-sites’ as colony scale are reported here. See Supplementary material Appendix 1 for a comparison of results using ‘sites’ vs ‘sub-sites’.

We also tried to overcome problems of statistical noise caused by small sample sizes by studying only species with at least 50 reported colonies. In this way we were able to analyse data for 20 of the 25 species surveyed during the Seabird 2000 project, that is, a total of 19 978 colonies and 3 779 919 nests. See Table 1 for the scientific names of the species studied and for the number of colonies and nests analysed.

Analyses

For each species we plotted their CSFD in three different ways, as exemplified in Fig. 1, to achieve a complete picture of the patterns involved and to be able to compare current practices in animal grouping research (common histograms; e.g. Fig. 1A) with more appropriate techniques for long-tailed frequency distributions such as log-linear (Bak and Meesters 1998; Fig. 1B) and log–log plots (Pueyo 2003, 2006, Jovani and Tella 2007; e.g. Fig. 1C). In common histograms, colony sizes (x-axis) were plotted in linear bins (intervals of e.g. [1,500], [501,1000], [1001,1500], ... [8001,8500] ... etc.) against the frequency of colonies (y-axis) of each size in a linear scale. For log-linear (Fig. 1B) and log–log plots (Fig. 1C) we used multiplicative bins of the form $[X^n, X^{n+1}-1]$ along the x-axis with $n = 0, 1, 2, 3 \dots$, e.g. for $X=2$ bins were [1,1], [2,3], [4,7], ... [8192,16384], ... etc., and we plotted the logarithmic midpoint of the bin $10^{(\log(\text{minimum colony size of the bin}) + \log(\text{maximum colony size of the bin}))/2}$. In the log-linear plots (Fig. 1B) we plotted the frequency of colonies as in common histograms on the y-axis. However, in log–log plots (Fig. 1C), zero values (a bin without representation in the population) cannot be plotted or used for function fitting as this may lead to spurious results (Pueyo and Jovani 2006). Hence, we used the lower X that produced no empty bin within the range of colony sizes of each species ($X=2$ was used for 17 species, and $X=3$ for three; this was also done in log-linear plots for consistency). In log–log plots the y-axis displayed the density function $p(x) = c/(C s)$, where c is the number of colonies within a bin, C is the total number of colonies analysed for the species and s is the length of the bin (i.e. the number of possible colony sizes within the bin: e.g. $s = 2^n$ for bins in powers of two). Note that logarithmic binning expands the information condensed in the first few bins of common histograms and at the same time condenses the scattered information in its large bins (Fig. 1).

From log-transformed colony size data we characterised CSFDs by retrieving their minimum, maximum, mean (i.e. geometric mean), coefficient of variation (as standard deviation/mean), kurtosis (a measure of peakedness), skewness (a measure of asymmetry: e.g. lesser black-backed gulls displayed more kurtosis and skewness than common guillemots, Fig. 2B) and a calculation of the fit of the CSFDs to a normal distribution with the Kolmogorov–Smirnov statistic (larger values denote a larger deviation from a normal distribution; the data was log-transformed and was thus a test of the fit of the distribution to a log-normal). Moreover, in log–log plots, we used mean squares to test the fit (R^2) of CSFDs, either to a log-normal or to a power law distribution. All species fitted one of these two

Table 1. Characteristics of the colony size frequency distributions (CSFDs) for the different species analysed, and the corresponding eigenvalues for the two first principal components of the principal component analysis (PCA). Minimum colony size was not introduced in the PCA because it was not a variable characteristic between species. c.s. means colony size.

Bird species	No. of colonies	No. of nests	Min c.s.	Max c.s.	Geometric mean	Kurtosis	Skewness	Coefficient of variation	Kolmogorov-Smirnov
Northern fulmar <i>Fulmarus glacialis</i>	2976	537 988	1	20 424	32.143	-0.404	0.187	0.530	0.035
Manx shearwater <i>Puffinus puffinus</i>	53	332 272	1	120 000	135.202	-0.754	0.406	0.624	0.127
European storm-petrel <i>Hydrobates pelagicus</i>	98	82 818	1	27 297	62.663	-0.356	0.330	0.559	0.063
Great cormorant <i>Phalacrocorax carbo</i>	267	13 628	1	675	19.736	-0.448	-0.325	0.511	0.083
European shag <i>Phalacrocorax aristotelis</i>	1392	32 222	1	1720	8.251	-0.285	0.297	0.648	0.070
Arctic skua <i>Stercorarius parasiticus</i>	650	2136	1	107	1.736	3.169	1.795	1.577	0.352
Great skua <i>Stercorarius skua</i>	726	9635	1	2293	2.656	2.861	1.518	1.256	0.243
Black-headed gull <i>Larus ridibundus</i>	677	141 871	1	14 575	19.270	-0.170	0.515	0.691	0.074
Common gull <i>Larus canus</i>	1554	49 728	1	11 219	6.530	0.520	0.623	0.747	0.090
Lesser black-backed gull <i>Larus fuscus</i>	995	116 640	1	19 487	8.470	0.576	0.834	0.868	0.125
Herring gull <i>Larus argentatus</i>	2845	148 869	1	10 129	10.223	-0.362	0.397	0.723	0.084
Great black-backed gull <i>Larus marinus</i>	2073	19 616	1	983	2.989	0.767	1.143	1.142	0.213
Black-legged kittiwake <i>Rissa tridactyla</i>	679	415 994	1	17 546	145.111	-0.178	-0.184	0.368	0.030
Common tern <i>Sterna hirundo</i>	424	13 859	1	1033	7.224	-0.388	0.554	0.825	0.119
Arctic tern <i>Sterna paradisea</i>	928	48 469	1	4000	13.914	-0.227	0.213	0.590	0.062
Little tern <i>Sterna albifrons</i>	119	2093	1	220	7.119	-0.716	0.253	0.688	0.104
Common guillemot <i>Uria aalge</i>	681	1 044 856	1	75 493	194.209	-0.426	-0.123	0.430	0.027
Razorbill <i>Alca torda</i>	870	144 743	1	11 384	28.196	-0.396	0.181	0.560	0.042
Black guillemot <i>Cephus grylle</i>	1481	21 731	1	323	7.014	-0.781	0.048	0.641	0.088
Atlantic puffin <i>Fratercula arctica</i>	490	600 751	1	59 471	40.317	0.380	0.664	0.633	0.063
Principal component analysis									
Eigenvalues for each characteristic of CSFD									
PC1				-0.622	-0.852	0.880	0.918	0.988	0.956
PC2				0.758	0.401	0.350	0.344	0.104	0.090

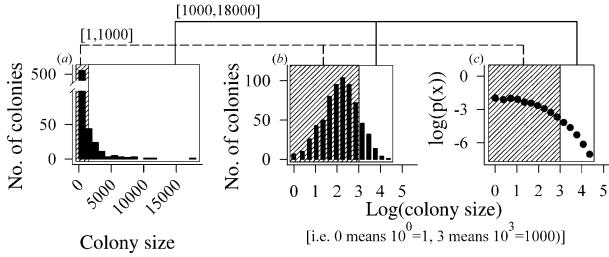


Figure 1. Comparison between the three kinds of plots discussed here using as an example the data for black-legged kittiwakes from Fig. 2. Note that in (B) and (C) logarithmic binning unfolds in ten bins the information hidden in the first bin of the common histogram (A) (hatched areas) and collapses into just five bins the few scattered information contained in the long-tail of (A) (unhatched areas).

distributions better than an exponential distribution, which in Fig. 2C displayed in all cases a much faster decay from small to large colony sizes. We performed a principal component analysis (PCA) to summarise all this information by collapsing all variables into just a few. For this analysis we used the maximum, mean, coefficient of variation, kurtosis, skewness and the Kolmogorov–Smirnov statistic. We did not use the minimum because it showed no variation between species (Table 1).

Results

As expected, all seabird species displayed a long-tailed (highly right skewed) CSFD when plotted in common histograms (Fig. 2A). However, when we binned colony sizes in multiplicative bins, plotting them in log-linear and log-log plots, contrasting patterns were revealed (Fig. 2B–C). Interestingly, due to the properties of logarithmic binning (Methods; Fig. 1) we found different, hidden CSFDs shapes within the small and large colony size bins (Fig. 2B–C). Within large bins, logarithmic binning revealed that the apparently unstructured ‘long-tails’ of CSFDs in common histograms were hiding an interesting structure in which maximum colony sizes are an integral part of the CSFD (Fig. 2B–C) and not just scattered points (Fig. 2A).

Different species displayed contrasting values in all the characteristics analysed (except in the minimum colony size, which was invariably one; Table 1). Fourteen out of 20 species showed a good fit to a log-normal distribution, five were fitted more closely to a power law, and a single species had a similar fit for both distributions (Fig. 2C).

The values of the different characteristics of the CSFDs were highly correlated (Fig. 3). The PCA supported the redundancy between the six characteristics measured; the first two components accounted for 93.54% of the original variance. 76.96% and 16.58% of variance were summarised in the first and second components, respectively, while the third component only increased the explained variance by 3.67% (see Table 1 for the eigenvalues of each component). Plotting species location in the plane created by these two new variables (the two first PCA components, Fig. 4) highlighted a gradual transition between the CSFDs of the

different species. For instance, in Fig. 4 if we go from the top-left (common guillemot *URIAAL*) to middle-bottom (little tern *STEALB*) species with similar log-linear CSFD plots lie close together and vice versa.

Discussion

Patterns of colony size variation

Current approaches to colony size variation in birds (i.e. common histograms; Götmark 1982, Brown et al. 1990, Brown and Brown 1996, Tella et al. 2001) have led to the idea that all species show long-tailed CSFDs which essentially differ in the length of this ‘tail’ (e.g. Fig. 2A). Here, we offer a rather different interpretation reached through the use of log-linear and log-log plots (Fig. 2B–C). We have shown that species not only differ in their extreme colony sizes (i.e. in the largest colony size bins), but also in the shape of the relative frequency between small-sized colonies (Fig. 2B–C). This result was impossible to extract from common histograms (compare Fig. 2A with 2B). Our analysis with logarithmized colony sizes has revealed that similar long-tailed CSFDs in common histograms do in fact hide contrasting patterns such as power laws and log-normal CSFDs (Fig. 2B–C).

Species differed greatly in the characteristics measured from their CSFDs (Table 1), which are a more accurate and quantitative way of describing the currently vague concept of ‘colony size variation’. The values for each characteristic in each species did not co-occur by chance; instead, there was a high correlation within species of the values of these characteristics (Fig. 3). This finding is important. For instance, a possible and potentially reasonable alternative (to be expected *a priori*) was that different species showed the same log-normal shape but with their geometric mean in different positions on the x-axis in Fig. 2B. That is, some species may have colonies with from one to 100 nests, others from 100 to 1000 nests and others even from 10 000 to 100 000 nests, but all, nevertheless, may have the same log-normal shape. In these cases the geometric mean and the coefficient of variation would not have shown any correlation between the species. Rather, the non-random co-occurrence of different CSFD characteristics leads to gradual transitions between CSFDs of the different species in the PCA plot (Fig. 4). This ranking highlighted here by the PCA provides a new integrated view of patterns of colony size variation and highlights two important features of CSFDs. Firstly, an interesting feature is whether species fit better to a log-normal or a power law distribution – note that the Kolmogorov–Smirnov statistic correlated strongly with all the other characteristics (Fig. 3). Power laws have been suggested by Schneider (2002) as a potential CSFD in birds, while Jovani and Tella (2007) have recently reported for the first time the occurrence of this kind of CSFD in a population of white storks, a terrestrial waterbird. Here, we confirm that power law CSFDs may be widespread in birds, as predicted by Schneider (2002). However, we also show that, according to our dataset, the occurrence of log-normal CSFDs could be even more frequent in (sea)birds.

Secondly, species following a log-normal distribution were ranked according to how large their colonies were (all

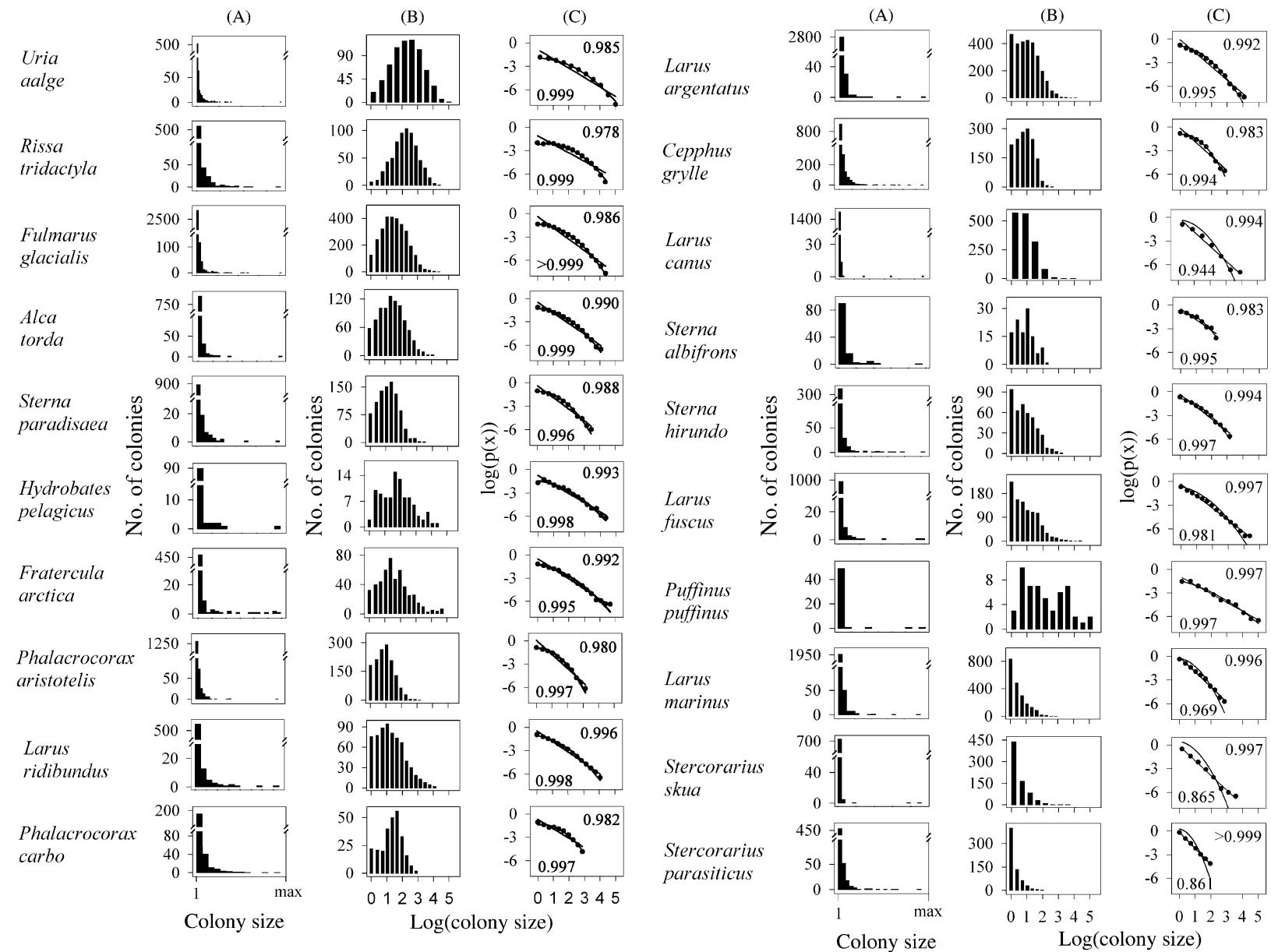


Figure 2 (Continued)

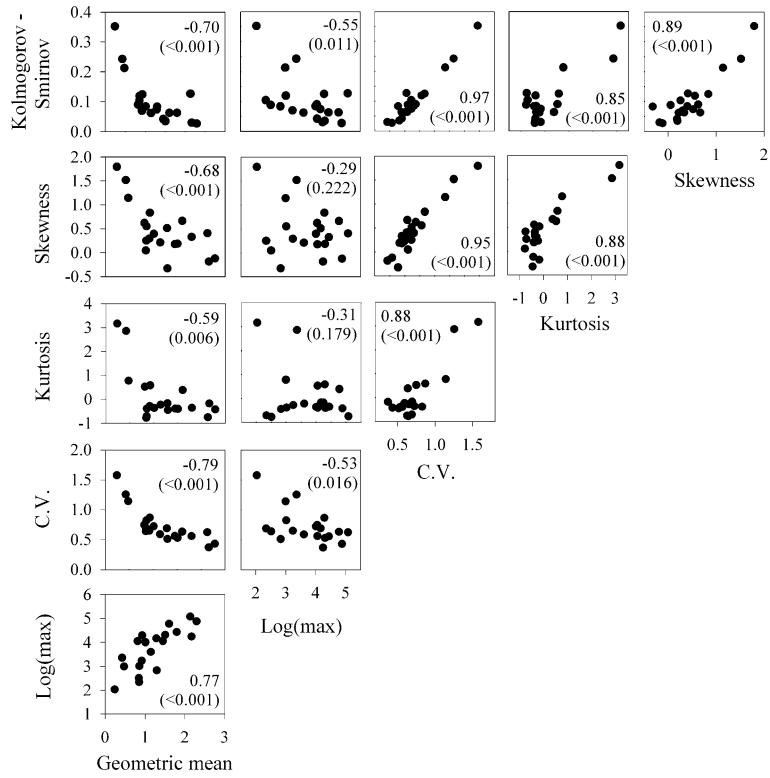


Figure 3. Pearson correlations between pairs of statistical characteristics of colony size frequency distributions (i.e. each point is a seabird species). Inset numbers indicate the Pearson correlation coefficient and the associated two-tailed significance value in parenthesis. C.V. means coefficient of variation.

species had very small colony sizes), that is, according to the value of their geometric mean and maximum colony size. The lowest coefficients of variation for those species with a larger geometric mean (Fig. 3) occurred because these species were the ones that best fitted a log-normal distribution (Fig. 2); this type of distribution is far less variable than power laws. In other words, the CSFDs of those species with larger colony sizes were ‘enlarged’ representations of those with smaller colony sizes. This occurred through a shifting of their geometric mean towards larger maximum colony sizes, while also having small colony sizes (Table 1).

The use of logarithmic transformations of colony sizes (Fig. 1) also gives us a fresh view on maximum colony sizes. This feature has been used in previous research into bird coloniality (Erwin 1978, Coulson 2002) since it provides a good estimation of the upper limit that colonies can achieve. Moreover, it captures the fact that some species exist in close-packed, huge colonies, while others have never been recorded in colonies of more than a few hundred breeding pairs. However, an obvious drawback of this statistic is that anything affecting the size of the largest colony will also seriously affect the shape of the ‘long-tail’ of

the CSFD. For instance, if the largest colony of a given species is lost through some catastrophic predation event, the next largest colony may be thousands of nests smaller.

Even so, maximum colony size does seem to be an informative characteristic for a species. For instance, it has been found that species that travel longer distances to find food for their nestlings (i.e. pelagic species) exist in larger colony sizes than inshore foragers (Erwin 1978, Coulson 2002). This, together with other studies (Ashmole 1963, Furness and Birkhead 1984, Lewis et al. 2001, Forero et al. 2002), offers strong evidence that food availability around colonies imposes an upper limit on colony growth through negative density-dependent processes.

Examining colony size after logarithmic transformation as we have done here strengthens the relevance of previous studies on maximum colony sizes. This is because we have found that maximum colony sizes should not be seen as unpredictable ‘outliers’, but rather as an integral part of CSFDs (Fig. 2B–C). Therefore, very large colonies are likely to be the outcome of the same processes (although with different strengths) that govern the dynamics of other colony sizes in a given species. This shows that we may not need to invoke special processes to explain the occurrence of

Figure 2. Colony size frequency distributions of breeding seabirds in Britain and Ireland ranked according to the value of their Kolmogorov–Smirnov statistic (i.e. by their fit to the log-normal distribution; Table 1). Species are ranked from top to bottom, from closer to a log-normal to closer to a power law. (A) common histograms with variable bin lengths according to colony size range of each species. Thus, ‘max’ in the x-axis indicates the maximum colony size for each species (Table 1). (B) log binned histograms, with logarithmic x-axis and linear y-axis. (C) log–log distribution of colony sizes: upper-right and lower-left numbers indicate the fit (R^2) of each distribution to a power law (the straight line) or to a log-normal (the curved one), respectively. See text for further details.

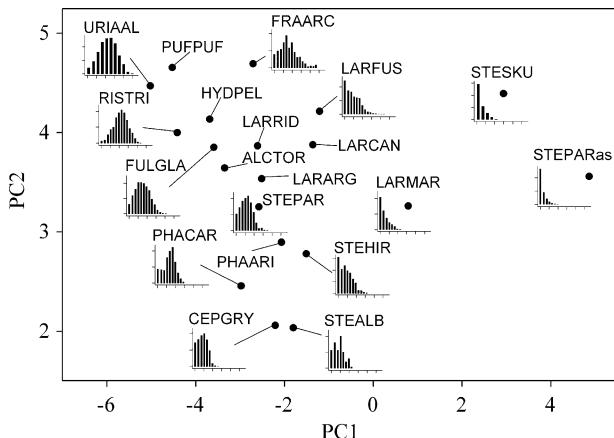


Figure 4. Ranking of colony size frequency distributions according to the values of the first two components of a principal component analysis (Table 1) done using six characteristics calculated from the colony size frequency distribution of each species. Some examples of the distributions (from Fig. 2B) are shown. Species are indicated by the first three letters of each part of their scientific names. STEPARas indicates *Stercorarius parasiticus* and STEPAR *Sterna paradisea*.

very large colony sizes and further study of this factor as an ecologically informative (and easy to calculate) characteristic of species is necessary.

Conclusions

Monitoring schemes may benefit from our logarithmic analysis of colony sizes. In the common histograms currently used all species show a long-tailed CSFD (e.g. Fig. 2A), with colonies clumping in the smallest colony bins. Whilst not hiding any of the information contained in the large tail of the histogram, the use of log(colony sizes) reveals the hidden (internal) structure within small colony sizes, which differs widely between species (Fig. 2B) in contrast to the apparent similarity observed in common histograms (Fig. 2A). This tells us that monitoring methods such as sampling all colonies above a certain size (or giving more priority to large colonies) may have different consequences for the population estimates of different species.

By analysing CSFDs logarithmically (Fig. 1), we have disclosed new patterns of colony size variation. Moreover, we have been able to quantitatively characterize species according to their different CSFD features and have found that in terms of many of these characteristics species differ greatly. Current hypotheses on the ultimate mechanisms that lead some species to breed in colonies (Siegel-Causey and Kharitonov 1990, Brown and Brown 2001, Coulson 2002) and on the proximate cues used by individuals of colonial species to choose different colony sizes (Brown et al. 1990) have never been tested as a means of explaining CSFDs (e.g. through pattern-oriented modelling; Grimm and Railsback 2005). There is still a big gap in our knowledge regarding bird coloniality, since even after decades of research colony size variation remains a major unexplained pattern, despite being the main focus of existing hypotheses. We have shown that different species

show very contrasting CSFDs and so tests of old (Lack 1967, Cairns 1989, Brown and Brown 2001, Coulson 2002) and new hypotheses (Danchin and Wagner 1997, Rolland et al. 1998, Wagner et al. 2000, Safran 2004, Serrano et al. 2004) of bird coloniality with the new patterns discovered here would seem now to be a priority on the agenda of bird coloniality research. Moreover, the patterns shown here have exhibited a high robustness to spatial scaling (Methods and Supplementary material Appendix 1). That means that data on colony sizes collected with different criteria to define colonies (e.g. collected by different teams in different parts of the world) would be comparable at the species level.

We have not been able to resolve here the reasons behind the huge variation in colony sizes found in the wild. However, we have shown that species with larger geometric means and maximum colony sizes (roughly speaking species with larger colonies) are those with a lower coefficient of variation in colony sizes (i.e. in colony size variation, Fig. 3). This is intimately linked to the fact that these are the species best fitted to log-normal distributions (Fig. 2, 3). Therefore, to understand the patterns of colony size variation it will first be necessary to explain why some species are better fitted to a power law and others to a log-normal CSFD. Why, for instance, are those species that are better fitted to a power law usually semi-colonial (e.g. skuas) and why are typically colonial species (e.g. common guillemot) better fitted to log-normals?

This will force us to reach a more formal (quantitative) description of the (gradual?) transition from territorial to semi-colonial and colonial species and return to the initial integrative approaches performed some decades ago on bird spacing patterns (Brown 1964, Lack 1967) that have now been abandoned in favour of more specialized hypotheses. The fact, for instance, that conspecific attraction and territoriality are increasingly accepted to be common behavioural traits in both colonial and territorial species (Stamps 1988, 1991, Serrano and Tella 2003, Ahlering and Faaborg 2006) is a good reason for revisiting and formalizing these old integrative approaches using current methods and large datasets such as the one used here.

Acknowledgements – José L. Tella, David Serrano, Carlos Rodríguez and Esperanza Ursúa made valuable comments on earlier drafts. Ian Mitchell kindly helped with the interpretation of the data from the Seabird 2000 project. Lidia López and Mike Lockwood helped with the English.

References

- Ahlering, M. A. and Faaborg, J. 2006. Avian habitat management meets conspecific attraction: if you built it, will they come? – Auk 123: 301–312.
- Ashmole, N. P. 1963. The regulation of numbers of tropical oceanic birds. – Ibis 103: 458–473.
- Bak, R. P. M. and Meesters, E. H. 1998. Coral population structure: the hidden information of colony size-frequency distributions. – Mar. Ecol. Progr. Ser. 162: 301–306.
- Bonabeau, E. et al. 1999. Scaling in animal group-size distributions. – Proc. Natl Acad. Sci. USA 96: 4472–4477.

- Brown, C. R. and Brown, M. B. 1996. Coloniality in the cliff swallow. The effect of group size on social behaviour. – Univ. of Chicago Press.
- Brown, C. R. and Brown, M. B. 2001. Avian coloniality. – Curr. Ornithol. 16: 1–82.
- Brown, C. R. et al. 1990. Choice of colony size in birds. – Trends Ecol. Evol. 5: 398–403.
- Brown, J. L. 1964. The evolution of diversity in avian territorial systems. – Wilson Bull. 76: 160–169.
- Cairns, D. K. 1989. The regulation of seabird colony size: a hinterland model. – Am. Nat. 134: 141–146.
- Coulson, J. C. 2002. Colonial breeding in seabirds. – In: Schreiber, E. A. and Burger, J. (eds), Biology of marine birds. CRC Press, pp. 87–113.
- Coulson, J. C. and Dixon, F. 1979. Colonial breeding in seabirds. – In: Larwood, G. and Rosen, B. R. (eds), Biology and systematics of colonial organisms. Academic Press, pp. 445–458.
- Cramp, S. et al. 1974. The seabirds of Britain and Ireland. – Taplinger.
- Crawford, R. J. M. 2003. Influence of food on numbers breeding, colony size and fidelity to localities of swift terns in South Africa's Western Cape, 1987–2000. – Waterbirds 26: 44–53.
- Danchin, E. and Wagner, R. H. 1997. The evolution of coloniality: the emergence of new perspectives. – Trends Ecol. Evol. 12: 342–347.
- Erwin, R. M. 1978. Coloniality in terns: the role of social feeding. – Condor 80: 211–215.
- Forero, M. G. et al. 2002. Conspecific food competition explains variability in colony size: a test in Magellanic penguins. – Ecology 83: 3466–3475.
- Furness, R. W. and Birkhead, T. R. 1984. Seabird colony distributions suggest competition for food supplies during the breeding season. – Nature 311: 655–656.
- Garrison, B. A. et al. 1987. Bank swallow distribution and nesting ecology on the Sacramento River, California. – Western Birds 18: 71–76.
- Gaston, A. J. 2004. Seabirds. A natural history. – Yale Univ. Press.
- Götmark, F. 1982. Coloniality in five *Larus* gulls: a comparative study. – Ornis Scand. 13: 211–224.
- Grimm, V. and Railsback, S. F. 2005. Individual-based modeling and ecology. – Princeton Univ. Press.
- Jovani, R. and Tella, J. L. 2007. Fractal bird nest distribution produces scale-free colony sizes. – Proc. R. Soc. Lond. B 274: 2465–2469.
- Kharitonov, S. P. and Siegel-Causey, D. 1988. Colony formation in seabirds. – Curr. Ornithol. 5: 223–272.
- Lack, D. 1967. Interrelationships in breeding adaptations as shown by marine birds. – In: Proc. XIV. Int. Ornithol. Congr. 1966, pp. 3–42.
- Lewis, S. et al. 2001. Evidence of intra-specific competition for food in a pelagic seabird. – Nature 412: 816–819.
- Mitchell, P. I. et al. 2004. Seabird populations of Britain and Ireland. – T and A D Poyser.
- Nelson, B. 1980 Seabirds. Their biology and ecology. – Hamlyn.
- Pueyo, S. 2003. Irreversibility and criticality in the biosphere. PhD thesis. – Publicacions Univ. de Barcelona.
- Pueyo, S. 2006. Diversity: between neutrality and structure. – Oikos 112: 391–404.
- Pueyo, S. and Jovani, R. 2006. Comment on 'A keystone mutualism drives pattern in a power function'. – Science 313: 1739c.
- Rolland, C. et al. 1998. The evolution of coloniality in birds in relation to food, habitat, predation, and life-history traits: a comparative analysis. – Am. Nat. 151: 514–529.
- Safran, R. J. 2004. Adaptive site selection rules and variation in group size of barn swallows: individual decisions predict population patterns. – Am. Nat. 164: 121–131.
- Schneider, D. C. 2002. Scaling theory: application to marine ornithology. – Ecosystems 5: 736–748.
- Serrano, D. and Tella, J. L. 2003. Dispersal within a spatially structured population of lesser kestrels: the role of spatial isolation and conspecific attraction. – J. Anim. Ecol. 72: 400–410.
- Serrano, D. et al. 2004. Dispersal and social attraction affect colony selection and dynamics of lesser kestrels. – Ecology 85: 3438–3447.
- Siegel-Causey, D. and Kharitonov, S. P. 1990. The evolution of coloniality. – Curr. Ornithol. 7: 285–330.
- Sjöberg, M. et al. 2000. Truncated power laws: a tool for understanding aggregation patterns in animals? – Ecol. Lett. 3: 90–94.
- Stamps, J. A. 1988. Conspecific attraction and aggregations in territorial species. – Am. Nat. 131: 329–347.
- Stamps, J. A. 1991. The effect of conspecifics on habitat selection in territorial species. – Behav. Ecol. Sociobiol. 28: 29–36.
- Tella, J. L. et al. 2001. Offspring body condition and immuno-competence are negatively affected by high breeding densities in a colonial seabird: a multiscale approach. – Proc. R. Soc. Lond. B 268: 1455–1461.
- Wagner, R. H. et al. 2000. Colonies as byproducts of commodity selection. – Behav. Ecol. 11: 572–573.
- Wittenberger, J. F. and Hunt, G. L. Jr. 1985. The adaptive significance of coloniality in birds. – In: Avian Biology. VIII. Academic Press, pp. 1–78.