Diversity and Distributions, (Diversity Distrib.) (2016) 1–12

A Journal of Conservation Biogeography
Diversity and Distributions
BIODIVERSITY RESEARCH

Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark

Johannes Kamp[1,2], Steffen Oppel[1]*, Henning Heldbjerg[3,4], Timme Nyegaard[3] and Paul F. Donald[1]

[1]RSPB Centre for Conservation Science, Royal Society for the Protection of Birds, The David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK, [2]Institute of Landscape Ecology, University of Münster, Heisenbergstr. 2, 48149 Münster, Germany, [3]Dansk Ornitologisk Forening (DOF), BirdLife Denmark, Vesterbrogade 140, DK-1620 København, Denmark, [4]Department of Bioscience, Aarhus University, Kalø, Grenåvej 14, DK-8410 Rønde, Denmark

*Correspondence: Steffen Oppel, RSPB Centre for Conservation Science, Royal Society for the Protection of Birds, The David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK.
E-mail: steffen.oppel@rspb.org.uk

ABSTRACT

Aim Long-term monitoring of biodiversity is necessary to identify population declines and to develop conservation management. Because long-term monitoring is labour-intensive, resources to implement robust monitoring programmes are lacking in many countries. The increasing availability of citizen science data in online public databases can potentially fill gaps in structured monitoring programmes, but only if trends estimated from unstructured citizen science data match those estimated from structured monitoring programmes. We therefore aimed to assess the correlation between trends estimated from structured and unstructured data.

Location Denmark.

Methods We compared population trends for 103 bird species estimated over 28 years from a structured monitoring programme and from unstructured citizen science data to assess whether trends estimated from the two data sources were correlated.

Results Trends estimated from the two data sources were generally positively correlated, but less than half the population declines identified from the structured monitoring data were recovered from the unstructured citizen science data. The mismatch persisted when we reduced the structured monitoring data from count data to occurrence data to mimic the information content of unstructured citizen science data and when we filtered the unstructured data to reduce the number of incomplete lists reported. Mismatching trends were especially prevalent for the most common species. Worryingly, more than half the species showing significant declines in the structured monitoring showed significant positive trends in the citizen science data.

Main conclusions We caution that unstructured citizen science databases cannot replace structured monitoring data because the former are less sensitive to population changes. Thus, unstructured data may not fulfil one of the most critical functions of structured monitoring programmes, namely to act as an early warning system that detects population declines.

Keywords

citizen science, common bird monitoring, JAGS, occupancy model, population trend, volunteer.

INTRODUCTION

Monitoring changes in species' populations is an essential element of biodiversity conservation. Objective quantification of population change allows problems to be identified and conservation responses to be developed. The performance of subsequent management can then be evaluated from continued monitoring. Dedicated population monitoring schemes for biodiversity have been running for decades in many countries, particularly in Europe and North America

DOI: 10.1111/ddi.12463
http://wileyonlinelibrary.com/journal/ddi

(Greenwood, 2003; Schmeller *et al.*, 2012). However, these structured schemes require considerable investment and organization, and usually rely on a large number of dedicated volunteers who are able and willing to apply standardized methods over large areas and long time periods (Schmeller *et al.*, 2009). Many biodiversity-rich countries, however, lack the resources for such schemes, necessitating the identification of other sources of data and methods to monitor biodiversity.

Casual observations collected without following a structured protocol by members of the public may potentially contribute to research and conservation, and a growing number of unstructured 'citizen science' databases have become available in recent years (Devictor *et al.*, 2010; Sullivan *et al.*, 2014; Theobald *et al.*, 2015). However, several sources of bias in unstructured data are well known and the information content of unstructured data can be highly variable (Dickinson *et al.*, 2010; Hochachka *et al.*, 2012; Isaac & Pocock, 2015). Field observations collected in an unstructured manner usually do not represent random samples and exhibit considerable spatial bias towards more densely populated regions (Boakes *et al.*, 2010; Lin *et al.*, 2015), protected areas, and areas rich in biodiversity and threatened species (Tulloch *et al.*, 2013a). Observation effort is not standardized as in structured monitoring schemes (Dickinson *et al.*, 2010), and there might be considerable reporting bias, as many observers tend to report only unusual or rare species (van Strien *et al.*, 2013). These characteristics of unstructured data make it difficult to assess how reliable they can be for biodiversity monitoring.

Several approaches have been developed to account for some of the bias inherent in unstructured data and extract more reliable information (van Strien *et al.*, 2010; Hochachka *et al.*, 2012; Isaac *et al.*, 2014). Correcting for varying observation effort in unstructured data has been achieved using the number of species reported per visit ('list length'; Szabo *et al.*, 2010) or, where recorded, the time spent per field visit (Kindberg *et al.*, 2009). A more significant challenge, the problem that an unknown proportion of those species that are present will not be detected during a given visit, has been addressed using site-occupancy models that account for imperfect detection and may simultaneously correct for reporting bias (Kéry *et al.*, 2010a,b; van Strien *et al.*, 2013). However, a robust validation of such approaches is necessary before unstructured data can be used with confidence for biodiversity monitoring (Isaac *et al.*, 2014).

The value of unstructured monitoring data can be assessed by comparing population trends derived from unstructured citizen science data against the best available independent, structured monitoring schemes. Previous comparisons have detected correlations between reporting rates from weakly structured atlas data and data from a standardized random-sampled survey that range from strong (Szabo *et al.*, 2012) to weak and inconsistent (Snäll *et al.*, 2011). Accounting for imperfect detection using occupancy models based on comprehensive species lists matched the trends of a robust monitoring scheme better than presence-only data (van Strien *et al.*, 2010), and strong trends in structured monitoring data may be recovered from unstructured data when analysed with occupancy models (van Strien *et al.*, 2013; Isaac *et al.*, 2014). The usefulness of unstructured data therefore clearly depends on how they are processed and analysed.

One major difference between many long-term structured monitoring programmes and unstructured citizen science data is that the former often provide counts or indices of abundance, whereas unstructured citizen science data often only provide detection/non-detection data because of highly varying recording intensity in space and over time (Isaac *et al.*, 2014). However, if unstructured data are to fulfil the role of structured monitoring programmes, then they need to be able to identify approximately the same population trends as structured monitoring despite this inherent difference in data quality. Because even simple detection/non-detection data can allow inference about the abundance of a population (Royle & Nichols, 2003), population trends should be detectable with unstructured citizen science data (van Strien *et al.*, 2013; Isaac *et al.*, 2014), but may be less reliable than trends derived from structured monitoring data with higher information content (Johnston *et al.*, 2015). The potential for unstructured data to recover trends could therefore possibly be improved using only records with higher information content (Roy *et al.*, 2012). A typical deficiency of many online public databases is the lack of differentiation between complete species lists (which allow inference about the non-detection of species) and incidental records of a subset of the species actually observed (Kéry *et al.*, 2010b; van Strien *et al.*, 2013; Tulloch *et al.*, 2013b). Using data sets with higher information content that allow the statistical modelling of detection probability can yield improved trend estimates (Kéry *et al.*, 2010a; Isaac & Pocock, 2015), but whether such filtering can overcome other deficiencies of unstructured citizen science data is unclear.

Here, we assess whether unstructured observation records can recover population trends derived with confidence from structured surveys, despite having data with inherently lower information content. We use unstructured bird monitoring data from a country-wide public online database containing more than 12 million records collected over 28 years in Denmark. We first estimated population trends of 103 bird species from unstructured data using occupancy models. We then correlated these trends with population trends estimated over the same period by a structured, standardized common bird monitoring programme in the same country. Finally, we compared trend estimates from both data sources and assessed whether mismatches in these estimates were a consequence of fundamentally different information content by: (1) reducing the information content of structured monitoring data from count to detection/non-detection data, and (2) applying multiple filtering criteria to retain only records in the unstructured data with increasing information content. Our study thus provides a thorough examination of the potential of unstructured citizen science data to detect population trends and identifies factors that may affect the correspondence between structured and unstructured data sources for biodiversity monitoring.

MATERIALS AND METHODS

Structured monitoring and unstructured observation data

The Common Bird Monitoring (CBM) scheme in Denmark was established in 1975 (Heldbjerg *et al.*, 2014) following standard guidelines for structured bird monitoring programmes (Gregory *et al.*, 2004). Birds were monitored once during the breeding season (1 May to 15 June) on observer-chosen (non-randomly placed) routes, with each route containing 10–20 points which were spaced at least 300 m apart and were visited each year by the same observer. At each point, all birds seen or heard within a 5-min interval were counted. This forms the key difference to unstructured observation data, which are generally obtained from random surveys of highly variable duration yielding only detection/non-detection information. Although no *a priori* stratification of routes was applied, the survey routes covered all main habitat types in Denmark and were distributed relatively evenly across the country with no obvious concentrations in urban areas. We used data from 1986 to 2013, during which the number of routes remained relatively stable at between 300 and 400 (Fig. 1; Heldbjerg *et al.*, 2014).

We used unstructured observation data from the online database 'DOFbasen' (http://www.dofbasen.dk, Nyegaard *et al.*, 2012), developed by the Dansk Ornitologisk Forening

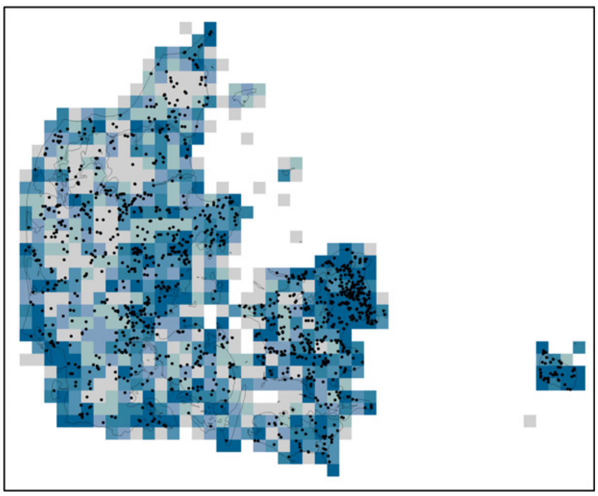

Figure 1 Map of Denmark showing the distribution of common bird monitoring census points used between 1986 and 2013 (black dots) and the distribution of records available from the online database 'DOFbasen' until 2013 aggregated over a grid of 10 × 10 km squares. Darker shading indicates higher density of records. The quantity of data from each data source per grid cell was positively correlated (r_S = 0.78).

(DOF). DOFbasen was launched in 2002, and most observations have been entered since then (Fig. S1 in Supporting Information). However, many observers have entered data retrospectively, and DOFbasen now holds sufficient historical data to compare trends with structured monitoring data from 1986 to 2013. All records include key fields such as species, date and location. As with many public online databases, DOFbasen did not differentiate until 2012 between complete bird lists (lists of all the species observed on each visit) and incidental records of a subset of the species observed (Kéry *et al.*, 2010b; van Strien *et al.*, 2013; Tulloch *et al.*, 2013b).

Observations recorded in DOFbasen were not distributed randomly across Denmark (Fig. 1). However, records covered all parts of the country and areas with a larger number of structured monitoring survey routes from the common bird monitoring overlapped with areas of high DOFbasen data density (Spearman's rank correlation between the number of contributed DOFbasen records and the number of common bird monitoring counts per 100 km^2 area, r_S = 0.78; Fig. 1). It is therefore reasonable to assume that trends derived from these unstructured data were as representative of a well-covered country such as Denmark as the structured monitoring data.

Population trend estimation

We compared population trends estimated from a standardized population monitoring scheme (CBM) and from unstructured observation data (DOFbasen) to assess the extent to which trends estimated from the two data sources matched. We estimated population trends for 103 species breeding in Denmark (Table S1) over a 28-year period (1986–2013) for which data were available from both data sources. However, because the citizen science database was launched only in 2002, we also estimated trends for a shorter 11-year period (2002–2013) corresponding to the period after the launch of the database when observers could enter contemporary records.

From the structured monitoring data, we estimated a population trend for each species using a generalized linear mixed model with a Poisson error distribution and a random 'route' effect to account for spatial and habitat differences at the route level. This approach is the standard analytical procedure for estimating trends from bird count data when no ancillary data (e.g. distance to detected birds, continuous covariates affecting detectability) or repeat visits are available to account for imperfect detection (Kéry & Schaub, 2012; Inger *et al.*, 2014). We implemented the GLMM in R package 'lme4' for each species with the generic formula glmer (Number ~ Year + (Year|Route_ID), family = 'poisson'). 'Year' was fitted as a continuous covariate.

From the unstructured citizen science data, we first extracted only breeding season records (May and June), corresponding to the recording period of the structured monitoring scheme. Every record in DOFbasen includes a location identifier, which served as our definition of a 'site' at which observations of species were recorded, but these sites are not of a defined size, which makes any abundance information difficult to interpret. Every visit to a site that was entered into the online database was treated as a 'list' of species, and species that were not recorded on the list were assigned a value of 0 (not observed or not reported). To estimate population trends from the unstructured data, we used a multi-year occupancy modelling framework to account for imperfect detection or recording (van Strien *et al.*, 2013; Isaac *et al.*, 2014). We used the detection/non-detection information contained in contributed species lists and considered that two covariates influenced the probability of detection: we included 'month' (May or June) as a covariate because for some birds, the probability of detection can vary over the course of the breeding season, and we included the number of species recorded during that visit at that site, as an indication of observer effort and quality (Franklin, 1999; Roberts *et al.*, 2007; Szabo *et al.*, 2010).

We fitted occupancy models using Markov Chain Monte Carlo (MCMC) methods in a Bayesian framework following the approach described by Isaac *et al.* (2014), including a random site effect to account for spatial differences. The trend model component of the multi-year occupancy model was therefore structurally similar to the trend GLMM used for the structured monitoring data described above. For each species, we ran three Markov chains each with 5000 iterations and discarded the first 2500 iterations as burn-in. From the remaining iterations we only used every second iteration for inference. Convergence was tested using the Gelman-Rubin diagnostic (Brooks & Gelman, 1998), and trend estimates were retained only if this diagnostic indicated convergence (R-hat < 1.02). We fitted all occupancy models in JAGS 3.3 (Plummer, 2012) via the R2JAGS package (Su & Yajima, 2012) called from R 3.1.1 (R Development Core Team, 2013).

Comparison of trends derived from structured and unstructured data

We used Spearman's rank correlation coefficient to assess whether population trends derived from the unstructured and the structured data across the 103 species were correlated using a significance threshold of $\alpha = 0.05$, and we performed separate correlations for the long (1986–2013) and the short (2002–2013) time series. Because the correlation does not account for the uncertainty in trend estimates, we also compared the direction of trend estimates between the two data sources taking uncertainty into account. We first classified trend direction as either increasing or decreasing if the 95% confidence interval of the estimated population growth rate was > 0 or < 0. Species for which the 95% confidence interval of the estimated population growth rate spanned 0 were considered to have stable or inconclusive trends. We then cross-tabulated the trend directions from both data sources and calculated the proportion of species that had matching and non-matching trend directions for the period 1986–2013 and 2002–2013.

Examining causes for mismatches in trend direction

Because the structured monitoring data have a higher information content than the unstructured data (abundance vs. detection/non-detection), and because the value of unstructured monitoring data may vary among species (van Strien *et al.*, 2013), we expected some discrepancies among trend estimates and examined whether these were due to the information content of the data or could be explained by species-specific traits such as abundance and migratory strategy.

To examine the information content of data, we first reduced the structured monitoring data to simple detection/non-detection data and estimated trends using a similar GLMM as described above but with a binomial rather than a Poisson error structure, which is analogous to the trend model used for the unstructured data. In a second step, we aimed to increase the quality of the unstructured data by retaining only selected records with high information content. We applied three hierarchical data filters to the unstructured data, discarded all records that did not meet these filtering criteria and estimated trends from the data remaining after each iteration.

The first filter was applied to the number of species recorded during a site visit to increase the likelihood that a list was complete and that the species missing from that list could therefore be considered as not observed in data analysis. We discarded all records that reported only a single species during a visit to one site on one day, and considered the remaining lists 'complete' if the number of species recorded exceeded a threshold that was scaled to the total number of species recorded at a given site to avoid bias due to spatial effects of species richness (Kéry *et al.*, 2010a). We explored three different thresholds, considering lists as 'complete' if the number of recorded species exceeded 5%, 10% or 25% of the cumulative total number of all species ever recorded at a particular site. We explored higher thresholds, but because the number of records that reported > 25% of all known species at a site was very small, it was rarely possible to estimate trends when a higher threshold was chosen.

The second data filtering step considered the number of reported visits to a given site during May and June in one year. Repeat visits during a period of demographic closure are necessary to account for imperfect detection in an occupancy modelling framework. The probability of observing a species increases with the number of visits to a site, and we therefore used thresholds of 3, 5 and 10 visits during the breeding season to include sites in the estimation of occupancy. We eliminated sites that had less than the various thresholds of site visits on the subset of 'complete' lists based on the criteria described in the first filtering step above. The third and final data filtering step considered the number of years during which sites were covered with a sufficient number of visits meeting the criteria for 'complete' lists (Roy *et al.*, 2012; Isaac *et al.*, 2014). Trend estimation is generally more reliable if the same sites are monitored over a longer period of time. We therefore eliminated sites if they had visits with 'complete' lists for < 3, < 5 or < 10 years. These different data filtering rules resulted in a total of 27 combinations (3 thresholds for list length, 3 thresholds for number of visits and 3 thresholds for number of years) of selected data for estimating annual occupancy and population trend for all our target species. For the estimation of trends from 2002 to 2013, we omitted the data filtering step that mandated sites with at least 10 years of monitoring data, because the monitoring interval included only 11 years and very few sites matched this criterion.

Besides the information content of the data, we also examined species-specific biological traits that explained statistical variation in the mismatches between population trend estimates derived from structured and unstructured data. We used the cross-tabulation of matching and non-matching trend estimates over the period 1986–2013 described above and linked this response (match/mismatch) to five explanatory variables: male body mass (as a proxy for body size), national population size in Denmark, breeding system (colonial, semi-colonial and territorial), habitat preference (marine, coastal, inland wetland, boreal and temperate forests, farmland and grassland, habitat generalists) and migration strategy (resident, partial migrant, migrant within Europe, short-distance migrant to North Africa or the Middle East, long-distance migrant to sub-Saharan Africa or Asia). We extracted body mass data, migration strategy and habitat preferences from standard references (Glutz von Blotzheim, 1985–1998; Tucker & Evans, 1997; Snow & Perrins, 1998). Population size was calculated as the geometric mean of the minimum and maximum population estimates for Denmark (BirdLife International, 2004).

We used a machine-learning algorithm based on ensembles of regression trees (RandomForest) to evaluate which of these five variables explained the most variation in mismatching trend estimates (Cutler *et al.*, 2007; Hochachka *et al.*, 2007).We used a random forest procedure with unbiased classification trees based on a conditional inference framework (package 'party' in R 3.1.1; Hothorn *et al.*, 2006) that allows to account for bias in variable importance measures among categorical variables with different numbers of levels (Strobl *et al.*, 2007; Boulesteix *et al.*, 2012). We constructed 1500 classification trees and used a random subset of 64% of the data without replacement to build single trees. We report the relative variable importance as the decrease in model accuracy after permutation scaled to 100% for the most important variable. The accuracy of the random forest model was assessed with a simple confusion matrix of the predicted and actual trend estimate matches.

RESULTS

Based on the structured monitoring for the full 28 years, 60 species showed significant long-term declines in abundance, 26 species increased significantly and the remaining 17 species showed either stable or fluctuating populations without a significantly positive or negative trend (Table 1). For the shorter time period (2002–2013), 48 species declined significantly, 25 species increased significantly and 30 were stable or the trend estimate was too imprecise to assign the trend as increasing or decreasing (Table S2). By contrast, the unstructured data identified only 20 species as declining over 28 years (19 species over 11 years), 49 species as increasing (48 over 11 years) and 34 species (35 over 11 years) that were either stable or where the trend estimate was too imprecise (Tables 1 & S2).

There was a general positive correlation between population trends estimated from structured and unstructured data sources (Table 2, Fig. 2). However, despite the positive correlation between trends derived from structured and unstructured data, the direction of trend estimates matched for < 50% of species when taking the uncertainty in trend estimates into account (Table 1). The majority of species that were in decline based on the structured monitoring were estimated to have a stable or increasing population trend in the unstructured data (Table 1). Conversely, population declines estimated from unstructured data were largely confirmed by the structured monitoring (Table 1).

Reducing the information content of the structured monitoring data to estimate trends in occupancy rather than abundance did not increase the strength of the trend correlation (Table 2, Fig. 2) or the proportion of matching trend directions between structured and unstructured data (Tables 1 & S2 for 2002–2013). Similarly, filtering the unstructured data to retain only data with higher information content did not improve the strength of the correlation; increasingly strict filters led to poorer correlations (Table 2). By contrast, trends in occupancy and abundance derived from the structured monitoring data were strongly correlated ($r_S = 0.87$, Fig. S2), and the trend direction matched for 82.5% of all species (Table S3).

The three most important variables explaining the mismatch of trend estimates between the structured monitoring and unstructured citizen science data were population size, body size and habitat preference (classification success of random forest model = 84.2%). Trend estimates did not match at all for very abundant species (blackbird *Turdus merula*, chaffinch *Fringilla coelebs* and skylark *Alauda arvensis*, all with population sizes > 1 million birds) and matched poorly for relatively small birds, especially in forest and inland wetland habitats (Fig. 3). When the structured monitoring data were reduced to detection/non-detection data, mismatches in trend direction were almost exclusively explained by male body size (classification success = 80.3%), with birds below 500 g body size having generally poorly matching trends, while trend estimates matched well for birds > 500 g. In both analyses, migration strategy and breeding system had no influence on the extent to which trends from the two data sources matched (both < 5% relative variable importance).

Table 1 Number of species with matching and non-matching population trend directions for 103 bird species between 1986 and 2013 in Denmark derived from an unstructured observation database and a structured monitoring scheme using either original abundance data or simple detection/non-detection data (occupancy)

			Unstructured data		
			Decreasing	Stable/inconclusive	Increasing
Structured monitoring	Abundance	Decreasing	16	23	21
		Stable/inconclusive	4	7	6
		Increasing	0	4	22
	Occupancy	Decreasing	18	25	24
		Stable/inconclusive	2	4	1
		Increasing	0	5	24

Trends were considered increasing or decreasing if the 95% confidence interval of the population growth rate estimate was > 0 or < 0, respectively, and stable or inconclusive if the interval spanned 0.

DISCUSSION

Population trends estimated from structured and unstructured data were generally positively correlated, but there was substantial variation among species, and the declines of many common species were not detected with unstructured citizen science data. This pattern was evident regardless of whether we used the abundance information in the structured monitoring data or reduced these data to simple detection/non-detection data. We therefore conclude that structured monitoring programmes are more powerful to detect population declines than unstructured citizen science data.

Many common European bird species are declining (Sanderson *et al.*, 2006; Inger *et al.*, 2014), and range retractions are also common (Balmer *et al.*, 2014). However, more than half of the species that showed significant long-term population declines in both abundance and occupancy based on our structured monitoring data were classified as either stable or even increasing by the unstructured data (Table 1). This discrepancy indicates that caution is needed when using unstructured data for estimating population trends, and that unstructured citizen science data cannot generally replace standardized monitoring schemes. While this mismatch may be explained by factors such as the reporting of complete lists which may not apply to all online databases, we caution that unstructured citizen science data may not fulfil one of the most critical functions of structured monitoring programmes, namely to act as an early warning system that detects population declines, especially of common and widespread species (Inger *et al.*, 2014).

The structured monitoring data yielded similar numbers of species declining, increasing or with stable or inconclusive trend regardless of whether we used abundance data or reduced the information content to use just detection/non-detection data. The mismatching trends derived from structured monitoring data and unstructured citizen science data are therefore not due to the inherently lower information content of unstructured data. Furthermore, our filtering to extract only records with the highest information content from the unstructured citizen science data did not improve the correlation between trend estimates. Stricter filter criteria led to a rapid decline in the amount of data that passed the filter, and trend estimates resulting from these smaller data sets were generally less reliable. Appropriate modelling of the various sources of bias in citizen science data may therefore be the best strategy to derive the most reliable trend information (van Strien *et al.*, 2013; Isaac *et al.*, 2014), but this information is nonetheless inferior to the power of standardized monitoring programmes in detecting species declines.

In situations where structured monitoring is not feasible or too costly, online databases might constitute the only data sources available. While such data can be informative, their value for trend monitoring could be improved by informing contributors about deficiencies (Sullivan *et al.*, 2014). For example, in our data set, the number of visits to sites every year was highly skewed, with some sites receiving > 500 visits per year, and others only single visits. Encouraging recorders to repeatedly contribute data from rarely visited areas might increase the suitability of data for trend analyses using recommended methods that account for imperfect detection (van Strien *et al.*, 2013; Isaac *et al.*, 2014; Isaac & Pocock, 2015). The skewed distribution of visits is likely a result of casual observers frequenting easily accessible, well-known and ‘interesting’ sites (e.g. sites with high diversity or rare species, Tulloch & Szabo, 2012; Tulloch *et al.*, 2013a). One solution to this problem could be to survey regions and habitats that are neglected by casual observers with professional observers (Tulloch *et al.*, 2013a), an approach used during fieldwork of the recent UK breeding bird atlas (Balmer *et al.*, 2014). There are many other approaches for enticing ‘citizen scientists’ to provide data that are of higher information content, but the effort required to do this may be better spent on designing a structured monitoring scheme and recruiting observers to participate in this scheme (Isaac & Pocock, 2015). After all, the structured monitoring data we analysed here were also collected by volunteer ‘citizen scientists’, who follow a certain set of standard protocols which renders trend estimation and inference more reliable.

Table 2 Correlation (Spearman's rank correlation coefficient r_S) between population trend estimates for 103 bird species derived from a structured monitoring scheme and an unstructured observation database filtered by certain criteria over a 28-year period (1986–2013) and an 11-year period (2002–2013) in Denmark

			Abundance		Occupancy	
Proportional list length	*n* visits	*n* years	r_S (28 years)	r_S (11 years)	r_S (28 years)	r_S (11 years)
Unfiltered data			**0.600**	**0.491**	**0.625**	**0.501**
0.25	10	10	0.147		0.175	
		5	0.327	0.184	0.376	0.199
		3	0.355	0.185	0.412	0.202
	5	10	0.353		0.346	
		5	0.314	0.197	0.374	0.121
		3	0.358	0.179	0.432	0.212
	3	10	0.363		0.350	
		5	0.402	0.246	0.459	0.259
		3	0.408	0.357	0.467	0.374
0.1	10	10	0.475		0.484	
		5	0.357	0.191	0.405	0.171
		3	0.388	0.222	0.444	0.211
	5	10	0.413		0.463	
		5	0.424	0.416	0.515	0.353
		3	0.437	0.438	0.510	0.394
	3	10	0.464		0.541	
		5	0.485	0.479	0.550	0.463
		3	0.501	0.491	0.574	0.482
0.05	10	10	0.384		0.444	
		5	0.390	0.348	0.469	0.336
		3	0.409	0.359	0.479	0.355
	5	10	0.470		0.545	
		5	0.485	0.435	0.537	0.419
		3	0.469	0.426	0.528	0.426
	3	10	0.517		0.570	
		5	0.503	0.465	0.567	0.434
		3	0.498	0.453	0.548	0.438

Data from the structured monitoring were either used as counts (*abundance*) or reduced to detection/non-detection (*occupancy*). Filters were applied in a hierarchical fashion based on the number of species recorded during each visit (as proportion of the total species number ever recorded at a given site, *prop. list length*), the number of visits with 'complete' lists in a given year, and the number of years with sufficient visits with 'complete' lists. Strongest correlations are highlighted in bold.

The most important biological variables explaining mismatching population trend estimates were population size for the abundance data and body mass for occupancy. Abundance trends estimated for extremely common and widespread species matched very poorly between the structured and unstructured data, which may be a consequence of the unstructured data yielding only information about occurrence: very common species may experience declines in abundance before completely disappearing from certain areas, and the unstructured data may therefore not be ideal to detect the declines of common species which are currently occurring across Europe (Inger *et al.*, 2014). When we reduced the structured monitoring data to occurrence data, the mismatches persisted but were better explained by male body size, indicating that trends of mostly small birds are very poorly captured by the unstructured citizen science data.

Another potential explanation for differences in trend estimates could be that the different analytical approaches used for both data sources account for variable amounts of uncertainty. In particular, imperfect detection is a well-known problem for the monitoring of wild animals (Royle & Nichols, 2003; Kéry *et al.*, 2009). We used occupancy modelling to correct for varying detection probability in the unstructured data, but we could not apply this method to the structured monitoring data as neither repeated visits nor ancillary data were available from our CBM programme, as is the case for a large number of standardized monitoring programmes (Schmeller *et al.*, 2012). While our method of trend estimation incorporates some uncertainty associated with spatial heterogeneity, the inability to account for imperfect detection may introduce bias into the structured monitoring data if detection probability changes systematically over time (Kéry *et al.*, 2010b).

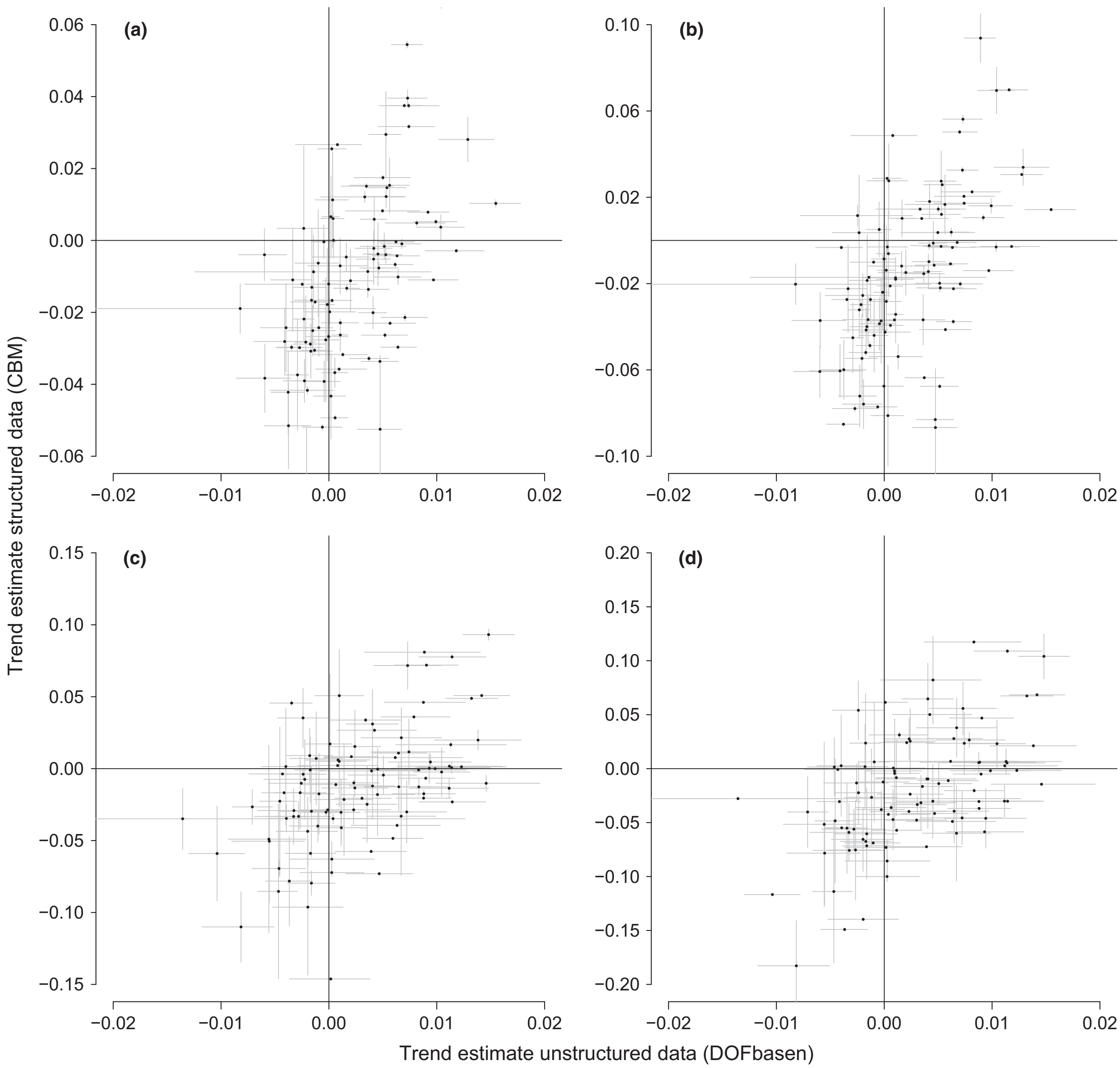

Figure 2 Correlation between population trend estimates (± 95% confidence interval) derived from structured monitoring data (CBM) and from unstructured observation records (DOFbasen) for 103 bird species in Denmark in 1986–2013 (a and b) or 2002–2013 (c and d); trends from structured monitoring data were either based on abundance data (a and c) or reduced to detection/non-detection data (b and d). Note that the scale of axes differs among plots for better clarity.

Non-matching trends from the two data sources could arise either because the unstructured data were inadequate for trend estimation, or the structured monitoring was inadequate for certain species that prefer habitats that are poorly covered by the routes used for the structured monitoring. Structured monitoring schemes are generally designed to cover a large number of common, widespread and territorial species (Newson *et al.*, 2005). Such schemes are therefore often unsuitable for species with localized breeding distributions such as some waterbird species, which may explain the poorly matching trends estimated for species preferring inland wetlands (Fig. 3). Non-matching trends for such species highlight the potential value of unstructured online databases even in countries where structured monitoring schemes exist: casual observations for some species may provide a better basis for population trend estimation than structured monitoring routes that are suboptimal for certain species. However, trends derived from unstructured data would have to be validated with relevant monitoring schemes such as specific wetland bird counts (Zbinden *et al.*, 2014). Identifying the species that are poorly covered by structured monitoring schemes and communicating this knowledge gap to casual observers may enhance the value of data contributed to online databases.

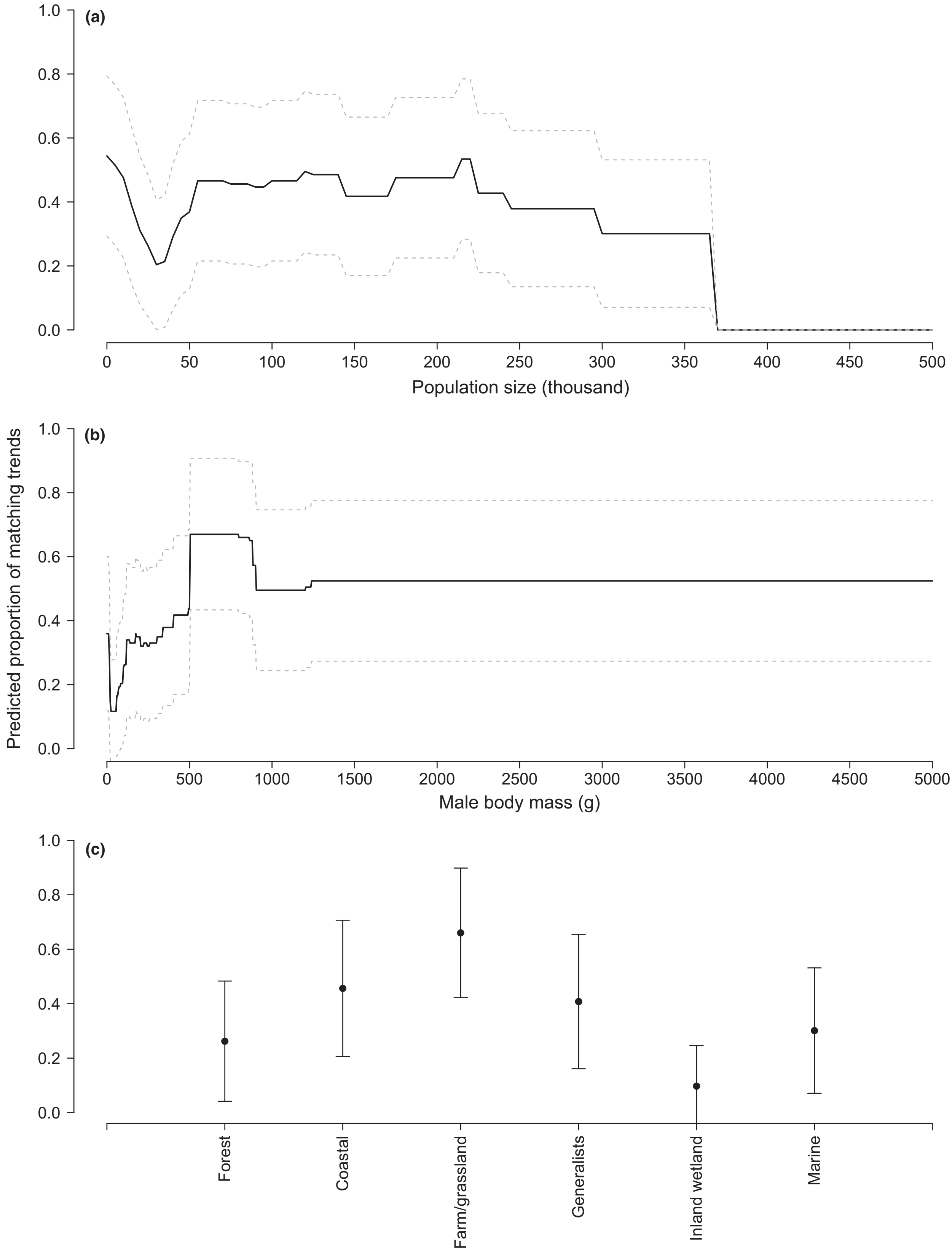

Figure 3 Predicted proportion of matching population trend estimates (± 1 standard deviation) derived from structured monitoring data and from unstructured citizen science data for 103 bird species depending on (a) their total population size in Denmark, (b) their body size measured as male body mass (in g) and (c) their preferred habitat. Predictions are based on a conditional random forest model classifying trend matches.

Further important causes of non-matching trends between structured and unstructured data are changes in reporting behaviour or the observer community over time (Snäll *et al.*, 2011). We found positive trends for several common species in the unstructured data, which declined according to the structured monitoring data. There are two potential explanations for this pattern: (1) the initial contributors to an online reporting scheme are likely to be experienced birdwatchers, who may tend to record mostly those species they consider 'interesting' (Isaac & Pocock, 2015). As a scheme becomes more publicized and widely known, an increasing number of citizens might join who may record also more common and widespread birds (Fig. S1). Our finding that trends did not match for the most common and smallest species is consistent with such an interpretation. (2) In addition to the change in the reporting community, declines of species revealed by structured monitoring schemes might be publicized and lead to a higher awareness among birdwatchers, resulting in changes in reporting behaviour and more contributed records of formerly common and underreported species (Snäll *et al.*, 2011). Examples in our data that are consistent with such explanations include the House Sparrow (*Passer domesticus*; widespread, heavily publicized declines of a familiar urban bird, Hole *et al.*, 2002; De Laet & Summers-Smith, 2007) and the Willow Warbler (*Phylloscopus trochilus*; flagship species for a suite of declining long-distance migrants, Morrison *et al.*, 2010). The best solution to tackle reporting bias is to offer recorders the possibility to submit 'complete' checklists, that is lists that contain all species recorded and allow inference about species that were not detected (Sullivan *et al.*, 2009; Kéry *et al.*, 2010a; van Strien *et al.*, 2013). This feature, which was absent from the Danish online database when we conducted our analysis, has in the meantime been launched in Denmark and other online databases and is considered a standard solution to address some of the biases inherent in citizen science data (Isaac & Pocock, 2015).

CONCLUSIONS

Our analyses suggest that citizen science data collected using unstructured methods may be useful for biodiversity monitoring for species or in areas where dedicated, structured survey data are not available, but that various sources of bias need to be considered in the interpretation of population trend estimates. We recommend retaining all data for analysis and encouraging database managers to distinguish between the reporting of complete and incomplete lists. We suggest that in countries currently without dedicated monitoring systems, encouraging observers to submit records to online databases could make a useful contribution to the monitoring of biodiversity. In countries where structured monitoring data are available, unstructured databases may play a useful role in public education and monitoring of areas or species that cannot be covered with a structured approach. However, our results warn against abandoning existing structured monitoring schemes in the hope that unstructured data contributed by volunteers would be able to fulfil the same purpose with the same power and precision.

ACKNOWLEDGEMENTS

We want to thank all field observers who spent considerable time and effort to enter their observations into DOFbasen, as well as all volunteers participating in the Danish Common Bird Census scheme. We appreciate the support and advice of Marc Kéry, Nick Isaac, Tom August, Arco van Strien and Adam Butler in implementing multi-year occupancy models. Ian Fisher assisted with processing DOFbasen data. The comments of Alison Johnston, Wesley Hochachka, Mark Eaton, Danaë Sheehan, Richard Gregory, Graeme Buchanan, Alison Beresford, Rob Robinson, Boris Schröder and an anonymous reviewer improved the manuscript considerably. The European Commission is acknowledged for financial support as part of the EuroGEOSS project (Grant Agreement 226487).

REFERENCES

Balmer, D., Gillings, S., Caffrey, B., Swann, B., Downie, I. & Fuller, R. (2014) *Bird Atlas 2007–11: the breeding and wintering birds of Britain and Ireland*. HarperCollins, Glasgow, UK..

BirdLife International (2004) *Birds in Europe: population estimates, trends and conservation status*. BirdLife International, Cambridge.

Boakes, E.H., McGowan, P.J.K., Fuller, R.A., Chang-qing, D., Clark, N.E., O'Connor, K. & Mace, G.M. (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology*, **8**, e1000385.

Boulesteix, A.L., Janitza, S., Kruppa, J. & König, I.R. (2012) *Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics*. Technical Report 129, pp. 1–31. Department of Statistics, University of Munich, Munich, Germany.

Brooks, S.P. & Gelman, A. (1998) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, **7**, 434–455.

Cutler, D.R., Edwards, T.C. Jr, Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007) Random forests for classification in ecology. *Ecology*, **88**, 2783–2792.

De Laet, J. & Summers-Smith, J. (2007) The status of the urban house sparrow *Passer domesticus* in north-western Europe: a review. *Journal of Ornithology*, **148**, 275–278.

Devictor, V., Whittaker, R.J. & Beltrame, C. (2010) Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, **16**, 354–362.

Dickinson, J., Zuckerberg, B. & Bonter, D.N. (2010) Citizen science as an ecological research tool: challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 149–172.

Franklin, D.C. (1999) Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. *Biological Conservation*, **90**, 53–68.

Glutz von Blotzheim, U.N. ed. (1985–1998). *Handbuch der Vögel Mitteleuropas*. Aula, Wiesbaden.

Greenwood, J.J.D. (2003) The monitoring of British breeding birds: a success story for conservation science? *Science of the Total Environment*, **310**, 221–230.

Gregory, R.D., Gibbons, D.W. & Donald, P.F. (2004) Bird census and survey techniques. *Bird ecology and conservation: a handbook of techniques* (ed. by W.J. Sutherland, I. Newton and R.E. Green), pp. 17–55. Cambridge University Press, Cambridge, UK.

Heldbjerg, H., Brandtberg, N. & Jørgensen, M.F. (2014) Overvågning af de almindelige fuglearter i Danmark 1975-2013. Årsrapport for Punkttællingsprogrammet. Dansk Ornitologisk Forening, Copenhagen.

Hochachka, W.M., Caruana, R., Fink, D., Munson, A.R.T., Riedewald, M., Sorokina, D. & Kelling, S. (2007) Data-mining discovery of pattern and process in ecological systems. *Journal of Wildlife Management*, **71**, 2427–2437.

Hochachka, W.M., Fink, D., Hutchinson, R.A., Sheldon, D., Wong, W.-K. & Kelling, S. (2012) Data-intensive science applied to broad-scale citizen science. *Trends in Ecology and Evolution*, **27**, 130–137.

Hole, D.G., Whittingham, M.J., Bradbury, R.B., Anderson, G.Q., Lee, P.L., Wilson, J.D. & Krebs, J.R. (2002) Widespread local house-sparrow extinctions. *Nature*, **418**, 931–932.

Hothorn, T., Hornik, K. & Zeileis, A. (2006) *party: A Laboratory for Recursive Part(y)itioning*. Available at: http://CRAN.R-project.org/ (accessed 13 December 2014).

Inger, R., Gregory, R., Duffy, J.P., Stott, I., Voříšek, P. & Gaston, K.J. (2014) Common European birds are declining rapidly while less abundant species' numbers are rising. *Ecology Letters*, **18**, 28–36.

Isaac, N.J.B. & Pocock, M.J.O. (2015) Bias and information in biological records. *Biological Journal of the Linnean Society*, **115**, 522–531.

Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P. & Roy, D.B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, **5**, 1052–1060.

Johnston, A., Fink, D., Reynolds, M.D., Hochachka, W.M., Sullivan, B.L., Bruns, N.E., Hallstein, E., Merrifield, M.S., Matsumoto, S. & Kelling, S. (2015) Abundance models improve spatial and temporal prioritization of conservation resources. *Ecological Applications*, **25**, 1749–1756.

Kéry, M. & Schaub, M. (2012) *Bayesian population analysis using WinBUGS: a hierarchical perspective*. Academic Press, Oxford, UK.

Kéry, M., Dorazio, R.M., Soldaat, L., van Strien, A., Zuiderwijk, A. & Royle, J.A. (2009) Trend estimation in populations with imperfect detection. *Journal of Applied Ecology*, **46**, 1163–1172.

Kéry, M., Gardner, B. & Monnerat, C. (2010a) Predicting species distributions from checklist data using site-occupancy models. *Journal of Biogeography*, **37**, 1851–1862.

Kéry, M., Royle, J.A., Schmid, H., Schaub, M., Volet, B., Hafliger, G. & Zbinden, N. (2010b) Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. *Conservation Biology*, **24**, 1388–1397.

Kindberg, J., Ericsson, G. & Swenson, J.E. (2009) Monitoring rare or elusive large mammals using effort-corrected voluntary observers. *Biological Conservation*, **142**, 159–165.

Lin, Y.-P., Deng, D., Lin, W.-C., Lemmens, R., Crossman, N.D., Henle, K. & Schmeller, D.S. (2015) Uncertainty analysis of crowd-sourced and professionally collected field data used in species distribution models of Taiwanese moths. *Biological Conservation*, **181**, 102–110.

Morrison, C.A., Robinson, R.A., Clark, J.A. & Gill, J.A. (2010) Spatial and temporal variation in population trends in a long-distance migratory bird. *Diversity and Distributions*, **16**, 620–627.

Newson, S.E., Woodburn, R.J., Noble, D.G., Baillie, S.R. & Gregory, R.D. (2005) Evaluating the Breeding Bird Survey for producing national population size and density estimates. *Bird Study*, **52**, 42–54.

Nyegaard, T., Heldbjerg, H. & Brølling, S. (2012) DOFbasen 10 years. *Bird Census News*, **25**, 13–16.

Plummer, M. (2012) JAGS version 3.3.0. Available at: http://sourceforge.net/projects/mcmc-jags/files/Manuals/ (accessed 12 August 2013).

R Development Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://cran.r-project.org/.

Roberts, R.L., Donald, P.F. & Green, R.E. (2007) Using simple species lists to monitor trends in animal populations: new methods and a comparison with independent data. *Animal Conservation*, **10**, 332–339.

Roy, H.E., Adriaens, T., Isaac, N.J., Kenis, M., Onkelinx, T., Martin, G.S., Brown, P.M., Hautier, L., Poland, R. & Roy, D.B. (2012) Invasive alien predator causes rapid declines of native European ladybirds. *Diversity and Distributions*, **18**, 717–725.

Royle, J.A. & Nichols, J.D. (2003) Estimating abundance from repeated presence-absence data or point counts. *Ecology*, **84**, 777–790.

Sanderson, F.J., Donald, P.F., Pain, D.J., Burfield, I.J. & van Bommel, F.P.J. (2006) Long-term population declines in Afro-Palaearctic migrant birds. *Biological Conservation*, **131**, 93–105.

Schmeller, D.S., Henry, P.-Y., Julliard, R., Gruber, B., Clobert, J., Dziock, F., Lengyel, S., Nowicki, P., Deri, E. & Budrys, E. (2009) Advantages of volunteer-based biodiversity monitoring in Europe. *Conservation Biology*, **23**, 307–316.

Schmeller, D., Henle, K., Loyau, A., Besnard, A. & Henry, P.-Y. (2012) Bird-monitoring in Europe – a first overview of practices, motivations and aims. *Nature Conservation*, **2**, 41–57.

Snäll, T., Kindvall, O., Nilsson, J. & Part, T. (2011) Evaluating citizen-based presence data for bird monitoring. *Biological Conservation*, **144**, 804–810.

Snow, D. & Perrins, C. (1998) *The birds of the western palearctic: concise edition*. Oxford University Press, Oxford.

van Strien, A.J., Termaat, T., Groenendijk, D., Mensing, V. & Kéry, M. (2010) Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists. *Basic and Applied Ecology*, **11**, 495–503.

van Strien, A.J., van Swaay, C.A.M. & Termaat, T. (2013) Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, **50**, 1450–1458.

Strobl, C., Boulesteix, A.-L., Zeileis, A. & Hothorn, T. (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, **8**, 25–45.

Su, Y.-S. & Yajima, M. (2012) *R2jags: A package for running jags from R. version 0.03-08*. Available at: http://CRAN.R-project.org/package=R2jags (accessed 12 August 2013).

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D. & Kelling, S. (2009) eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation*, **142**, 2282–2292.

Sullivan, B.L., Aycrigg, J.L., Barry, J.H. *et al.* (2014) The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation*, **169**, 31–40.

Szabo, J.K., Vesk, P.A., Baxter, P.W.J. & Possingham, H.P. (2010) Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis. *Ecological Applications*, **20**, 2157–2169.

Szabo, J.K., Fuller, R.A. & Possingham, H.P. (2012) A comparison of estimates of relative abundance from a weakly structured mass-participation bird atlas survey and a robustly designed monitoring scheme. *Ibis*, **154**, 468–479.

Theobald, E.J., Ettinger, A.K., Burgess, H.K., DeBey, L.B., Schmidt, N.R., Froehlich, H.E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M.A. & Parrish, J.K. (2015) Global change and local solutions: tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, **181**, 236–244.

Tucker, G.M. & Evans, M.I. (1997) *Habitats for birds in Europe: a conservation strategy for the wider environment*. BirdLife International, Cambridge, UK.

Tulloch, A.I. & Szabo, J.K. (2012) A behavioural ecology approach to understand volunteer surveying for citizen science datasets. *Emu*, **112**, 313–325.

Tulloch, A.I., Mustin, K., Possingham, H.P., Szabo, J.K. & Wilson, K.A. (2013a) To boldly go where no volunteer has gone before: predicting volunteer activity to prioritize surveys at the landscape scale. *Diversity and Distributions*, **19**, 465–480.

Tulloch, A.I., Possingham, H.P., Joseph, L.N., Szabo, J. & Martin, T.G. (2013b) Realising the full potential of citizen science monitoring programs. *Biological Conservation*, **165**, 128–138.

Zbinden, N., Kéry, M., Häfliger, G., Schmid, H. & Keller, V. (2014) A resampling-based method for effort correction in abundance trend analyses from opportunistic biological records. *Bird Study*, **61**, 506–517.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1 Number of records in DOFbasen from 1986 to 2013.

Figure S2 Correlation between population trends from occupancy and abundance data.

Table S1 List of all species and trend estimates derived from structured and unstructured monitoring data.

Table S2 Number of species with matching population trends between 2002 and 2013.

Table S3 Number of species with matching population trends derived from structured occupancy and abundance data.

BIOSKETCH

Johannes Kamp, **Steffen Oppel** and **Paul Donald** are conservation scientists at the University of Münster and the RSPB Centre for Conservation Science. Their work focuses on monitoring populations of threatened species, identifying causes for decline and developing conservation solutions.

Henning Heldbjerg and **Timme Nyegaard** are conservation scientists coordinating citizen science based monitoring of bird populations as well as identifying and communicating causes for population changes.

Author contributions: J.K. and P.D. conceived the idea and coordinated the project. H.H. and T.N. coordinated surveys and accumulated and managed the data. S.O. analysed the data. J.K. and S.O. wrote the manuscript, and all authors contributed to the text.

Editor: Boris Schröder