

ORIGINAL ARTICLE

Long-term effective population size dynamics of an intensively monitored vertebrate population

A-K Mueller, N Chakarov², O Krüger and JI Hoffman

Long-term genetic data from intensively monitored natural populations are important for understanding how effective population sizes (N_e) can vary over time. We therefore genotyped 1622 common buzzard (*Buteo buteo*) chicks sampled over 12 consecutive years (2002–2013 inclusive) at 15 microsatellite loci. This data set allowed us to both compare single-sample with temporal approaches and explore temporal patterns in the effective number of parents that produced each cohort in relation to the observed population dynamics. We found reasonable consistency between linkage disequilibrium-based single-sample and temporal estimators, particularly during the latter half of the study, but no clear relationship between annual N_e estimates (\hat{N}_e) and census sizes. We also documented a 14-fold increase in \hat{N}_e between 2008 and 2011, a period during which the census size doubled, probably reflecting a combination of higher adult survival and immigration from further afield. Our study thus reveals appreciable temporal heterogeneity in the effective population size of a natural vertebrate population, confirms the need for long-term studies and cautions against drawing conclusions from a single sample.

Heredity (2016) **117**, 290–299; doi:10.1038/hdy.2016.67; published online 24 August 2016

INTRODUCTION

One of the most fundamental and oldest objectives of ecology is to understand the factors and mechanisms causing fluctuations in the number of animals in a given area (Elton, 1942; Andrewartha and Birch, 1954; Lack, 1954; Royama, 1992; Coulson *et al.*, 2004). This search for underlying mechanisms has recently acquired a special importance as many natural populations have suffered dramatic declines (Hunter, 2002) because of persecution, exploitation or habitat loss (Beissinger and Snyder, 1992; Casey and Myers, 1998; Krüger *et al.*, 2001, 2010; Ferrer *et al.*, 2003; Pimm *et al.*, 2014).

Traditionally, animal populations have been assessed by means of direct census counts for conservation and management purposes (Luikart *et al.*, 2010). However, an influential contribution of evolutionary theory to conservation biology has been the development of a framework for predicting the fate of small populations (Palstra and Ruzzante, 2008). Central parameters to this framework are the population census size (N_c) and effective population size (N_e). N_e , which is defined as the number of individuals in an ideal population experiencing the same rate of random genetic change over time as the actual population (Crow *et al.*, 1970), is a particularly important quantity as it is inversely proportional to the loss of genetic diversity because of inbreeding and genetic drift in finite, randomly mating populations (Nunney and Elam, 1994; Frankham, 2005; Charlesworth, 2009). Thus, low N_e values are often interpreted as providing an indication of increased extinction risk (Newman and Pilson, 1997).

Although ascertaining N_e is crucial to conservation and management, it is still rather difficult to obtain reliable estimates in real populations because of constraints on collecting enough demographic data to directly measure N_e (Waples, 2005; Luikart *et al.*, 2010).

Consequently, molecular genetic approaches have become increasingly popular in recent years (Waples and Do, 2010). A variety of different approaches are now available to estimate N_e from genetic marker data (reviewed by Leberg, 2005; Wang, 2005; Palstra and Ruzzante, 2008; Luikart *et al.*, 2010). These can be broadly classified into single-sample and temporal approaches. The former estimate the effective population size from properties of a single sample of individuals, most often using the unbiased linkage disequilibrium (LD) method of Hill (1981) that is based on the premise that in small populations with few parents, random genetic drift generates nonrandom associations between alleles at different loci. In contrast, temporal approaches exploit the fact that drift is higher in smaller populations by quantifying differences in allele frequencies between two or more samples collected at different time points. Applied to discrete cohorts of offspring, these approaches strictly estimate the effective number of parents that produced the sample, or N_b (Schwartz *et al.*, 1999; Leberg, 2005; Waples and Do 2010; Waples and England, 2011; Waples *et al.*, 2014).

Both sets of approaches have advantages and disadvantages. Temporal approaches are often considered to be the most accurate, assuming that samples can be gathered sufficiently far apart in time (Leberg, 2005; Luikart *et al.*, 2010). However, they require at least two non-overlapping generations to be sampled, and this is often impossible for species with long generation times such as many large vertebrates (Luikart *et al.*, 2010). Single-sample approaches have therefore increased in popularity as they can provide a snapshot of a population without the need to collect multiple samples. Precision also appears to be reasonably high, at least for the LD method, when using the 10–20 genetic markers and ~ 50 samples typical of most studies, as

Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany

Correspondence: A-K Mueller, Department of Animal Behaviour, Bielefeld University, Morgenbreede 45, Konsequenz 45, 33501 Bielefeld, Germany.

E-mail: annakatharinamueller@web.de

²Current address: Molecular Ecology and Evolution Lab, Lund University, Lund, Sweden.

Received 31 October 2015; revised 2 June 2016; accepted 3 June 2016; published online 24 August 2016

long as N_e is less than ~ 500 (Waples and Do, 2010). However, single-sample approaches can be sensitive to gene flow (Gilbert and Whitlock, 2015) and there is also a need for studies that are long enough to quantify temporal variability when applying temporal methods (Palstra and Ruzzante, 2008).

A long-term study of common buzzards (*Buteo buteo*) provides an excellent opportunity to explore the temporal dynamics of the effective population size of an intensively studied vertebrate population. This species is a common, medium-sized bird of prey that breeds across the entire Palearctic from southwestern Europe to Japan, where it preys mainly on microtine rodents (del Hoyo *et al.*, 1994). A population in Eastern Westphalia, Germany, has been intensively monitored since 1989, with nests having been climbed and all chicks marked to allow individual recognition since 2002. Our census data suggest that this population was relatively stable at ~ 200 breeding adults since the early 2000s, but began to increase in 2010 towards a peak of 516 in 2012 (see Results).

Common buzzards also exhibit a striking plumage polymorphism (Cramp and Simmons, 1980; Krüger *et al.*, 2001) that serves as a phenotypic marker and thus provides an interesting additional dimension (Chakarov *et al.*, 2008; Chakarov *et al.*, 2013). Three colour morphs have been described that differ in their levels of plumage melanisation, termed 'light', 'intermediate' and 'dark'. Plumage morph is fixed throughout an individual's lifetime, follows

Mendelian expectations for a single locus with two alleles (Krüger *et al.*, 2001) and is associated with a mutation in the melanocortin-1 receptor (MA Pointer *et al.*, unpublished data). The intermediately melanised morph, which is presumed to be heterozygous at the colour locus, is on average longer lived and has the greatest lifetime reproductive success (Chakarov *et al.*, 2008). However, the plumage morphs do not differ significantly in genome-wide heterozygosity measured using 18 microsatellites (Boerner *et al.*, 2013), suggesting that heterozygote advantage is not a genome-wide phenomenon. The three morphs are also genetically undifferentiated, despite buzzards tending to mate with partners of the same morph as their mother (Boerner *et al.*, 2013).

Here, we genotyped 1622 common buzzard chicks comprising 12 complete cohorts at 15 polymorphic microsatellite loci. The resulting data were used to evaluate the comparability of single-sample and temporal estimators, as well as to explore temporal patterns in relation to the observed demography of the population.

MATERIALS AND METHODS

Study site

This study was conducted in a 300 km² study area (8°25'E and 52°6'N) in Eastern Westphalia, Germany (Figure 1). It consists of two 125 km² grid squares and 50 km² of edge areas. The dominant land cover in the study area is the Teutoburger Wald, a low-level forested mountain region reaching a height

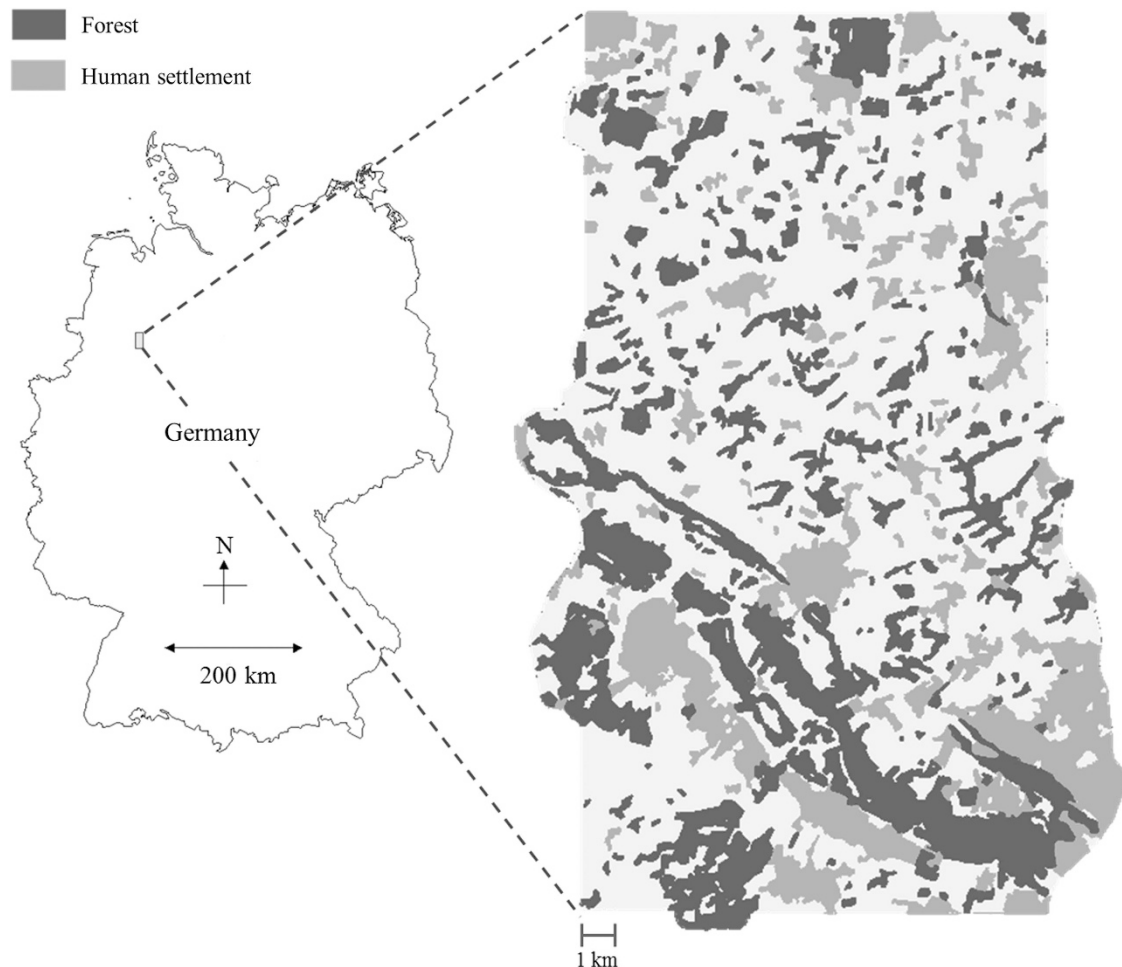


Figure 1 Location of the study area in Germany together with a detailed map (insert) of the 300 km² study area, with human settlements shown in orange, forest patches in green and agricultural areas in cream. A full colour version of this figure is available at the *Heredity* journal online.

of 315 m above sea level. The second most abundant land cover type is a cultivated landscape to the north and south.

Collection of census and breeding success data

Census dynamics of common buzzards were monitored from 1989 to 2013 inclusive. All forest patches were visited in late winter to look for territorial pairs and nests. During the breeding season, these forest patches were visited again and checked for activity of the study species. This includes breeding pairs (occupying a nest and showing signs of egg-laying activity) as well as nonbreeding pairs that just occupy a territory. Hence, our buzzard census data include breeding and nonbreeding pairs but the number of ‘floating’ individuals that do not hold a territory cannot be reliably counted and recorded.

Each active nest was visited at least 3 and up to 10 times a year to determine breeding success (success or failure) and brood size (number of chicks fledged) for successful breeding attempts. From 1989 to 2001 inclusive, data were collected through careful and intensive observation from the ground. In subsequent years, between 85 and 99% of all successful nests were climbed and the chicks were ringed, normally in late May and early June.

Blood sampling

Buzzard nests were climbed with a rope-climbing technique and once the climber was at the nest, chicks were lowered to the ground using another rope. On the ground, they were ringed, biometric measures were recorded and a 0.5 ml blood sample was taken from the brachial vein with a syringe or needles and capillaries. Sample sizes are shown in Table 1. Blood was transferred into 1.5 ml screw-cap tubes filled with 1.0 ml ethanol or phosphate-buffered saline–EDTA buffer. Back at the laboratory, all tubes were stored at –20 °C.

Microsatellite genotyping

Total genomic DNA was extracted from 10–20 µl of each sample using a standard chloroform extraction protocol and genotyped at 15 previously developed microsatellite loci (Johnson *et al.*, 2005). All but two of these loci map to different loci in the zebra finch (*Taeniopygia guttata*) and are therefore unlikely to be physically linked (Table 2). The microsatellites were PCR amplified in a single multiplexed reaction using a Type It Kit (Qiagen GmbH, Hilden, NW, Germany). The following PCR profile was used: one cycle of 5 min at 94 °C; 24 cycles of 30 s at 94 °C, 90 s at 56 °C and 30 s at 72 °C; and one final cycle of 15 min at 72 °C. Fluorescently labelled PCR products were then resolved by electrophoresis on an ABI 3730xl capillary sequencer (Applied Biosystems, Carlsbad, CA, USA) and allele sizes were scored automatically using GeneMarker version 2.6.2 (Softgenetics, State College, PA, USA). To ensure high genotype quality, all traces were manually inspected and any obvious scoring errors were adjusted accordingly.

Genetic data analyses

Tests for deviation from Hardy–Weinberg equilibrium and LD were implemented using Genepop version 4.3 (Rousset, 2008), specifying 10 000 dememorisations, 1000 batches and 10 000 iterations per batch. Adjustment of *P*-values for the false discovery rate with an α -level of 0.05 was carried out on all tabulated results using the program *q*-value version 1.38.0 (Storey, 2002). Genepop was also used to calculate observed and expected heterozygosities at each of the microsatellite loci.

To test for population structure, we used two complementary approaches. First, Structure version 2.3.4 (Pritchard *et al.*, 2000) was used to test for the presence of distinct genetic clusters without prior knowledge of the sampling locations of individuals. This program uses a maximum-likelihood approach to determine the most likely number of genetically distinct clusters in a sample (*K*) by subdividing the data set in a way that maximises Hardy–Weinberg equilibrium and minimises LD within the resulting clusters. We ran 20 independent runs for *K* = 1–10 using 1 000 000 Markov chain Monte Carlo iterations after a burn-in of 500 000 with the correlated allele frequencies model and assuming admixture. The most likely number of groups was evaluated using the maximal average value of Ln *P*(*D*), a model choice criterion that estimates the posterior probability of the data. Second, we used hierarchical

analyses of molecular variance within GenALEX version 6.5 (Peakall and Smouse, 2012) to test for genetic differences between buzzards sampled to the north and south of the ridge of the Teutoberger Wald.

Table 1 Sample sizes of common buzzards with the number of analysed samples (*N*), census population size (*N*_c), number of breeders (*N*_b), estimated effective number of parents that produced each cohort (\hat{N}_e), minimal allele frequency cutoff (*P*_{crit}) and 95% confidence intervals (CIs)

	Sample size	<i>N</i> _b	<i>N</i> _c	<i>P</i> _{crit}	\hat{N}_e	Lower 95% CI	Upper 95% CI
Dark	159	79	159	0.01	101.4	82.1	128.3
				0.02	99.0	78.5	128.5
				0.05	83.1	63.9	111.6
Intermediate	739	268	739	0.01	199.0	176.7	224.7
				0.02	231.7	201.9	267.3
				0.05	260.6	214.1	320.8
Light	502	208	502	0.01	177.9	155.3	204.7
				0.02	158.9	137.2	184.9
				0.05	179.8	146.3	223.9
2002	79	110	79	0.01	125.7	84.6	220.2
				0.02	263.2	132.4	1997.3
				0.05	180.6	95	765.3
2003	109	140	109	0.01	107.8	80.4	154.2
				0.02	95.1	71.3	134.4
				0.05	113.1	78.1	183.1
2004	83	164	83	0.01	65.2	49.9	89.1
				0.02	59.6	46	80.3
				0.05	51.5	37.6	74.1
2005	180	200	180	0.01	64	56.8	72.1
				0.02	62.4	54.5	71.5
				0.05	70.3	59.2	84.2
2006	23	48	23	0.01	19.4	14.5	26.9
				0.015	19.4	14.5	26.9
				0.02	21.3	15.7	30.1
2007	39	46	39	0.01	15.9	11.7	22.4
				0.013	37.8	26.4	59.6
				0.02	40.1	27.6	64.7
2008	38	66	38	0.01	29.5	21.5	42.8
				0.02	25.5	18.1	38.1
				0.05	21.9	17.3	28.4
2009	34	54	34	0.013	21.9	17.3	28.4
				0.02	25.5	19	35.6
				0.05	48	29.6	105.4
2010	298	228	298	0.01	53.6	33.3	110.7
				0.015	27.3	20.2	38.9
				0.02	45.6	29.2	86.3
2011	138	182	138	0.05	80	37.1	566.1
				0.01	141.7	118.4	171.9
				0.02	127	105	155.6
2012	262	220	262	0.05	120.2	95.5	154.1
				0.01	517	256.9	5071.1
				0.02	367.2	202.3	1286.9
2013	126	148	126	0.05	236.4	134.7	645.7
				0.01	288.8	213.6	421.8
				0.02	286.2	208.8	427.1
				0.05	554.4	298.3	2097.2
				0.01	152.8	108	241.9
				0.02	110	80.6	161.6
				0.05	94.5	66.8	145.4

Table 2 Details of the 12 microsatellite loci used in this study together with their polymorphism characteristics in 1419 common buzzards

Locus	Chromosome in		Observed heterozygosity	Expected heterozygosity
	<i>Taeniopygia guttata</i>	Allele number		
Bbu03	15	8	0.70	0.73
Bbu06	5	7	0.50	0.53
Bbu11	5	10	0.49	0.50
Bbu14	7	12	0.55	0.59
Bbu17	1	9	0.55	0.57
Bbu22	1	4	0.13	0.13
Bbu30	2	5	0.23	0.23
Bbu33	3	10	0.79	0.82
Bbu34	1	17	0.81	0.83
Bbu38	2	5	0.11	0.11
Bbu42	4	25	0.87	0.92
Bbu53	19	4	0.23	0.26

Single-sample estimators

We used the software NeEstimator version 2.01 (Do *et al.*, 2014) to implement the LD approach of Waples and Do (2008). As the inclusion of rare alleles can upwardly bias LD-based estimates (Waples and Do, 2010), we followed the authors' recommendation of choosing the minor allele frequency threshold (P_{crit}) to be the larger of 0.02 or a value that screens out single copy alleles (Waples and Do, 2010). To explore sensitivity, we also repeated the analysis with $P_{crit} = 0.01$ and 0.05 while similarly applying the same criterion as above to ensure that single copy alleles were not counted. The 95% confidence intervals (CIs) were derived using the 'parametric' option that implements χ^2 approximation (Waples, 2006).

Temporal estimators

We used the temporal approach of Wang and Whitlock (2003) as this is the least sensitive of the temporal estimators to immigration (Gilbert and Whitlock, 2015). The software MLNe 1.0 (Wang and Whitlock, 2003) was used to generate both the maximum-likelihood and moment estimators, assuming that the population is not at equilibrium. For this analysis, we specified a maximum N_e of 5000, a monitor value of four and six threads. We estimated the generation time as the mean age at maturity plus the mean reproductive lifespan (IUCN Standards and Petitions Subcommittee, 2016). As the breeding lifespan of an adult female buzzard is on average 2.75 years and most buzzards recruit as breeding adults at ~2 years of age (Krüger and Lindström, 2001), our estimate is 4.75 years. For the temporal analysis, we therefore took the allele frequencies from 2002 and 2013 and assumed two generations between samples. To explore sensitivity to the number of generations assumed to separate the samples, we also repeated this analysis specifying between one and five generations.

RESULTS

We genotyped 1622 common buzzard individuals sampled over 12 years at 15 microsatellite loci. Individuals that failed to produce interpretable genotypes at three or more loci were discarded, leaving a total of 1419 individuals for the data analysis (Table 1). The loci were moderately variable, carrying an average of 9.75 alleles (Table 2). Tests for deviations of each locus from Hardy–Weinberg equilibrium in each of the years revealed the number of deviations that remained significant following table-wide false discovery rate correction for multiple tests (Supplementary Table S1). Two of the loci (Bbu 35 and 46) deviated significantly from Hardy–Weinberg equilibrium in 9 out of 12 years and a further locus (Bbu 26) deviated significantly in 7 years. We therefore took the conservative measure of excluding all three of these loci from further analyses. Tests for LD among the

remaining 12 loci revealed a number of significant associations (Figure 2). However, no pairs of loci were consistently in LD across multiple years, the extent of LD varied from year to year (Figure 2) and all but two of the loci mapped to different chromosomes in the zebra finch (Table 2), suggesting that these associations are unlikely to be due to physical linkage.

Buzzard morph effective population sizes

To test for differences in the effective population sizes of the three buzzard plumage morphs, we implemented the LD method (Hill, 1981; Waples, 2006; Waples and Do, 2010). For this analysis, it was necessary to pool individuals across years because of the low frequency of the dark morph. Sensitivity to the minor allele frequency cutoff was analysed by generating effective population size estimates for $P_{crit} = 0.01$, 0.02 and 0.05. The resulting estimates, which were reasonably robust to the P_{crit} value used, were lowest for the dark morph, intermediate for the light morph and highest for the intermediate morph (Figure 3 and Table 1), reflecting their frequencies in the wider population.

Temporal estimators

Based on allele frequencies from 2002 and 2013 and assuming two generations between the samples, MLNE produced likelihood and moment-based estimates of 185.7 and 122.5 respectively (Figure 4 and Table 1). Exploring sensitivity to the number of generations assumed to separate the samples, we found that \hat{N}_e increased gradually from one to five generations (Figure 4).

Temporal patterns

With a 12-year time series and the majority of breeding buzzards observed and their offspring sampled, we could explore temporal changes in \hat{N}_e in relation to the observed dynamics of the study population (Figure 5 and Table 1). The results were reasonably insensitive to P_{crit} and showed appreciable variation over the course of the study. In particular, \hat{N}_e was typically in the order of 25–100 for the period leading up to and including 2009, then increased towards a peak of ~250–500 in 2011–2012. This increase broadly coincides with a period of rapid population growth in which N_c more than doubled between 2009 and 2012 (Figure 5b and Table 1) and the frequency of the light morph also increased (Figure 5c; $F_{1,8} = 26.26$, $P < 0.001$ for the period 2002–2013 inclusive).

For every year of the study, \hat{N}_e was smaller than the corresponding number of breeders (N_b) that, in turn, was smaller than the census size (N_c). This is reflected in the N_b/N_c ratios shown in Figure 5d that were consistently < 1 and fell to between 0.83 and 0.55 for the period 2004–2009, when the population experienced particularly low breeding success. Consequently, the \hat{N}_e/N_c and \hat{N}_e/N_b ratios (shown for $P_{crit} = 0.02$ in Figure 5d) were also at their lowest during this period. Analysing all of the years together, we observed positive but nonsignificant correlations between \hat{N}_e (based on $P_{crit} = 0.02$) and both N_c ($F_{1,8} = 1.93$, $P = 0.20$) and N_b ($F_{1,8} = 2.14$, $P = 0.18$). The ratio of \hat{N}_e to N_c did not correlate significantly with N_c ($F_{1,8} = 0.10$, $P = 0.76$).

Population structure

As population structure can cause inaccuracies in N_e estimation (Waples and England, 2011; Neel *et al.*, 2013; Gilbert and Whitlock, 2015), we used two approaches to test for the presence of cryptic population structure within the study area. Arguably the most versatile tests of population structure need not rely on knowledge of where individuals were sampled. Consequently, we first implemented a

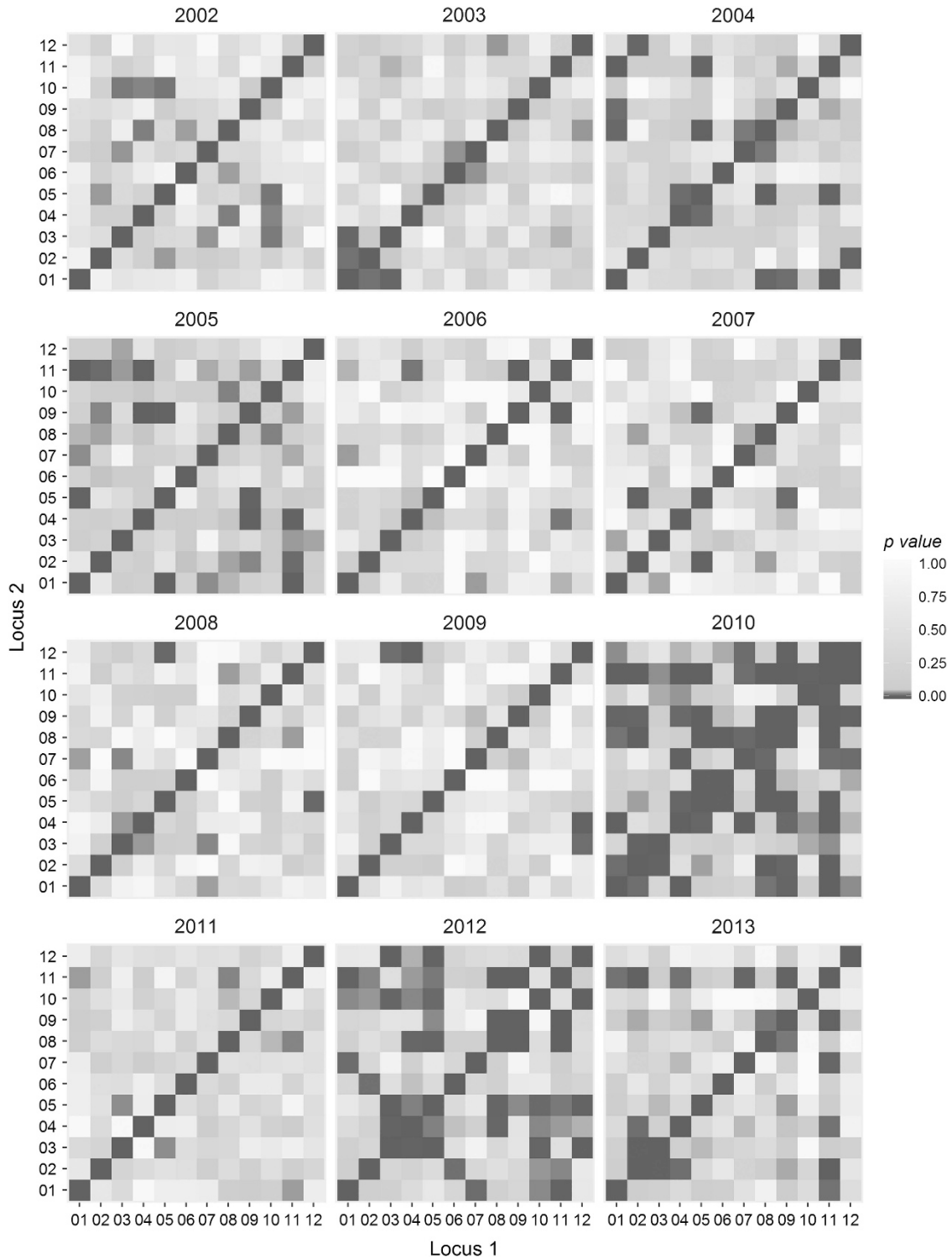


Figure 2 Summary of pairwise LD tests conducted within Genepop (Rousset, 2008) for each of the 12 successive years. Locus numbers are given in the same order as shown in Table 2.

Bayesian cluster analysis using the program STRUCTURE (Pritchard *et al.*, 2000) to determine whether any genetic structure could be detected in the absence of *a priori* geographic data. The resulting posterior probabilities were highly concordant among replicate runs, with the highest average value indicating the most likely number of clusters, K . The average log likelihood value climbed steadily with

increasing K to peak at $K=20$ (Supplementary Figure S3). However, when this analysis was repeated separately for each of the twelve years, the maximal average log likelihood values were mainly associated with $K=1$, suggesting a lack of detectable population structure within years (Supplementary Figure S4). The only exceptions were 2005 and 2013, for which the most likely genetic structure consisted of six and four

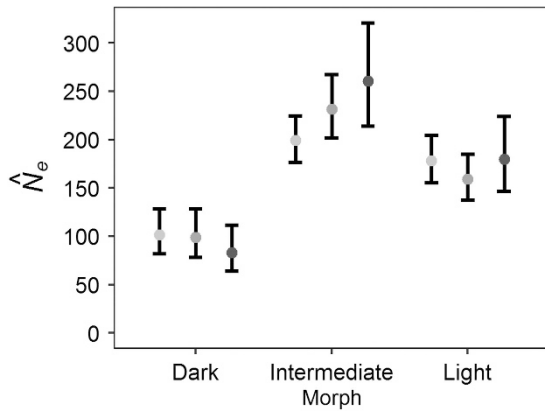


Figure 3 Single sample N_e estimates and their associated 95% confidence intervals based on the LD approach of Waples and Do (2008) implemented in NeEstimator (Do *et al.*, 2014) for P_{crit} values of 0.01, 0.02 and 0.05 respectively.

clusters respectively. Visual inspection of individual cluster memberships for these years revealed considerable admixture and no clear evidence for the presence of distinct sub-populations.

As a further test of population structure that exploits prior information on sampling locations, we implemented an analysis of molecular variance. For this analysis, we compared animals sampled to the north and south of the Teutoburger Wald, a low mountain region that bisects the study area. Around 1% of the variance was partitioned between the north and south ($F=0.03$, $P=0.001$) indicating the presence of weak but statistically significant population substructure. To test whether this has an effect on the effective population size estimates, we repeated the temporal analysis shown in Figure 5a separately for the northern and southern sub-populations. The same overall trend was observed for the northern sub-population, whereas \hat{N}_e values were both lower and less variable over time for the southern sub-population (Supplementary Figure S1A). This is consistent with census data (Supplementary Figure S1B) and suggests that the overall pattern shown in Supplementary Figure S1 is driven by changes in the larger northern sub-population.

Sensitivity to sample size

Sample sizes were generally lower in the early part of the study, which appears to be reflected to some extent in the corresponding effective population size estimates. We therefore tested for a relationship between annual \hat{N}_e and sample size. This was not statistically significant ($F_{1,10}=2.76$, $P=0.13$), suggesting that sample size does not have a strong effect on the magnitude of the estimates. To further explore whether our conclusions could be affected by variation in sample size, we generated \hat{N}_e values for pooled data corresponding to the periods 2002–2009 and 2010–2013 inclusive. To mimic sampling effects, we then randomly selected differently sized subsets of individuals each 10 times. The resulting \hat{N}_e values were consistently greater for the latter period (Supplementary Figure S2) and no clear relationship was found between \hat{N}_e and sample size for either period.

DISCUSSION

Long-term genetic studies are essential for understanding temporal variation in the effective population size of natural populations. We therefore generated a large microsatellite data set for an intensively monitored common buzzard population in northern Germany. The 12-year duration of the study allowed us to use both single-sample and

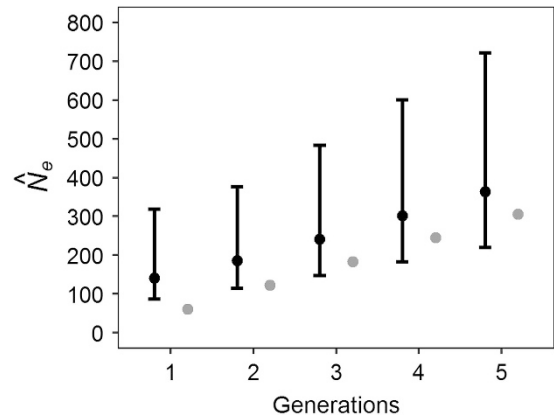


Figure 4 Likelihood and moment-based temporal estimators of Wang and Whitlock (2003) respectively based on the comparison of 2002 with 2013, with an increasing number of generations from one to five assumed to separate the samples.

temporal approaches that yielded comparable estimates, at least for the latter part of the study. Analysis of 12 successive cohorts also uncovered appreciable temporal heterogeneity, with \hat{N}_e varying by a factor of 14 over the course of the study.

We used multiple approaches to estimate the effective population size of our study population of common buzzards. However, because of the fact that the population has expanded in recent years, we took the conservative measure of focussing on those estimators that are least sensitive to immigration into the focal population (Gilbert and Whitlock, 2015). We found that the results of the LD method were reasonably consistent across a range of P_{crit} values from 0.01 to 0.05. One relevant feature of common buzzards is the presence of three distinct colour morphs that differ markedly in their frequency in the study population (Krüger *et al.*, 2001). Accordingly, \hat{N}_e was smallest for the dark morph, intermediate for the light morph and largest for the intermediate morph. Although heterozygote advantage operates at the colour locus in this population, with the intermediate morph having the greatest average lifetime reproductive success (Chakarov *et al.*, 2008), we recently showed that the three morphs do not differ significantly in their genome-wide heterozygosity, nor are they genetically differentiated from one another (Boerner *et al.*, 2013). Variation in the effective population sizes of the morphs is therefore unlikely to be an artefact of differences in genome-wide heterozygosity or population substructure and instead appears to be a reflection of the relative frequencies of the morphs in the population.

We also used the LD method to generate annual effective population size estimates. These generally had rather small 95% confidence intervals, consistent with a simulation study (Waples and Do, 2008) showing that the LD method can generate precise estimates with ~ 10 – 20 microsatellites and 50 individuals sampled where N_e is less than ~ 500 . The estimates themselves varied by a factor of 14 over the duration of the study and were consistently lower during 2002–2009. The smallest estimate of 25.5 was obtained for 2008, a year in which breeding success was particularly low. Afterwards, the estimates steadily increased towards a peak of 367.2 in 2011. Although the exact causes of this increase are not known, census data indicate that the population more than doubled from 2009 to 2012, whereas over the same period the frequency of the white morph increased from 30 to 41%. Hence, temporal patterns in \hat{N}_e appear to capture underlying population processes that are also reflected in the census data.

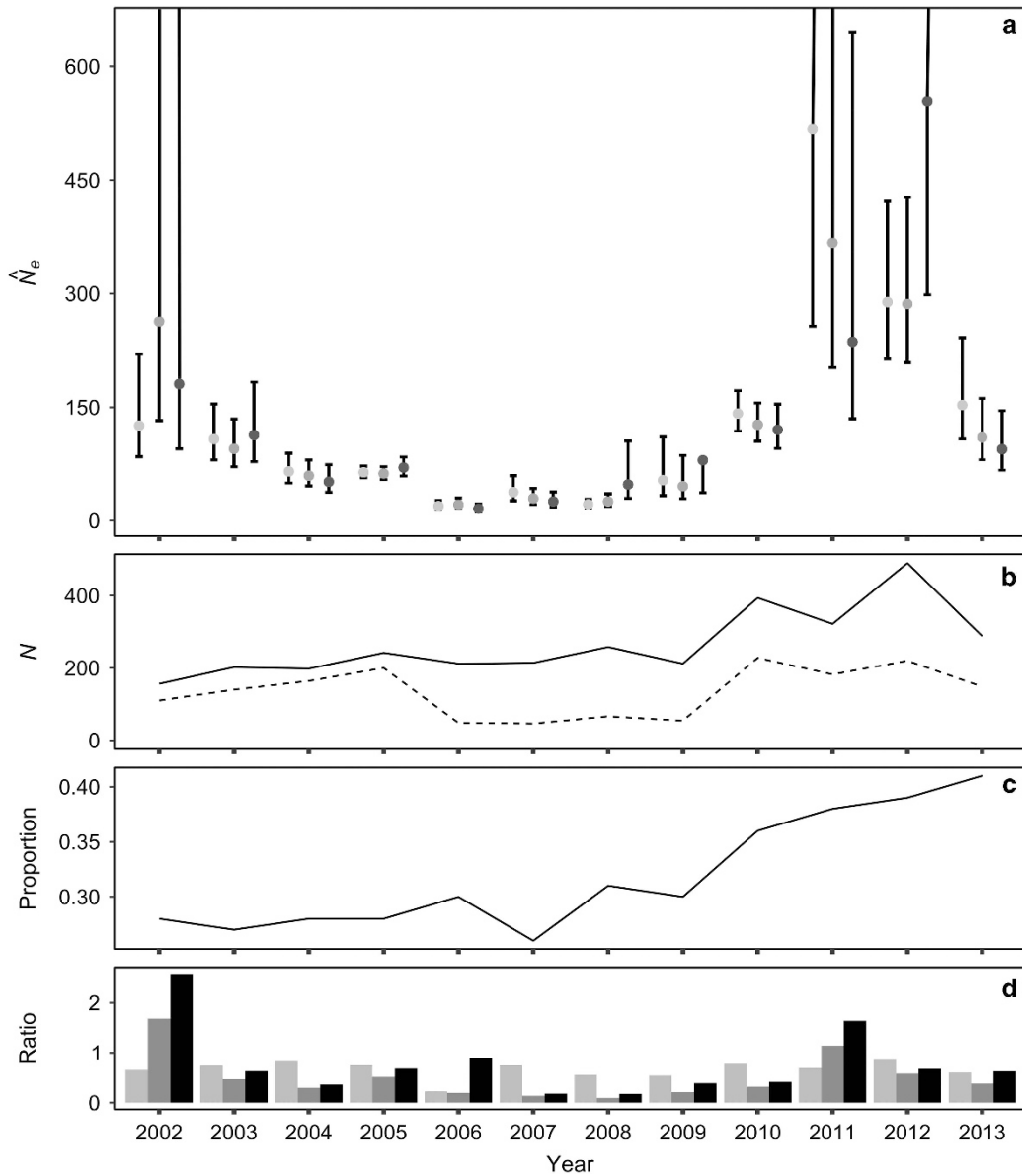


Figure 5 Temporal variation in the common buzzard study population between 2002 and 2013. (a) Annual \hat{N}_e values with 95% confidence intervals based on the LD method (Waples and Do, 2008) with P_{crit} values of 0.01, 0.02 and 0.05, respectively. (b) Observed number of parents (N_b , black line) and census population size (N_c , dashed line). (c) Proportion of the light morph in the population. (d) Ratios of N_b to N_c (light grey bars), \hat{N}_e to N_c (dark grey bars) and \hat{N}_e to N_b (black bars) based on a P_{crit} value of 0.02.

Possible causes for population growth include increased survival and local recruitment, and immigration from further afield. However, the former is unlikely to have played a major role as, despite having fitted almost all fledged chicks from the study area with clearly visible wing tags, we have not observed an increase in the number of local birds recruiting into the population (Chakarov *et al.*, 2013). In fact, many new territories have been found by individuals without wing tags (O Krüger, personal communication). As we are confident that these individuals are not local, the only explanation remaining is that population growth is attributable at least in part to immigration. This is also consistent with changes over time both in the frequency of the light morph and in the proportion of pairwise comparisons among loci yielding significant LD values (see below).

As the LD approach will be affected by the presence of physically linked markers, we checked for LD among the 12 microsatellites in

each of the years. We did not find consistent patterns of association between particular pairs of loci, and this agrees with the fact that most of the markers map to different chromosomes in the zebra finch. However, the proportion of pairwise comparisons yielding significant test statistics varied from year to year, being mostly low but reaching highs of ~53% in 2010 and 38% in 2012 (Figure 2). Although this could partly reflect larger sample sizes for these two particular years, over the same period we found very little variation in either observed heterozygosity (Table 2) or standardised allelic richness (Table 3), suggesting that genetic diversity has not altered appreciably. Similarly, the results of the Structure analysis suggest that there has been no apparent increase in the amount of population structure over time in parallel with the N_e estimates. Thus, it seems likely that NeEstimator is detecting changes in LD within the population. These changes could potentially be a consequence of gene flow into the focal population, as

Table 3 Standardised allelic richness with s.d. for the 12 microsatellite loci used in this study for each year

Years	Allelic richness	s.d.
2002	5.25	0.236
2003	4.75	0.236
2004	4.73	0.199
2005	4.92	0.312
2006	5.08	0
2007	5.17	0.349
2008	4.95	0.24
2009	5.05	0.385
2010	4.63	0.336
2011	4.87	0.326
2012	5.25	0.317
2013	4.58	0.257

this is known to create LD when allele frequencies are unequal among populations exchanging migrants (Gilbert and Whitlock, 2015).

Although it has been estimated that 80% of buzzards recruit to within 20 km of their natal territories (Zang *et al.*, 1989) and 96% within 100 km (Walls and Kenward, 1998), our genetic data are not consistent with localised immigration from the vicinity of our study population, as the high mobility of common buzzards should reduce population structure over this scale. However, longer-distance dispersal does happen regularly in this species (Cramp and Simmons, 1980; Kenward *et al.*, 2001) and buzzard dispersal is strongly influenced by weather patterns (Walls *et al.*, 2005). In particular, harsh winters are known to severely affect buzzards (Cramp and Simmons, 1980) and the winter of 2009–2010 was among the harshest in recent years that might have induced significant dispersal events. On current evidence, it remains speculative where the influx of light birds may have come from, as there is no clear evidence for clinal variation in buzzard morph frequency.

One caveat to the LD method is that LD can be generated by many different phenomena, from inbreeding through population structure to immigration (Luikart *et al.*, 2010). Inbreeding is unlikely to be important in this system, partly because the species is monogamous but also because a previous study found no evidence for inbreeding (Boerner *et al.*, 2013). However, population structure could potentially be present as the study area is bisected by the Teutoburger Wald. To check for the presence of discrete populations within the study area, we therefore used Bayesian cluster analysis and analysis of molecular variance. The Structure results hinted at the presence of multiple clusters when all of the data were analysed together but generally indicated a lack of structure when the years were analysed separately. Such a pattern could potentially arise because of the fact that buzzard pairs often breed across multiple years, meaning that a stronger signal of family structure may be present in the full data set relative to individual years.

In contrast, analysis of molecular variance uncovered a small but significant genetic difference between buzzards breeding to the north and south of the Teutoburger Wald. This most likely reflects differences in habitat suitability, as similar differences between the northern and southern parts of our study area have been documented for buzzard survival in interaction with local weather patterns (Jonker *et al.*, 2014). This is probably a reflection of the predominance of scots pine forest on sandy soils in the southern area that is known to be a suboptimal habitat for buzzards (Krüger, 2004).

Sensitivity of effective population size estimators to violation of the assumption of discrete, non-overlapping generations is an issue that hampers many studies of natural populations (Waples *et al.*, 2014; Kamath *et al.*, 2015). Although samples that combine multiple cohorts (strictly, as many cohorts as there are in a generation) estimate N_e , single-cohort samples are usually thought of as providing information about the effective number of parents that produced the sample, that is, N_b (Schwartz *et al.*, 1999; Leberg, 2005; Waples and Do, 2010; Waples and England, 2011; Waples *et al.*, 2014). With knowledge of the ratio of N_b/N_e , which can be estimated from life history information (Waples *et al.*, 2013), it is possible to directly estimate N_e . However, this was not possible for our study population because the majority of breeding individuals are of unknown age.

Another important caveat is that for the temporal method the number of generations between samples affects both precision and bias (Waples and Yokota, 2007; Gilbert and Whitlock, 2015). Precision can be reasonably good with as little as two generations separating samples, but age structure can cause bias by violating the assumption of discrete generations (Jorde and Ryman, 1995; Wang *et al.*, 2010). Although the magnitude and direction of bias depends on the species' life history and the type of samples taken, as a guideline it has been suggested that at least 3–5 generations are needed to minimise bias (Waples and Yokota, 2007). In the case of our study population, wing-tag data indicate that local recruitment is negligible, suggesting that there will be little if any overlap between individuals sampled as chicks at the beginning of the study and adults that contributed chicks towards the population in the later part of the study. However, to quantify the magnitude of bias will require at least one more decade of genetic sampling.

Despite the above caveats, we believe that our main conclusions are unlikely to be strongly affected by the presence of overlapping generations for three main reasons. First, we are more interested in the relative values of the estimates rather than the absolute values. This is why we estimated N_e based on pooled samples for the different colour morphs, as the dark morph in particular is too infrequent to be able to generate meaningful estimates separately for each of the years. Second, although the results of the temporal analysis may be subject to bias (see above) and vary with the precise estimator used and the number of generations assumed to separate the samples, arguably the temporal estimates are reasonably consistent with the single-sample N_e estimates obtained for the latter years of the study, as also shown by Miller and Waits (2003) and Rowe and Beebe (2004). Third, any temporal variation in \hat{N}_e across the study will tend to be dampened by individuals who contribute towards estimates in successive years, making our temporal analysis if anything somewhat conservative.

One potential issue with our study is that sample sizes varied from year to year and were generally lower in the first half of the study where the corresponding \hat{N}_e values were also smaller. We therefore tested whether the conclusion of a temporal increase in N_e was robust to sample size variation. For this analysis, we pooled data for the periods 2002–2009 and 2010–2013 inclusive, and then calculated \hat{N}_e for each period using differently sized random subsamples. Regardless of sample size, consistently larger estimates were obtained for the latter period, in support of a genuine temporal increase in N_e . Consistent with the observation that annual \hat{N}_e did not correlate significantly with sample size, we also found no indication of an increase in \hat{N}_e with the number of randomly selected individuals.

A handful of previous studies have used similar approaches to explore temporal patterns in the effective population size of natural populations, although for species like bears with relatively long generation times this requires the analysis of museum specimens

(see, for example, Miller and Waits, 2003). In one study, 20 years of archived chinook salmon scales were analysed to reveal a long-term decrease in N_e despite the census size having increased by a factor of five (Shrimpton and Heath, 2003). Another study used a time series of European brown bear samples to document a temporal increase in N_e (Skrbinšek *et al.*, 2012) that is very similar to the one we observe. They concluded that this could be related to population growth, but lacked census data with which to test this.

An important quantity in conservation genetics is the ratio of N_e to N_c (Palstra and Ruzzante, 2008) and there is considerable interest in whether this might change over time. In natural salmonid populations, species with lower census sizes tend to have higher N_e to N_c ratios, a finding that has been attributed to ‘genetic compensation’, a buffering mechanism that could help to retain genetic diversity in small populations (Palstra and Ruzzante, 2008). The same has also been reported within species using time series data from wild Atlantic salmon and steelhead trout populations (Ardren and Kapuscinski, 2003). We found no evidence of compensation in this common buzzard population, but this makes sense given the species has very low reproductive skew. In contrast, salmonids have much higher reproductive skew and thus compensation can occur, for instance, if juvenile parr have increased reproductive success at low densities (Palstra and Ruzzante, 2008).

Finally, it is worth considering the practical implications of our findings. Estimating N_e separately for each of the years resulted in a range of values that mostly fell above the inbreeding avoidance criterion of 50 proposed by Franklin (1980), Lande (1988), Franklin and Frankham (1998) as well as Lynch and Lande (1998). This is consistent with recent work on the same buzzard population that found no evidence of inbreeding (Boerner *et al.*, 2013). However, estimates for 2006–2008 inclusive fell below the proposed threshold and thus very different conclusions could be reached depending on the year in question. Taken together with the documented increase in census size, our results suggest that our buzzard population is not at imminent risk of extinction, and argue that caution is warranted when drawing firm conclusions on the basis of a single sample.

CONCLUSIONS

Effective population size monitoring is advocated in conservation and management programmes (Schwartz *et al.*, 1999; Leberg, 2005; Schwartz *et al.*, 2007) yet collecting long-term observational and genetic data from natural populations presents a major challenge. We generated a microsatellite data set spanning over a decade for an intensively monitored buzzard population that uncovered marked temporal variation in N_e . Further long-term studies of natural populations are needed in order to generalise our findings to other species and ecological contexts.

DATA ARCHIVING

Data sets used in the analysis of this paper are available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.jr107>.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to Robin Waples, Kimberly J Gilbert, an anonymous reviewer and the special issue editors of *Heredity* for helpful comments on the manuscript. This study was supported by the Deutsche Bundesstiftung Umwelt (to A-KM), the Volkswagen Foundation (to NC) and a Marie Curie FP7-Reintegration Grant (to JIH) within the Seventh European Community Framework Programme (PCIG-GA-2011-303618).

- Andrewartha HG, Birch LC (1954). The innate capacity for increase in numbers. *The Distribution and Abundance of Animals*. University of Chicago: Chicago. pp 31–54.
- Ardren WR, Kapuscinski AR (2003). Demographic and genetic estimates of effective population size (N_e) reveals genetic compensation in steelhead trout. *Mol Ecol* **12**: 35–49.
- Beissinger SR, Snyder NFR (1992). *New World Parrots in Crisis: Solutions from Conservation Biology*. Smithsonian Institution Press: Washington, DC, USA.
- Boerner M, Hoffman JI, Amos W, Chakarov N, Krüger O (2013). No correlation between multi-locus heterozygosity and fitness in the common buzzard despite heterozygote advantage for plumage colour. *J Evol Biol* **26**: 2233–2243.
- Casey JM, Myers RA (1998). Near extinction of a large, widely distributed fish. *Science* **281**: 690–692.
- Chakarov N, Boerner M, Krüger O (2008). Fitness in common buzzards at the cross-point of opposite melanin-parasite interactions. *Funct Ecol* **22**: 1062–1069.
- Chakarov N, Jonker RM, Boerner M, Hoffman JI, Krüger O (2013). Variation at phenological candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird of prey. *Mol Ecol* **22**: 5430–5440.
- Charlesworth B (2009). Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**: 195–205.
- Coulson T, Rohani P, Pascual M (2004). Skeletons, noise and population growth: the end of an old debate? *Trends Ecol Evol* **19**: 359–364.
- Cramp S, Simmons KEL (eds) (1980). *Handbook of the Birds of Europe, the Middle East and North Africa. The Birds of the Western Palearctic*. Oxford Univ. Press: Oxford UK.
- Crow JF, Kimura M *et al.* (1970). An introduction to population genetics theory. *An Introduction to Population Genetics Theory*.
- del Hoyo J, Elliott A, Christie D (1994). Handbook of the birds of the world. *New World Vultures to Guinea-fowl*, Vol. 2. Lynx: Barcelona.
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* **14**: 209–214.
- Elton C (1942). *Voles, Mice and lemmings. Problems in Population Dynamics*. Clarendon Press: Oxford.
- Ferrer M, Penteriani V, Balbontin J, Pandolfi M (2003). The proportion of immature breeders as a reliable early warning signal of population decline: evidence from the Spanish imperial eagle in Doñana. *Biol Conserv* **114**: 463–466.
- Frankham R (2005). Genetics and extinction. *Biol Conserv* **126**: 131–140.
- Franklin IR (1980). Evolutionary change in small populations. In: Soule ME, BA Wilcox (eds). *Conservation Biology: An Evolutionary-Ecological Perspective*. Sinauer Associates: Sunderland, MA, USA, pp 135–149.
- Franklin IR, Frankham R (1998). How large must populations be to retain evolutionary potential? *Anim Conserv* **1**: 69–70.
- Gilbert KJ, Whitlock MC (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution* **69**: 2154–2166.
- Hill WG (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet Res* **38**: 209–216.
- Hunter ML (2002). *Fundamentals of Conservation Biology*, 2nd edn. Blackwell Science: Cambridge, MA, USA.
- IUCN Standards and Petitions Subcommittee (2016). Guidelines for Using the IUCN Red List Categories and Criteria. Prepared by the Standards and Petitions Subcommittee. Available from <http://www.iucnredlist.org/documents/RedListGuidelines.pdf>.
- Johnson PCD, Fowlie MK, Amos W (2005). Isolation of microsatellite loci from the common buzzard, *Buteo buteo* (Aves: Accipitridae). *Mol Ecol Notes* **5**: 208–211.
- Jonker RM, Chakarov N, Krüger O (2014). Climate change and habitat heterogeneity drive a population increase in Common Buzzards *Buteo buteo* through effects on survival. *Ibis* **156**: 97–106.
- Jorde PE, Ryman N (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* **139**: 1077–1090.
- Kamath P, Haroldson MA, Luikart G, Paetkau D, Whitman C, van Manen FT (2015). Multiple estimates of effective population size for monitoring a long-lived vertebrate: an application to Yellowstone grizzly bears. *Mol Ecol* **24**: 5507–5521.
- Kenward RE, Walls SS, Hodder KH (2001). Life path analysis: scaling indicates priming effects of social and habitat factors on dispersal distances. *J Anim Ecol* **70**: 1–13.
- Krüger O (2004). The importance of competition, food, habitat, weather and phenotype for the reproduction of Buzzard *Buteo buteo*. *Bird Study* **51**: 125–132.
- Krüger O, Grünkorn T, Struwe-Juhl B (2010). The return of the white-tailed eagle (*Haliaeetus albicilla*) to northern Germany: modelling the past to predict the future. *Biol Conserv* **143**: 710–721.
- Krüger O, Lindström J (2001). Lifetime reproductive success in common buzzard, *Buteo buteo*: from individual variation to population demography. *Oikos* **93**: 260–273.
- Krüger O, Lindström J, Amos W (2001). Maladaptive mate choice maintained by heterozygote advantage. *Evolution* **55**: 1207–1214.
- Lack D (1954). Cyclic mortality. *J Wildl Manag* **18**: 25–37.
- Lande R (1988). Genetics and demography in biological conservation. *Science* **241**: 1455–1460.
- Leberg P (2005). Genetic approaches for estimating the effective size of populations. *J Wildl Manag* **69**: 1385–1399.
- Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet* **11**: 355–373.
- Lynch M, Lande R (1998). The critical effective size for a genetically secure population. *Anim Conserv* **1**: 70–72.

- Miller CR, Waits LP (2003). The history of effective population size and genetic diversity in the Yellowstone grizzly (Ursus arctos): implications for conservation. *Proc Natl Acad Sci USA* **100**: 4334–4339.
- Neel MC, McKelvey K, Ryman N, Lloyd MW, Short Bull R, Allendorf FW *et al.* (2013). Estimation of effective population size in continuously distributed populations: there goes the neighborhood. *Heredity* **111**: 189–199.
- Newman D, Pilson D (1997). Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution* **51**: 354–362.
- Nunney L, Elam DR (1994). Estimating the effective population size of conserved populations. *Conserv Biol* **8**: 175–184.
- Palstra FP, Ruzzante DE (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol Ecol* **17**: 3428–3447.
- Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics (Oxford, England)* **28**: 2537–2539.
- Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN *et al.* (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science* **344**: 1246752.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Rousset F (2008). Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Mol Ecol Resour* **8**: 103–106.
- Rowe G, Beebee TJ (2004). Reconciling genetic and demographic estimators of effective population size in the anuran amphibian *Bufo calamita*. *Conservation Genetics* **5**: 287–298.
- Royama T (1992). Basic properties and structure of population processes. In: *Analytical Population Dynamics*. 1. publ. in paperback. Chapman & Hall: London, pp 3–46.
- Schwartz MK, Luikart G, Waples RS (2007). Genetic monitoring as a promising tool for conservation and management. *Trends Ecol Evol* **22**: 25–33.
- Schwartz MK, Tallmon DA, Luikart G (1999). Using genetics to estimate the size of wild populations: many methods, much potential, uncertain utility. *Anim Conserv* **2**: 321–323.
- Shrimpton JM, Heath DD (2003). Census vs. effective population size in chinook salmon: large- and small-scale environmental perturbation effects. *Mol Ecol* **12**: 2571–2583.
- Skrbinšek T, Jelenčič M, Waits L, Kos I, Jerina K, Trontelj P (2012). Monitoring the effective population size of a brown bear (*Ursus arctos*) population using new single-sample approaches. *Mol Ecol* **21**: 862–875.
- Storey JD (2002). A direct approach to false discovery rates. *J R Stat Soc B* **64**: 479–498.
- Walls SS, Kenward RE (1998). Movements of radio-tagged Buzzards *Buteo buteo* in early life. *Ibis* **140**: 561–568.
- Walls SS, Kenward RE, Holloway GJ (2005). Weather to disperse? Evidence that climatic conditions influence vertebrate dispersal. *J Anim Ecol* **74**: 190–197.
- Wang J (2005). Estimation of effective population sizes from data on genetic markers. *Philos Trans R Soc Lond B Biol Sci* **360**: 1395–1409.
- Wang J, Brekke P, Huchard E, Knapp LA, Cowlishaw G (2010). Estimation of parameters of inbreeding and genetic drift in populations with overlapping generations. *Evolution* **64**: 1704–1718.
- Wang J, Whitlock MC (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**: 429–446.
- Waples RS (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Mol Ecol* **14**: 3335–3352.
- Waples RS (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. *Conserv Genet* **7**: 167–184.
- Waples RS, Antao T, Luikart G (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* **197**: 769–780.
- Waples RS, Do C (2008). Idne: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* **8**: 753–756.
- Waples RS, Do C (2010). Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol Appl* **3**: 244–262.
- Waples RS, England PR (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics* **189**: 633–644.
- Waples RS, Luikart G, Faulkner JR, Tallmon DA (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proc Biol Sci* **280**: 20131339.
- Waples RS, Yokota M (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics* **175**: 219–233.
- Zang H, Heckenroth H, Knolle F (1989). Die Vögel Niedersachsens und des Landes Bremen - Greifvögel. Niedersächs. Landesverwaltungsamt, Naturschutz, Landschaftspflege, Vogelschutz. Hannover.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)