Analysing bird population trends from monitoring data with highly structured sampling designs

Mirjam R. Rieger[1], Christoph Grüneberg[2], Michael Oberhaus[2], Sven Trautmann[3], Madalin Parepa[1], Nils Anthes[1]

[1] Institute of Evolution and Ecology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

[2] State Agency for Nature, Environment and Consumer Protection in North Rhine-Westphalia (LANUV), 40208 Düsseldorf

[3] Dachverband Deutscher Avifaunisten e.V., An den Speichern 2, 48157 Münster

nils.anthes@uni-tuebingen.de (corresponding author)

Summary

Population trends derived from systematic monitoring programmes are essential to identify species of conservation concern and to evaluate conservation measures. However, monitoring data pose several challenges for statistical analysis, including spatial bias due to an unbalanced sampling of landscapes or habitats, variation in observer expertise, frequent observer changes, and overdispersion or zero-inflation in the raw data. An additional challenge arises from so-called 'rolling' survey designs, where each site is only visited once within each multi-year rotation cycle. We developed a GAMM-based workflow that addresses these challenges and exemplify its application with the highly structured data from the Ecological Area Sampling (EAS) in the German federal state North-Rhine Westphalia (NRW). First, we derive a routine that allows informed decisions about the most appropriate combination of distribution family (Poisson or negative binomial), model covariates (e.g., habitat characteristics), and zero inflation formulations to reflect species-specific data distributions. Second, we develop a correction factor that buffers population trend estimates for variation in observer expertise as reflected in variation in total bird abundance. Third, we integrate model weights that adjust for between-year variation in the representation of habitat or landscape types within the yearly subset of sampled sites. In a consistency check, we found good match between our GAMM-based EAS trends and TRIM-based trends from the standard German common Bird monitoring scheme. The study provides a template script for *R* statistical software so the workflow can be adapted to other monitoring programmes with comparable survey designs and data structures.

Keywords: breeding bird monitoring, trend analysis, rolling surveys, spatial bias, observer effects

1 Introduction

Bird populations decline worldwide (Klvaňová et al. 2009, Sauer et al. 2017, Burns et al. 2021a, Burns et al. 2021b), with the European Union alone facing an approximate loss of 560–620 million bird individuals (17–19%) within 40 years (Burns et al. 2021a, Burns et al. 2021b). Such knowledge of population sizes and their trends heavily relies on standardized monitoring programmes to provide reliable population density estimates. The resultant population trend estimates allow allocating conservation resources to species of highest conservation concern (Niemelä 2000, Buckland and Johnston 2017), provide feedback on the efficiency of conservation efforts (Johnston et al. 2015), and raise awareness for the value and state of biodiversity among decision makers and the public (Jennings 2021). However, many existing monitoring programmes pose substantial challenges for unbiased statistical analyses and can reduce the reliability and robustness of the estimated trends if not treated with care (Buckland and Johnston 2017).

Analytical challenges typically arise from two prime sources. First, the structure and implementation of a given monitoring programme may (i) require analytical tools for hierarchical or non-representative sampling designs (van Strien et al. 2004, Buckland and Johnston 2017), (ii) introduce random and systematic estimation errors due to variation in experience within and among observers (Sauer et al. 1994, Link and Sauer 1998, Jiguet 2009, Farmer et al. 2014), and (iii) provide small sample sizes (Svensson 1978, Nielsen et al. 2009). Second, count data such as the number of bird territories typically require consideration of (iv) properties of the underlying data distribution (Blasco-Moreno et al. 2019, Campbell 2021), (v) temporal (= serial) and spatial autocorrelation among measurements (van Strien et al. 2004, Korner-Nievergelt et al. 2015, Kéry and Royle 2020), and (vi) imperfect detection of existing territories (Kéry et al. 2005, Kéry and Royle 2020).

This work focuses on three of these aspects: spatial bias due to non-representative sampling designs, observer effects and specificities of data distributions. First, spatial bias arises when landscape types or habitats of the investigated region are not represented according to their spatial coverage in the monitoring sample (van Strien et al. 2004). As a result, when population trends vary between landscapes or habitats, the estimated overall population trend can be biased towards overrepresented landscapes, or reflect changes in sample composition rather than a true change in abundance (Buckland and Johnston 2017, Bowler et al. 2022).

Second, the quality of trend estimates can vary with the observer (Johnston et al. 2018). Between-observer variation in expertise or other qualifications that affect survey results primarily increase (random) variation between sites and years, as long as average observer qualification remains constant (Sauer et al. 1994, Johnston et al. 2018). Within-observer learning effects, instead, can introduce (systematic) biases in trend estimates. The required familiarization with any new study site and the

specificities of the monitoring programme usually cause underestimated bird abundances in initial survey years (Link and Sauer 1998, Jiguet 2009). Moreover, observer expertise usually increases with the time since active engagement in bird territory surveys, while species detectability declines with age if hearing impairment reduces the detection of high-pitch bird vocalisations (Farmer et al. 2014).

Third, the distribution of breeding bird census data is often characterised by overdispersion and zero-inflation. Overdispersion arises when the variance of count data exceeds their mean value. It is typically resolved by modelling trends with a negative binomial error distribution instead of a Poisson distribution (Blasco-Moreno et al. 2019, Campbell 2021). Zero-inflation occurs when datasets contain so-called structural zeros. This implies a disproportionally large fraction of unoccupied sites that cannot be explained solely by sampling variation among sites considered suitable, but rather arises from survey sites that are unsuitable. Where present, structural zeros are modelled separately from the count distribution in zero-inflated model components (Korner-Nievergelt et al. 2015, Blasco-Moreno et al. 2019, Campbell 2021, Tirozzi et al. 2021).

The software TRIM (Pannekoek and van Strien 2001) represents the current analytical standard for the analysis of monitoring data and is used, for example, in the Pan European Common Bird Monitoring Scheme (Vorisek et al. 2008) and the national bird monitoring programmes of Sweden (Jiguet et al. 2013), Finland (Pöysä et al. 2013), or Germany (Kamp et al. 2021). TRIM computes trends and annual population indices from loglinear Poisson regressions. It corrects for overdispersion, serial correlation and missing values (van Strien et al. 2004) and allows weights to account for spatial bias (van Strien et al. 2004). Yet, designed for monitoring schemes with almost-yearly surveys per site by the same observer, TRIM reaches its limits with substantial turnover in survey sites between years leading to a fraction of missing values exceeding approx. 60 % (van Strien et al. 2001, Bogaart et al. 2020, Dakki et al. 2021). Moreover, TRIM is restricted to categorical covariates, requiring climate or landscape composition covariates to be transformed into categories (Bogaart et al. 2020). Finally, long-term trend analyses are restricted to linear trends, while more recent developments favour the integration of trend smoothers, e.g. with general additive (mixed) models (GA(M)Ms) (e.g., Fewster et al. (2000), Knape (2016)). Using a smoothing function, GA(M)Ms can also capture non-linear short- and long-term trends while still allowing to identify periods of strongly increasing or decreasing trends as well as changes in (linear) trend direction (Fewster et al. 2000, Zuur 2012, Knape 2016, Wood 2021).

Here, we develop an analytical approach for monitoring designs that rely on rolling (or 'rotating') surveys with intervals of several years between repetitive surveys per site (Buckland and Johnston 2017). Rolling surveys can cover more study sites and thus a wider range of landscape types across years, but they also accept frequent observer changes and produce a high fraction of study sites with missing values each year. Given a lack of adequate modelling approaches for such monitoring programmes, we develop

a user-friendly tool that integrates high observer turnover, long time gaps between successive surveys, spatial bias, as well as overdispersion and zero-inflation, to produce reliable and robust estimates of breeding bird trends. The approach is exemplified with data from the Ecological Area Sampling (EAS), a monitoring programme of the German federal state North-Rhine Westphalia (NRW) with a six-year rolling sampling scheme. As a quality check, we compared the predicted population trends with West German trends derived from the German Common Bird Monitoring scheme (MhB) with yearly surveys per site. The statistical approach is provided as an *R* script (Rieger et al. 2024) and package (Rieger 2024) that can be adjusted to other datasets with a comparable structure.

2 Material and Methods

2.1 The Ecological area sampling (EAS) dataset

Our case study dataset originates from the Ecological Area Sampling (EAS), a long-term biodiversity monitoring programme of the State Agency for Nature, Environment and Consumer Protection in North Rhine-Westphalia (LANUV) with 170 study sites that represent the average landscape of North-Rhine Westphalia (NRW) (LANUV 2016b). Sites are distributed across two biogeographic regions (atlantic lowlands, continental highlands) and six natural regions (Fig. 1) proportional to their spatial coverage. Another 21 metropolitan sites were added in 2011 to represent landscape characteristics of the Rhine-Ruhr metropolis region (Weiss and Schulze-Hagen 2014). The sites are surveyed in an alternating six-year cycle resulting in a sample size of typically 25 to 36 (exceptionally up to 43 or 47) sites per year, two to four replicated surveys per site, and 614 surveys in total for the survey period 2002-2020. Each site is visited at least nine times per sampling year (two to eight hours per date) between February and July with complete coverage of its 1 km² area (LANUV 2016a). These repetitive surveys are used to derive the number of territories for each species per km², which represents the response variable *abundance* in all statistical analyses (LANUV 2016a). For further details about the methodology, see appendix section 1.

2.2 Site characterization

Site-specific categories included *biogeographic region* (atlantic, continental), *natural region* (A, B, KB, KM, SB, ST, cf. Fig. 1), and localisation in the *metropolitan area* (no, yes). Quantitative site characteristics covered the *altitude* above sea level (ATKIS 2013), ten *climate parameters* and three *landscape compositions* to account for biases that arise from variation in mean environmental attributes between the surveyed site subsets per year. Climate parameters included long-term averages between 1981 and 2010 (CDC 2010) of minimum, mean, and maximum temperatures, precipitation, and sunshine duration,

each for spring and winter. Landscape variables were the coverage of forest, arable land, and settlements (ATKIS 2013). Given substantial collinearity between these 14 attributes we conducted a Principal Component Analysis and integrated all principal components (PC) with eigenvalues exceeding 1 into the statistical models (three PCs in our case – for further details see Appendix section 2).

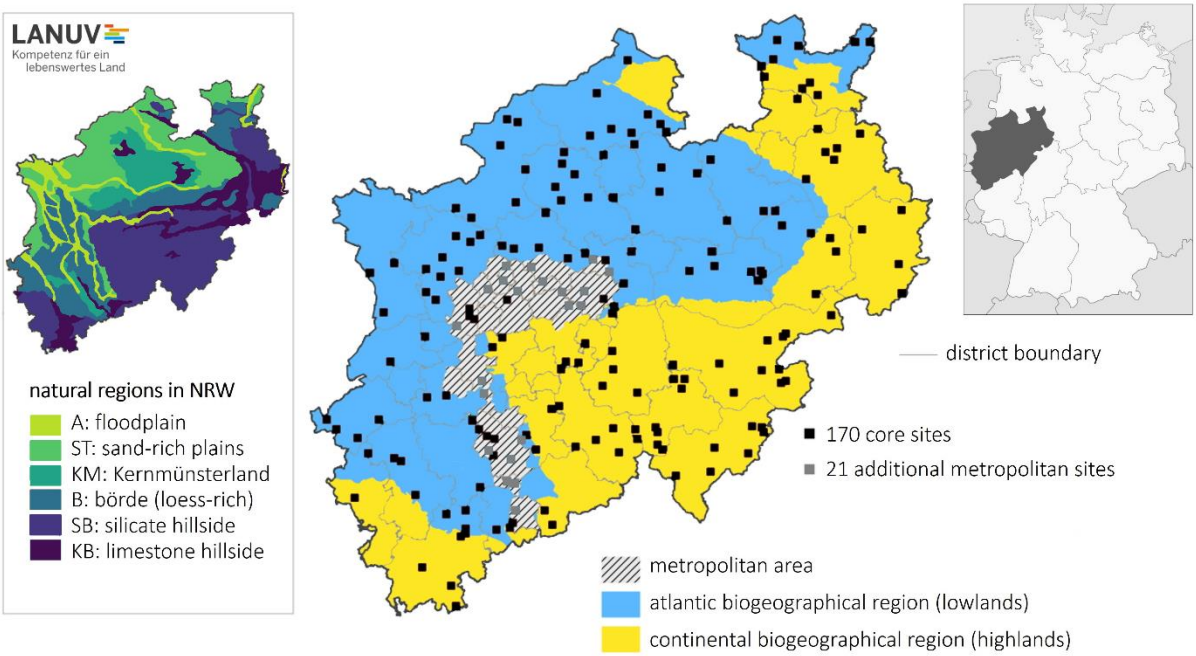

Figure 1: Sites and study area of EAS coloured by six natural regions (left) and two biogeographic regions including the Rhine-Ruhr metropolis (middle). Map source: adapted from LANUV (2016b).

2.3 Spatial bias

The initial sample of 170 EAS monitoring sites almost perfectly represented the proportional coverage of landscape types in NRW, but bias arose from two sources. First, the monitoring programme additionally included 21 metropolitan sites without balanced sampling across landscape types. Second, the required yearly subsampling of sites was highly unbalanced with respect to the combination of natural regions and metropolitan area (Fig. 2) in the early phase of the EAS programme, with improvements since 2007 and almost perfect balance since 2013 (Fig. 2).

To account for the uneven allocation in early programme years and for the overrepresentation of metropolitan areas since 2011, we weighted response measures accordingly. Weights per site and year were calculated as

$$\text{weight} = \frac{TS_l}{AS_{lj}}$$

with TS_l the target share per landscape l (= proportional landscape cover in NRW) and AS_{lj} the actual share (= proportion of sites of the given landscape l) among all surveyed sites in year j (Rieger 2024). A

match between sample share and landscape coverage in NRW results in weight = 1, over- or underrepresented sample shares result in weights <1 or >1, respectively (see Appendix section 3).

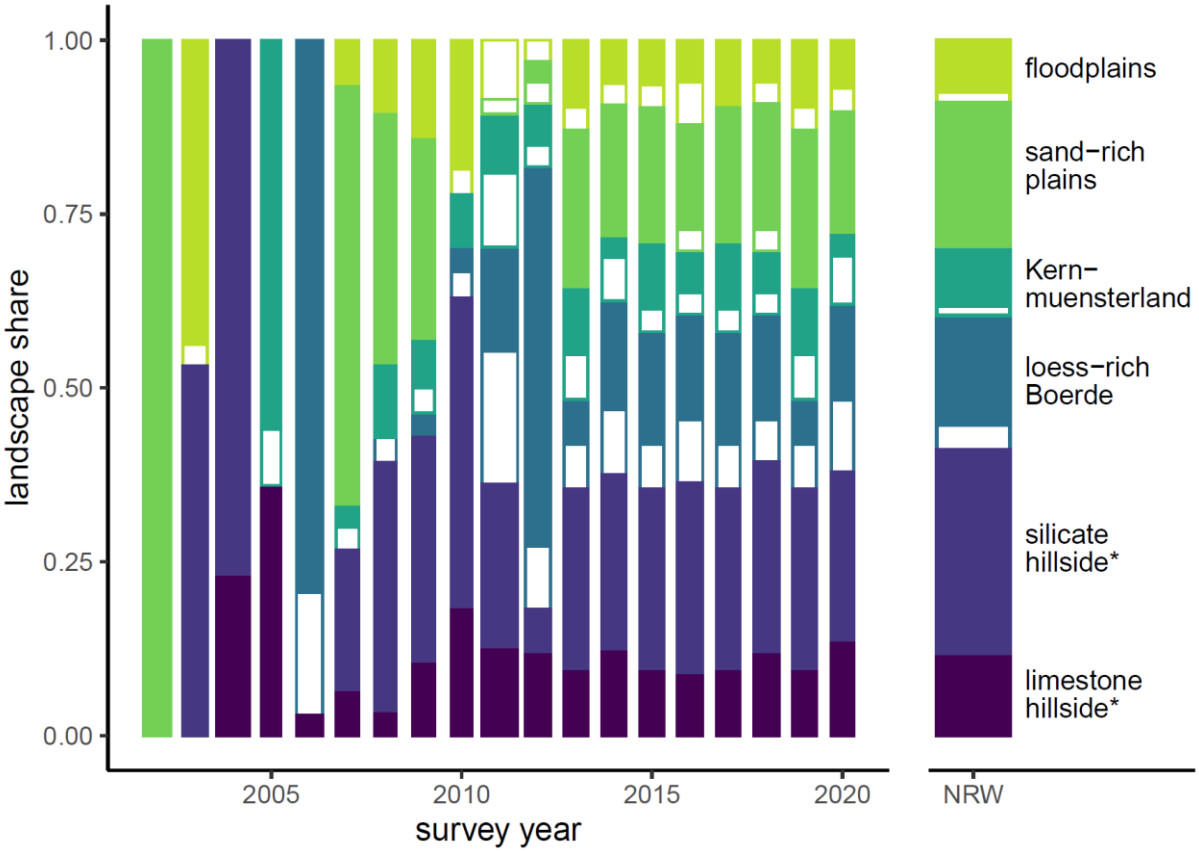

*Figure 2: Landscape shares in annual EAS samples (main panel) and their spatial coverage in NRW (right panel) for all combinations of natural regions and metropolitan area. Bar widths are proportional to the number of surveyed sites per year. White boxes at the lower end of each landscape refer to sample sites in metropolitan areas. *metropolitan area of landscape not present in EAS due to low shares in NRW (SB = 0.4%, KB = 0.6%).*

2.4 Observer effects

The delineation of breeding territories is prone to under- and overestimation arising from variation in observer expertise, territory delineation, and field survey quality (Südbeck et al. 2005, Johnston et al. 2018), even in programmes such as EAS that hire skilled ornithologists (König 2020). Observer effects add further noise to monitoring data when observers change within site. Excluding data from first year surveys, an established standard in monitoring data with yearly site coverage (e.g., German Common Bird Monitoring (Kamp et al. 2021)), is unfeasible for rolling surveys: In the EAS dataset, 434 out of 614 breeding bird surveys between 2002 and 2020 (71 %) were such first-year combinations of observer and site. Instead, we delineated surveys with suspiciously high or low territory counts. The procedure assumes that severe observer effects typically manifest in a general under- or overestimation of territory numbers (= abundance) across species, so that total abundance for a given site and year stands out against average total abundance for that site. Based on this logic, we calculated

$$\text{Observer effect} = \frac{AB_{js}}{\overline{AB_s}}$$

with AB_{js} the abundance sum across species on site s in survey year j, and $\overline{AB_s}$ the mean abundance sum on site s across all survey years (Rieger 2024). The observer effect is a ratio, so that a value of 1 indicates no deviation from the mean, and a value of 0.75 (1.25) a deviation of -25% (+25%) from the observed mean. We integrated the *observer effect* as a categorical predictor variable classified as 'negative' (or 'positive') when the abundance sum of a given survey was at least 25% smaller (or larger) than the mean per site, and as 'none' otherwise. As a sensitivity analysis, we also checked observer effect thresholds of 20% and 15%.

As an alternative way to characterise observer expertise and learning effects, we categorized observers as (i) unfamiliar with the survey programme and site, (ii) familiar with survey programme but unfamiliar with the site, or (iii) familiar with both. Yet, we excluded this predictor in final models since it did not improve model fit, most likely because observer turnover rates were high and thus site familiarization effects unlikely to introduce relevant bias into the EAS trend estimates.

2.5 Distribution characteristics and the statistical analysis of EAS

We analysed species-specific population trends between 2002 and 2020 for the entire federal state for all species with a proportion of zero-sightings below 90% in both biogeographic regions (61 out of 148 species), and for just the Atlantic or continental region for species that met this criterion only in one of the two biogeographic regions (24 out of 148). For clarity and conciseness, this study focuses on a subset of 14 species selected to cover the full range of abundance distributions and proportions of zero counts (Table 1).

Breeding bird *abundance* trends were analysed using generalized additive mixed models (GAMMs) within a Bayesian framework as implemented in the package *brms* (v.2.21.0 (Bürkner 2017)) of the statistical software R (v.4.3.3 (R Core Team 2024)). The *brms* package fits Bayesian models using *Stan* via the package *cmdstanr* (v.0.7.1 (Gabry and Češnovar 2020)) and allows to more flexibly integrate the required error families and model structures than current frequentist GAMM packages such as *mgcv* (Wood 2021) or *gamm4* (Wood and Scheipl 2020)). Initial models implemented a Poisson error structure with a log-link (Fig. 3, Table 4). Where species-specific model assessment implied excessive zero frequencies or overdispersion of the response variable *abundance*, we chose zero-inflated negative binomial (zinb), zero-inflated Poisson (zip) or negative binomial (nb) models as needed (Fig. 3, Table 2, Appendix section 5). Since the underlying survey method can lead to 0.5 abundances due to peripheral territories, decimal raw abundances were rounded to the next integer (see Appendix section 1). Given

non-representative sampling across landscapes, we weighted the *abundance* by the factor defined in section 2.3.

All abundance models had *Survey year* as the main continuous predictor modelled with a smoothing function (thin plate regression splines with an automatic selection of the degree of smoothing) to (i) account for non-linear changes in abundance over time, while (ii) still allowing adjacent years to be related to each other (Fewster et al. 2000, Zuur 2012, Knape 2016, Wood 2021). Thus, smoothers buffer against the property of rolling surveys that sampled site subsets differ between successive years, so that year-to-year changes in abundance cannot differentiate among-site variation from true short-term change in population size (Zuur 2012). To estimate regional population densities and trends as well as their differences, we added *biogeographic region* as a categorical main effect to obtain separate smoothers for the Atlantic and continental region (Wood 2021). Due to a non-overlap of natural regions between biogeographical regions (the Atlantic region mainly consisted of floodplains, sand-rich plains, loess-rich Boerde, and chalk-rich Kernmuensterland, whereas the continental region mainly consisted of silicate and limestone hillside), we refrained from adding natural regions as a further predictor. Models therefore cannot provide separate trend and density estimates per natural region but integrate corrections for non-representative sampling of natural regions through the weighting factors introduced above. *Observer effects* as outlined in section 2.4 were added as a categorical covariate. Numeric covariates included linear and quadratic terms of site-specific environmental attributes captured in the first *three principal components* (sections 2.2 and Appendix section 2) to account for bias that arises from variation in mean environmental attributes between the surveyed site subsets per year. Finally, we added *site ID* as a random intercept to reflect repeated measures per site and model variation in mean abundance between the different site subsets per year.

The zero-inflation model component, where needed (Table 1), represents a binomial (Bernoulli) model with a logit link to estimate the additional occurrence of structural zeros. We kept this component as simple as possible, adding *environmental PCs*, *biogeographic region*, and the random intercept *site ID* as needed to improve model convergence and better reflect the observed distribution of zero counts (Fig. 3, Table 2).

There was no need to integrate temporal autoregressive structures into models since visual inspection of variograms plotted with the *variog* function of the package *geoR* (v.1.9-3 (Ribeiro Jr and Diggle 2020)) revealed no issues with temporal autocorrelation, probably because the six-year gaps between repeated measures (usually accompanied by a change in observer) added more variation to model residuals than expected from the remaining temporal pattern, rendering any remaining temporal autocorrelation negligible.

240 *Table 1: Overview of model types (main model family, coefficients in the optional zi-binomial model) and number of case study species per model type. Species names are given for the 14 exemplary species shown in this paper. For further details of all species, see Appendix section 9.*

model family	binomial model coefficients	# species (%)	exemplary species
pois		19 (22.4%)	Common Kestrel (*Falco tinnunculus*), Eurasian Jay (*Garrulus glandarius*)
nb		4 (4.7%)	Common Blackbird (*Turdus merula*)
zip	~ 1	4 (4.7%)	Common Chaffinch (*Fringilla coelebs*)
	~ R	1 (1.2%)	Willow Tit (*Poecile montanus*)
	~ PCs	11 (12.9%)	--
	~ (1|ID)	5 (5.9%)	Great Spotted Woodpecker (*Dendrocopos major*)
	~ R + (1|ID)	2 (2.4%)	Common House Martin (*Delichon urbicum*)
	~ PCs + (1|ID)	8 (9.4%)	White Wagtail (*Motacilla alba*)
	~ PCs + R + (1|ID)	1 (1.2%)	
zinb	~ 1	3 (3.5%)	Eurasian Blackcap (*Sylvia atricapilla*)
	~ R	3 (3.5%)	Common Chiffchaff (*Phylloscopus collybita*)
	~ PCs	7 (8.2%)	Eurasian Magpie (*Pica pica*)
	~ (1|ID)	2 (2.4%)	--
	~ R + (1|ID)	6 (7.1%)	Great Tit (*Parus major*)
	~ PCs + (1|ID)	5 (5.9%)	Carrion Crow (*Corvus corone*), Marsh Tit (*Poecile palustris*)
	~ PCs + R + (1|ID)	4 (4.7%)	

The main model formula was
~ s(survey year, by = R) + survey year + poly(PC1, 2) + poly(PC2, 2) + poly(PC3, 2) + OE + (1|ID).
pois = Poisson, nb = negative binomial, zip = zero-inflated Poisson, zinb = zero-inflated negative binomial,
PCs = PC1 + PC2 + PC3, R = biogeographic region, OE = observer effect, (1|ID) = random effect of site ID.

Each model ran eight Markov chain Monte Carlo chains with 1500 warm-up and 4000 post-warm-up samples per chain. The 32000 post-warm-up samples were used to assess model fit via posterior predictor checks along four criteria that describe the ability to simulate the observed data distribution.
245 Criteria included the proportion of zero counts (propZ), mean and maximum (max) abundance, and the standard deviation of abundance (SD) as an indicator of dispersion. The proportion of simulations exceeding the observed value (based on raw data) is given as a so-called Bayesian p-value (Korner-Nievergelt et al. 2015). A good model performance thus results in Bayesian p-values close to 0.5, implying that approximately half the simulations exceeded the observed value, while values
250 approaching 0 or 1 would flag models that clearly under- or overestimated the observed value. Based on these Bayesian p-values we additionally calculated a selection value:

$$\text{Selection value} = \sum(0.5 - B_{fun})^2$$

with B the Bayesian p-value for the four different functions *propZ, mean, max*, and *SD* (Rieger 2024). Values close to 0 indicate good model performance with only minor deviations from the ideal Bayesian p-values, values close to 1 indicate poor model performance (see also section 3.1 and Appendix section 5).

The post-warm-up samples were also used to derive mean coefficient estimates for all covariates with their 95% credible intervals (CrI) as well as the posterior probability (postP) of the estimate to exceed 0. Posterior probabilities approaching 0 or 1 indicate rather robust coefficient estimates.

Population trend

Model predictions ± 95% CrI were used to display trends graphically. When estimating trends across survey years and differences between biogeographic regions, values of the remaining model covariates were set to their sample mean (continuous covariates) and to the level 'none' (for 'observer effect') (Korner-Nievergelt et al. 2015). We derived overall trends across biogeographic regions by combining region-specific trends according to the region's share of NRW (55.5% Atlantic region and 44.5% continental region). Based on pairwise differences of simulated abundances, we estimated trends as changes in abundance between the start and end date of any desired time period as

$$\text{Pairwise difference} = ab_{ij} - ab_{i(j-\Delta t)}$$

with ab the predicted abundance for the ith simulation at times j and $j - \Delta t$, where Δt is the desired time period. If the 95% CrI of these simulated differences excluded 0, the trend estimate was considered robust. We specifically produced pairwise differences to estimate fine-scale changes in trend slopes (Δt = 0.001), annual changes (Δt = 1), and a 12-year trend between 2008 to 2019 (Δt = 12, j = 2019).

Population index

For trend comparisons with other monitoring programmes, we integrated the option to transform predicted mean abundances per km² into an index relative to a user-selected baseline year (here: 2006). Even though the first few years of the monitoring programme suffer from poor representativity (see section 2.3), weights and smoothing-function enabled accurate estimates in these years without distorting following estimates (Buckland and Johnston 2017). Abundance indices were calculated as

$$\text{Abundance index} = \frac{ab_{ij}}{\overline{ab_J}}$$

with ab_{ij} the predicted abundance of the ith simulation in year j, and $\overline{ab_J}$ the mean predicted abundance of all simulations in the baseline year J (here 2006). Based on 32000 simulations, we calculated the mean index and its 95% CrI for each time step j.

285

Data exclusion

Based on an internal validation process of the LANUV, some abundance data were classified as implausible and therefore excluded from the analysis. This process was species-dependent, leading to sample sizes of 608 to 614 per species (full dataset: 614, Appendix section 9). Weights were derived
290 from these revised datasets.

2.6 Consistency check

We checked our EAS trends for NRW for consistency with population trends in all West Germany according to the German Common Bird Monitoring scheme (MhB), coordinated by the German
295 association for field ornithologists (DDA). MhB study sites are ideally surveyed each year by the same observer (full programme details in Kamp et al. (2021)).

The consistency analysis included 59 widely distributed species spanning a broad range of habitat associations and breeding strategies for which we had trend estimates available from the MhB as well as for both biogeographic regions in the EAS. We expected that even such structured monitoring
300 programmes should generate largely consistent trend estimates when sufficiently robust, except for well justified cases. We calculated two parameters for each species: first, a Pearson 'correlation' coefficient between the annual TRIM-based MhB indices and the annual EAS-indices extracted from model smoothers between 2005 and 2020; second, an index of 'coincidence' as the fraction of annual MhB index values that fell within the 95% CrI of EAS-derived indices.

305 It is obviously difficult to compare trend estimates between programmes with different methodologies and survey efforts (nine vs. four surveys per year), with different target scales (NRW vs. West Germany) coupled with just a partial match in study sites and regional coverage, and with such fundamentally different analytical approaches (trend smoothers vs. the far more variable yearly point estimates). Nevertheless, largely consistent estimates from EAS and MhB would reassure that the analytical
310 approach developed for rolling sampling strategies produces reliable trends that are robust to the underlying survey design.

3 Results

Breeding bird trends were analysed for 85 out of 148 bird species (full detail in Appendix section 8, 9), but we focus the description of model assessment results on 14 exemplary species (tab. 1). Exemplified data for five species, script, and code are available from Github (https://github.com/m-rieger/EAS_bird) (Rieger et al. 2024) as well as a package with helper functions (https://github.com/m-rieger/EAS) (Rieger 2024).

3.1 Distribution characteristics: Species-specific optimisation

Model optimisation resulted in different best-performing model families and a species-specific selection of coefficients added to the binomial part of zero-inflated models (Table 1, species-specific overview in Appendix section 9). When several models resulted in comparably good performance, we continued with the most parsimonious model, i.e., the model with the lower number of covariates or model components. We illustrate below the selection of model structure on four species with striking differences in overall mean abundance, abundance dispersion, and the proportion of sites with zero counts: Common Kestrel, Eurasian Blackbird, White Wagtail, and Common Chiffchaff. The four measures that inform model optimisation (propZ, mean, SD, max) can – but do not necessarily need to – vary with each other, so that an over- or underestimation of zero counts often goes along with an under- or overestimation of the population mean, and an over- or underestimation of SD with an over- or underestimation of maximum abundance.

Common Kestrel (Table 2, Fig. 3a) represents species with typically low territory numbers. Such patterns are well captured with a Poisson model, so that the addition of zero-inflated or negative binomial distribution families did not improve model fit. Nevertheless, data variation (SD, max) was overestimated by all models, indicating slight underdispersion (and thus conservative coefficient estimates).

Eurasian Blackbird (Table 2, Fig. 3b) represents abundant and widespread species that occur in nearly all study sites. Here, zero-inflation is no issue. However, compared to Kestrel, abundance varied substantially between sites and clearly more than expected by a Poisson distribution, an issue solved using a negative binomial model family.

White Wagtail (Table 2, Fig. 3c) is widespread but at typically low abundance, resulting in a larger fraction of unoccupied sites than expected under a Poisson distribution. A negative binomial model family was insufficient to capture this zero-bias, so the final model added a zero-inflation component to the Poisson model, with environmental PCs and site ID as binomial model coefficients.

Common Chiffchaff (Table 2, Fig. 3c) has a broad abundance distribution comparable to Blackbird but with a larger fraction of unoccupied sites. These two problems were solved by using zero-inflated negative binomial models.

Table 2: Examples for model assessment and procedure to select species-specific models from six alternative model structures (for further species and models see Appendix section 5). Models differ in family and the presence of a zero-inflated model component (zinb and zip), the latter also in binomial model coefficients. Model assessment quantifies whether simulated data from the model is consistent with the observed response variable (abundance) with respect to its proportion of zeros (propZ), mean and maximum value, and standard deviation (SD). Bayesian p-values close to 0.5 indicate that the observed values are central within the posterior distribution of simulated raw datasets (32000 in total), indicating good model fit (bold values range between 0.3 and 0.7). Values close to 0 and 1 (grey and italic below 0.025 and above 0.975) indicate substantial under- or overestimation and thus poor model fit. The selection value is a proxy for model fit with values close to 0 indicating good model fit. Abbreviations see Table. 1.

Species	model family	binomial model coefficients	Bayesian p-value				selection value
			propZ	mean	SD	max	
Common Kestrel *Falco tinnunculus*	**pois**		0.71	**0.58**	0.91	*0.98*	**0.46**
	nb		0.72	**0.58**	0.92	*0.99*	**0.46**
	zip	R	0.72	**0.58**	0.91	*0.99*	**0.46**
	zinb	R	0.73	**0.57**	0.92	*0.99*	**0.47**
	zip	PCs + (1|ID)	0.73	**0.64**	0.96	*0.99*	0.53
	zinb	PCs + (1|ID)	0.75	**0.64**	0.96	*0.99*	0.53
Eurasian Blackbird *Turdus merula*	pois		**0.52**	**0.46**	0.08	0.13	0.32
	nb		**0.52**	**0.5**	**0.39**	0.29	**0.06**
	zip	R	0.80	**0.38**	0.08	0.13	0.42
	zinb	R	0.80	**0.46**	**0.41**	0.30	0.14
	zip	PCs + (1|ID)	0.80	**0.39**	0.08	0.12	0.42
	zinb	PCs + (1|ID)	0.80	**0.46**	**0.40**	0.29	0.15
White Wagtail *Motacilla alba*	pois		0.05	**0.59**	0.92	0.85	0.51
	nb		0.05	**0.60**	0.94	0.87	0.54
	zip	R	0.09	**0.57**	0.93	0.84	0.47
	zinb	R	0.10	**0.58**	0.95	0.87	0.50
	zip	**PCs + (1|ID)**	**0.47**	**0.57**	0.88	0.76	**0.21**
	zinb	PCs + (1|ID)	**0.50**	**0.57**	0.89	0.79	0.24
Common Chiffchaff *Phylloscopus collybita*	pois		*0.01*	**0.62**	0.07	0.27	0.49
	nb		*0.01*	**0.66**	**0.39**	**0.52**	0.28
	zip	R	0.27	**0.56**	0.09	0.26	0.28
	zinb	**R**	0.29	**0.60**	**0.41**	**0.50**	**0.06**
	zip	PCs + (1|ID)	0.26	**0.57**	0.09	0.26	0.29
	zinb	PCs + (1|ID)	0.27	**0.60**	**0.41**	**0.51**	**0.07**

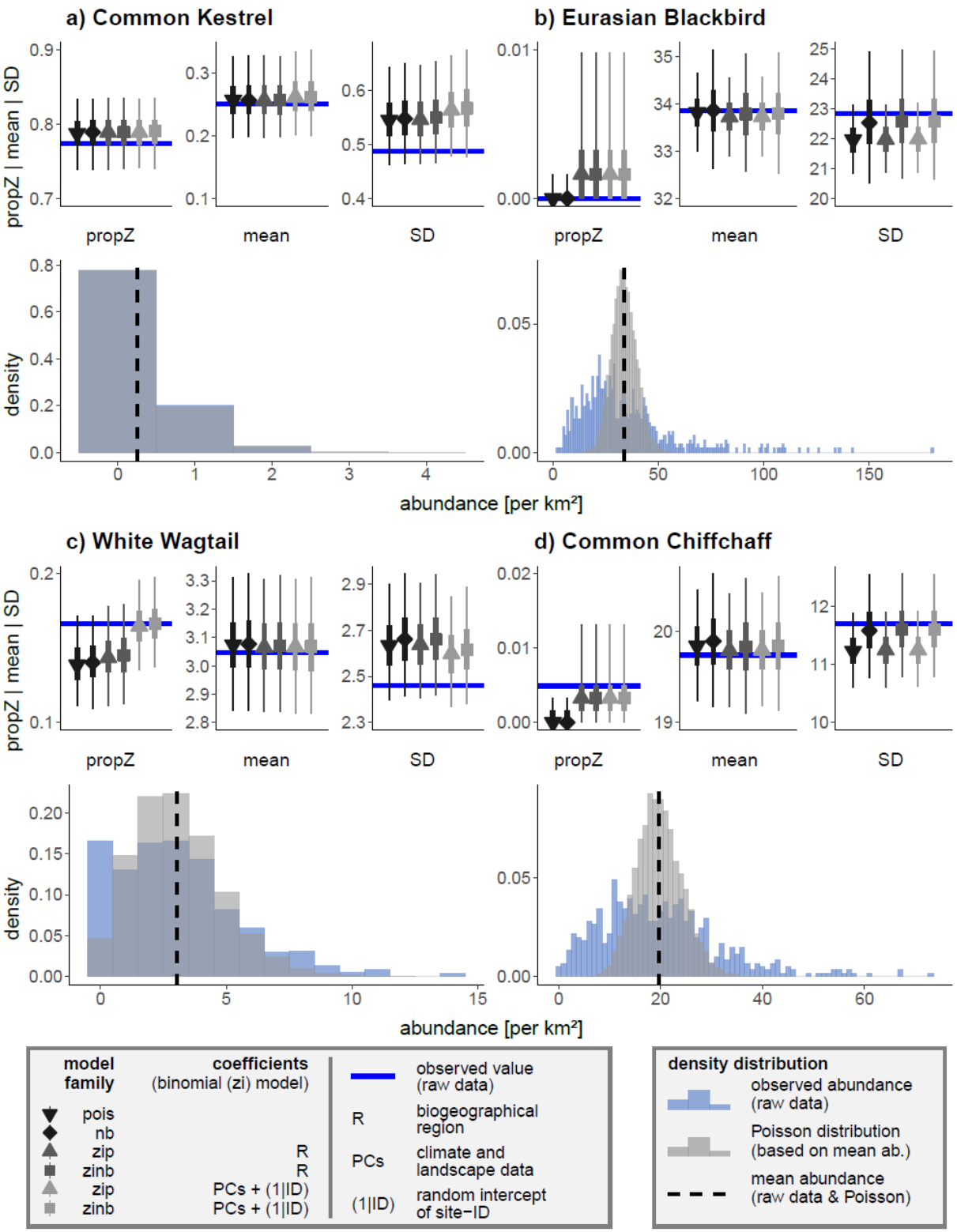

*Figure 3: Reasoning for selecting appropriate model structures for four species. **Upper panels**: Model performance with respect to the proportion of zero counts (propZ), mean, and standard deviation (SD) of the response variable 'abundance'. Symbols with error bars display the median ± 50% CrI (thick bars) and 95% CrI (thin bars) for 32000 posterior raw datasets simulated from the selected models as given in the legend. Optimal models have a close match between observed and model-predicted data. **Lower panels**: Density distributions of observed abundances (blue) and simulated Poisson distributions (grey) based on the observed mean abundance (dashed line). In many cases, the observed distribution is clearly wider than predicted by Poisson, indicating overdispersion. Y-axes differ in scale.*

3.2 Spatial bias

Due to added weights, especially the early years in the EAS-period had less influence on the model than more recent years with a more balanced sampling distribution (Fig. 4, Fig. 2). Therefore, only recent years achieved a sufficient coverage (= mean weight) of almost 1, whereas the first five years did not even approach a coverage of 50% (Fig. 4). Additionally, there were single high weights between 2006 and 2012 (Fig. 4).

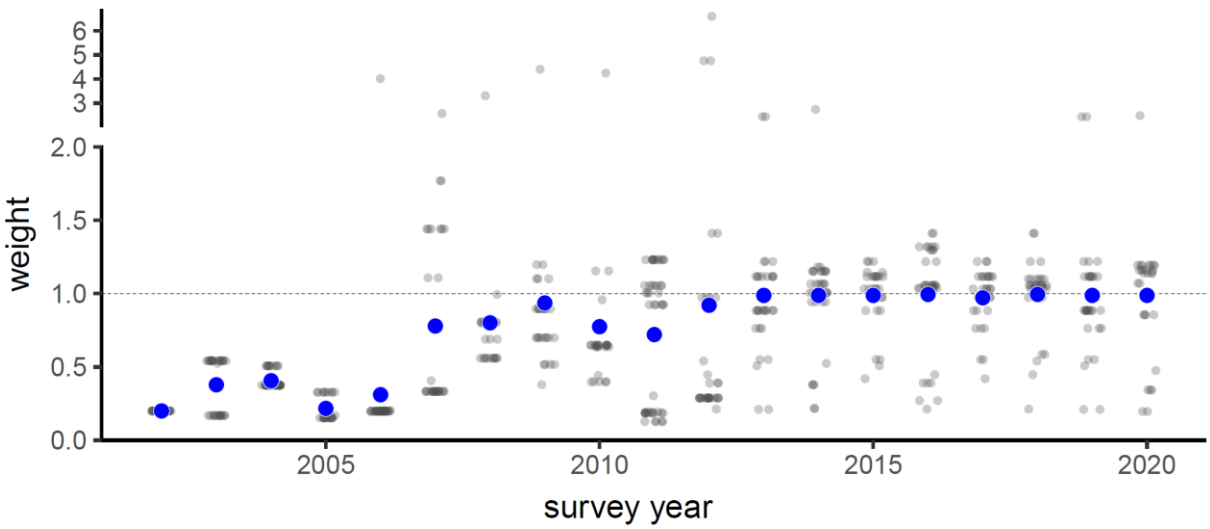

Figure 4: Weights per year and site (grey dots) as well as mean weight per year (blue dots). Sites within a landscape share identical weights within years, resulting in bandings of dots. Note y-axis break.

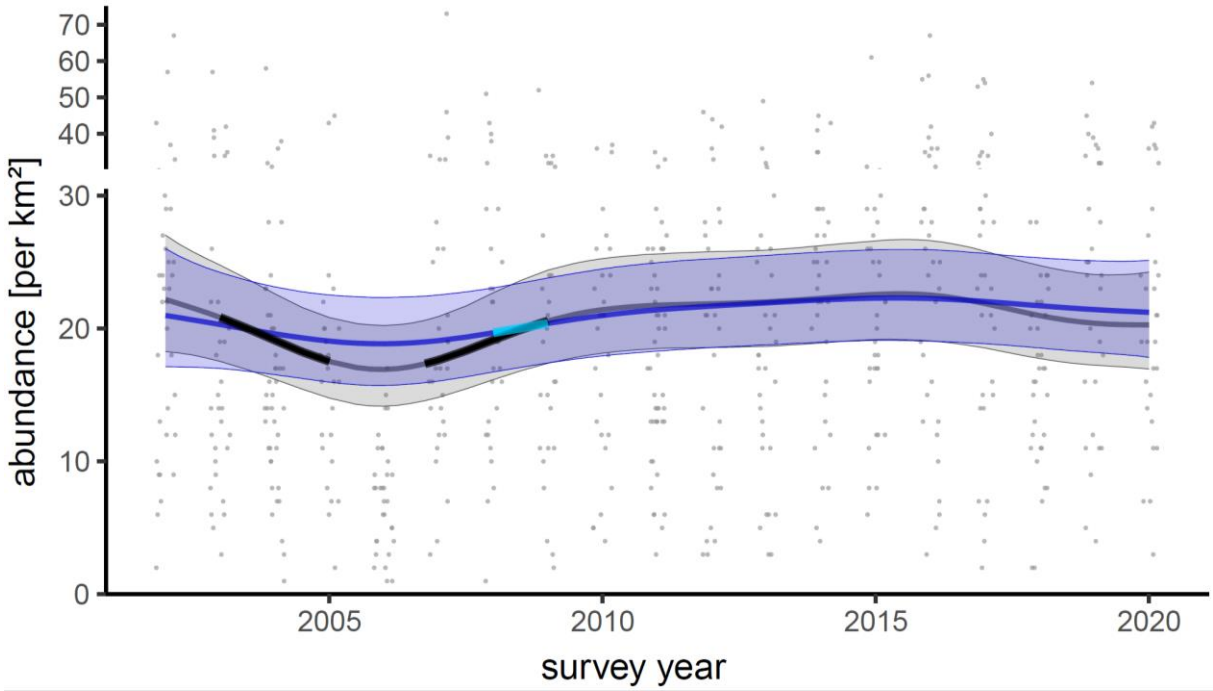

Figure 5: Estimated mean abundance ± 95% CrI for models with (blue) and without weights (grey) added to the response variable abundance, including raw data (grey dots) for Chiffchaff. Slopes substantially differing from 0 (see section 2.5) are highlighted with black (model without weights) and light blue (model with weights). Note y-axis break.

To exemplify the effect of weighting, we compare the predicted population trends for Chiffchaff between models with and without weights (Fig. 5). The latter predicted a decline in abundance until 2006 and an increase afterwards, whereas the former produced more stable abundance estimates with only a mild increase between 2008 and 2009 (Fig. 5). Most sites surveyed in 2006 were located in the 'loess-rich Boerden' (Fig. 2), a landscape with intense farming and typically low Chiffchaff abuandance (Fig. 5). The distorted trend estimated in the unweighted model (Fig. 5, grey trend) was successfully treated with by the weighted model (Fig. 5 blue trend).

3.3 Observer effects

We found a strikingly inhomogeneous distribution of observer effects across years, bearing substantial potential for biased population trend estimates when unaccounted for. High total abundances occurred disproportionally often since 2015, and strikingly low abundances before 2006 (Fig. 6, left), irrespective of the chosen threshold value (Appendix section 6). Models ignoring observer effects generally produced more positive 12-year trend estimates, i.e., resulted in steeper increases or less steep declines than models with the covariate *observer effect* (Fig. 6, right). For Chaffinch, correcting for observer effects even reversed the increase in abundance in the non-corrected model to a decrease in the corrected model (Fig. 6, right). OE-corrected trend estimates were generally more consistent with West German trend estimates form the MhB programme (Appendix section 6).

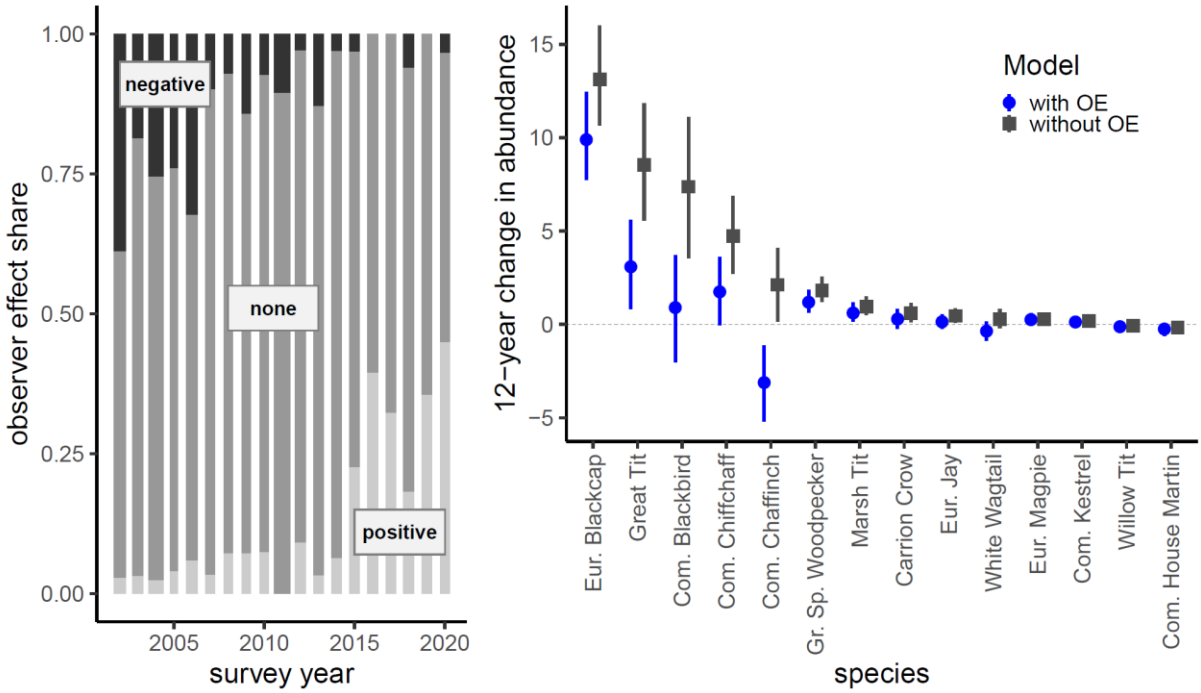

Figure 6: Left: Proportion of surveyed EAS sites per observer effect level per year. Bar widths are proportional to the number of surveyed sites per year. Right: 12-year change in abundance (2008-2019) estimated from models with (circles, blue) and without (squares, grey) the covariate observer effects (OE) for 14 species (example trend curves in Appendix section 6). The y-axis displays the absolute change in abundance across 12 years (and not per year).

3.4 Consistency check

Correlation coefficients between annual EAS indices and West German MhB indices revealed solid matches, i.e., with correlation coefficients and their 95% CI > 0, for 37 out of 59 species (e.g., Blackbird and Chiffchaff in Fig. 7a, b). Only one species (Common Swift, Appendix section 7) showed negatively correlated yearly indices, with correlation coefficient and 95% CI < 0. Likewise, coincidence between index estimates was reasonably high, with 55 (47) out of 59 species having ≥ 75% (90%) of the annual mean MhB indices within the 95% CrI of the respective annual EAS index. High correlation not necessarily goes along with a high coincidence index, because the latter is sensitive to the index baseline year as illustrated by the Chiffchaff data (Fig. 7b). Here, EAS- and MhB-analyses generate a nicely consistent trend pattern, but our coincidence index fails because the exceptionally low MhB index in the baseline year 2006 distorts almost all other annual MhB-index values to exceed the EAS credible intervals. In contrast, Great Tit and Eurasian Jay exemplify how poor correlation between annual index values can go along with high index coincidence (Fig. 7c, d), here because a rather stable vs. more fluctuating trend (Great Tit) that is identified by a weak correlation is hidden within the broad credible intervals of the EAS trend estimate or largely overlapping but slightly diverging trends (Eurasian Jay, especially due to year 2005) fail to correlate. Correlation coefficients, coincidence values and trend displays for all 59 species are in Appendix section 7.

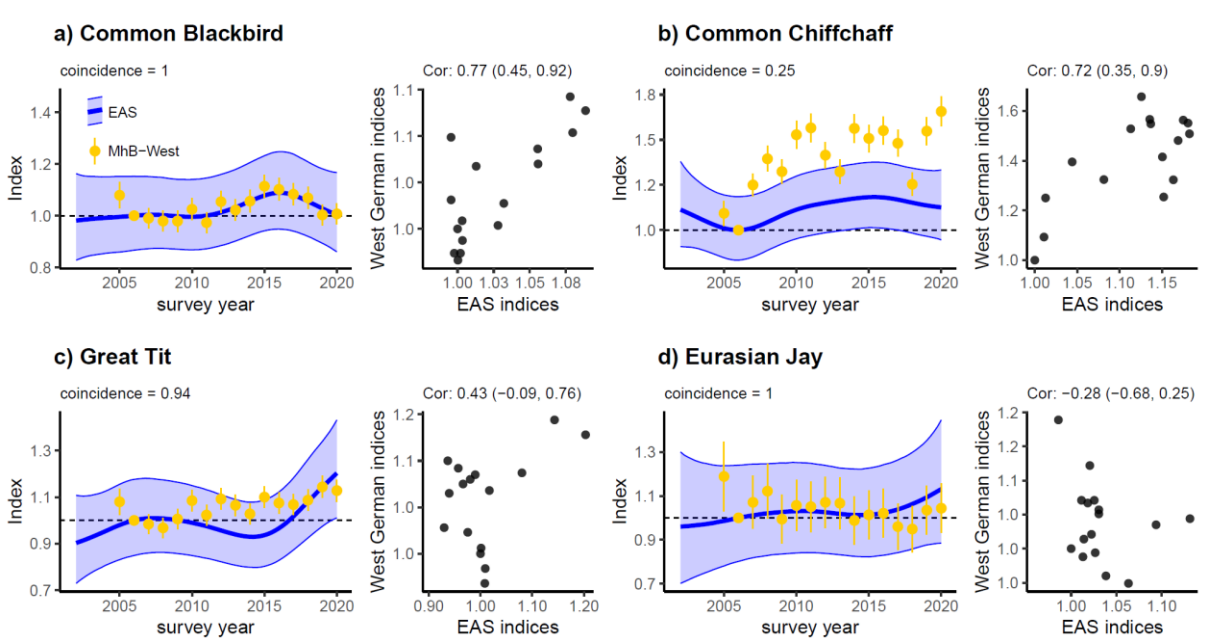

Figure 7: Trend comparison of EAS monitoring programme and West German MhB for four species. Left (per panel): Coincidence of mean annual indices ± 95% CrI (EAS) or CI (MhB-West). Right (per panel): correlation and correlation coefficients (± 95% CI) of annual indices from 2005 - 2020.

4 Discussion

Count data are the raw unit for many population trend analyses, but their distribution poses several challenges during statistical analyses to account for overdispersion (34 out of 85 species in our current dataset) or zero inflation (62 out of 85 species). Moreover, many biodiversity monitoring programmes rest on 'rolling' sampling designs where each study site is visited only every so many years and observer identity frequently changes between years. We developed a workflow that takes these challenges into account (Fig. 8) (Rieger 2024, Rieger et al. 2024). First, we derive a routine that allows informed decisions about the most appropriate combination of residual families (Poisson or negative binomial), model covariates (e.g., habitat characteristics), and zero-inflation formulations to reflect species-specific data distributions. Second, we develop a correction factor that buffers population trend estimates for variation in observer expertise. Third, we integrate model weights that adjust for between-year variation in the representation of habitat or landscape types within the yearly subset of sampled sites. Finally, by using generalized additive mixed models we account for missing values due to rolling surveys. With these corrections in place, we find – within limits of the comparison – good consistency between the EAS trends and trends from the standard German common Bird monitoring scheme.

Challenges	Solutions
zero-inflation & overdispersion	**species-specific model optimization:** → **overdispersion**: nb, zinb → **zero-inflation** (including optimal covariate combination): zip, zinb nb = negative binomial, zinb = zero-inflated neg. binomial, zip = zero-inflated Poisson
observer effects (varying with time)	assess **observer effects** using: $\frac{\textbf{total abundance}\text{ per site and year*}}{\textbf{mean total abundance}\text{ per site*}}$ → *negative* (< 0.75), *none* (0.75-1.25), *positive* (>1.25) *summed up across all species
spatial bias (non-representative habitat coverage)	**weighted abundance** using: $\frac{\textbf{habitat target share}}{\text{(annual) }\textbf{actual share}}$ → prevents bias in predicted abundance caused by shifted habitat composition of surveyed sites
rolling surveys (many missing values)	**generalized additive mixed models (GAMMs):** → smoothing function (*additive*)* → site ID as random intercept (*mixed*) *based on thin plate regression splines

Figure 8: Workflow to address several challenges that arise when analysing breeding bird trends

4.1 Distribution characteristics: Species specific optimisation

We found substantial variation in model structures that best reflected the empirical abundance distributions and fraction of structural zeros across the 85 study species. Model formulations spanned from simple Poisson or negative binomial models (e.g., Common Kestrel and Blackbird; 27.1 % of all modelled species) to models with elaborate zero inflation formulae to account for variation in the fraction of structural zeros across biogeographic regions, study sites, or environmental covariates (e.g., Great Spotted Woodpecker or Marsh Tit; 72.9 % of all modelled species, tab. 1). This variation often tied in with breeding ecologies (cf. examples in section 3.1). Widespread and abundant species with intermediate territory counts in many study sites typically required rather simple count models but no correction for structural zeroes. In contrast, heterogeneously distributed species with a high fraction of unoccupied sites or colonial breeders with often large territory numbers at the few occupied sites almost exclusively required the integration of zero inflation formulae, possibly connected with controls for variation in zero frequencies between regions, sites or environments. This diversity in distribution and abundance patterns is typical for most (bird) monitoring programmes independent of the chosen modelling approach, with similar choices documented elsewhere (Etterson et al. 2009, Tirozzi et al. 2021, 2022, Hernández-Navarro et al. 2023).

4.2 Spatial bias

We found more realistic EAS trend estimates after adding weights that reflect the degree of spatial bias among natural regions and metropolitan areas per year. This was exemplified by woodland species such as Chiffchaff, Chaffinch or Willow Warbler, which showed striking dips in uncorrected abundance estimates in 2006 where survey sites focused on agricultural landscapes with low woodland cover (Fig. 2). Weighting successfully buffered against yearly variation in the coverage of landscape types and thus provided more plausible trend estimates. Similar distortions of trends or species distributions due to spatial bias often occur when using citizen science data since the sampled sites are selected by the observer and thus biased by its preferences (e.g., urban sites, sites of ecological interest (Johnston et al. 2020, Bowler et al. 2022))

Monitoring schemes should obviously aim to avoid the extreme disbalance in sampling that characterizes the early phase of the EAS monitoring, where each year focused on a subset of landscape types. Such bias is difficult to correct *a posteriori* and bears the risk that high weights for rare habitat types overly emphasise few surveys with exceptionally high or low population densities. However, when the yearly sampling subsets are representative, as in the EAS programme since 2013 (Fig. 2), weighting can unfold its full capacity by accounting for minor bias in habitat coverage or the overrepresentation of municipalities among study sites.

4.3 Observer effects

Observer effects caused substantial shifts in abundance estimates. The unfavourable combination of negative observer effects in early years of the EAS programme and positive effects in recent introduced a systematic bias towards an overestimation of positive trends unless observer effects were considered. In severe cases, adding observer effects as covariate successfully buffered, as confirmed by comparison with West German breeding bird trends from the MhB programme of the DDA (Appendix section 6).

Our approach to categorize observer effects from site-specific abundance data risks to mask drastic increases or declines in true bird abundance, e.g., after a substantial change in habitat quality or land-use. Such extreme shifts, however, will typically have diverging species-specific effects on abundance with winners and losers (Lemoine et al. 2007), and thus appears unlikely to massively affect total abundance across species. Moreover, steady increases or declines in total abundance with habitat quality will also change mean total abundance across survey years, so that the threshold beyond which survey years are flagged as 'extreme' also changes accordingly. Our categorization fails, however, to identify cases of species-specific observer effects, such as insufficient expertise in discriminating 'difficult' species (e.g., songs of Garden Warbler vs. Blackcap), or hearing impairment for high pitched vocalisations (e.g., Goldcrest and Firecrest). For EAS, such cases are marked as implausible via an internal validating process and excluded from analysis (see section 2.5 'Data exclusion').

Alternative indices may capture components of observer effort beyond our rather simplistic representation of expertise. A widespread approach integrates survey duration or transect length as model covariates, or combines both measures into survey duration per unit length (Kéry et al. 2005, Kéry and Royle 2020). Kéry et al. (2005) found higher survey effort to increase the detection rates of many bird species, resulting in higher abundance estimates. Others integrated observer age as a (quadratic) model covariate to correct for age-related changes in survey completeness (Farmer et al. 2014). Citizen Science data, e.g. those reported through online platforms such as ornitho (https://www.ornitho.de/) or ebird (https://ebird.org/), further allow estimating observer- and species-specific detection rates from species checklists, providing even more precise correction factors (Johnston et al. 2018). So far, the EAS monitoring programme did not provide data required to extract any of these covariates, but we highly recommend their assessment to correct for possible observer effects.

Some of the observer effects inherent to the EAS dataset likely arise from individual differences in the approach to delineate breeding territories. In future refinements of the program, this could be mitigated by using point observations per survey instead of territories as the basis for trend analyses, possibly in hierarchical models comparable to those used in other monitoring programmes, e.g., the Swiss Breeding Bird survey (Kéry et al. 2005, Kéry and Royle 2020, Strebel et al. 2020). Alternatively, territory

delineation could be automated from raw point observations, e.g., using the AutoTerri algorithm from TerriMap online developed by the Swiss Ornithological Institute (Wechsler 2018).

4.4 Consistency check

Population trend estimates for 59 species revealed a reasonably close match between the GAMM-based trend smoothers applied on the EAS dataset and TRIM-based yearly point estimates applied in the West German common breeding bird census. Despite substantial differences in the chosen modelling approach and spatial coverage, most trend estimates had solid positive correlations and/or a good coincidence between annual indices (Fig. 7, Appendix section 7). In many cases, smoothed GAMM-trends for the EAS data captured even short-term population fluctuations at 3–7-year scale, similar to the – by definition – more fine-grained yearly MhB indices (e.g., Common Blackbird, Eurasian Wren, Firecrest, Appendix section 7). Neutral or even negative correlations between yearly EAS- and MhB-indices occurred primarily in two species subsets. The first concerns colonial birds (e.g., Common Swift, Tree Sparrow, Common House Martin, Barn Swallow, Appendix section 7) where trend estimates are highly sensitive to the share of study sites with colonies present, the influence of few large colonies on the trend estimates, the precision of nest counts, and observer access to a given colony (Südbeck et al. 2005, LANUV 2016a). The second subset concerns species with pronounced affinity to urban areas as derived from low or high coefficient estimates for the linear effect of PC3 (Appendix section 8), where affinity can be positive (e.g., Eurasian Magpie, Dunnock, Common House Martin, Common Kestrel and Black Redstart) or negative (e.g., Common Chaffinch, Eurasian Jay, Marsh Tit or Crested Tit). Given an exceptionally high share and density of urban areas in NRW compared to other West German federal states, NRW trends are plausibly more positive than West German MhB trends for urban specialists (e.g., Eurasian Magpie, Dunnock, Common Kestrel, Black Redstart, Appendix section 7) and more negative for non-urban species (e.g. Common Chaffinch and Crested Tit, Appendix section 7).

5 Conclusion

We developed a modelling routine for bird population trend analyses that avoids several common weaknesses in existing approaches. Our approach accounts for unbalanced sampling of landscape types, optimises species-specific model structure with respect to zero-inflation and overdispersion, partially accounts for variation in observer expertise, and integrates GAM-based smoothing to bridge even extensive gaps between replicated surveys in a ‘rolling’ sampling design. Our consistency check against trends from the Germany common breeding bird programme indicates that our routine produces reliable and robust trends for most species. The R script and package provided with this study can be adapted to other monitoring programmes with comparable survey designs and data structures.

Acknowledgements

We thank the many observers for contributing to the EAS and MhB monitoring programmes, Matteo Santon for statistical and computational advice, and Jakob Katzenberger, Johannes Kamp, and Johannes Wahl for comments on earlier versions of this workflow.

Funding

The development of this workflow has been financially supported by State Agency for Nature, Environment and Consumer Protection in North Rhine-Westphalia (LANUV) to University of Tübingen.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, scripts, code and supplementary information availability

A data example for five species, R script, and code are available from Github (https://github.com/m-rieger/EAS_bird) (Rieger et al. 2024) as well as a package with helper functions (https://github.com/m-rieger/EAS) (Rieger 2024). Further figures, model outputs and detailed descriptions are provided in the electronic supplementary material ('EAS-bird_Appendix.html').

References

ATKIS. 2013. Amtliches Topographisch-Kartographisches Informationssystem. https://www.adv-online.de/AdV-Produkte/Geotopographie/ATKIS/

Blasco-Moreno, A., M. Pérez-Casany, P. Puig, M. Morante, and E. Castells. 2019. What does a zero mean? Understanding false, random and structural zeros in ecology. Methods in Ecology and Evolution **10**: 949-959. https://doi.org/10.1111/2041-210X.13185

Bogaart, P., M. van der Loo, and J. Pannekoek. 2020. Package 'rtrim' - Trends and Indices for Monitoring Data, version 2.1.1. R package. https://github.com/markvanderloo/rtrim

Bowler, D. E., C. T. Callaghan, N. Bhandari, K. Henle, M. Benjamin Barth, C. Koppitz, R. Klenke, M. Winter, F. Jansen, and H. Bruelheide. 2022. Temporal trends in the spatial bias of species occurrence records. Ecography **2022**: e06219. https://doi.org/10.1111/ecog.06219

Buckland, S., and A. Johnston. 2017. Monitoring the biodiversity of regions: Key principles and possible pitfalls. Biological Conservation **214**: 23-34. https://doi.org/10.1016/j.biocon.2017.07.034

Bürkner, P.-C. 2017. brms: An R package for Bayesian multilevel models using Stan. Journal of statistical software **80**: 1-28. https://doi.org/10.18637/jss.v080.i01

Burns, F., M. A. Eaton, I. J. Burfield, A. Klvaňová, E. Šilarová, A. Staneva, and R. Gregory. 2021a. Abundance decline in the avifauna of the European Union conceals complex patterns of biodiversity change. Authorea Preprints. https://doi.org/10.22541/au.162557488.83915072/v1

Burns, F., M. A. Eaton, I. J. Burfield, A. Klvaňová, E. Šilarová, A. Staneva, and R. D. Gregory. 2021b. Abundance decline in the avifauna of the European Union reveals cross-continental similarities in biodiversity change. Ecology and Evolution **11**: 16647-16660. https://doi.org/10.1002/ece3.8282

Campbell, H. 2021. The consequences of checking for zero-inflation and overdispersion in the analysis of count data. Methods in Ecology and Evolution **12**: 665-680. https://doi.org/10.1111/2041-210X.13559

CDC. 2010. Deutscher Wetterdienst Climate Data Center: Vieljährige Stationsmittelwerte für die Klimareferenzperiode 1981-2010. https://www.dwd.de/DE/klimaumwelt/cdc/cdc_node.html

Dakki, M., G. Robin, M. Suet, A. Qninba, M. A. El Agbani, A. Ouassou, R. El Hamoumi, H. Azafzaf, S. Rebah, and C. Feltrup-Azafzaf. 2021. Imputation of incomplete large-scale monitoring count data via penalized estimation. Methods in Ecology and Evolution **12**: 1031-1039. https://doi.org/10.1111/2041-210X.13594

Etterson, M. A., G. J. Niemi, and N. P. Danz. 2009. Estimating the effects of detection heterogeneity and overdispersion on trends estimated from avian point counts. Ecological applications **19**: 2049-2066. https://doi.org/10.1890/08-1317.1

Farmer, R. G., M. L. Leonard, J. E. Mills Flemming, and S. C. Anderson. 2014. Observer aging and long-term avian survey data quality. Ecology and Evolution **4**: 2563-2576. https://doi.org/10.1002/ece3.1101

Fewster, R. M., S. T. Buckland, G. M. Siriwardena, S. R. Baillie, and J. D. Wilson. 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology **81**: 1970-1984. https://doi.org/10.1890/0012-9658(2000)081[1970:AOPTFF]2.0.CO;2

Gabry, J., and R. Češnovar. 2020. Cmdstanr: R interface to'CmdStan'. https://github.com/stan-dev/cmdstanr/

Hernández-Navarro, A. J., F. Robledano, M. V. Jiménez-Franco, J. A. Royle, and J. F. Calvo. 2023. Long-term trends of local bird populations based on monitoring schemes: are they suitable for justifying management measures? Journal of Ornithology: 1-13. https://doi.org/10.1007/s10336-023-02114-3

Jennings, K. 2021. Deploying data for nature conservation: How monitoring data is used to inform conservation policy and practice [Keynote address]. BOUsci21 - Developments in monitoring science, United Kingdom, November 24-25.

Jiguet, F. 2009. Method learning caused a first-time observer effect in a newly started breeding bird survey. Bird Study **56**: 253-258. https://doi.org/10.1080/00063650902791991

Jiguet, F., M. Barbet-Massin, V. Devictor, N. Jonzén, and Å. Lindström. 2013. Current population trends mirror forecasted changes in climatic suitability for Swedish breeding birds. Bird Study **60**: 60-66. https://doi.org/10.1080/00063657.2012.733337

Johnston, A., D. Fink, W. M. Hochachka, and S. Kelling. 2018. Estimates of observer expertise improve species distributions from citizen science data. Methods in Ecology and Evolution **9**: 88-97. https://doi.org/10.1111/2041-210X.12838

Johnston, A., D. Fink, M. D. Reynolds, W. M. Hochachka, B. L. Sullivan, N. E. Bruns, E. Hallstein, M. S. Merrifield, S. Matsumoto, and S. Kelling. 2015. Abundance models improve spatial and temporal prioritization of conservation resources. Ecological applications **25**: 1749-1756. https://doi.org/10.1890/14-1826.1

Johnston, A., N. Moran, A. Musgrove, D. Fink, and S. R. Baillie. 2020. Estimating species distributions from spatially biased citizen science data. Ecological Modelling **422**: 108927. https://doi.org/10.1016/j.ecolmodel.2019.108927

Kamp, J., C. Frank, S. Trautmann, M. Busch, R. Dröschmeister, M. Flade, B. Gerlach, J. Karthäuser, F. Kunz, and A. Mitschke. 2021. Population trends of common breeding birds in Germany 1990–2018. Journal of Ornithology **162**: 1-15. https://doi.org/10.1007/s10336-020-01830-4

Kéry, M., and J. A. Royle. 2020. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 2: Dynamic and Advanced Models. Academic Press. https://doi.org/10.1016/C2015-0-04070-9

Kéry, M., J. A. Royle, and H. Schmid. 2005. Modeling avian abundance from replicated counts using binomial mixture models. Ecological applications **15**: 1450-1461. https://doi.org/10.1890/04-1120

Klvaňová, A., P. Voříšek, R. D. Gregory, A. van Strien, and A. Gmelig Meyling. 2009. Wild birds as indicators in Europe: latest results from the Pan-European Common Bird Monitoring Scheme (PECBMS). Avocetta **33**: 7-12.

Knape, J. 2016. Decomposing trends in Swedish bird populations using generalized additive mixed models. Journal of Applied Ecology **53**: 1852-1861. https://doi.org/10.1111/1365-2664.12720

König, H. 2020. Standards und Indikatorartengruppen - Erfahrungen aus der ÖFS in Nordrhein-Westfahlen. Pages 17-32 *in* W. Züghart, S. Stenzel, and B. Fritsche, editors. Umfassendes bundesweites Biodiversitätsmonitoring. BfN Schriften. https://www.bfn.de/sites/default/files/BfN/service/Dokumente/skripten/skript585.pdf

Korner-Nievergelt, F., T. Roth, S. Von Felten, J. Guélat, B. Almasi, and P. Korner-Nievergelt. 2015. Bayesian data analysis in ecology using linear models with R, BUGS, and Stan. Academic Press. https://doi.org/10.1016/C2013-0-23227-X

LANUV. 2016a. Arbeitsanleitung für Brutvogel-Revierkartierungen im Auftrag des LANUV NRW. Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany: 1-79. https://methoden.naturschutzinformationen.nrw.de/methoden/web/babel/media/Arbeitsanleitung_fuer_Brutvogel_Revierkartierungen_NRW.pdf

LANUV. 2016b. Daten zu Natur in Nordrhein-Westfalen 2016 - LANUV-Fachbericht 83. Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany: 1-216. https://www.lanuv.nrw.de/fileadmin/lanuvpubl/3_fachberichte/Daten_zur_Natur_in_NRW_2016.pdf

Lemoine, N., H.-G. Bauer, M. Peintinger, and K. Böhning-Gaese. 2007. Effects of climate and land-use change on species abundance in a central European bird community. Conservation Biology **21**: 495-503. https://doi.org/10.1111/j.1523-1739.2006.00633.x

Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological applications **8**: 258-268. https://doi.org/10.1890/1051-0761(1998)008[0258:EPCFCD]2.0.CO;2

Nielsen, S. E., D. L. Haughland, E. Bayne, and J. Schieck. 2009. Capacity of large-scale, long-term biodiversity monitoring programmes to detect trends in species prevalence. Biodiversity and Conservation **18**: 2961-2978. https://doi.org/10.1007/s10531-009-9619-1

Niemelä, J. 2000. Biodiversity monitoring for decision-making. Pages 307-317 *in* Annales Zoologici Fennici. JSTOR. http://www.jstor.org/stable/23735723

Pannekoek, J., and A. J. van Strien. 2001. TRIM 3 manual (Trends & Indices for Monitoring Data). CBS Statistics Netherlands, Voorburg, The Netherlands.

Pöysä, H., J. Rintala, A. Lehikoinen, and R. A. Väisänen. 2013. The importance of hunting pressure, habitat preference and life history for population trends of breeding waterbirds in Finland. European Journal of Wildlife Research **59**: 245-256. https://doi.org/10.1007/s10344-012-0673-8

R Core Team. 2024. R: a language and environment for statistical computing. R Foundation for Statistical Computing website. https://www.R-project.org/

Ribeiro Jr, P. J., and P. J. Diggle. 2020. Package 'geoR' - Analysis of Geostatistical Data. R package. http://www.leg.ufpr.br/geoR/

Rieger, M. R. 2024. EAS - Module to analyse population trends from monitoring data with highly structured sampling designs. Zenodo. https://doi.org/10.5281/zenodo.12517388

Rieger, M. R., C. Grüneberg, M. Oberhaus, S. Trautmann, M. Parepa, and N. Anthes. 2024. EAS_bird - Data and script of 'Analysing bird population trends from monitoring data with highly structured sampling designs'. Zenodo. https://doi.org/10.5281/zenodo.12518872

Sauer, J. R., K. L. Pardieck, D. J. Ziolkowski Jr, A. C. Smith, M.-A. R. Hudson, V. Rodriguez, H. Berlanga, D. K. Niven, and W. A. Link. 2017. The first 50 years of the North American breeding bird survey. The Condor: Ornithological Applications **119**: 576-593. https://doi.org/10.1650/CONDOR-17-83.1

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American breeding bird survey. The Auk **111**: 50-62. https://doi.org/10.2307/4088504

Strebel, N., R. Dröschmeister, H. Schmid, I. Stützle, S. Trautmann, and J. Wahl. 2020. Ein Vergleich zwischen revier-und individuenbasierten Verfahren zum Abschätzen von Brutbestandsindizes im Monitoring häufiger Brutvögel. Die Vogelwelt **140**: 183-206.

Südbeck, P., H. Andretzke, K. Gedeon, T. Schikore, K. Schröder, S. Fischer, and C. Sudfeldt. 2005. Methodenstandards zur Erfassung der Brutvögel Deutschlands. Max-Planck-Institut für Ornithologie. Vogelwarte Radolfzell.

Svensson, S. E. 1978. Census efficiency and number of visits to a study plot when estimating bird densities by the territory mapping method. Journal of Applied Ecology **16**: 61-68. https://doi.org/10.2307/2402728

Tirozzi, P., V. Orioli, O. Dondina, L. Kataoka, and L. Bani. 2021. Species Traits Drive Long-Term Population Trends of Common Breeding Birds in Northern Italy. Animals **11**: 3426. https://doi.org/10.3390/ani11123426

Tirozzi, P., V. Orioli, O. Dondina, L. Kataoka, and L. Bani. 2022. Population trends from count data: handling environmental bias, overdispersion and excess of zeroes. Ecological Informatics **69**: 101629. https://doi.org/10.1016/j.ecoinf.2022.101629

van Strien, A., J. Pannekoek, W. Hagemeijer, and T. Verstrael. 2004. A loglinear Poisson regression method to analyse bird monitoring data. Bird **482**: 33-39. https://api.semanticscholar.org/CorpusID:15188304

van Strien, A. J., J. Pannekoek, and D. W. Gibbons. 2001. Indexing European bird population trends using results of national monitoring schemes: a trial of a new method. Bird Study **48**: 200-213. https://doi.org/10.1080/00063650109461219

Vorisek, P., R. D. Gregory, A. J. Van Strien, and A. G. Meyling. 2008. Population trends of 48 common terrestrial bird species in Europe: results from the Pan-European Common Bird Monitoring

Scheme. Revista Catalana d'Ornitologia **24**: 4-14. https://pecbms.info/wp-content/uploads/2022/01/vorisek-etal-2008.pdf

Wechsler, S. 2018. Automating the analysis of territory mapping data in bird monitoring. UNOGIS-Paris Lodron-Universität Salzburg. https://api.semanticscholar.org/CorpusID:53652000

Weiss, J., and K. Schulze-Hagen. 2014. Viele Menschen–viele Vögel? Ornithologie und Vogelschutz in Nordrhein-Westfalen. Charadrius **50**: 3-22.

Wood, S. 2021. mgcv - Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. R package: 1-323. https://doi.org/10.32614/CRAN.package.mgcv

Wood, S., and F. Scheipl. 2020. gamm4 - Generalized Additive Mixed Models using lme4 and mgcv. R package. https://doi.org/10.32614/CRAN.package.gamm4

Zuur, A. F. 2012. A beginner's guide to generalized additive models with R. Highland Statistics Limited Newburgh, NY, USA.