Spatially varying population indices

Jonas Knape

jonas.knape@slu.se
Swedish University of Agricultural Sciences
Box 7044, 750 07 Uppsala, Sweden

Abstract

Large scale monitoring is fundamental for reliably tracking the fate of animal populations under changing environments and land-use practices. A common application of large scale population monitoring data is to produce indices of temporal change in species abundances, which are used in environmental policy assessments of species and biodiversity statuses. For index estimation, spatio-temporal models can be used to take advantage of the spatial component of large scale data in order to better capture and understand spatial variation in population change. This paper presents a generalized approach to estimating indices of relative population change across different spatial and temporal scales from fits of spatio-temporal models to population monitoring data. Using flexible specifications of baselines for indices, the approach can be used for a range of different comparisons of abundance across space and time, aggregated at small as well as large spatial and short as well as long term temporal scales. This is illustrated in an application to Swedish monitoring data of the common cuckoo, for which we estimate a range of national, county-wise and fine scale indices. An R-package, spotr, that aids computation of indices from fitted models accompanies the paper.

Keywords

population index; abundance; model-based; spatio-temporal; population monitoring; biodiversity index

1 Introduction

Estimating the magnitude and nature of population change is of primary interest for assessing the status of species and biodiversity, particularly in light of current and past land-use and climate change. Long-term wildlife monitoring programs covering large spatial areas form a backbone for such assessments at regional and national levels. Data from such programs are often used to produce indices of absolute or relative population size over time. These indices may be used as official indicators of species status, as building blocks for biodiversity indicators, or may be used among a set of other nature indicators with the more general purpose of assessing ecosystem health (Gregory & Strien, 2010).

Indices of species abundance over time are typically derived from statistical models of the original data. These models sometimes assume that population change is constant over large spatial areas such as entire countries or large strata. In practice population change is likely to vary also within large areas (Conn et al., 2015), and there has been an increasing recent interest in models that can deal with spatial variation in more detail (Johnson et al., 2024; Smith et al., 2019; Wikle, 2003). Ignoring spatial variation in population change when it is present could lead to poor model fits, possibly adversely affecting inference of large scale trends, but more importantly to missed opportunities for assessing change at finer spatial resolutions. Mitigation measures, management plans, or grassroots initiatives are often implemented at local scales and may be better informed by status assessments at those scales (Davey et al., 2010). Responses of populations to climate change may require more fine scale spatial estimation (Barnett et al., 2021) as such changes can be slow and occur mainly at edges of species ranges. Similarly, responses to land-use change may be localized to areas undergoing change. Inference about when and where changes in populations have occurred can further provide clues for underlying mechanisms that may be obscured at the level of regional or national status assessments (Bowler et al., 2021).

Spatio-temporal analyses of population change include models assuming separate trends across different discrete spatial areas (Sauer & Link, 2011) or habitat types (Newson et al., 2009), models that link trends to multiple environmental covariates via machine learning (Smith et al., 2019), models with trends as an explicit spatio-temporal statistical process (Breivik et al., 2021; Vanhatalo et al., 2017), models assuming linear but spatially varying trends (Thorson et al., 2023), and models including spatio-temporal smoothing (Harrison et al., 2014). They range in generality from being customised and tai-

lored to certain species (Wikle, 2003) or monitoring programs (Sauer & Link, 2011) to specifications that can be estimated using more generic model fitting tools, such as general purpose R packages for GLMM or GAMs.

Abundance indices are usually defined from annual aggregates of (possibly weighted) abundance estimates across the whole of a study area or over a smaller administrative area (Harrison et al., 2014; Sauer & Link, 2011), or from aggregate abundance estimates over the season (Dennis et al., 2013). To estimate an index of change, these annual aggregates are then compared to an aggregate over the same area in a baseline year or a mean over a period of years (Gregory et al., 2019; Knape, 2023). Here, we use more flexible specifications of both the numerator and baseline (denominator) to present a generalized definition of abundance indices. By allowing the numerator and baseline to be defined from arbitrary time periods as well as arbitrary spatial configurations, a wide range of comparisons of absolute or relative abundances or densities across both space and time fits in under this definition. In a case study, we fit four spatio-temporal models based on hierarchical GAMs (Pedersen et al., 2019) to data from the Swedish bird survey. We show how different specification of indices for these models can be used to visualise an array of aspects of spatio-temporal changes and their uncertainties at various spatial and temporal scales. An R-package, `spotr`, assisting extraction of indices from spatio-temporal models fitted to population monitoring data via the R packages `mgcv` or `brms`, or directly from posterior simulations, is provided with the paper.

2 Methods

2.1 Data

We consider data in the form of observed abundances, $y_{i,t}$, of some organism at a set of sites $i = 1, \ldots, S$ spread out across a spatial area (e.g. a country) and for times $t = 1, \ldots, T$. We will sometimes refer to observed abundances as counts since for birds, the target of the case study below, monitoring is usually conducted by counting all individuals heard or seen. However, other abundance measures such as presence absence or non-integer data such as biomass also fit into the general framework (section 2.2.2). Similarly we will often refer to time points as years, but time can also be measured in weeks, days, or other units and need not be evenly spaced in time.

Counts at the same site i are repeated across time so that there is a longitudinal structure to the data. Typically counts from some sites will be missing in some years.

Locations of sites should preferably be randomly or systematically selected. However, to facilitate volunteer participation some survey designs let observers choose sites opportunistically. In this case existing sites may disappear and new sites enter the survey as time passes. To mitigate bias due to systematic appearance or disappearance of survey sites with lower or higher abundances, site effects are often included in models fitted to data (Thomas, 1996; van Strien et al., 2000).

2.2 Estimating population indices

We first review a general approach for computing population indices for spatial areas from large scale survey data. We focus on relative abundance indices since absolute abundance can often not be reliably estimated from population monitoring data. When data are informative about absolute abundances, the approach can however be adapted by dropping the denominator.

A standard base model for estimating overall population indices is built around site and time effects (Fewster et al., 2000; ter Braak et al., 1994),

$$E(y_{i,t}) = \exp(\alpha_i + \beta_t)$$

where α_i are the site and β_t the time effects. The stochasticity of the response is typically modelled via some appropriate distribution, such as a Poisson or negative binomial for counts. To compute global population indices for the whole study area in a target year t from a fit of such a model, one may predict population size across a set of sites for year t and divide by a predicted baseline population size in a reference year (here the first year):

$$I_t = \frac{\sum_i \hat{N}_{i,t}}{\sum_i \hat{N}_{i,1}} = \exp(\beta_t - \beta_1)$$

The site effects thus cancel when computing the indices and only the time effects are needed to compute the index. This is due to the fundamental assumption of the model that the time effects are the same across the surveyed area.

2.2.1 Indices over geographical areas

More generally, models may include variation in population trends across space, and there may be interest in estimating local as well as global abundance indices. Generalising the model above to encompass spatially varying trends, as well as nuisance effects, we may write

$$E(y_{i,t}) = \exp(\mu(\mathbf{s}_i, t) + \eta(i, t))$$

where $\mu(\mathbf{s}_i, t)$ is a predictor of population size at the spatial location $\mathbf{s}_i$ of site i and $\eta(i, t)$ contains nuisance effects that capture variation in counts that are not of primary biological interest, such as observer effects or effects capturing variation in effort or detectability etc. The idea is that $\mu(\mathbf{s}_i, t)$ will be used in computing indices while $\eta(i, t)$ will be controlled for in estimation but not included when predicting abundance.

Now consider computing an index for a spatial area S, which could be the full area that the survey is targeting such as a country, a subregion of that area, for example a county or region, the area covered by a specific habitat type, or just a single small-scale pixel. Ideally, the index would be computed by integrating the predicted relative abundance over all spatial locations in the area:

$$\hat{N}_{S,t} = \int_{\mathbf{s} \in S} \exp(\mu(\mathbf{s}, t))$$

and the index for S formed from

$$I_{S,t} = \frac{\hat{N}_{S,t}}{\hat{N}_{S,1}}$$

Depending on the model, these integrals may be more or less easily computed. For example, if $\mu(\mathbf{s}, t)$ can be decomposed into additive temporal and spatial effects over S (but not necessarily across the full spatial extent) then the spatial effects may be ignored in computing $I_{S,t}$ because the spatial effects

cancel in the same way as they did for the base model (e.g. Sauer & Link, 2011).

In general it may be difficult to exactly calculate the integrals. Instead, they may be approximated by finite sums across a set of prediction points, $\mathbf{p}_1, \ldots, \mathbf{p}_K$, in area S (e.g. Breivik et al., 2021). The prediction points do not have to be identical to the locations of the observed data. Assuming that the prediction at point $\mathbf{p}_k$ represent an area A_k of S, with $\sum A_k =$ area of S, the predicted abundance in S may be approximated by the area weighted sum

$$\hat{N}_{S,t} \approx \sum_k A_k \hat{N}_{\mathbf{p}_k,t}$$

A good selection of prediction points is important for ensuring that the approximation is accurate, but will be model dependent. For example, if abundances vary discretely in space, i.e. they are constant within each of several discrete geographic units or habitat types within S, a single prediction point within each discrete unit may be selected and the A_k set to the area of each discrete unit divided by the area associated with the prediction (e.g. Sauer & Link, 2011). On the other hand if abundances vary continuously in space, a grid of prediction points whose density depends on the spatial resolution of the spatial variation of $\mu(\mathbf{s}_i, t)$ is a better choice (e.g. Bled et al., 2013; Breivik et al., 2021; Harrison et al., 2014).

The unit of measurement of A_k does not affect relative indices. It is then sufficient to define A_K as relative weights, for example they could be scaled to sum to 1 across the prediction points.

2.2.2 General indices

Indices can be extended by including more general summaries of predictions, both in the target (numerator) and the baseline (denominator), as well as more general weights.

For this, we let P be a set of times over which we want to aggregate abundance, for example a set of years or days of a season in a specific year. Weights $w_{k,t}$ represent the amount that the predicted abundance in point $\mathbf{p}_k$ at time t contribute to abundance in S across the times in P so that

$$\hat{N}_{S,P} = \sum_{\mathbf{p}_k \in S, t \in P} w_{k,t} \hat{N}_{\mathbf{p}_k,t}$$

A general relative index can now be computed by using the aggregated abundance over a set of sites W, which may or may not be the same as S, and a time period B as the reference. For the reference, abundance is estimated as

$$\hat{N}_{W,B} = \sum_{\mathbf{p}_k \in W, t \in B} w^*_{k,t} \hat{N}_{\mathbf{p}_k,t}$$

where the weights $w^*_{k,t}$ correspond to the amount that predicted abundance in site $\mathbf{p}_k$ at time t contribute to the aggregate abundance in the reference. The index targeting abundance in S over the times in P relative to abundance in W over times in B is then defined as

$$I_{S,P} = \frac{\hat{N}_{S,P}}{\hat{N}_{W,B}}.$$

Often each site has the same proportional contribution for all time points and it may be suitable to aggregate over time by taking the mean abundance over all the time points in P. The weights for the target then simplify to

$$w_{k,t} = \frac{1}{|P|} w_k$$

and the weights for the reference to

$$w^*_{k,t} = \frac{1}{|B|} w^*_k$$

where $|P|$ and $|B|$ are the number of time points in P and B. If S and W are identical, different weights w_k and w^*_k are typically not needed and the superscript may be dropped. Weights proportional to the number of time points can be used for instance to compare average abundance between time periods, or to compare abundance across years for seasonal organism like insects.

The generalization to using different sets of sites in the numerator and denominator (S and W) can be used to explore spatio-temporal changes in abundance, rather than local changes within S. This may be relevant for example for species undergoing spatial expansions, where the baseline might be set to the population density in a source area at the beginning of the expansion.

Note that the weights used here are weights for predictions from previously fitted models, they should not be confused with weights used during model fitting.

2.3 Case study

We use data on common cuckoo (*Cuculus canorus*) from the Swedish bird survey (Lindström & Green, 2020) to illustrate how indices at different spatial scales can be computed and visualised using the approach described above. Worked R-code for the case study is included in Appendix A.

The Swedish bird survey scheme started in 1996 and uses a design where survey sites are systematically placed on a regular grid across Sweden. Sites are surveyed once per year by volunteers doing line transect counts of all species heard or seen along the edges of a 2x2 km square. Not all sites are surveyed every year, particularly in the northern parts of Sweden that are sparsely populated and where sites are generally less accessible. On average, between 400 and 500 out of 716 sites are surveyed annually but fewer sites were surveyed in the first few years.

We fit four different GAMM models to annual site-wise counts of the common cuckoo between 2000 and 2020. The data are restricted to be within the convex hull of all sites where the species has been found at least once. All models use a negative binomial response distribution with a log-link and have the same overall structure,

$$y_{i,t} \sim \text{NegBin}(\exp(\mu_0 + f_1(t) + f_2(\mathbf{s}_i) + f_3(t, \mathbf{s}_i) + \alpha_i + \gamma_t)),$$

with an intercept μ_0, a smooth temporal component f_1, a smooth spatial component f_2 and a spatio-temporal interaction f_3. In addition there are random site effects, α_i, and random year effects, γ_t. The four models share the same specification of the temporal and spatial main effects (f_1 and f_2) and random effects (α_i and γ_t), and differ only in how the spatio-temporal interaction term (f_3) is specified. The factorization into main effects and a spatio-temporal interaction facilitates model comparisons, e.g. making it easy to compare models with and without a spatio-temporal component, but

more importantly allows modelling the different components with different degrees of complexity. With spatio-temporal bird data, there is usually large spatial variation but less pronounced temporal variation. For the spatial main effect we therefore use a smooth two dimensional function with a high dimensional basis (k = 60) to allow capturing fine scale spatial variation in the intercept, for the temporal main effect we use a smooth function with a moderate basis dimension (k = 9) to allow non-linear that are not overly wiggly (Fewster et al., 2000), and for spatio-temporal interactions we use moderate basis dimensions for both space and time (k = 25 and k = 9, respectively). The four models treat the spatio-temporal interaction as either missing or as smooth or discrete in time and space:

- Model M_0 has no spatio-temporal interaction ($f_3 = 0$).
- Model M_{ss} fits a smooth spatio-temporal surface. The spatio-temporal interaction is constructed from a tensor product of temporal and spatial bases using an `mgcv` interaction of the form `ti(yr, lat, lon, k = c(9, 25), d = c(1,2))`. The model is a variant of the GAM of Harrison et al. (2014), but separating the spatio-temporal variation into main effects and interaction and including random site and year effects, α_i and γ_t.
- Model M_{sd} fits temporal smooth curves that vary discretely in space among counties. The spatio-temporal interaction f_3 is here modelled as county specific deviations from the overall trend f_1, assuming similar smoothness for each county. This is done using an `mgcv` interaction of the form `s(yr, county, k = 9, bs = "fs")`. The area specific deviations have penalized first derivatives. This may be thought of as random smooth curves with random slopes, i.e. county specific trends are shrunk towards the overall trend f_1 (Pedersen et al., 2019). On a technical note the `bs = "fs"` term includes random county intercepts which we force to zero to reduce confounding with the random site effects and make the models more easily comparable.
- Model M_{ds} fits smooth spatial surfaces discretely in time using an `mgcv` interaction of the form `s(factor(yr), lon, lat, k = c(25), bs = "fs")`. Here, the spatio-temporal interaction is modelled as smooth deviation surfaces for each year, assuming similar smoothness for all years. First derivatives are penalized so that the model may be viewed as random surfaces shrunk towards the overall smooth surface f_2.

We fit the models using the `gam` function from the `mgcv` package (Wood, 2017) with smoothing penalties selected using REML. For all smooth components we use reduced rank thin-plate regression spline bases, which are the default in `mgcv`. These bases do not depend on knots and therefore avoid the need to determine their placement. Full details of model specification and fitting are given in Appendix A.

An assumption behind the models is that detection rates do not change considerably over time or space. A trend in detection rates over time would bias trend estimates if it cannot be controlled for via covariates. Spatial variation in detection that cannot be controlled for could cause bias in indices comparing abundance between areas, but could also lead to bias in temporal trends when abundance trends vary across space. Detection rates are not known for the Swedish data, but counts follow a standard protocol aimed to reduce the risk of trends in these rates.

2.3.1 Other models

Many variations on and alternatives to the above model types can be fit with `mgcv`, or alternatively `brms`, including generalized linear models, models with spatially varying slopes, models with random slopes, and models with discrete spatial structure interacting with temporal components. These packages also support bases for the smooths other than thin plate regression splines, including cyclic smooths, gaussian processes, and markov random fields. For an overview of some of the possibilities of `mgcv`, see Wood (2017) and Pedersen et al. (2019).

2.3.2 Assessing model fit

We assessed model fit using standard tools available in the `mgcv` package, and using additional checks of randomized quantile residuals (Dunn & Smyth, 1996) and random effects. The checks include investigating the negative binomial response distribution, and the geographical and temporal patterns of residuals and random effects (Appendix B).

2.3.3 Index estimation

To compute indices for the fits of the models to the cuckoo data, we use a grid of points across Sweden with distances of 25 kilometers to construct the prediction points, with one point per grid cell and year. The grid is restricted to be within the convex hull of all the routes where the species has been observed in the same way that the data was restricted.

We first compute indices for the whole country. I.e. letting S be the full spatial extent of the data and using all the grid points to carry out the numerical integration. Second, we compute indices for each of the 21 counties of Sweden. For this we let S be the individual counties and use all the points in the grid within the counties for the numerical integration. Third, we compute indices for each grid point to map changes at a higher spatial resolution.

In prediction for computing indices, we treat site effects as nuisance parameters (part of η). Year effects are treated either as nuisance parameters, producing less variable and typically more smooth estimates of long-term population change, or are included in predictions, which will produce non-smooth estimates of annual indices (Knape, 2016). Weights are simply set to 1 as each grid point represents the same spatial area. For an example of use of non-constant areal weights, see Appendix A.

To illustrate use of non-trivial weights and comparisons across space we compute indices of how population density in the Swedish counties has changed relative to the overall density in year 2000. We then use the inverse of the number of prediction points of counties as weights for the target, $w_{k,t} = \frac{1}{n_S}$ where n_S is the number of prediction points in the county S. For the baseline we use the average density across all of Sweden so that W is the whole study area and with weights $w^*_{k,1} = \frac{1}{n}$ where n is the total number of prediction points.

Index computations were done using the `index` function of the `spotr` package (section 2.4).

2.4 spotr

The R-package `spotr` accompanying this paper implements computation of indices via post-processing of models fitted via frequentist or Bayesian approaches.

To estimate an index, the user first needs to fit an abundance model to the data using the R-packages `mgcv` (Wood, 2017) or `brms` (Bürkner, 2018). Next a set of prediction points and any associated covariates that the indices should be computed over needs to be defined. The `spotr` function `index` uses the fitted model and the prediction points to compute indices for spatial areas. Uncertainty estimates for indices are approximated via simulation. In the case of `mgcv`, a normal approximation is used for simulation (Wood, 2006). Multiple random samples are first drawn from an approximate normal posterior distribution of parameter values (including spline coefficients). The simulated parameter values for each sample are then used to compute simulated indices. For `brms`, simulated indices are computed from samples of the posterior distribution. In both cases, uncertainty is estimated from the variation in indices across simulations. Estimates of indices can alternatively be computed directly from simulations across prediction points. This may be used for models that are not directly supported by `spotr`.

Posterior simulation can be computationally intensive if there are many prediction points and/or time points. The implementation therefore uses customised C-code for computational efficiency and to avoid numerical issues when computing sums of exponentiated terms. For the case study, estimating one type of index for a single model took up to 20 seconds on a standard desktop computer.

The package may be used to compute indices over geographical or environmental space (e.g habitat categories or environmental gradients), or over combinations of environmental and geographic space.

3 Results

All four models showed similar fit, which was deemed acceptable, while not perfect (Appendix B). The main discrepancy the checks identified was that the random site and year effects deviated from a normal distribution, with somewhat wider tails for the site random effects, and a few years with lower counts than expected for the year effects. Importantly, however, there were no clear temporal or spatial patterns to residuals or random effects.

3.1 National indices

There was strong support for spatio-temporal interaction terms for all three models with such terms (M_{ss}, M_{ds}, and M_{sd}, Appendix A). Population indices at the national scale were similar for all four models (Fig. 1). Indices treating random year effects as nuisance show increases during the first half and decreases during the second half.

Including year effects in the index suggests strong inter-annual variation overlaying the long-term trends. In particular, there were a few years with lower counts than otherwise expected (Fig. 1), as also indicated by checks of the random year effect residuals (Appendix B).

3.2 County indices

County-wise indices differ more among the different models than nation-wide indices, but all spatio-temporal models suggest strong increases in the first half for many counties, particularly in southern and central Sweden, and more stable or declining trends in the northern counties, followed by stable or declining trends in the second half (Fig 2).

To map change across counties we consider two options. The first is to select two time points and estimate the change between them (using the first of the time points as the baseline for the index of the second time point). Using the first and last years as the time points shows that long-term change over the entire study period (2000-2020) cuckoos declined in the northernmost counties but increased in many of the southern counties, and with unclear changes in central and southernmost Sweden (Fig. 3, upper panels). Such estimates can be sensitive to the selection of time points (Knape, 2023), especially if random year effects are included in the computation. Random year effects were therefore treated as nuisance.

The second options is to estimate change in average abundance between two time periods. Lower panels in Fig. 3 show the estimated change from the first to the second decade of the study. That is, indices are computed from averages over 2010-2019 relative to averages over 2000-2009. Here, random year effects were included in the prediction since averages across multiple years are less sensitive to short term variation. The results show less strong change but also less uncertainty than when using the first and last years, which may be expected as the time period for the comparison is shorter and averages improving the estimation (although the additional variation from the year effects could have countered this effect).

As a concise option for illustrating uncertainty, a difficulty for maps, we plot the smallest possible change compatible with confidence intervals (Fig 3). In other words, if confidence intervals for the change are entirely positive (at the log scale) we show the lower confidence limit, and if confidence intervals are entirely negative we show the upper limit.

3.3 Indices at higher resolution

Change can also be investigated at finer spatial resolution, particularly for models M_{ss} and M_{ds} that have spatio-temporal interactions that do not follow county borders. As for the county indices, one can map the change between two time points but instead using all the points on the prediction grid (Fig. 4). Maps of change for model M_{sd} follows county borders as its interaction term depends on space only via the county.

Static maps ignore some information as they do not show when changes occurred, which is often of interest. One possibility is to show the index map for each year, or a sub-sequence of years. Another is to visualise change against a one dimensional gradient (e.g. von Brömssen et al., 2021). Because of the elongated shape of Sweden, we use the latter approach to show how the estimated index changes across a latitudinal gradient (Fig. 5).

One may also use this approach to illustrate changes between consecutive years (Fig. 6), these can be seen as estimates of annual growth rates. The two models with small scale temporal smoothing (M_{ss} and M_{sd}) show similar inter-annual changes, while the model with only global temporal smoothing (M_{sd}) indicate erratic but mostly uncertain changes in specific years.

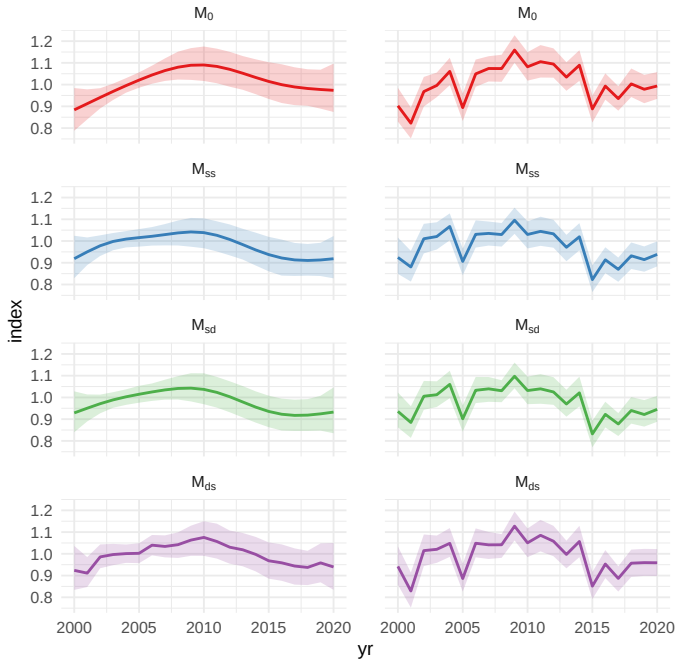

Figure 1: National popupulation indices for the four models. The baseline for the index is the first ten years (2000-2009). Random year effects are treated as nuisance in left column but included in index computation in right column.

Figure 2: Population indices for three select counties in southern (Skåne), central (Dalarna) and northern (Norrbotten) Sweden for the four models. The baseline for the index is the first ten years (2000-2009). Random year effects were treated as nuisance.

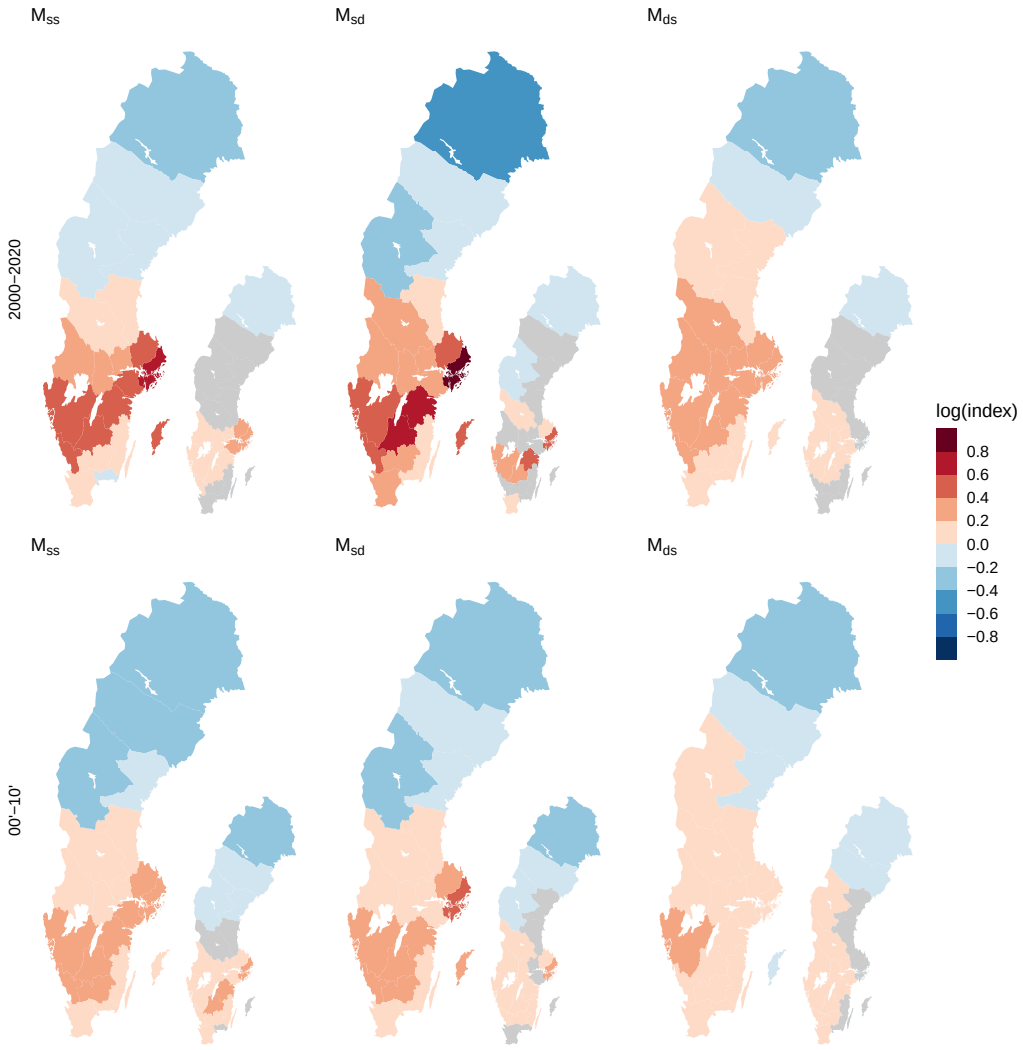

Figure 3: Estimated relative change (log scale) per county between year 2000 and year 2020 (upper panels) and from the first and second decade (lower panels) for the three models with spatio-temporal interactions. Inset plots show uncertainty via the smallest change within 95% confidence intervals. Counties for which confidence intervals overlap 0 are shown in grey. Random year effects were treated as nuisance in upper panels but included in lower panel estimates.

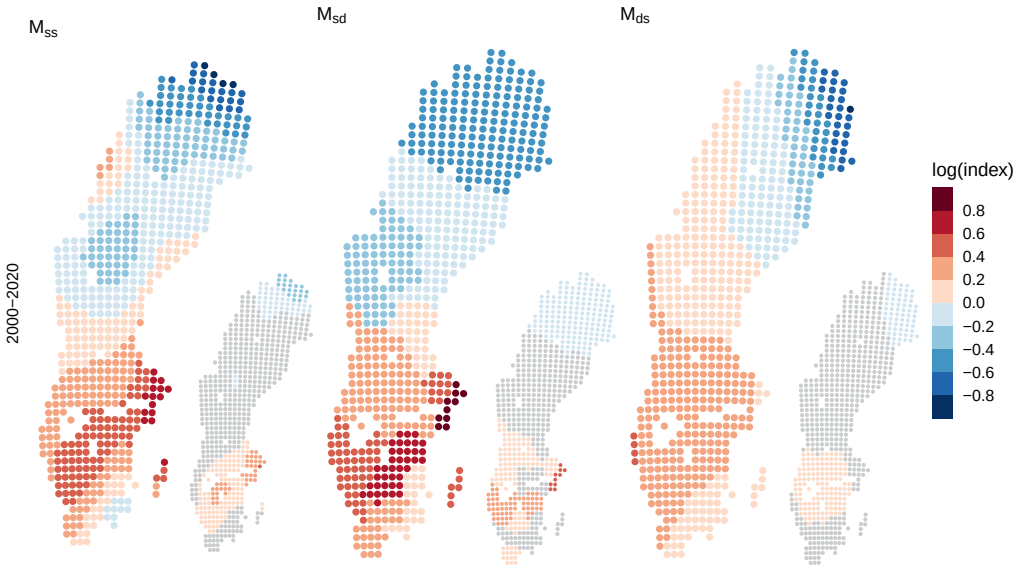

Figure 4: Estimated relative change (log scale) from 2000 to 2020 across a grid for the three models with spatio-temporal interactions. Inset plots show uncertainty via the smallest change within 95% confidence intervals, with grey representing points where confidence intervals overlap zero. Random year effects were treated as nuisance.

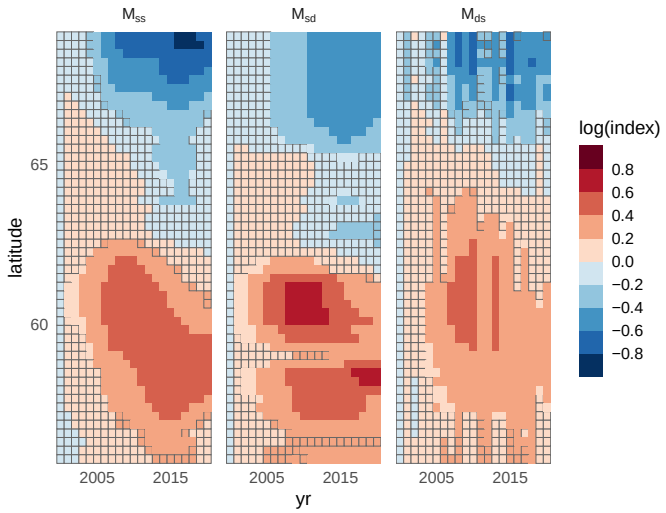

Figure 5: Estimated index from 2000 to 2020 across a latitudinal gradient for the three models with spatio-temporal interactions, using the first year as the reference. Borders around squares indicate that 95% confidence intervals for change relative to year 2000 include 0. Random year effects were treated as nuisance.

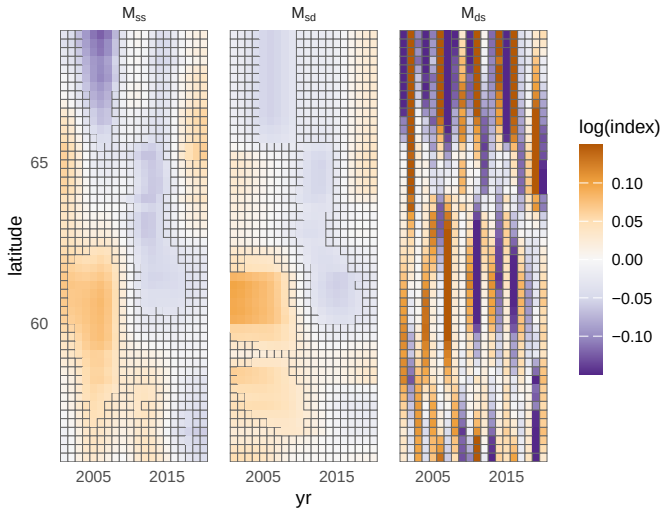

Figure 6: Estimated annual growth rates from 2000 to 2020 across a latitudinal gradient for the three models with spatio-temporal interactions. Borders around squares indicate that 95% confidence intervals of growth rates include 0. Random year effects were treated as nuisance.

3.4 Population density

The above indices have all been local, contrasting population size over an area to population size in the same area but another time point or period. Information about how abundance is distributed across space is then lost. Visualising this information (Fig. 7) shows that the high densities in the north of Sweden have been replaced by high densities in the central and south-eastern parts of Sweden.

4 Discussion

The general definition of abundance indices across space and time presented here relies on aggregating abundance through sums of weighted model predictions. The aggregations may be done at small or large spatial scales, and at short or long time scales, to highlight different features of population change. Annual population indices for birds produced from spatio-temporal models (Harrison et al., 2014; Sauer & Link, 2011), as well as annual indices from data with seasonal structure, similar to official indices derived from butterfly counts (Dennis et al., 2013), are included under this definition. The approach however also extends to comparisons across space, such as comparing population densities or totals across over space and time as illustrated in the case study (Fig. 7).

The population indices could be computed from a wide class of models fitted to survey data. For GAMs, models can include seasonal and/or spatial and spatio-temporal non-linear components (Bürkner, 2018; Pedersen et al., 2019; Wood, 2017). For insect data collected annually over the season, tensor product terms of seasonal curves interacting with year could be used to accommodate changes in phenology over time, or models with seasonal curves as smooth random effects could be used to incorporate variability in insect emergence among years. Models could alternatively have trends that vary by habitat types (Newson et al., 2009), across environmental gradients, or have combinations of geographical and environmental components. Indices could also be computed from fits of model types other than GAMs, e.g. from spatial and spatio-temporal random effects models (Vanhatalo et al., 2017), or from models estimated by machine learning approaches (Smith et al., 2019).

The indices are computed by aggregating weighted predicted abundances across space and/or time. To estimate uncertainty for these non-linear transformations of linear predictors, one can use a matrix of predictions whose rows correspond to prediction points and whose columns represent uncertainty via random draws. The random draws can be simulations from the posterior predictive distribution of Bayesian models, or randomised representations of uncertainty of frequentist models (Fewster et al., 2000; Wood, 2006). Some care is required in the computation of the weighted sums, particularly when log-links are used, as sums of exponential terms are sensitive to numerical overflow. Computations handling these issues are implemented in the `spotr` package for `mgcv` and `brms` models, or for a matrix of random predictions representing uncertainty.

Using weighted predictions to estimate indices can be seen as a variant of post-stratification, a technique to adjust estimates of population level quantities from unbalanced samples (Anganuzzi & Buckland, 1993). With this technique, predictions are weighted and averaged after fitting to correct for the sample imbalance. In our case study, data are more sparse in the northern parts of Sweden, particularly in the first years of the survey program. This imbalance, could potentially lead to biased trends. Post-stratifying by allowing trends to vary geographically, then using prediction across a spatially representative grid can lead to more accurate estimates. However, while our models suggest differences in trends between northern and southern Sweden, we found little difference in national scale indices between models allowing spatial variation and a model assuming identical relative change across space. Indicating that geographical imbalance is not a main issue for the cuckoo indices.

The connection to post-stratification could potentially be taken further than considered here using ideas behind multilevel regression with post-stratification (Gelman & Little, 1997) for inference from samples lacking a proper sampling design (Boyd et al., 2023). In the context of estimating indices from unbalanced data, this approach could use a set of factors that are thought to affect trends and are measured at sampled sites and whose total amount are known across the target area.

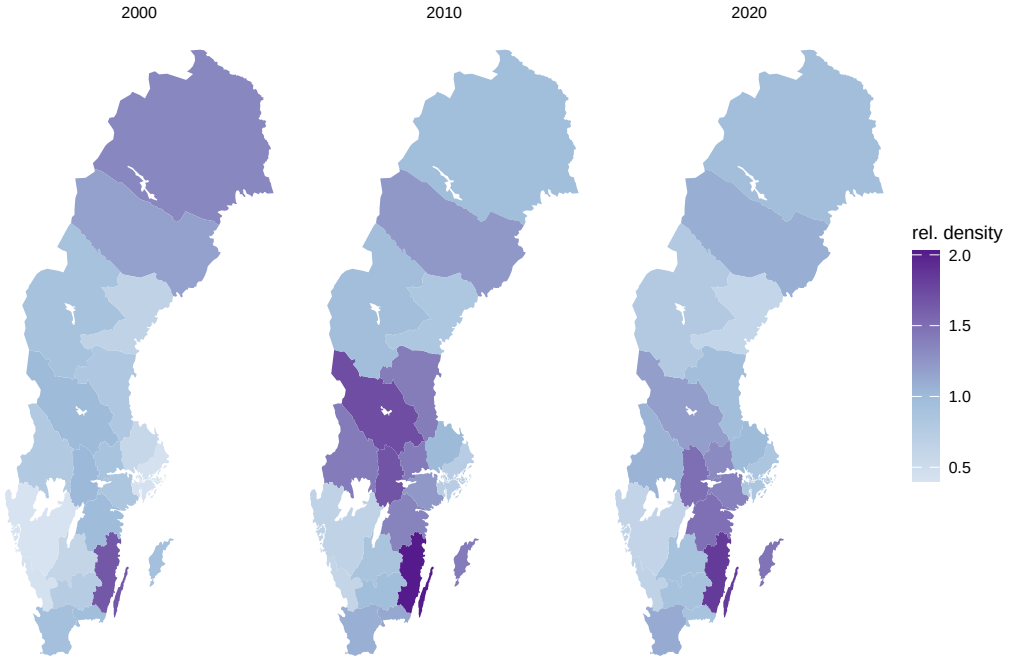

Figure 7: Estimated population density per county for year 2000, 2010 and 2020 relative to overall density in 2000. Estimates are based on model M_{ss}, random year effects were treated as nuisance.

These could be included in models using random effects and smoothing to deal with small sample sizes, and corrected at the aggregate level via weights (Authier et al., 2021). Factors could for example be habitat and land-use types, including temporal changes in such variables. As with most techniques to adjust for imbalance, the success of such an approach to correct for imbalance hinges on important factors affecting species trends being identified and measured (Anganuzzi & Buckland, 1993; Conn et al., 2015), and is not a replacement for careful sampling design.

The increased flexibility of spatio-temporal models compared to models assuming static trends across space opens up for the possibility of more accurately capturing population change and makes it possible to estimate local change at higher resolution (Harrison et al., 2014). However, as sample size decreases at smaller spatial scales, there is a greater risk of noise being mistaken for signal and for bias due local sampling inconsistencies. I.e. generally larger contributions of stochastic events affecting estimates and stronger sensitivity to model assumptions. The extent of improvements gained from spatio-temporal models will therefore depend on the ability of models to capture the true spatial structure of variation in population change. An important component in analysis is therefore to assess that models can reasonably capture important features of the data. Strategies for checking models used to produce abundance indices has received fairly little attention in the literature. A perfectly well fitting model is usually hard to achieve for large data sets, as seen in the case study, but not all lack of fit will have consequences for derived indices. Future research could focus on identifying what model assumptions are most crucial for estimating reliable indices under specific circumstances, and develop targeted checks of such assumption.

4.1 Conclusions

Many types of indices of absolute or relative population abundances, or of biodiversity, can be formed from weighted sums of predictions from spatio-temporal abundance models. Importantly, indices at small scales can form an evidence base for local population and biodiversity management and conservation actions. At these local scales data are usually limited and it is too costly to carry out local monitoring with sufficient precision. The model based approach uses information throughout the study area to improve precision. The success of this hinges on model assumptions being reasonably accurate, but if well specified, spatio-temporal models could mitigate some of the sparseness of local data. Future work could look into the gains in precision at local scales that spatio-temporal models can provide under various scenarios, and what types of models best provide them.

Acknowledgement

I am grateful to Claudia von Brömssen, Fränzi Korner-Nievergelt, and three anonymous reviewers for comments that improved the paper.

Data availability

The `spotr` package is available at https://github.com/jknape/spotr. It includes data for the case study.

References

Anganuzzi, A. A., & Buckland, S. T. (1993). Post-Stratification as a Bias Reduction Technique. *The Journal of Wildlife Management*, *57*(4), 827–834. https://doi.org/10.2307/3809085

Authier, M., Rouby, E., & Macleod, K. (2021). Estimating Cetacean Bycatch From Non-representative Samples (I): A Simulation Study With Regularized Multilevel Regression

and Post-stratification. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.719956

Barnett, L. A. K., Ward, E. J., & Anderson, S. C. (2021). Improving estimates of species distribution change by incorporating local trends. *Ecography*, *44*(3), 427–439. https://doi.org/10.1111/ecog.05176

Bled, F., Sauer, J., Pardieck, K., Doherty, P., & Royle, J. A. (2013). Modeling Trends from North American Breeding Bird Survey Data: A Spatially Explicit Approach. *PLOS ONE*, *8*(12), e81867. https://doi.org/10.1371/journal.pone.0081867

Bowler, D., Richter, R. L., Eskildsen, D., Kamp, J., Moshøj, C. M., Reif, J., Strebel, N., Trautmann, S., & Voříšek, P. (2021). Geographic variation in the population trends of common breeding birds across central Europe. *Basic and Applied Ecology*, *56*, 72–84. https://doi.org/10.1016/j.baae.2021.07.004

Boyd, R. J., Powney, G. D., & Pescott, O. L. (2023). We need to talk about nonprobability samples. *Trends in Ecology & Evolution*, *38*(6), 521–531. https://doi.org/10.1016/j.tree.2023.01.001

Breivik, O. N., Aanes, F., Søvik, G., Aglen, A., Mehl, S., & Johnsen, E. (2021). Predicting abundance indices in areas without coverage with a latent spatio-temporal Gaussian model. *ICES Journal of Marine Science*, *78*(6), 2031–2042. https://doi.org/10.1093/icesjms/fsab073

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. https://journal.r-project.org/archive/2018/RJ-2018-017/index.html

Conn, P. B., Johnson, D. S., Hoef, J. M. V., Hooten, M. B., London, J. M., & Boveng, P. L. (2015). Using spatiotemporal statistical models to estimate animal abundance and infer ecological dynamics from survey counts. *Ecological Monographs*, *85*(2), 235–252. https://doi.org/10.1890/14-0959.1

Davey, C., Vickery, J., Boatman, N., Chamberlain, D., Parry, H., & Siriwardena, G. (2010). Regional variation in the efficacy of Entry Level Stewardship in England. *Agriculture, Ecosystems & Environment*, *139*(1-2), 121–128. https://doi.org/10.1016/j.agee.2010.07.008

Dennis, E. B., Freeman, S. N., Brereton, T., & Roy, D. B. (2013). Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern. *Methods in Ecology and Evolution*, *4*(7), 637–645. https://doi.org/10.1111/2041-210X.12053

Dunn, P. K., & Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, *5*(3), 236–244. https://doi.org/10.2307/1390802

Fewster, R. M., Buckland, S. T., Siriwardena, G. M., Baillie, S. R., & Wilson, J. D. (2000). Analysis of population trends for farmland birds using generalized additive models. *Ecology*, *81*(7), 1970–1984. https://doi.org/10.1890/0012-9658(2000)081%5B1970:AOPTFF%5D2.0.CO;2

Gelman, A., & Little, T. C. (1997). Poststratification into many categories using hierarchical logistic regression. *Survey Methodlogy*, *23*, 127--135.

Gregory, R. D., Skorpilova, J., Vorisek, P., & Butler, S. (2019). An analysis of trends, uncertainty and species selection shows contrasting trends of widespread forest and farmland birds in Europe. *Ecological Indicators*, *103*, 676–687. https://doi.org/10.1016/j.ecolind.2019.04.064

Gregory, R. D., & Strien, A. van. (2010). Wild Bird Indicators: Using Composite Population Trends of Birds as Measures of Environmental Health. *Ornithological Science*, *9*(1), 3–22. https://doi.org/10.2326/osj.9.3

Harrison, P. J., Buckland, S. T., Yuan, Y., Elston, D. A., Brewer, M. J., Johnston, A., & Pearce-Higgins, J. W. (2014). Assessing trends in biodiversity over space and time using the example of British breeding birds. *Journal of Applied Ecology*, *51*, 1650--1660. https://doi.org/10.1111/1365-2664.12316

Johnson, T. F., Beckerman, A. P., Childs, D. Z., Webb, T. J., Evans, K. L., Griffiths, C. A., Capdevila, P., Clements, C. F., Besson, M., Gregory, R. D., Thomas, G. H., Delmas, E., & Freckleton, R. P. (2024). Revealing uncertainty in the status of biodiversity change. *Nature*, *628*(8009), 788–794. https://doi.org/10.1038/s41586-024-07236-z

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

Knape, J. (2023). Effects of choice of baseline on the uncertainty of population and biodiversity indices. *Environmental and Ecological Statistics*, *30*(1), 1–16. https://doi.org/10.1007/s10651-022-00550-7

Lindström, Å., & Green, M. (2020). *Swedish Bird Survey: Fixed routes (Standardrutterna)*. https://doi.org/10.15468/hd6w0r

Newson, S. E., Ockendon, N., Joys, A., Noble, D. G., & Baillie, S. R. (2009). Comparison of habitat-specific trends in the abundance of breeding birds in the UK. *Bird Study*, *56*(2), 233–243. https://doi.org/10.1080/00063650902792098

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ*, *7*, e6876. https://doi.org/10.7717/peerj.6876

Sauer, J. R., & Link, W. A. (2011). Analysis of the North American Breeding Bird Survey Using Hierarchical Models. *The Auk*, *128*(1), 87–98. https://doi.org/10.1525/auk.2010.09220

Smith, A., Hofner, B., Lamb, J. S., Osenkowski, J., Allison, T., Sadoti, G., McWilliams, S. R., & Paton, P. (2019). Modeling spatiotemporal abundance of mobile wildlife in highly variable environments using boosted GAMLSS hurdle models. *Ecology and Evolution*, *9*(5), 2346–2364. https://doi.org/10.1002/ece3.4738

ter Braak, C. J. F., van Strien, A. J., Meijer, R., & Verstrael, T. J. (1994). Analysis of monitoring data with many missing values: which method? *Bird Numbers 1992. Distribution, Monitoring and Ecological Aspects. Proceedings of the 12th International Conference of IBCC and EOAC.*, 663–673.

Thomas, L. (1996). Monitoring Long-Term Population Change: Why are there so Many Analysis Methods? *Ecology*, *77*(1), 49–58. http://www.jstor.org/stable/2265653

Thorson, J. T., Barnes, C. L., Friedman, S. T., Morano, J. L., & Siple, M. C. (2023). Spatially varying coefficients can improve parsimony and descriptive power for species distribution models. *Ecography*, *2023*(5), e06510. https://doi.org/10.1111/ecog.06510

van Strien, A., Pannekoek, J., Hagemeijer, W., & Verstrael, T. (2000). A loglinear Poisson regression method to analyse bird monitoring data. *Proceedings of the International Conference and 13th Meeting of the European Bird Census Council. Bird Numbers 1995. Bird Census News.*, 33–39.

Vanhatalo, J., Hosack, G. R., & Sweatman, H. (2017). Spa-

tiotemporal modelling of crown-of-thorns starfish outbreaks on the Great Barrier Reef to inform control strategies. *Journal of Applied Ecology*, *54*(1), 188–197. https://doi.org/10.1111/1365-2664.12710

von Brömssen, C., Betnér, S., Fölster, J., & Eklöf, K. (2021). A toolbox for visualizing trends in large-scale environmental data. *Environmental Modelling & Software*, *136*, 104949. https://doi.org/10.1016/j.envsoft.2020.104949

Wikle, C. K. (2003). Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology*, *86*, 1382–1394. http://www.esajournals.org/doi/abs/10.1890/0012-9658%282003%29084%5B1382%3AHBMFPT%5D2.0.CO%3B2

Wood, S. N. (2006). On Confidence Intervals for Generalized Additive Models Based on Penalized Regression Splines. *Australian & New Zealand Journal of Statistics*, *48*(4), 445–464. https://doi.org/10.1111/j.1467-842X.2006.00450.x

Wood, S. N. (2017). *Generalized additive models: an introduction with R* (2nd ed.). CRC press.