Incorporating sampling error in the estimation of autoregressive coefficients of animal population dynamics using capture-recapture data

Pedro G. Nicolau*[1], Sigrunn H. Sørbye[1], and Nigel G. Yoccoz[2]

[1]Department of Mathematics and Statistics, Faculty of Science and Technology, UiT The Arctic University of Norway

[2]Department of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics, UiT The Arctic University of Norway

Author Note

Corresponding author. E-mail: pedro.nicolau@uit.no. Address: Institutt for Matematikk og Statistikk, UiT Norges arktiske universitet, Postboks 6050 Langnes, 9037 Tromsø

Abstract

Population dynamics models combine density-dependence and environmental effects. Ignoring sampling uncertainty might lead to biased estimation of the strength of density-dependence. This is typically addressed using state-space model approaches, which integrate sampling error and population process estimates. Such models seldom include an explicit link between the sampling procedures and the true abundance, which is common in capture-recapture settings. However, many of the models proposed to estimate abundance in the presence of heterogeneity lead to incomplete likelihood functions and cannot be straightforwardly included in state-space models.

We assessed the importance of estimating sampling error explicitly by taking an intermediate approach between ignoring uncertainty in abundance estimates and fully specified state-space models for density-dependence estimation based on autoregressive processes. First, we estimated individual capture probabilities based on a heterogeneity model, using a conditional multinomial likelihood, followed by a Horvitz-Thompson estimate for abundance. Second, we estimated coefficients of autoregressive models for the log abundance. Inference was performed using the methodology of integrated nested Laplace approximation (INLA). We performed an extensive simulation study to compare our approach with estimates disregarding capture history information, and using R-package VGAM, for different parameter specifications. The methods were then applied to a real dataset of gray-sided voles *Myodes rufocanus* from Northern Norway.

We found that density-dependence estimation was improved when explicitly modelling sampling error in scenarios with low innovation variances, in which differences in coverage reached up to 8% in estimating the coefficients of the autoregressive processes. In this case, the bias also increased assuming a Poisson distribution in the observational model. For high innovation variances, the differences between methods were small and it appeared less important to model heterogeneity.

KEYWORDS: density-dependence, capture-recapture, population dynamics, sampling error, heterogeneity, INLA.

Incorporating sampling error in the estimation of autoregressive coefficients of animal population dynamics using capture-recapture data

1 Introduction

Models used to analyze population dynamics include a combination of density-dependence and environmental effects. Ignoring the uncertainty in abundance estimates biases estimates of the strength of density-dependence, and different approaches exist to achieve better accuracy (see Lebreton and Gimenez (2012) for a review). In particular, state-space models combining an observation model – linking the observations such as counts to the true abundance - and a process model – describing the processes driving population dynamics – have become a standard approach in many analyses (Dennis & Taper, 1994). However, these models rarely include an explicit model of the link between how counts were obtained and true abundance, often relying on a non-specific observation model such as log-normal or Poisson distribution (for instance, Ono, Langangen, and Stenseth (2019), but see below).

Capture-recapture methods have been extensively used to estimate abundance, and many methods have been developed to take different sources of variability in capture probabilities into account. Estimating abundance is a challenging statistical problem (Link, 2003), and heterogeneity in capture probabilities can lead to large biases in abundance estimates when using models assuming no heterogeneity (Carothers, 1973; Otis, Burnham, White, & Anderson, 1978). However, many of the models that have been proposed to estimate abundance in the presence of heterogeneity do not lead to observation models that can be included in state-space models as they do not lead to likelihood functions in a closed form (Chao & Huggins, 2006; Huggins & Hwang, 2011).

Many studies investigating density dependence have used simple process models such as the Gompertz model - i.e. a model which is a first order autoregressive model on a log scale (Ono et al., 2019; Thibaut & Connolly, 2019). However, ecological processes such as trophic interactions (Bjørnstad, Falck, & Stenseth, 1995) or intrinsic ecological properties

such as age structure (Lande, Engen, & Sæther, 2002) may lead to more complex process models such as a second-order autoregressive model (AR(2)). An important case is the population cycles observed in many small mammal populations, particularly in northern environments (Elton, 1924; Stenseth, 1999). These quasi-periodic fluctuations are quite well approximated by AR(2) models on a logarithmic scale (Bjørnstad et al., 1995). Whereas most analyses have ignored the uncertainty in abundance estimates (Bjørnstad et al., 1995), some have used state-space models (Cornulier et al., 2013; Ims, Yoccoz, & Killengreen, 2011; Kleiven, Henden, Ims, & Yoccoz, 2018; Stenseth et al., 2003). However, none of these approaches used a capture-recapture approach dealing with capture heterogeneity, an issue that had been emphasized in earlier reviews (Otis et al., 1978; Yoccoz, Ims, & Stenseth, 1993).

Here we investigated the performance of an intermediate approach between ignoring uncertainty in abundance estimates (i.e. using the raw population counts) and fully specified state-space models. Specifically, we first used a multinomial observation model to estimate capture probabilities followed by estimating abundance at each time point using the Horvitz-Thompson estimator (Horvitz & Thompson, 1952). Second, we fitted an AR(2) process model to the log-abundance to estimate direct and delayed density-dependence given by the first and second coefficient of the AR(2) model, respectively. Both estimation steps were performed in a unified way, incorporating the models within the general class of latent Gaussian models (Rue, Martino, & Chopin, 2009). Full Bayesian inference was then obtained using the methodology of integrated nested Laplace appproximation (INLA) (Rue et al., 2009, 2017).

We based our analyses on a large-scale study of population dynamics of the dominant small mammal species in northern Fennoscandia, the grey-sided vole *Myodes rufocanus* (Ims et al., 2011). This species shows large fluctuations with a 4 to 5 year periodicity (Ims et al., 2011; Marolla et al., 2019). We monitored populations of grey-sided voles along a 200 km gradient from coast to inland, using live capture-recapture methods, starting in 2000.

Previous analyses have shown that there was large heterogeneity in capture probabilities (Yoccoz & Ims, 2004). Our goal was to understand spatial patterns of population dynamics by going further than analyzing spatial correlations in seasonal or annual population growth (Bjørnstad, Stenseth, & Saitoh, 1999), doing so by including spatial variation in the autoregressive parameters as well as spatial environmental effects (i.e. the residual or stochastic process term). INLA provides a convenient approach to analyze such complex spatial patterns, but we needed first to assess the robustness of using an approach based on estimated abundances but without implementing a full state-space model. In this paper, we therefore use a simulation study built around the case study (adaptable to other situations from the code provided) to assess the estimation accuracy of the density-dependence, both including and excluding capture-history information.

The structure of this paper is as follows. Section 2 provides our methodological background to analyse capture-recapture data and describes the Bayesian framework to perform parameter estimation. This includes using INLA to estimate individual capture probabilities and the direct and delayed density-dependence given by the coefficients of AR(2) models. Section 3 contains an extensive simulation study, investigating how density-dependence estimates are influenced when individual capture probabilities are taken into account. In section 4, we study the population cycles of grey-sided voles. We first compare different observation models in estimating individual capture probabilities and then assess whether incorporation of individual capture probabilities influence density-dependence estimates. A summary and concluding remarks are given in Section 5.

2 Methodology

Capture-recapture experiments are important to assess heterogeneity in individual capture probabilities. This section describes our approach to incorporate capture-recapture information in the estimation of density-dependence. First, we define an observation model in which capture probabilities are modelled in terms of individual features and then used to

estimate abundance. Second, we fit an AR(2) process model to the estimated log-abundance to assess density-dependence. When using state-space approaches, the parameters of the observation and process model are estimated simultaneously. This is not possible in our case as the capture probabilities are estimated based on a conditional multinomial likelihood, due to individuals that were not observed. Instead, we apply a sequential approach, first estimating the capture probabilities and then the AR(2) coefficients. This allows us to use an explicit sampling model to estimate capture probabilities, instead of assuming that the observed counts have a Poisson or log-normal distribution. The given sequential approach is computationally efficient using the `R-INLA` package which is freely available at www.r-inla.org.

2.1 Statistical background on capture-recapture data

Assume a closed population with a total of N individuals and a capture-recapture experiment with τ capture sessions. Let

$$\boldsymbol{w}_i' = (w_{i1}, \ldots, w_{i\tau}), \quad i = 1, \ldots, N,$$

denote the capture history for the ith individual. If $w_{ij} = 1$, the individual was captured at the jth capture session while $w_{ij} = 0$ otherwise, i.e. $w_{ij} \sim \text{Bernoulli}(p_{ij})$, $j = 1, \ldots, \tau$. For each individual, the probability of a given capture history is then

$$p_{\boldsymbol{w}_i} = \prod_{j=1}^{\tau} p_{ij}^{w_{ij}} (1 - p_{ij})^{1 - w_{ij}}, \qquad i = 1, \ldots, N.$$

Assuming that all individuals are captured independently, the complete likelihood becomes

$$L(N, \{p_{ij}\} \mid \{w_{ij}\}) = \prod_{i=1}^{N} \prod_{j=1}^{\tau} p_{ij}^{w_{ij}} (1 - p_{ij})^{1 - w_{ij}}$$

where both N and the set of probabilities $\{p_{ij}\}$ are unknown. Due to the unknown number of non-captured individuals, computation of the likelihood is unfeasible. This is a well-known problem (Huggins & Hwang, 2011) and requires alternative strategies to perform parameter estimation.

A commonly applied approach is to maximise the conditional likelihood for the n individuals that were captured at least once. Let c_{ik}, $k = 0, \ldots, 2^\tau - 1$, denote the probability that the capture history of individual i is equal to category k. The different categories are defined by all possible permutations of the capture session vector, giving a total of $m = 2^\tau - 1$ categories for the captured individuals.

From here onwards we will refer to data sets with only two capture events, in which mortality and emigration are disregarded considering capture events on adjacent days. The event that an individual is never captured is then defined as category 0, while the categories 1, 2 and 3 refer to the capture histories $(1,0), (0,1)$ and $(1,1)$, respectively. To perform parameter estimation, we need to make realistic assumptions on the capture probabilities for different capture sessions. Otis et al. (1978) propose a total of eight different models characterising capture probabilities for different sessions depending on time, behaviour and homogeneity of the individuals, also including combinations of these three factors. Here, we consider a heterogeneity model including a temporal effect, M_{th}. This implies that the capture probabilities depend on different features of the individuals. Further, we assume that the capture probability on the first and second capture sessions are independent. The probabilities for the different categories are then specified as

$$c_{i0} = (1-p_{i1})(1-p_{i2}), \;\; c_{i1} = p_{i1}(1-p_{i2}), \;\; c_{i2} = (1-p_{i1})p_{i2}, \;\; c_{i3} = p_{i1}p_{i2}, \qquad i = 1, \ldots, N. \tag{1}$$

To estimate abundance based on individuals that were captured, we use the Horvitz-Thompson estimator (Horvitz & Thompson, 1952)

$$\hat{N} = \sum_{i=1}^{n} (1 - \hat{c}_{i0})^{-1} \tag{2}$$

where $\hat{c}_{i0}$ denotes the estimated probability that individual i was not captured. This probability is estimated using a regression model as explained in the next section.

2.2 A multinomial capture-recapture regression model including a Poisson transformation

An important question in analysing population processes from capture-recapture data is whether features of the captured individuals give valuable information in further analysis of density-dependence. To estimate the probabilities in (2), it is natural to assume a multinomial regression model for the captured individuals, incorporating covariate information which helps to separate different capture categories. Define the vector $\boldsymbol{Y}_i' = (Y_{i1}, \ldots, Y_{im})$ where $Y_{ik} = 1$ for an individual classified to category k, while the remaining elements of $\boldsymbol{Y}_i$ are 0. Each of the vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ has a multinomial distribution. Based on (1), probabilities for the $m = 3$ observed categories are defined by $\tilde{c}_{ik} = c_{ik}/(1 - c_{i0})$, $k = 1, \ldots, m$, ensuring that the probabilities sum to 1. These probabilities can then be modelled in terms of observed individual features like weight, sex, age, etc.

We denote the individual features or covariates by $\boldsymbol{z}_r' = (z_{1r}, \ldots, z_{nr})$. Further, define the linear predictor

$$V_{ik} = \sum_{r=1}^{p} \gamma_{kr} z_{ir}, \quad i = 1, \ldots, n, \quad k = 1, \ldots, m. \tag{3}$$

where the coefficient γ_{kr} is specific for category k and covariate r, while p is the number of covariates. The scaled probabilities for the captured individuals are then expressed as

$$\tilde{c}_{ik} = \frac{e^{V_{ik}}}{\sum_{k=1}^{m} e^{V_{ik}}}, \quad i = 1, \ldots, n, \quad k = 1, \ldots, m \tag{4}$$

The resulting multinomial likelihood is

$$L_M(\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \propto \prod_{i=1}^{n} \prod_{k=1}^{m} (\tilde{c}_{ik})^{y_{ik}} \tag{5}$$

where $\boldsymbol{\gamma}_r' = (\gamma_{1r}, \ldots, \gamma_{mr})$, $r = 1, \ldots, p$. Notice that in maximizing (5), the denominator of $\tilde{c}_{ik}$ does not simplify using the ordinary logarithmic transformation. It is therefore common to apply the well-known multinomial-Poisson transformation (Baker, 1994) in which the

likelihood is rewritten as

$$L_P(\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p, \boldsymbol{\beta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \propto \prod_{i=1}^{n} \prod_{k=1}^{m} e^{-\mu_{ik}} \mu_{ik}^{y_{ik}}.$$

Here, $\mu_{ik} = e^{V_{ik}+\beta_i}$ represents the rate of a Poisson distributed random variable Y_{ik}. The given transformation from a multinomial likelihood to the Poisson likelihood introduces auxiliary parameters $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_n)$, in which β_i is proportional to $\ln(\sum_{k=1}^{m} e^{V_{ik}})$. This is just a technical detail to make the approximation work correctly. The likelihood $L_P(.)$ is proportional to $L_M(.)$ and gives the same maximum likelihood estimates for the coefficient vectors $\boldsymbol{\gamma}_r$. The resulting regression model is then summarized in terms of linking the expectation of the Poisson variables to the linear predictor using the log-transform, i.e.

$$\ln(E(Y_{ik})) = \ln(\mu_{ik}) = \sum_{r=1}^{p} \gamma_{kr} z_{ir} + \beta_i + \epsilon_i, \quad i = 1, \ldots, n, \quad k = 1, \ldots, m, \tag{6}$$

where $\epsilon_i \sim N(0, \kappa^{-1})$ denotes small independent random error terms.

In fitting the given model to a data set, the vectors $\{\boldsymbol{\gamma}_r\}_{r=1}^{p}$ will not be identifiable. However, in our case we only need estimates of the differences of these coefficients as these represent ratios of log-probabilities between the different categories. For categories k and l, we notice that

$$\ln\left(\frac{\tilde{c}_{ik}}{\tilde{c}_{il}}\right) \;=\; V_{ik} - V_{il} = \sum_{r=1}^{p} (\gamma_{kr} - \gamma_{lr}) z_{ir}.$$

In estimating the parameters of the model, this implies that the auxiliary parameters and error terms disappear, but these are still included in fitting (6) to a data set. In the case of assuming (1), the estimated individual probabilities are then given by

$$\ln\left(\frac{\hat{p}_{i1}}{1-\hat{p}_{i1}}\right) = \sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{2r}) z_{ir} \tag{7}$$

$$\ln\left(\frac{\hat{p}_{i2}}{1-\hat{p}_{i2}}\right) = \sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{1r}) z_{ir}. \tag{8}$$

or equivalently

$$\hat{p}_{i1} \;=\; \frac{e^{\sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{2r}) z_{ir}}}{1 + e^{\sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{2r}) z_{ir}}} \tag{9}$$

$$\hat{p}_{i2} \;=\; \frac{e^{\sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{1r}) z_{ir}}}{1 + e^{\sum_{r=1}^{p} (\hat{\gamma}_{3r} - \hat{\gamma}_{1r}) z_{ir}}}. \tag{10}$$

These probabilities are then used to estimate $\hat{c}_{i0}$ in (2).

2.3 Implementation using a Bayesian framework

To fit (6) to a dataset and estimate the capture probabilities, we choose to apply a Bayesian approach. This implies that all parameters in (6) are viewed as random variables. Specifically, the resulting regression model can be incorporated within the computational framework of latent Gaussian models. This is a flexible class of three-stage hierarchical models, which can be analysed in a unified way using INLA. The model in (6) is then reformulated in terms of having conditionally independent observations, given a latent field and hyperparameters.

The three stages of a latent Gaussian model are expressed as follows, where $\pi(.)$ is generic notation for probability densities:

1. The first stage specifies the likelihood where the observations are assumed conditionally independent given a latent field $\boldsymbol{x}$ and hyperparameters $\boldsymbol{\theta}$. In our case, let $\boldsymbol{y}' = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_n')$ denote the stacked vector of the m categories for the n individuals. The likelihood is then expressed as

$$L(\boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y}) = \prod_{i=1}^{nm} \pi(y_i \mid x_i, \boldsymbol{\theta}).$$

2. The latent field $\boldsymbol{x}$ collects all random variables of the linear predictor

$$\boldsymbol{x} = \{\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p, \boldsymbol{\beta}, \boldsymbol{\epsilon}\} \tag{11}$$

where we could also include the predictor itself. The latent field models the dependency structure of the observations and is assigned a multivariate Gaussian prior

$$\pi(\boldsymbol{x} \mid \boldsymbol{\theta}) \sim N(\boldsymbol{0}, \boldsymbol{Q}^{-1}(\boldsymbol{\theta})).$$

The precision (inverse covariance) matrix $\boldsymbol{Q}$ is typically sparse such that $\boldsymbol{x}$ has Markov properties and is then referred to as a Gaussian Markov random field.

3. The hyperparameters $\boldsymbol{\theta}$ of a latent Gaussian model are usually assigned non-Gaussian priors. Here, we only have one hyperparameter being the precision parameter of the random error terms, $\theta = \kappa$. This parameter is assigned a penalised complexity prior (Simpson, Rue, Riebler, Martins, & Sørbye, 2017), implying that $\kappa^{-1/2}$ has an exponential density.

The joint posterior for all elements of the latent field and the additional hyperparameter is then described as

$$\pi(\boldsymbol{x}, \theta \mid \boldsymbol{y}) \propto \prod_{i=1}^{nm} \pi(y_i \mid x_i, \theta)\pi(\boldsymbol{x} \mid \theta)\pi(\theta).$$

The main interest is to calculate the marginal posteriors for each of the latent field components and each of the hyperparameters.

For the multinomial model, INLA is used to estimate the marginal posteriors for all the coefficients

$$\pi(\gamma_{kr} \mid \boldsymbol{y}), \quad k = 1, \ldots, m, \ r = 1, \ldots, p.$$

These provide posterior mean estimates of the differences $\gamma_{kr} - \gamma_{lr}$ which are used to estimate the individual capture probabilities and the abundance by (2).

2.4 Estimating density-dependence

Our final step is to fit a process model to study population dynamics of a species. Specifically, we focus on estimating density-dependence by fitting an AR(2) model to a given time series, reflecting the population cycle for the relevant species. Let $\ln(N_t)$ denote the true log-abundance at time t. The AR(2) model is then defined by

$$\ln(N_t) = \ln(K) + \phi_1 \ln(N_{t-1}) + \phi_2 \ln(N_{t-2}) + \epsilon_t, \quad t = 1, \ldots, T \tag{12}$$

where $\ln(K)$ denotes an offset while the noise terms are independent Gaussian variables, $\epsilon_t \overset{\text{iid}}{\sim} N(0, \sigma_\epsilon^2)$. T denotes the length of the time series while the coefficients ϕ_1 and ϕ_2 characterize the direct and delayed density-dependence of the series. The given process is stationary when $-1 \leq \phi_2 \leq 1 - |\phi_1| < 1$ and has pseudo-periodic behaviour when

$\phi_1^2 + 4\phi_2 \leq 0$. Estimation of the coefficients of AR(2) is not influenced by the offset $\ln(K)$. This implies that if the number of captured individuals at different time points are proportional to the underlying true abundance, we would get identical parameter estimates. The AR(2) model is fitted within the framework of latent Gaussian models using INLA. In this case the model has three hyperparameters, including $\kappa = \sigma_\epsilon^{-2}$ and the coefficients ϕ_1 and ϕ_2. These parameters are all assigned PC priors (Simpson et al., 2017; Sørbye & Rue, 2017). Of main interest is to study how the estimates of ϕ_1 and ϕ_2 vary when capture heterogeneity is accounted for using the multinomial observational model.

Often, simplifying assumptions regarding the data generating process are made, e.g by assuming a Poisson process (Stenseth et al., 2003) or a Log-normal distribution (Santin-Janin et al., 2014) for the observed counts. These assumptions can be implicit while defining the observation models in state-space approaches. We study the Poisson distribution assumption in an additional step also fitted using INLA. The log-rate of the assumed underlying Poisson process for the abundance is expressed in terms of the linear predictor

$$\lambda_t = \ln(E(N_t)) = \beta_0 + e_t, \qquad t = 1, \ldots, T. \tag{13}$$

Here β_0 denotes an intercept while $e_1, \ldots, e_T$ denote independent and identically distributed random variables, $e_i \sim N(0, \kappa_e^{-2})$. These error terms are included to model random variation as a function of time. As detailed in the next section, the AR(2) model will be fitted either to the estimated log-abundance $\ln(\hat{N}_1), \ldots, \ln(\hat{N}_T)$ or to the posterior means of the log-rates of the corresponding Poisson process, denoted $\hat{\lambda}_1, \ldots, \hat{\lambda}_T$.

3 Simulation study comparing methods to estimate density-dependence

This section provides an extensive simulation study to assess how the inclusion of capture history information influences estimation of density-dependence. We start by simulating data to approximate a realistic capture-recapture sampling scenario. The underlying log-population of the sampled species is generated as an AR(2) process in time, using

different fixed combinations of the coefficients (ϕ_1, ϕ_2) and the innovation variance σ_ϵ^2. Each resulting individual is then assigned a random weight, and a two-day capture history according to a multinomial model with probabilities defined by (1). We then fit an AR(2) process model to the estimates of log-abundance or log-rates obtained by different methods. These different methods are described in Section 3.1, while Section 3.2 specifies the simulation procedure and the method performance criteria used. Finally, Section 3.3 provides simulation results and an evaluation of the different methods.

3.1 Estimation methods

An overview of the different estimation methods used in the simulation study is given in Figure 1. The left-hand side of the figure shows the additional steps needed to implement the observation model, incorporating sampling error in terms of capture-history information. We employ two methods of estimating individual capture probabilities. The first is described in Sections 2.2 using INLA (method: CR-INLA), and corresponds to our suggested approach. The second, for comparison, estimates individual capture probabilities using the R-package VGAM (Yee, 2019). Among other utilities, the VGAM (vector generalised additive model) framework can be used to analyse closed population capture-recapture data, allowing the incorporation of individual covariates while using the conditional likelihood (Yee, Stoklosa, & Huggins, 2015). This application of VGAM allows for a flexible and efficient estimation of capture probabilities for all of the eight heterogeneity models given by Otis et al. (1978) (method: CR-VGAM). From the estimated capture probabilities from either of the two methods, we proceed to estimate the true log-abundance using the Horvitz-Thompson estimator in (2). At this point, we have two possible variants in estimating density-dependence: we either fit the AR(2) model to the times series of estimated log-abundance $\{\ln(\hat{N}_t)\}_{t=1}^T$ (A variant); or we fit the AR(2) model to the corresponding estimated log-rate of a Poisson process, $\{\hat{\lambda}_t\}_{t=1}^T$ (P variant). The right-hand side of Figure 1 illustrates the approach disregarding capture history,

fitting the AR(2) model directly to the observed log-counts, or to the log-rate of the corresponding Poisson process (method: ObsCount). Finally, the performance of the different estimation methods is compared with the optimal results, fitting the AR(2) model to the true generated log-abundance or estimated log-rate (method: Baseline).

3.2 Simulation procedure

For each combination of AR(2) coefficients, (ϕ_1, ϕ_2), we generated $M = 200$ time series. Specifically, we chose $\phi_1 \in (-1, -0.5, 0, 0.5, 1)$ and $\phi_2 \in (-0.8, -0.5, -0.2)$, giving a total of fifteen combinations of the coefficients. These combinations ensure that the resulting generated time series were stationary, also having pseudo-periodic behaviour. To investigate the effect of varying the innovation variance of the AR(2) process, we further compared density-dependence estimates for the values $\sigma_\epsilon^2 \in (0.04, 0.08, 0.16, 0.32, 0.64)$. The details of the simulation procedure can be described in the following steps:

1. Generate the series $\{\ln N_t\}_{t=1}^{T}$ according to (12) where $T = 20$, using different fixed combinations of (ϕ_1, ϕ_2). To remove the effect of sample size on the estimation of capture probability, we assumed that $E(N_t) = 20$ by using an offset $\ln(K) = \ln(20) - \frac{1}{2}\mathrm{Var}(ln(N_t))$. The series was rounded to give integer-values for $\{N_t\}_{t=1}^{T}$, representing the abundance of an animal population. The total number of individuals generated for one simulated AR(2) process was then $\tilde{N} = \sum_{t=1}^{T} N_t$.

2. For each of the $\tilde{N}$ individuals, we generated a random weight

$$z_{it}|\mu_t \sim \text{Lognormal}(\ln(\mu_t), \ln(\sigma_w))$$

where $\sigma_w = 1.2$ while $\mu_t \sim \text{Lognormal}(\ln(30), \ln(5))$. The weight was then scaled by the sample standard deviation of the generated weights to make it dimensionless. The resulting variable was used as an individual specific covariate in (3). In this context, weight is a proxy for detectability. We varied the expected value of weight with time to model varying detectability, reflecting changes in the composition of the

297 population at different time points. Thus, the varying mean reflects biological
298 variation which we considered more realistic than assuming constant capture
299 probabilities for different time points. The parameters relating to the weight
300 distribution were here chosen to illustrate this biological variation.

3. Assume a temporal effect M_{th} for the capture-recapture process with $\tau = 2$. To assign a capture history to each individual, we first assumed that the capture probabilities for day 1 and 2 were $p_{i1} \equiv p_1 = 0.55$ and $p_{i2} \equiv p_2 = 0.75$ for the total generated population. These probabilities were used to find reasonable values for the specific coefficients for the observed categories in terms of

$$\gamma_{31} - \gamma_{21} = \ln\left(\frac{p_1}{1-p_1}\right) \quad \text{and} \quad \gamma_{31} - \gamma_{11} = \ln\left(\frac{p_2}{1-p_2}\right).$$

301 The final individual capture probabilities were then computed according to (9) - (10)
302 including the generated random weight as a covariate, implying $p = 1$.

303 4. Remove individuals with capture history according to category 0 (undetected).

304 5. Estimate abundance using each of the methods described in Section 3.1, and fit an
305 AR(2) model to the resulting time series including both the A and P variants.

306 The choices made in this simulation study intended to approximate the characteristics of a
307 real ecological data set. Specifically, we have chosen to simulate rather short time series,
308 having similar length as the real data set used in Section 4. Also, the initial capture
309 probabilities for day 1 and day 2 were close to the proportions of captured individuals in
310 the real data set (being 0.55 and 0.77, respectively).

Our next step was to apply INLA and fit the AR(2) process model to the generated time series. This provided estimates of the marginal posterior distributions for the two AR-coefficients ϕ_1 and ϕ_2, for all approaches. Based on the posterior distributions, we could then calculate summary statistics, including the posterior mean of the coefficients, the standard deviations and credible intervals. To evaluate and compare the quality of the

different density-dependence estimates, we computed the estimated root mean squared error (RMSE), defined by

$$\text{RMSE}(\hat{\phi}_k) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\hat{\phi}_k - \phi_k)^2}, \quad k = 1, 2.$$

Here $\hat{\phi}_k$ denotes the posterior mean estimate of the kth AR coefficient while M is the number of simulations. We also compared the frequentist coverage properties using the different approaches. This corresponded to finding the proportion of times the true AR coefficient was inside the M estimated 95% equi-tailed credible intervals.

3.3 Simulation results

Table 1 displays the average performance in terms of coverage and RMSE for the different methods used to estimate density-dependence, including the two variants A and P. The averages were computed across all the given combinations of (ϕ_1,ϕ_2) and for each of the five fixed values of σ^2_ϵ. Due to the short time series length, coverage using the Baseline method will not achieve the nominal level of 0.95 (nominal for the A variant). It is well-known that estimators for the coefficients of AR processes are biased for small sample sizes (Shaman & Stine, 1988). Thus the Baseline method provides the practical optimal results for this length.

The differences for the different methods were rather small, except for the two lowest innovation variance levels where there was a clear benefit from including capture history. CR-INLA provided the highest coverage, followed by CR-VGAM and ObsCount. Using CR-INLA, the coverage was within the range (0.83 – 0.89) for ϕ_1 and within the range (0.80 – 0.86) for ϕ_2. Further, the results indicated that fitting the AR(2) model to the log-rate of a Poisson process (P variant) provided generally higher coverage than using the A variants. When comparing the different methods using RMSE, which considers both bias and variance, we see that CR-INLA had the smallest error in all cases, while the method ObsCount had the largest error. Again, the differences between the methods were very small except for the lowest levels of the innovation variance. In general, RMSE was

reduced for all methods as the innovation variance increased. Moreover, RMSE was higher for the P variants compared to the A variants at the two lowest innovation variance levels, using all methods.

This was due to both an increased variance and bias, which explains why the P variants had higher coverage.

The estimation bias of the different methods can be assessed explicitly in Fig. 2, containing the posterior mean estimates $(\hat{\phi}_1, \hat{\phi}_2)$ for each of the fixed combinations. The figure includes point estimates both using the A variant (left-hand side) and P variant (right-hand side) of the different methods. Here, the results refer to $\sigma_\epsilon^2 = 0.08$ (upper panels) and $\sigma_\epsilon^2 = 0.32$ (lower panels). The corresponding results using the other variance levels are given in the supplementary material (Figs A1–A9). For the two lowest levels of innovation variance, the estimation bias using CR-INLA was slightly lower than using the other methods for all combinations of (ϕ_1, ϕ_2). When the innovation variance was increased, the different methods gave approximately the same point estimates. The bias was slightly larger using the P variants compared with the A variants. This was in correspondence with the higher averages of the RMSE-values for the P variants, as already observed.

To further study coverage and RMSE for each of the 15 combinations, we computed a joint coverage being the proportion of times both of the estimated 95% credible intervals contained ϕ_1 and ϕ_2, respectively. We also computed a joint RMSE for both parameters, defined by

$$\text{RMSE}(\hat{\phi}_1, \hat{\phi}_2) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\sum_{k=1}^{2}(\hat{\phi}_k - \phi_k)^2}.$$

The results for coverage and RMSE are shown in Fig. 3 and Fig. 4, respectively. The results indicated that coverage decreases and RMSE increases as the direct density-dependence measured by $|\phi_1|$ gets stronger. This pattern might become more clear, if the number of simulations is increased. CR-INLA was seen to give the highest coverage and lowest RMSE for most of the combinations when $\sigma_\epsilon^2 = 0.08$, at least for the A variants. When $\sigma_\epsilon^2 = 0.32$, the results were very similar for all methods. In summary, we can conclude that including

capture-history information improved the estimation of density-dependence in process models having low innovation variance. Out of the tested method, our suggested approach CR-INLA, performed best, followed by CR-VGAM. For the given simulated data, the additional step of estimating log-rates of a Poisson process resulted in larger RMSE. Finally, we notice that both of the two AR coefficients were underestimated, and this bias increased with the absolute values of the coefficients.

The given simulation study was based on certain choices to illustrate a capture-recapture scenario using an AR(2) process model. Here we have assumed independent capture probabilities for the two capture sessions. The given approach could have easily been adapted to other models given by Otis et al. (1978), such as to also include a behavioural effect. Longer time series would have improved the estimation results using all of the suggested methods, albeit being less realistic from an ecological point of view.

4 Estimating density-dependence for a real data set

In this section, we estimated density-dependence for a real capture-recapture data set of small mammals, collected at 20 different spatial locations over a period of 18 years. Our main focus was to assess density-dependence estimates, studying how inclusion of capture history influenced the estimation. Using the CR-INLA approach, we estimated capture probabilities by the regression model in (6), including individual-specific covariate information and random effects. We proceeded to estimate the true abundances at each time point for each spatial location using (2). Finally, we fitted the AR(2) model to estimate density-dependence and compared the results with using the methods CR-VGAM and ObsCount. For all three methods, we assessed both the A and P variants.

4.1 Data description

The data included a total of 3090 grey-sided voles, captured alive in the Porsanger region (latidude 70°N), in Northern Norway. The data were collected at 20 different stations, evenly spaced in a transect of 200 km in the period 2000-2017. Sampling was conducted

twice a year, in spring and fall, and each capture session consisted of two visits. Two individual-specific variables were recorded, including *weight* and *sex*. Animals captured dead during the first trapping session were excluded from the analysis.

4.2 Observation model selection, estimating capture probabilities

To estimate individual capture probabilities, we used the whole data set across time points and stations. Our first step was to select a reasonable observation model. Fitting the regression model in (6), we considered inclusion of the following variables

1. *Weight* (continuous standardized variable);
2. *Sex* (categorical variable for male or female);
3. *Season* (categorical variable for spring or fall);
4. *Station* (index variable for the evenly spaced stations);
5. *Time* (index variable for year)

To select which variables should be included, we evaluated different models using various information criteria. When applying CR-INLA, we used the estimates for the Deviance Information Criterion (DIC) (Spiegelhalter, Best, Carlin, & van der Linde, 2002) and Watanabe-Akaikes Information Criterion (WAIC) (Watanabe, 2010). When using CR-VGAM, we used the estimates of Akaike's Information Criterion (AIC) (Akaike, 1973) and the Bayesian Information Criterion (BIC) (Schwarz, 1978).

An overview of the different models and the estimated information criteria are shown in Table 2, comparing the two methods for a total of 8 different models. The VGAM package does not allow for inclusion of random effect terms (Yee et al., 2015), which implies that *Time* could not be included in the CR-VGAM model explicitly. Using INLA, we can straightforwardly include nonlinear effects of covariates. Applying the method CR-INLA, we chose to model *Time* as a first-order random walk process (rw1) (Rue & Held, 2005;

Sørbye & Rue, 2014). Also, we considered to include *season* as a categorical covariate, both using CR-INLA and CR-VGAM. However, using the CR-INLA approach, *season* is not included simultaneously with *time* to avoid confounding.

The resulting optimal observation model for CR-INLA, minimizing both DIC and WAIC, included all variables except *season*. The linear predictor as defined by (6), is here given by

$$\ln(E(Y_{ik})) = \gamma_{k1} weight_i + \gamma_{k2} sex_i + \gamma_{k3} station_i + f(time_i) + \beta_i + \epsilon_i, \quad i = 1, \ldots, n, \quad k = 1, 2, 3,$$

where $f(time_i)$ denotes the rw1-model, specifying a non-linear random effect of time. In selecting an observation model for the CR-VGAM approach, we observed rather small differences in the values of the goodness-of-fit criteria for the different models. The optimal observation model according to AIC included *weight* and *sex*, while BIC was minimized when only *weight* was included. In the case of vole species, *sex* is known to have an effect on detectability (Bryja et al., 2005), so we chose to include both *weight* and *sex* in estimating the capture probabilities.

Fig. 5 illustrates the distributions of the estimated capture probabilities for the two capture sessions, $\{\hat{p}_{i1}\}_{i=1}^n$ and $\{\hat{p}_{i2}\}_{i=1}^n$, using both CR-INLA and CR-VGAM. The mean capture probability is seen to increase on the second day using both methods. CR-VGAM gave higher estimates of the capture probabilities, having a low variance for both days. Using CR-INLA, the estimated individual capture probabilities showed more heterogeneity, having a larger variance for both days. Using the given estimated capture probabilities for the observed categories, we can estimate the probability that an individual is never captured, corresponding to category c_{i0} in (1). The resulting 95% percentile interval for c_{i0} was (0.19–0.32) using CR-INLA. Using CR-VGAM, the corresponding interval was (0.08–0.12).

4.3 Fitting the AR(2) process model to estimate density-dependence

Given the estimates of the capture probabilities for each individual, we used the Horvitz-Thompson estimator to compute abundance at each time point for each station.

We then fitted the AR(2) model to the resulting estimated log-abundance, providing estimates of both direct and delayed density-dependence. We split the time series into spring and fall, to account for a possible seasonal influence in the parameter estimation. This resulted in two time series of length $T = 18$ for each of the 20 stations. The AR(2) model was fitted using the three presented methods (CR-INLA, CR-VGAM and ObsCounts) using both the A and P variants. Station 9 did not have enough observations for the parameters to be estimated, and was thus not included in the results.

The main results are displayed in Fig. 6, showing the posterior mean estimates of the AR coefficients for the two seasons, for variants A and P. The estimates of both direct and delayed density-dependence were very similar using all the given methods, and were thus lumped together (see Figs B1–B2 for detailed values). Interestingly, the differences seen in the capture probability estimates between CR-INLA and CR-VGAM do not seem to have influenced the density-dependence estimates. This is in correspondence with the simulation study in Section 3, as the innovation variance σ_ϵ^2 for all of the stations was quite high, with the overall average being $\sigma_\epsilon^2 = 0.9$. In both spring and fall, the estimates of ϕ_1 varied from around -0.25 to 0.6, whereas the estimates of ϕ_2 ranged from around 0 to -0.8. For all stations, except 3 and 13, the estimated time series showed a semi-periodic behaviour. We also notice that the AR(2) coefficients varied with season for the same station, which suggests a seasonal effect in the density-dependence. Additionally, during both seasons, the results indicate a decreasing trend in the value of ϕ_2 along the given transect (from coast to inland).

5 Discussion

The main goal of this paper was to assess the importance of including capture history information (individual heterogeneity) in the estimation of density-dependence, thus incorporating sampling error in the observation model. To investigate this, we performed an extensive simulation study in which we generated AR(2) time series, representing the

true log-abundance of an animal population, and simulated a CR sampling scenario from that population. We then tested the performance of different methods, both including capture history information and disregarding it. For the first method, CR-INLA, we defined an observation model to estimate individual capture probabilities through a Multinomial likelihood, and followed it with a Horvitz-Thompson estimate of the true abundance. The second method, CR-VGAM, used the existing VGAM methodology to estimate abundance from CR data, establishing it as a control method. Finally, we compared these two methods with a simple (yet common) approach, disregarding the capture history information (effectively assuming a homogeneous capture process), to estimate the true autoregressive coefficients from the observed counts directly. We further investigated the assumption of using a Poisson distribution for the capture data, fitting the AR(2) process to the estimated log-rates. This was chosen as an example of an observation model used in the ubiquitous state-space models, where the observation model typically assumes some type of homogeneous capture process, such as Poisson or log-normal.

We found that incorporating capture history information was important when modelling density-dependence in AR(2) settings with low innovation variance. In such scenarios, both methods including capture-history outperformed the method disregarding it, with reduced estimation bias and improved parameter coverage (8% higher in CR-INLA (A) compared to ObsCount (A) for the lowest tested innovation variance, see Table A1). However, in scenarios with a large innovation variance, the methods which estimated capture probability did not stand out, producing extremely similar results compared to the observed counts approach. Furthermore, parameter estimates for both AR coefficients were generally biased towards 0, using all the methods, increasingly underestimating the absolute values of the parameters. In the context of quasi-periodic dynamics described by an AR(2) process, this means underestimating the strength of direct ϕ_1 and delayed ϕ_2 density-dependence, and overestimating the innovation variance of the AR(2) model (see Fig. A10).

The data collected in Porsanger showcased vole populations with very large fluctuations in abundance, as is typical of such systems (Cornulier et al., 2013; Henttonen & Hanski, 2000). Moreover, the estimated capture probabilities were relatively high, resulting in a relatively small bias when comparing the observed counts and the estimated abundance. This resulted in all methods, and respective variants, producing similar results - this could have been expected given the observation variance is, in that case, only a minor component of the total variance. Other populations, such as large mammals, may show much smaller abundance fluctuations and therefore a larger contribution of the observation error to the total variance (e.g. Besbeas and Morgan (2019)). In the case of other animal populations, such as snakes, contrary, capture probabilities are often very low (Rose, Wylie, Casazza, & Halstead, 2018). The difference between the different approaches could then be substantial. Extending our approach to other observation process models (e.g. spatial capture-recapture models (Royle, Fuller, & Sutherland, 2017), including individual heterogeneity (Efford & Mowat, 2014), would provide a general approach to reducing biases in population dynamic models. One disadvantage of the CR-INLA method is that it would be cumbersome to apply in CR data sets with more than 3 days, given the data expansion necessary to fit Multinomial likelihoods in INLA, where all the category combinations, observed and not, must be present. This could potentially be automatised as in Bayesian fitting of capture-mark-recapture models (McCrea, 2014).

In summary, we have found that using capture-recapture information contributes to improve density-dependence estimates in low innovation variance processes. At least with such processes, we recommended that individual heterogeneity is accounted for in the observation model, as it can constitute an important part of the total sampling error.

6 Acknowledgements

This study was supported by the COAT Tools project. We thank Prof. Rolf Ims for taking part in the collection of the field data used to test the studied methods.

7 Author's Contributions

All three authors conceptualised the study and its objectives. PN and SS developed and implemented the statistical methods and analyses, together with NY's guidance. NY took part in the field data collection. All authors were active in the writing of the manuscript, contributing critically to the drafts and provided final approval for publication. The authors have no conflict of interest to declare.

8 Data Accessibility

The model code used to support the results in this paper, together with the real data used to test the methods, shall be submitted onto Dryad upon acceptance.

9 References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proc. 2nd inter. symposium on information theory* (pp. 267–281). Budapest. doi: 10.1007/978-1-4612-1694-0_15

Baker, S. G. (1994). The multinomial-Poisson transformation. *The Statistician*, *43*(4), 495–504.

Besbeas, P., & Morgan, B. (2019, April). Exact inference for integrated population modelling. *Biometrics*, *75*(2), 475–484. doi: 10.1111/biom.13045

Bjørnstad, O. N., Falck, W., & Stenseth, N. C. (1995). A geographic gradient in small rodent density fluctuations: a statistical modelling approach. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *262*(1364), 127–133. doi: 10.1098/rspb.1995.0186

Bjørnstad, O. N., Stenseth, N. C., & Saitoh, T. (1999). Synchrony and scaling in dynamics of voles and mice in Northern Japan. *Ecology*, *80*(2), 622–637. doi: 10.1890/0012-9658(1999)080[0622:sasido]2.0.co;2

Bryja, J., Nesvadbová, J., Heroldová, M., Jánová, E., Losík, J., Trebatická, L., & Tkadlec, E. (2005, 11). Common vole *Microtus arvalis* population sex ratio: Biases and process variation. *Canadian Journal of Zoology-revue Canadienne De Zoologie - CAN J ZOOL*, *83*, 1391-1399. doi: 10.1139/z05-133

Carothers, A. D. (1973). The effects of unequal catchability on jolly-seber estimates. *Biometrics*, *29*(1), 79. doi: 10.2307/2529678

Chao, A., & Huggins, R. M. (2006). Four modern closed-population capture–recapture models. In B. J. Manly, T. L. McDonald, & S. C. Amstrup (Eds.), *Handbook of capture-recapture analysis* (p. 58-87). New Jersey, USA: Princeton University Press.

Cornulier, T., Yoccoz, N. G., Bretagnolle, V., Brommer, J. E., Butet, A., Ecke, F., ... Lambin, X. (2013). Europe-wide dampening of population cycles in keystone herbivores. *Science*, *340*(6128), 63–66. doi: 10.1126/science.1228992

Dennis, B., & Taper, M. L. (1994). Density dependence in time series observations of natural populations: Estimation and testing. *Ecological Monographs*, *64*(2), 205–224. doi: 10.2307/2937041

Efford, M. G., & Mowat, G. (2014, May). Compensatory heterogeneity in spatially explicit capture–recapture data. *Ecology*, *95*(5), 1341–1348. doi: 10.1890/13-1497.1

Elton, C. S. (1924). Periodic fluctuations in the numbers of animals: Their causes and effects. *Journal of Experimental Biology*, *2*(1), 119–163.

Henttonen, H., & Hanski, I. (2000). Population dynamics of small rodents in northern fennoscandia. In *Chaos in real data* (pp. 73–96). Springer Netherlands. doi: 10.1007/978-94-011-4010-2_4

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, *47*(260), 663–685. doi: 10.1080/01621459.1952.10483446

Huggins, R., & Hwang, W.-H. (2011). A review of the use of conditional likelihood in capture-recapture experiments. *International Statistical Review*, *79*(3), 385–400. doi: 10.1111/j.1751-5823.2011.00157.x

Ims, R. A., Yoccoz, N. G., & Killengreen, S. T. (2011). Determinants of lemming outbreaks. *Proceedings of the National Academy of Sciences*, *108*(5), 1970–1974. doi: 10.1073/pnas.1012714108

Kleiven, E. F., Henden, J.-A., Ims, R. A., & Yoccoz, N. G. (2018). Seasonal difference in temporal transferability of an ecological model: near-term predictions of lemming outbreak abundances. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-33443-6

Lande, R., Engen, S., & Sæther, B. (2002, September). Estimating density dependence in time–series of age–structured populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1425), 1179–1184. doi: 10.1098/rstb.2002.1120

Lebreton, J.-D., & Gimenez, O. (2012). Detecting and estimating density dependence in

wildlife populations. *The Journal of Wildlife Management*, *77*(1), 12–23. doi: 10.1002/jwmg.425

Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics*, *59*(4), 1123–1130. doi: 10.1111/j.0006-341x.2003.00129.x

Marolla, F., Aarvak, T., Øien, I. J., Mellard, J. P., Henden, J.-A., Hamel, S., ... Ims, R. A. (2019). Assessing the effect of predator control on an endangered goose population subjected to predator-mediated food web dynamics. *Journal of Applied Ecology*, *56*(5), 1245–1255. doi: 10.1111/1365-2664.13346

McCrea, R. (2014). *Analysis of capture-recapture data*. Boca Raton, FL: CRC Press/Taylor and Francis Group.

Ono, K., Langangen, Ø., & Stenseth, N. C. (2019). Improving risk assessments in conservation ecology. *Nature Communications*, *10*(1). doi: 10.1038/s41467-019-10700-4

Otis, D. L., Burnham, K. P., White, G. C., & Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*(62), 3–135.

Rose, J. P., Wylie, G. D., Casazza, M. L., & Halstead, B. J. (2018, August). Integrating growth and capture-mark-recapture models reveals size-dependent survival in an elusive species. *Ecosphere*, *9*(8), e02384. doi: 10.1002/ecs2.2384

Royle, J. A., Fuller, A. K., & Sutherland, C. (2017, August). Unifying population and landscape ecology with spatial capture-recapture. *Ecography*, *41*(3), 444–456. doi: 10.1111/ecog.03170

Rue, H., & Held, L. (2005). *Gaussian Markov random fields*. Chapman & Hall/CRC, Boca Raton.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(2), 319–392. doi:

10.1111/j.1467-9868.2008.00700.x

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian computing with INLA: A review. *Annual Review of Statistics and its Application*, *4*, 395–421.

Santin-Janin, H., Hugueny, B., Aubry, P., Fouchet, D., Gimenez, O., & Pontier, D. (2014). Accounting for sampling error when inferring population synchrony from time-series data: A bayesian state-space modelling approach with applications. *PLoS ONE*, *9*(1), e87084. doi: 10.1371/journal.pone.0087084

Schwarz, G. (1978, March). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. doi: 10.1214/aos/1176344136

Shaman, P., & Stine, R. A. (1988). The bias of autoregressive coefficient estimators. *Journal of the American Statistical Association*, *83*(403), 842-848.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *232*(1), 1-28.

Sørbye, S. H., & Rue, H. (2014). Scaling intrinsic gaussian markov random field priors in spatial modelling. *Spatial Statistics*, *8*, 39-51.

Sørbye, S. H., & Rue, H. (2017). Penalised complexity priors for stationary autoregressive processes. *Journal of Time Series Analysis*, *38*(6), 923-935.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. doi: 10.1111/1467-9868.00353

Stenseth, N. C. (1999). Population cycles in voles and lemmings: Density dependence and phase dependence in a stochastic world. *Oikos*, *87*(3), 427. doi: 10.2307/3546809

Stenseth, N. C., Viljugrein, H., Saitoh, T., Hansen, T. F., Kittilsen, M. O., Bolviken, E., & Glockner, F. (2003). Seasonality, density dependence, and population cycles in hokkaido voles. *Proceedings of the National Academy of Sciences*, *100*(20),

11478–11483. doi: 10.1073/pnas.1935306100

Thibaut, L. M., & Connolly, S. R. (2019, November). Hierarchical modeling strengthens evidence for density dependence in observational time series of population dynamics. *Ecology*, *101*(1). doi: 10.1002/ecy.2893

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*, 3571—3594.

Yee, T. W. (2019). VGAM: Vector generalized linear and additive models [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=VGAM` (R package version 1.1-2)

Yee, T. W., Stoklosa, J., & Huggins, R. M. (2015). The VGAM Package for capture-recapture data using the conditional likelihood. *Journal of Statistical Software*, *65*(5). doi: 10.18637/jss.v065.i05

Yoccoz, N. G., & Ims, R. A. (2004). Spatial population dynamics of small mammals: some methodological and practical issues. *Animal Biodiversity and Conservation*, *27.1*, 427–435.

Yoccoz, N. G., Ims, R. A., & Stenseth, N. C. (1993). Estimating demographic parameters and the population size: an updated methodological survey. In N. C. Stenseth & R. A. Ims (Eds.), *The biology of lemmings* (p. 565-587). London: Academic Press.

Table 1

The estimated average coverage and RMSE for all combinations of (ϕ_1, ϕ_2) in the four methods, using five levels of σ_ϵ^2. The AR(2) process was either fitted to the log-abundance (A) or the log-rate of the corresponding Poisson process (P).

		Coverage				*RMSE*			
		ϕ_1		ϕ_2		ϕ_1		ϕ_2	
Method	σ_ϵ^2	*A*	*P*	*A*	*P*	*A*	*P*	*A*	*P*
Baseline	0.04	0.91	0.85	0.88	0.86	0.21	0.40	0.20	0.36
CR-INLA		0.83	0.87	0.80	0.85	0.27	0.38	0.27	0.35
CR-VGAM		0.80	0.83	0.77	0.83	0.29	0.40	0.28	0.37
ObsCount		0.77	0.81	0.75	0.82	0.31	0.42	0.29	0.38
Baseline	0.08	0.91	0.92	0.89	0.90	0.20	0.26	0.20	0.25
CR-INLA		0.87	0.89	0.85	0.87	0.25	0.27	0.24	0.26
CR-VGAM		0.86	0.88	0.84	0.86	0.26	0.29	0.25	0.27
ObsCount		0.84	0.87	0.82	0.85	0.27	0.31	0.26	0.29
Baseline	0.16	0.92	0.91	0.88	0.88	0.20	0.22	0.20	0.21
CR-INLA		0.89	0.89	0.86	0.86	0.23	0.24	0.22	0.23
CR-VGAM		0.88	0.88	0.86	0.85	0.24	0.25	0.23	0.24
ObsCount		0.87	0.88	0.85	0.85	0.24	0.25	0.23	0.24
Baseline	0.32	0.91	0.91	0.88	0.88	0.21	0.21	0.20	0.21
CR-INLA		0.89	0.89	0.86	0.87	0.23	0.23	0.22	0.22
CR-VGAM		0.88	0.88	0.87	0.86	0.23	0.23	0.22	0.23
ObsCount		0.88	0.88	0.86	0.86	0.23	0.24	0.23	0.23
Baseline	0.64	0.91	0.90	0.89	0.87	0.21	0.22	0.20	0.21
CR-INLA		0.88	0.88	0.85	0.85	0.23	0.23	0.22	0.22
CR-VGAM		0.88	0.87	0.84	0.84	0.23	0.23	0.23	0.23
ObsCount		0.87	0.87	0.84	0.84	0.23	0.24	0.23	0.23

Table 2

Observation model selection for CR-INLA and CR-VGAM, using the selected information criteria.

		CR-INLA		**CR-VGAM**	
Model	**Covariates**	***DIC***	***WAIC***	***AIC***	***BIC***
1	intercept	19400	19613	6568	6580
2	weight	19251	19455	6560	6578
3	weight+sex	19201	19396	6556	6580
4	weight+sex+season	19177	19371	6557	6588
5	weight+sex+station	19187	19381	6558	6587
6	weight+sex+time	19146	19335	-	-
7	weight+sex+season+station	19151	19344	6559	6595
8	weight+sex+station+time	19125	19313	-	-

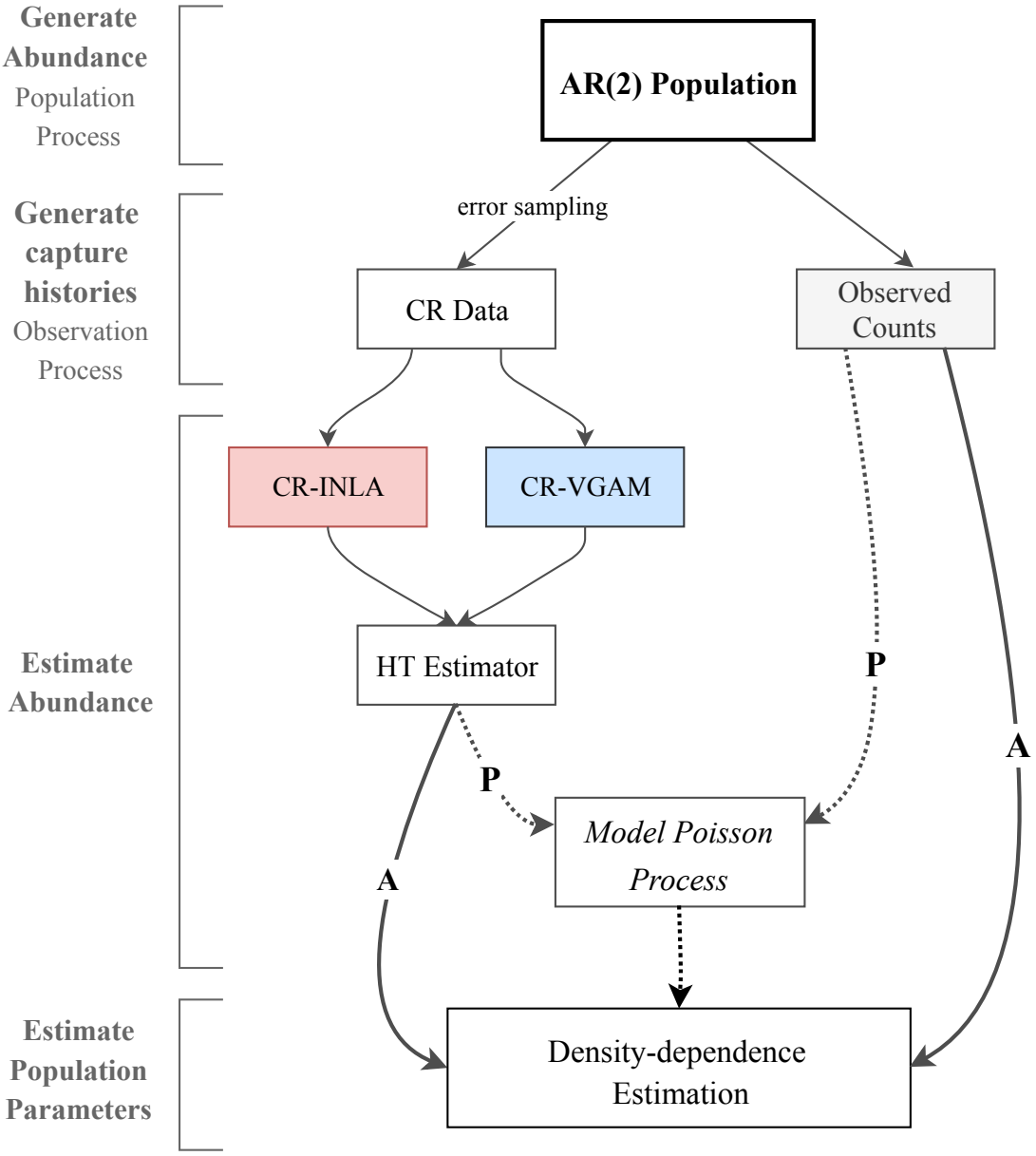


Figure 1. Methodological flowchart.

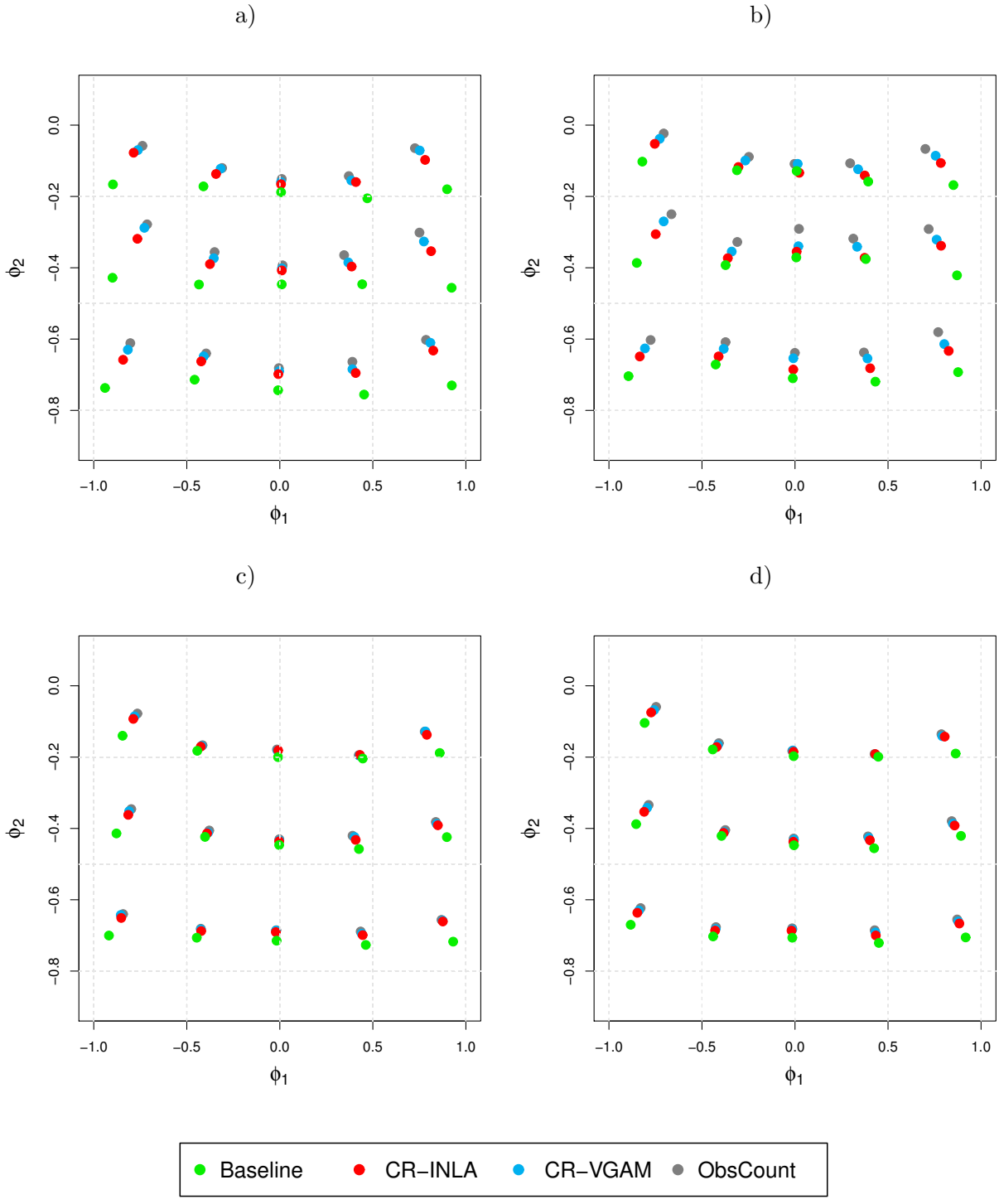


Figure 2. Posterior mean estimates of ϕ_1 and ϕ_2, for the A variants on the right panels, a and c, and P variants on the left, b and d. The points of intersection of the dotted grey lines correspond to the true parameter values. The intersections, at which each set of dots lean to, correspond to the true value of that given set. Panels a and b show results when $\sigma^2_\epsilon = 0.08$, whereas c and d correspond to $\sigma^2_\epsilon = 0.32$.

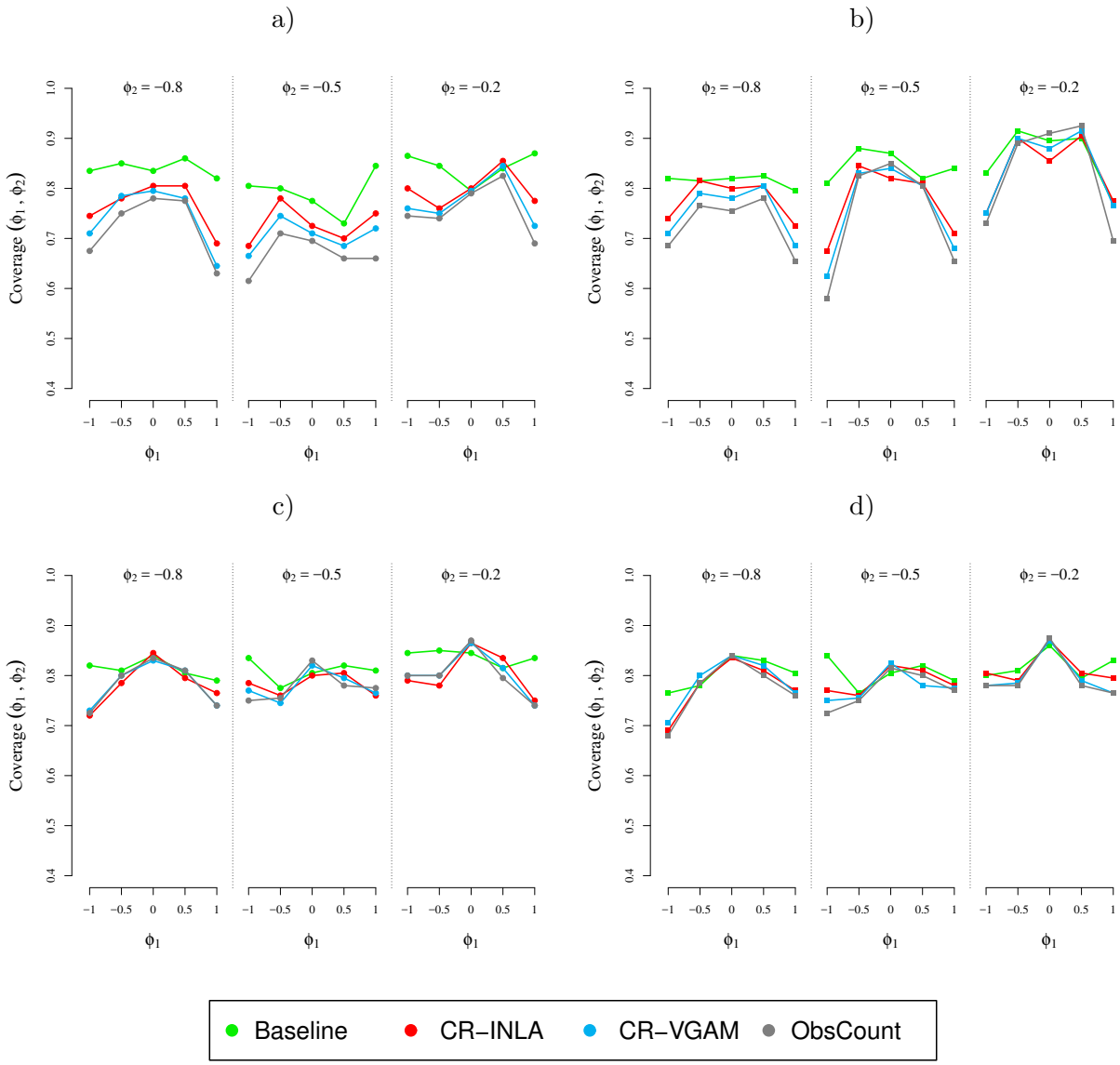


Figure 3. Joint coverage for different combinations of (ϕ_1,ϕ_2) for $\sigma^2_\epsilon = 0.08$ (panels a and b) and $\sigma^2_\epsilon = 0.32$ (panels c and d). A variants are represented on the left (panels a and c) and P variants on the right (panels b and d). The results were split into 3 sets ($\phi_2 \in (-0.8, -0.5, -0.2)$, where each set includes the coverage results for $\phi_1 \in (-1, -0.5, 0, 0.5, 1)$.

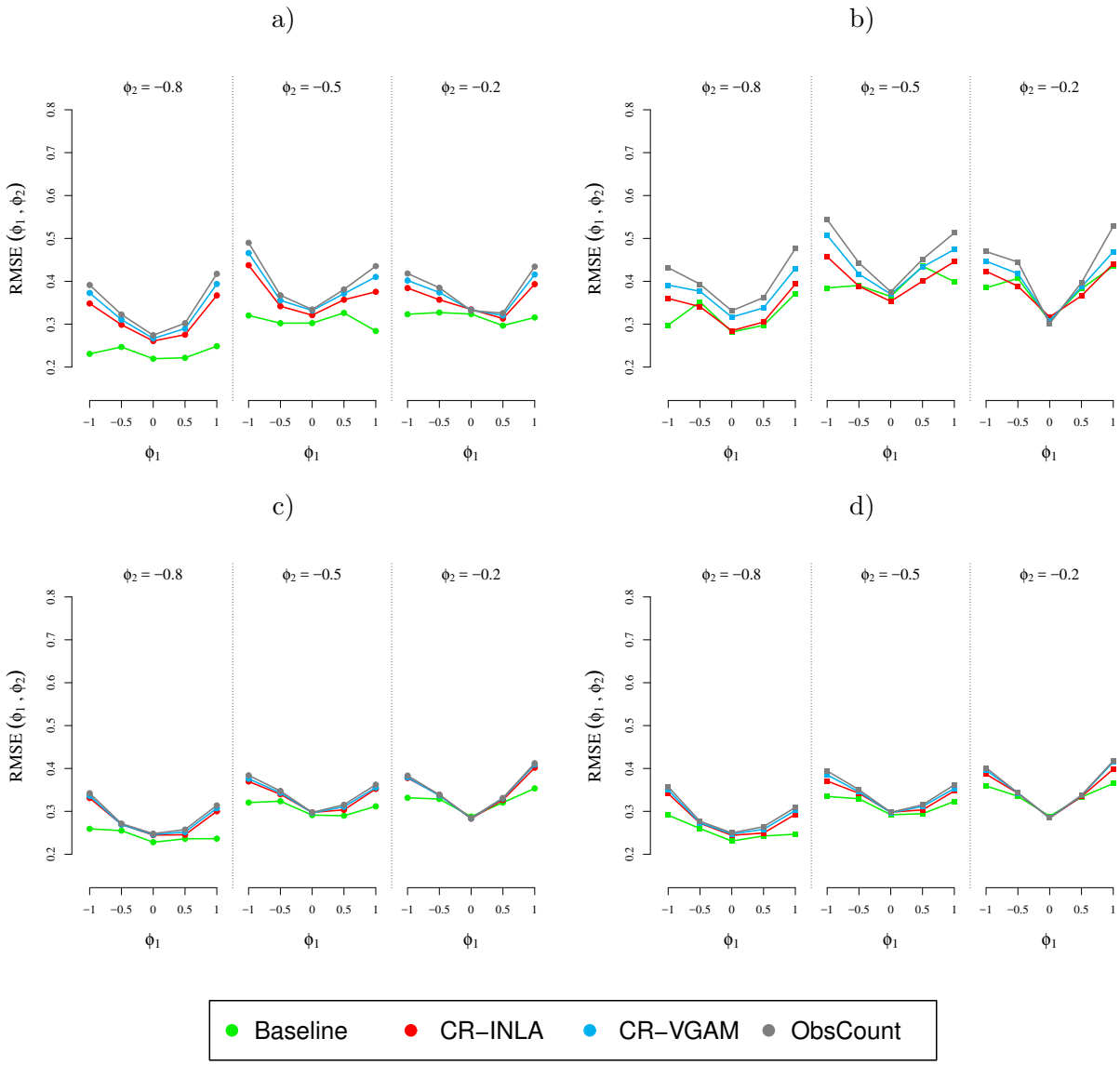


Figure 4. Joint RMSE for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.08$ (panels a and b) and $\sigma_\epsilon^2 = 0.32$ (panels c and d). A variants are represented on the left (panels a and c) and P variants on the right (panels b and d). The results were split into 3 sets (ϕ_2 $\in (-0.8, -0.5, -0.2)$, where each set includes the RMSE results for $\phi_1 \in (-1, -0.5, 0, 0.5, 1)$.

a) CR-INLA

b) CR-VGAM

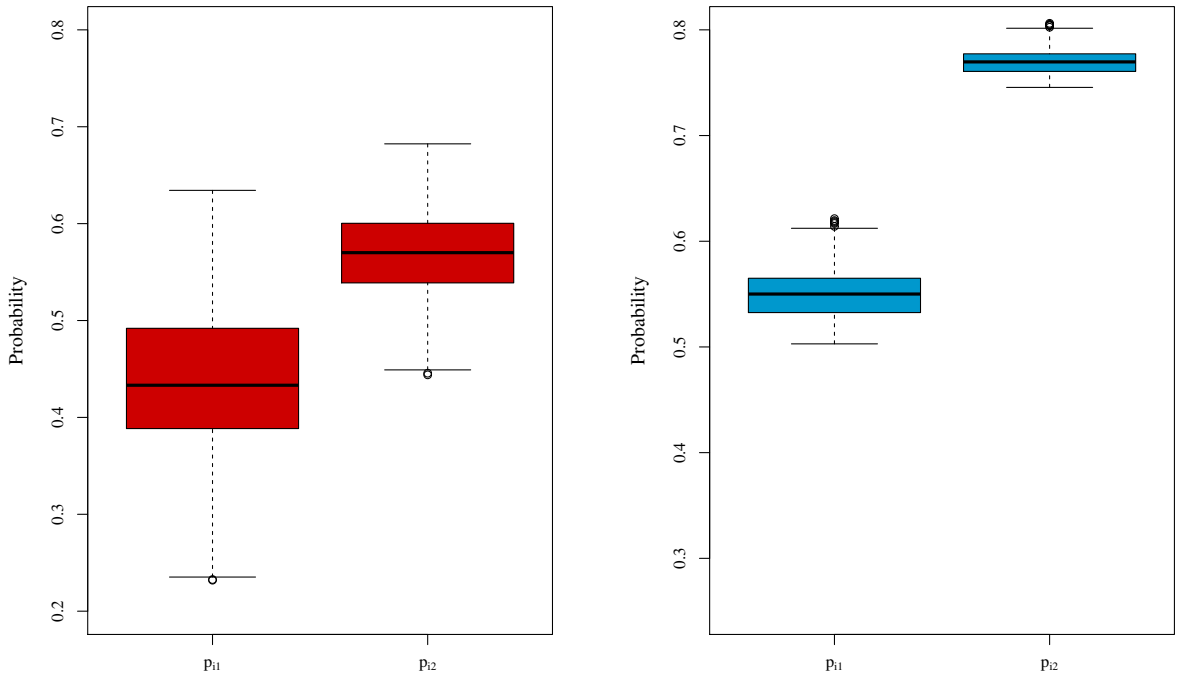
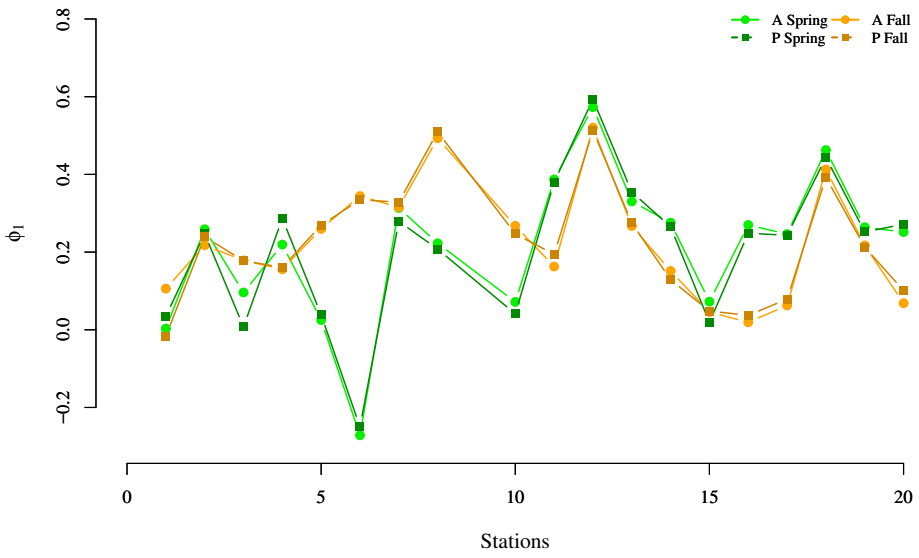


Figure 5. Estimates for p_1 and p_2 for the CR-INLA (panel a) and CR-VGAM (panel b) models.

a)



b)

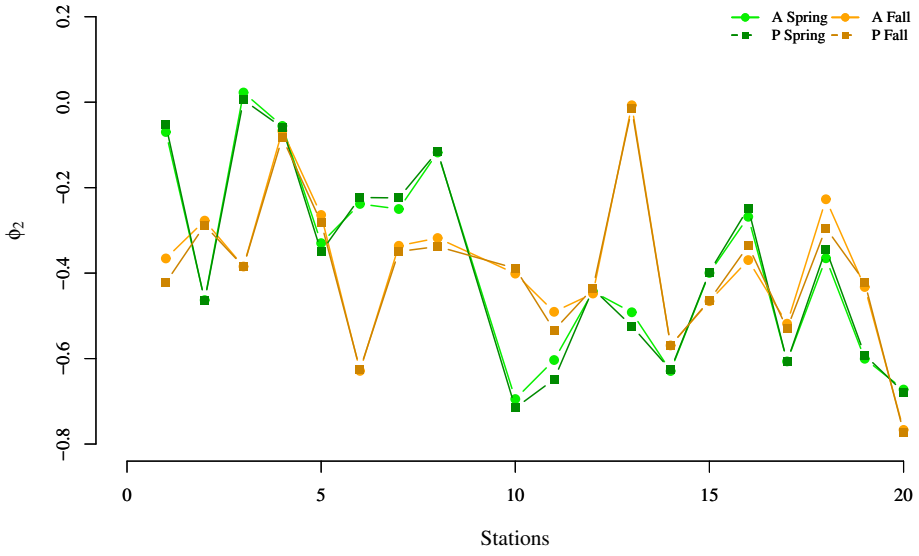


Figure 6. Estimates for ϕ_1 (panel a) and ϕ_2 (panel b) for the mean coefficients of both A and P variants, for the spring (green hue) and fall (orange hue) seasons separately.

Appendix A

Simulation results

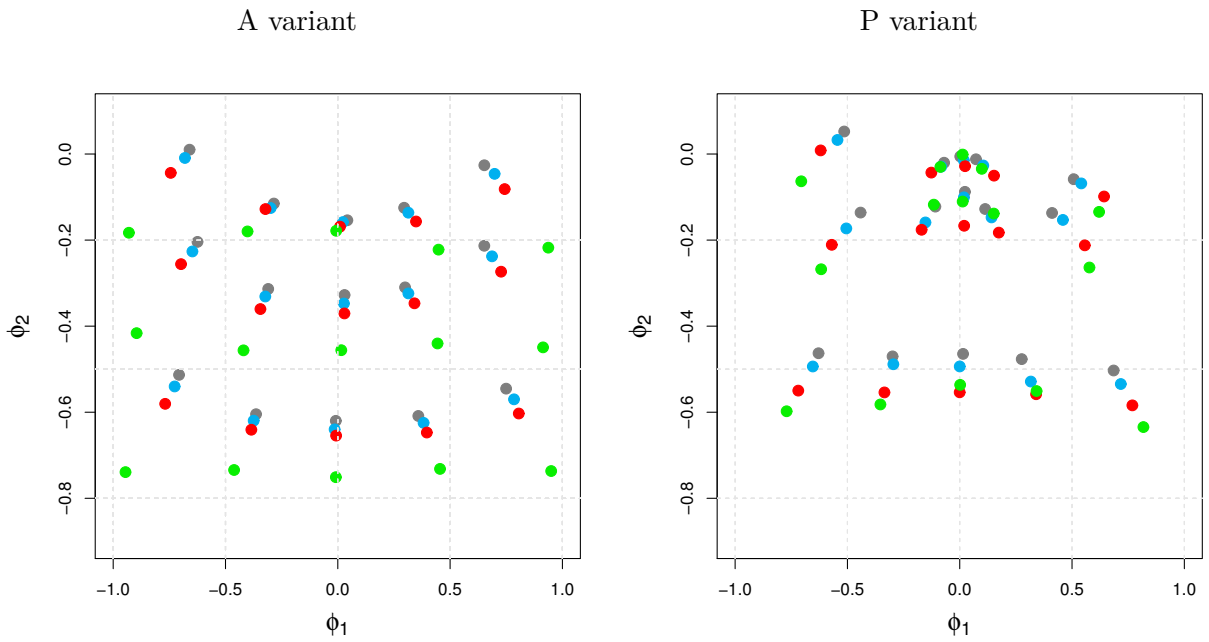


Figure A1. Estimates of the different methods for $\sigma_\epsilon^2 = 0.04$.

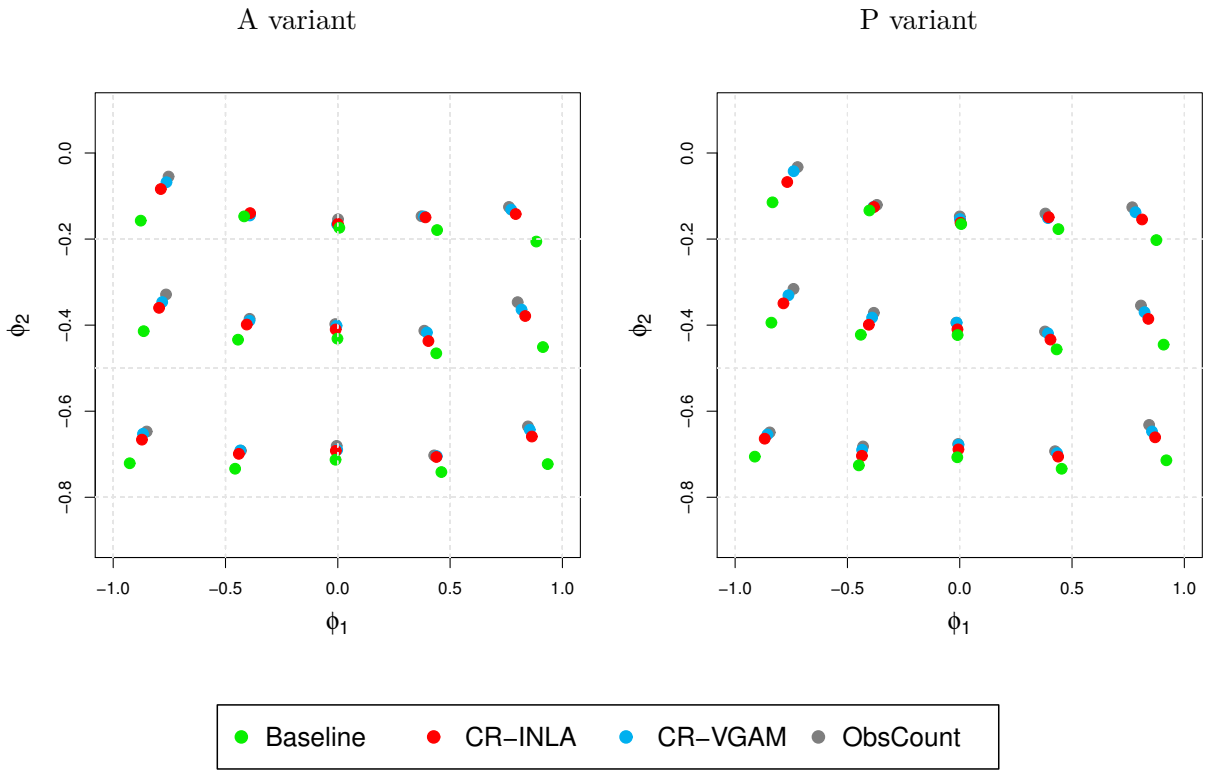


Figure A2. Estimates of the different methods for $\sigma_\epsilon^2 = 0.16$.

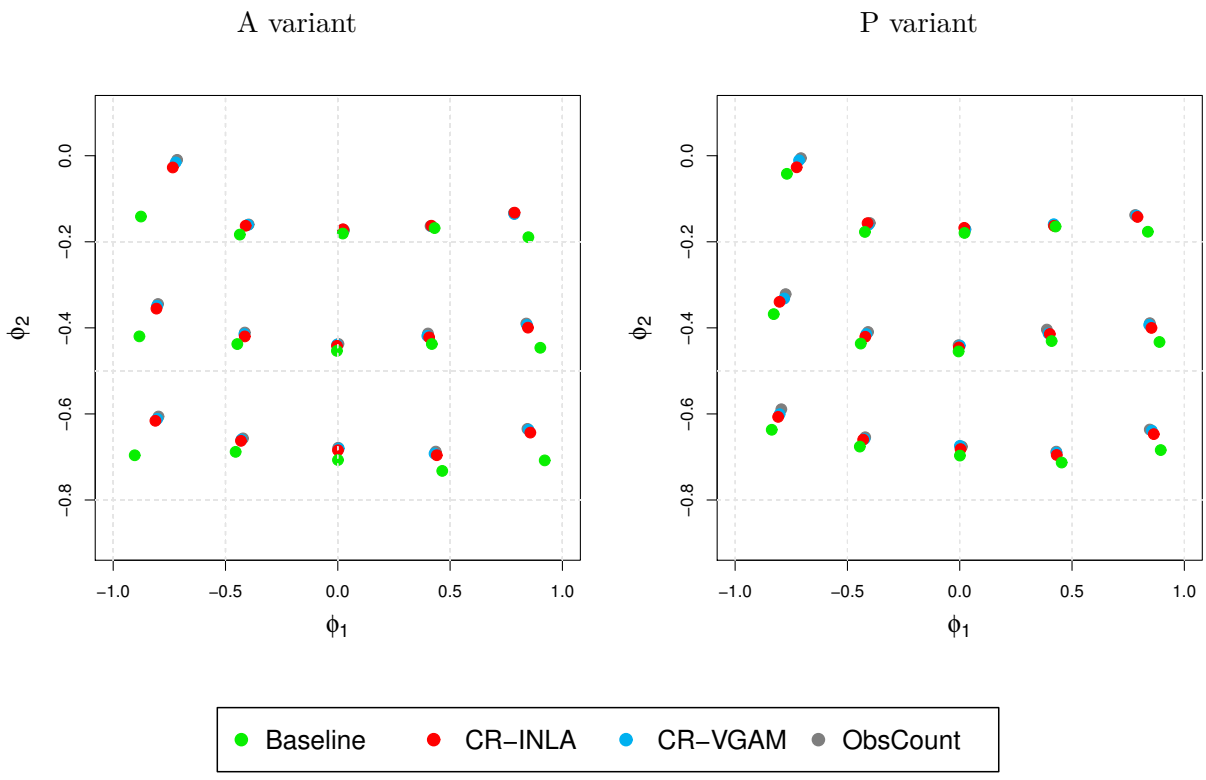


Figure A3. Estimates of the different methods for $\sigma_\epsilon^2 = 0.64$.

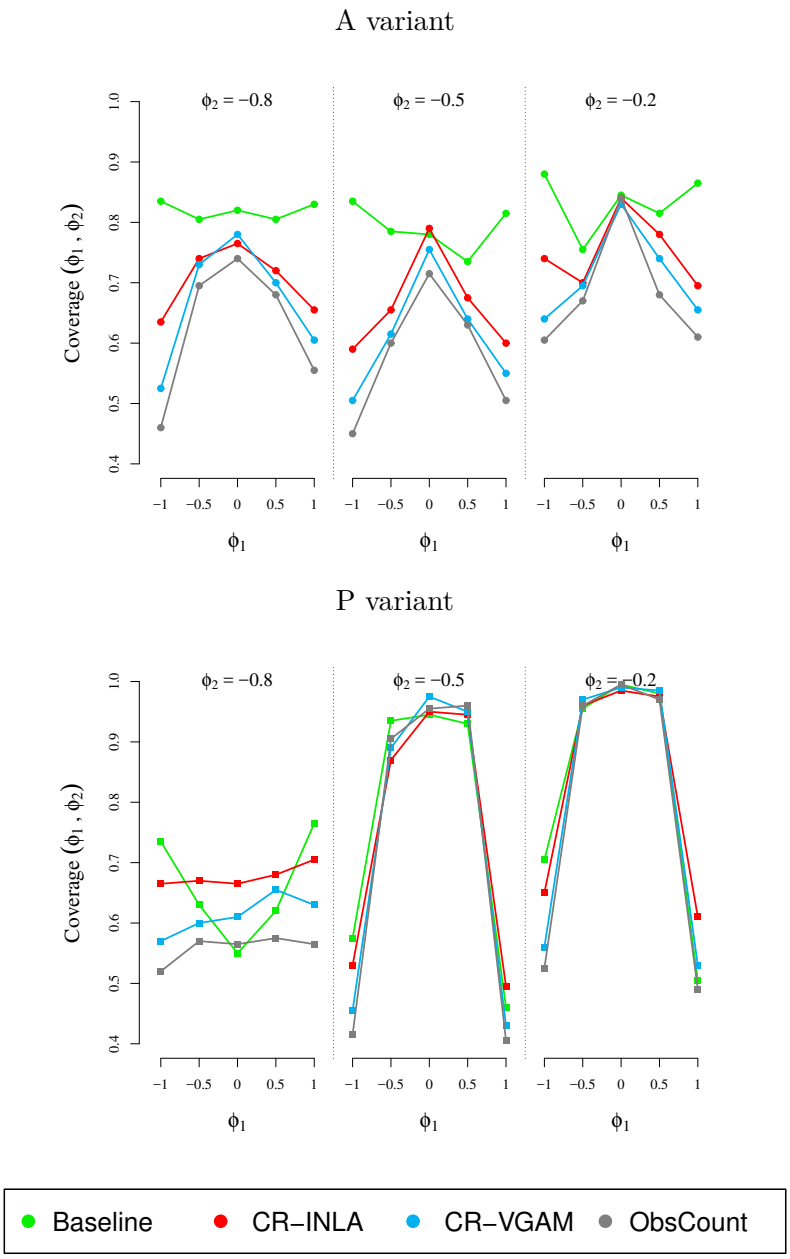


Figure A4. Coverage for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.04$ in both variants.

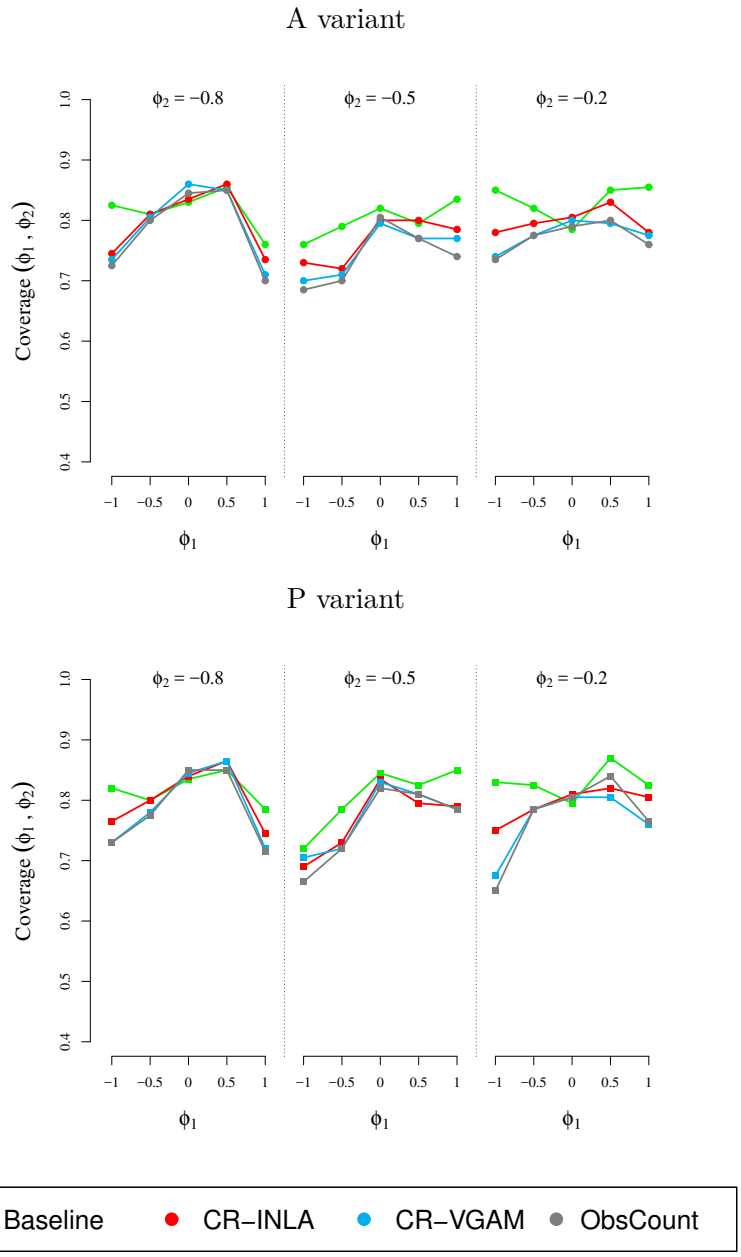


Figure A5. Coverage for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.16$ in both variants.

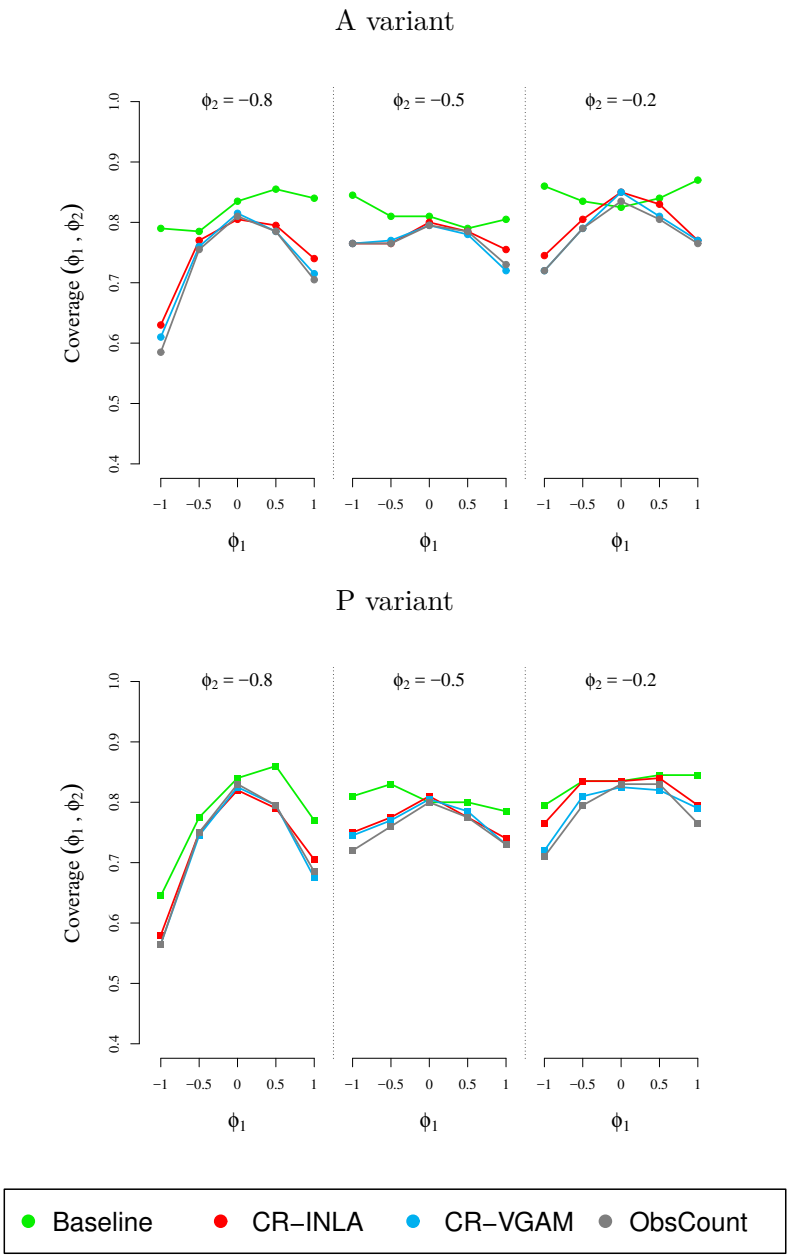


Figure A6. Coverage for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.64$ in both variants.

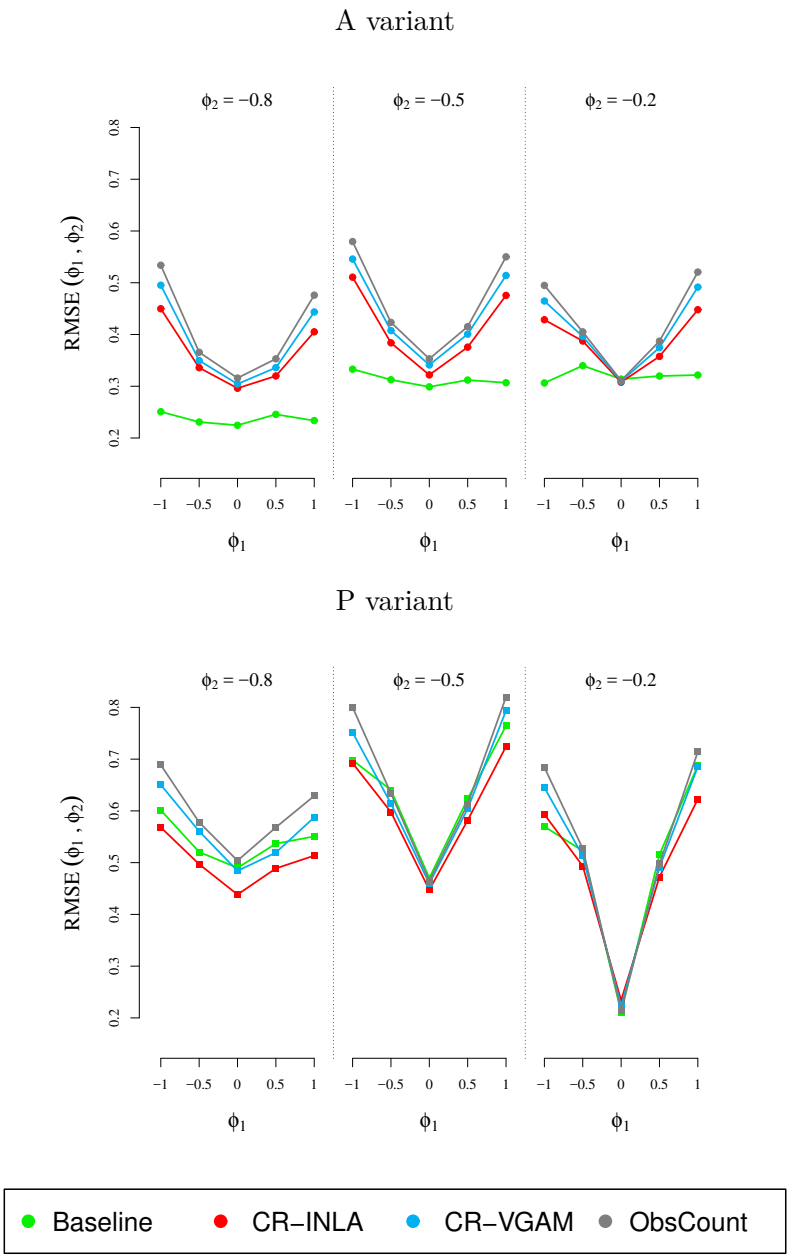


Figure A7. RMSE for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.04$ in both variants.

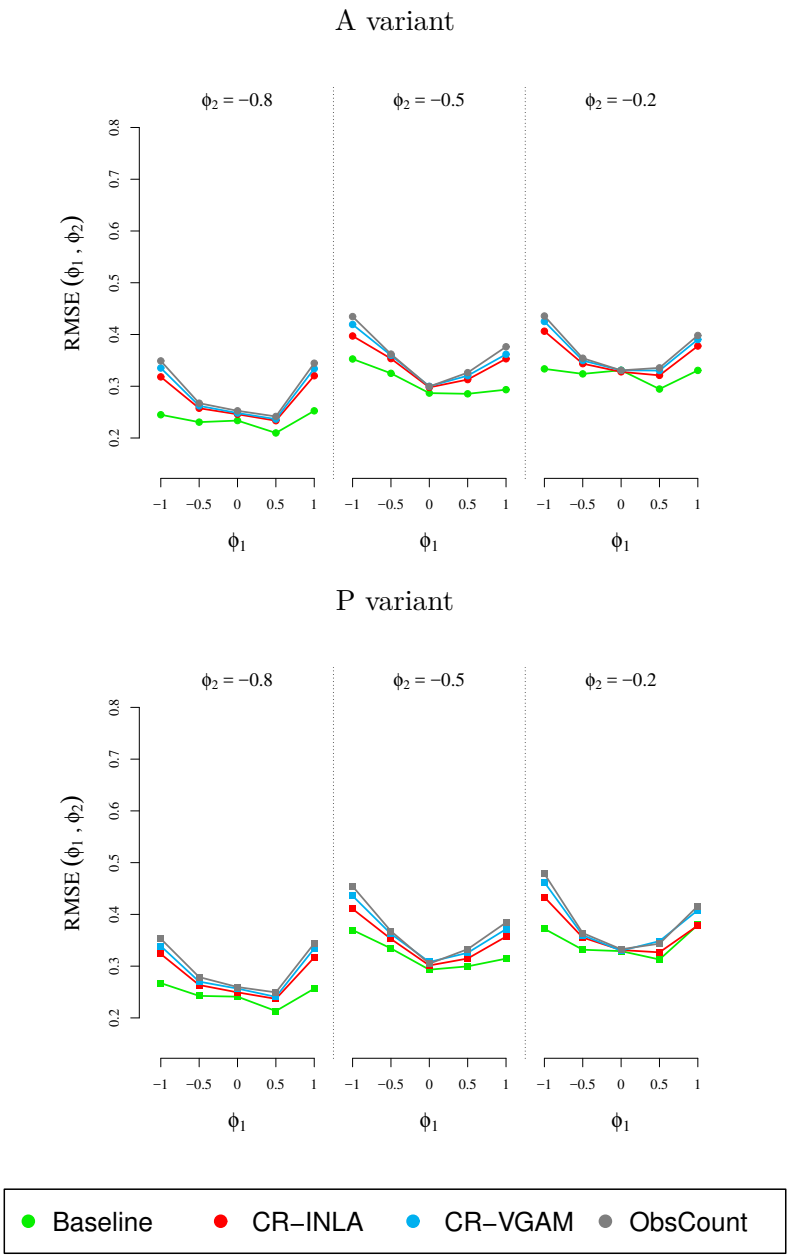


Figure A8. RMSE for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.16$ in both variants.

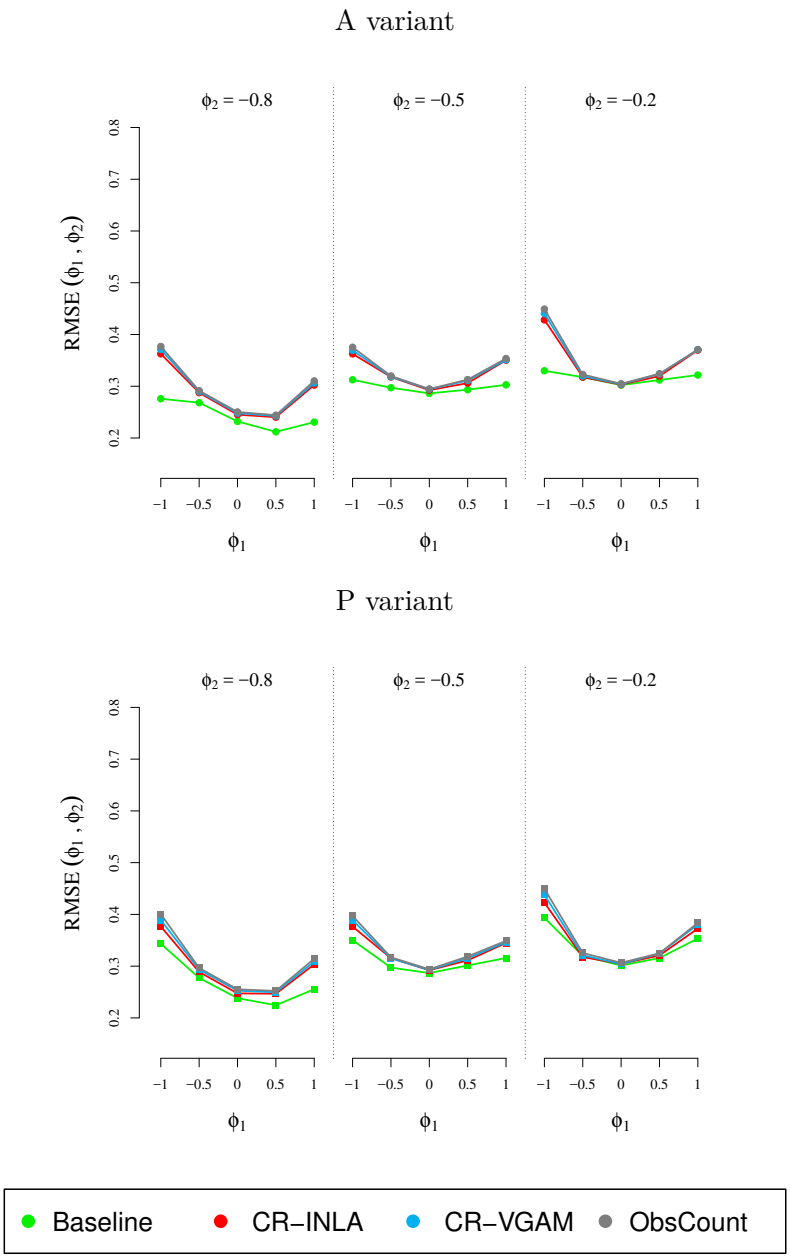


Figure A9. RMSE for different combinations of (ϕ_1,ϕ_2) for $\sigma_\epsilon^2 = 0.64$ in both variants.

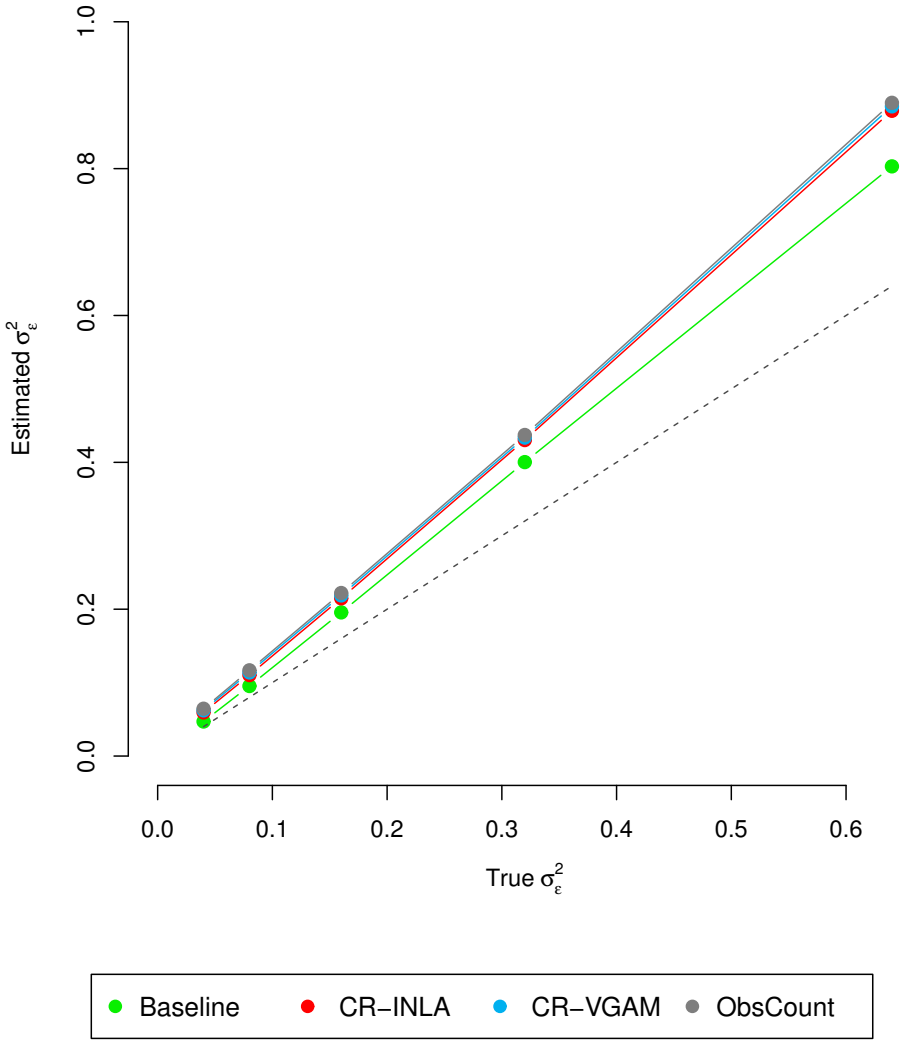


Figure A10. Mean innovation variance estimate for the A variants in the simulation exercise, across all combinations of parameters. The dashed grey line corresponds to the theoretical optimum. The green line provides the practical optimum. The three methods provide similar estimates of the innovation variance, increasingly overestimating it for larger σ^2_ϵ values.

Table A1

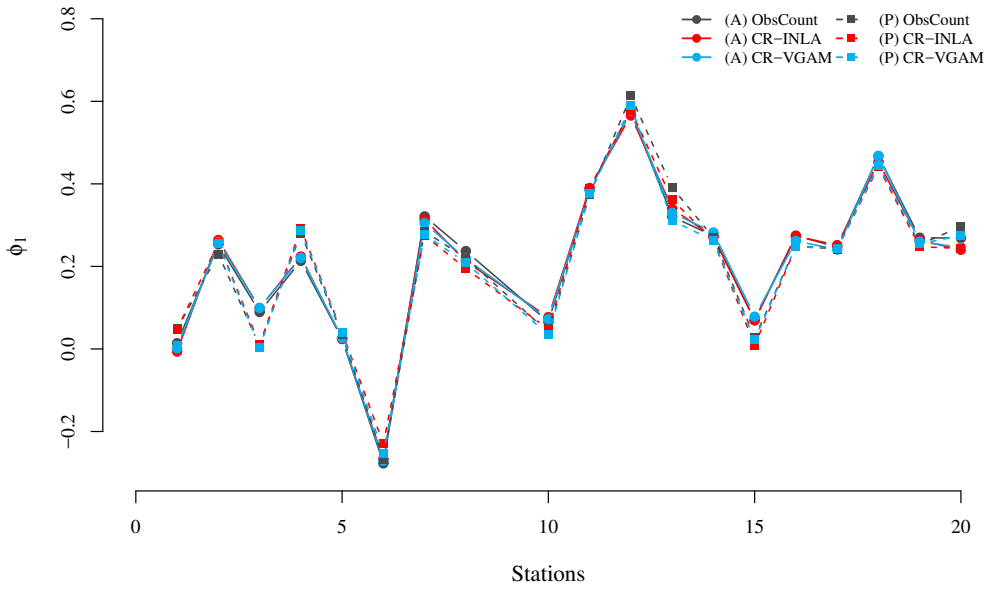
Values of average joint coverage and joint RMSE for all combinations of simulations for the four different methods, in 5 levels of variance σ^2_ϵ. (A) columns represent the log-Abundance variants, while (P) columns show values for the log-Poisson rate variants.

		Joint Coverage		*Joint RMSE*	
Method	σ^2_ϵ	*A*	*P*	*A*	*P*
Baseline	0.04	0.81	0.75	0.41	0.75
CR-INLA		0.71	0.76	0.54	0.72
CR-VGAM		0.66	0.72	0.58	0.77
ObsCount		0.63	0.69	0.61	0.80
Baseline	0.08	0.82	0.84	0.40	0.51
CR-INLA		0.76	0.80	0.48	0.53
CR-VGAM		0.74	0.78	0.51	0.57
ObsCount		0.72	0.77	0.53	0.60
Baseline	0.16	0.82	0.82	0.41	0.43
CR-INLA		0.79	0.79	0.46	0.46
CR-VGAM		0.77	0.77	0.47	0.48
ObsCount		0.77	0.77	0.48	0.49
Baseline	0.32	0.82	0.81	0.41	0.43
CR-INLA		0.79	0.79	0.45	0.45
CR-VGAM		0.79	0.79	0.45	0.46
ObsCount		0.79	0.78	0.46	0.47
Baseline	0.64	0.83	0.80	0.40	0.43
CR-INLA		0.77	0.77	0.45	0.45
CR-VGAM		0.76	0.76	0.46	0.46
ObsCount		0.76	0.76	0.46	0.47

Appendix B

Real data results

Spring



Fall

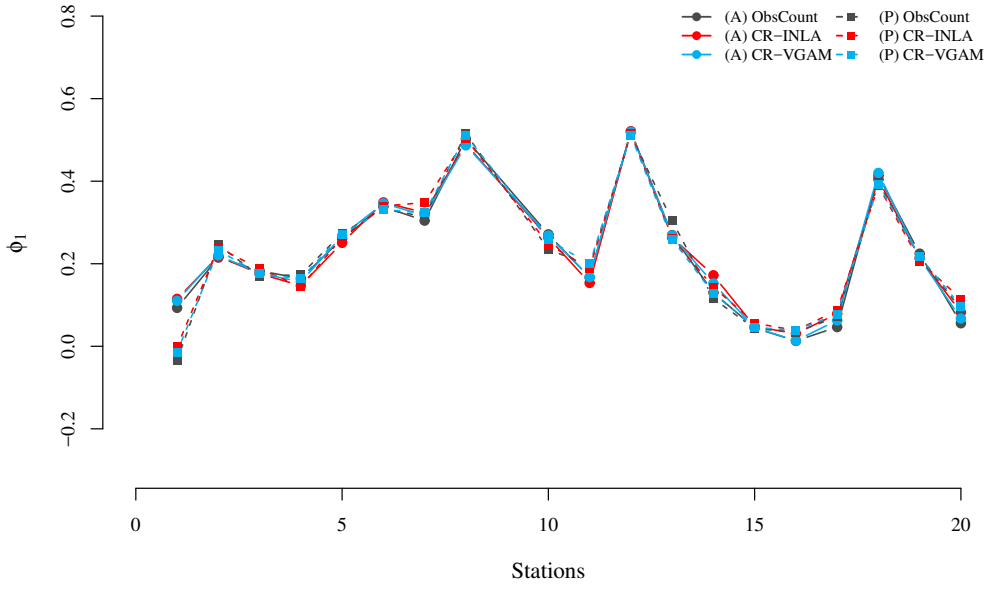
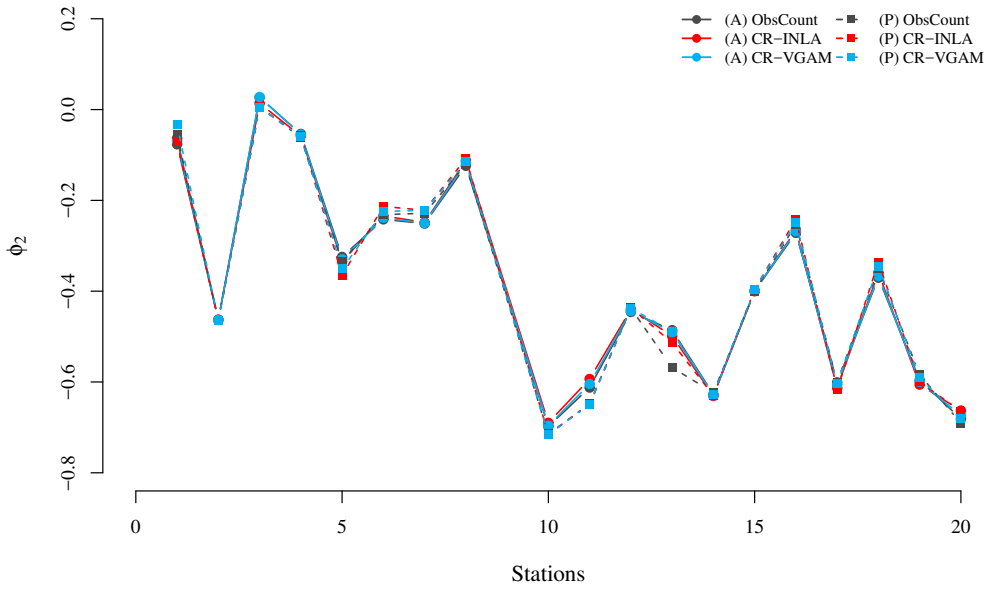


Figure B1. Estimates for ϕ_1 for the different methods in both variants, per season.

Spring



Fall

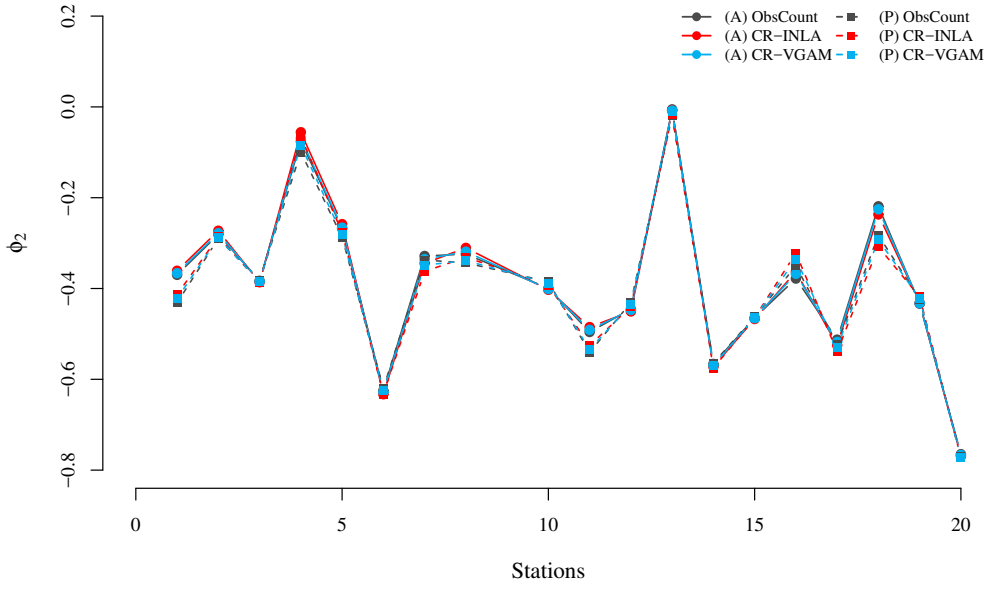


Figure B2. Estimates for ϕ_2 for the different methods in both variants, per season.