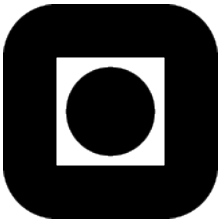
NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Animal models and Integrated Nested Laplace Approximations

by

Anna Marie Holland, Ingelin Steinsland, Sara Martino and Henrik Jensen

PREPRINT
STATISTICS NO. 4/2011



NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This preprint has URL
http://www.math.ntnu.no/preprint/statistics/2011/S4-2011.pdf
Ingelin Steinsland has homepage: http://www.math.ntnu.no/∼ingelins
E-mail: ingelins@math.ntnu.no
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Animal models and Integrated Nested Laplace Approximations

Anna Marie Holand* & Ingelin Steinsland† & Sara Martino‡ & Henrik Jensen§

May 12, 2011

Abstract

Animal models are generalized linear mixed model (GLMM) used in evolutionary biology and animal breeding to identify the genetic part of traits. Integrated Nested Laplace Approximation (INLA) is a methodology for making fast non-sampling based Bayesian inference for hierarchical Gaussian Markov models. In this paper we demonstrate that the INLA methodology can be used for many versions of Bayesian animal models. We analyse animal models for both synthetic case studies and house sparrow population case studies with Gaussian, Binomial and Poisson likelihoods using INLA. Inference results are compared with results using Markov Chain Monte Carlo (MCMC) methods. We also introduce an R package, `AnimalINLA`, for easy and fast inference for Bayesian Animal models using INLA.

Keywords: Additive genetic effects; Approximate Bayesian inference; Heritability; House Sparrow; Natural populations; Quantitative genetics.

1 Introduction

Quantitative genetics is the study of quantitative traits, and is a cornerstone in both evolutionary biology and animal breeding. Examples of quantitative traits are the height of adult humans, the amount of milk a cow produces and the litter size in sheep. Quantitative traits often display a continuous Gaussian distribution of phenotypic (observed) values, e.g. adult height for humans, and bill depth of house sparrows. However, not all quantitative traits follow a continuous Gaussian distribution. Examples are the presence or absence of intramammary infection in cows and litter size in sheep. Most life-history traits such as lifespan (how long an individual lives) and reproductive success (the

*Centre for Conservation Biology, Department of Biology, NTNU, NO-7491 Trondheim, Norway, e-mail: Anna.Marie.Holand@bio.ntnu.no

†Department of Mathematical Sciences, NTNU, NO-7491 Trondheim, Norway, e-mail: ingelins@math.ntnu.no

‡Department of Mathematical Sciences, NTNU, NO-7491 Trondheim, Norway, e-mail: martino@math.ntnu.no

§Centre for Conservation Biology, Department of Biology, NTNU, NO-7491 Trondheim, Norway, e-mail: henrik.jensen@bio.ntnu.no

number of offspring) are also non-Gaussian. These are important traits in an evolutionary perspective as they are major components of individuals' fitness (see e.g. Freeman and Herron, 2004).

The quantitative genetic theory is based on the assumption that the character is determined by a large number of genes with multiple alleles as well as by environmental conditions. An important quantitative genetic parameter is the heritability of a trait; the proportion of the phenotypic (observed trait) variance in a population which is explained by additive genetic effects. Genetic effects are additive if the effect of two or more gene loci are equal to the sum of their individual effects. See, e.g., Lynch and Walsh (1998), Simm (1998) and Sorensen and Gianola (2002) for an introduction to quantitative genetics.

To estimate the additive genetic variance (and thus the heritability) of different kinds of traits, biologists and animal breeders often use a generalized linear mixed model (GLMM) called an *animal model*. In an animal model individual i's trait, y_i has a genetic part, u_i. The value u_i is known as the breeding value of individual i. From the assumption that the breeding value is the sum of effects of many genes and from the central limit theorem, the breeding values are assumed to have a Gaussian distribution with a dependence structure given by the pedigree.

Inference for any stochastic model can either be done in a frequentist or in a Bayesian framework, see Blasco (2001) for a discussion of the animal model. Since early 1960's animal breeders have successfully used the frequentist approach with Restricted Maximum Likelihood (REML), to for example increase meat or milk yield in cattle (Simm, 1998). However, this approach has its limitations when it comes to calculating breeding values (Wilson et al., 2009) and for non-Gaussian traits (Tempelman and Gianola, 1994; Sorensen and Gianola, 2002; Fong et al., 2010). In a Bayesian framework parameters are considered random variables. This solves the problems of calculating breeding values and making inference for non-Gaussian traits (Tempelman and Gianola, 1994; Fong et al., 2010). This flexibility of the Bayesian framework has made Bayesian animal models increasingly popular. They have been used in animal breeding since late 1970's (Dempfle, 1977; Blasco, 2001; Sorensen and Gianola, 2002), while they only recently have been introduced to evolutionary biology (O'Hara et al., 2008; Ovaskainen et al., 2008; Hadfield, 2010; Steinsland and Jensen, 2010).

Markov Chain Monte Carlo (MCMC) methods is the traditional way of doing inference for Bayesian animal models (Sorensen and Gianola, 2002). MCMC is a very flexible methodology that can be used to make inference for any Bayesian model, and we can get posterior estimates for any random variable / parameter, marginally, jointly or functions of them. However, it has its pitfalls: for models with many variables MCMC is computationally expensive, i.e. it can take hours to do the inference. Further, the quality of the results can be difficult to assess as the Markov chain may suffer from undetected slow convergence or poor mixing. Setting up a good MCMC algorithm (quick convergence, good mixing and computationally fast) is challenging for a non-specialist. Recently this has improved for animal models as there are now packages available for doing inference for animal models with MCMC both in R (`MCMCglmm`; Hadfield,

2010) and in BUGS (Lunn et al., 2000).

For hierarchical Gaussian Markov random field models a non-sampling based alternative to MCMC, the Integrated Nested Laplace Approximations (INLA) has recently been introduced, (Rue et al., 2009). Using INLA we can calculate marginal posteriors for all parameters and each random effect, as well as the posterior for linear combinations of random effects. Because INLA is based on direct numerical integration instead of simulations, it is much faster than MCMC and more accurate for a given computation time (Rue et al., 2009). Steinsland and Jensen (2010) used a similar approach for doing non-sampling based inference for a Gaussian trait. INLA has been used in several fields of statistics, e.g. survival analysis (Martino et al., 2010), for spatial GLMM (Eidsvik et al., 2009) and in disease mapping (Schrödle et al., 2011).

This paper contributes to easier and faster Bayesian inference for both Gaussian and several non-Gaussian animal models by demonstrating that these models fit the INLA-framework and by providing an R-package, `AnimalINLA`, for doing the inference.

In Section 2 the data used in the case studies are introduced. Section 3 briefly revise relevant requirements for using INLA and the possibilities INLA gives. The animal models we use are fully specified in Section 4. In section 5 and 6 results from the synthetic case studies and house sparrow case studies are presented, respectively. Inference is done with INLA and for some cases results are compared with MCMC. The results of the house sparrow case study is discussed in Section 7, and Section 8 ends the article with a conclusion.

2 Data

For the case studies we have used data from a natural insular metapopulation of house sparrow (*Passer domesticus*) on five islands off the coast of Helgeland in Northern Norway (66°N, 13°E).

From adults and juveniles (i.e. birds born the same summer) a small blood sample was collected and from adults several morphological traits were measured (including bill depth). The blood samples were used to determine genetic parentage, and a genetic pedigree for the birds on the study islands could be established. For a more thorough description of the field work, study area and genetic parenthood analyses, see (Ringsby et al., 2002; Jensen et al., 2008; Pärn et al., 2009) and references therein.

This study system has many qualities for providing data on morphology and fitness-related traits as more than 90% of all birds on the five main study islands were individually ringed. Further, intensive observation and capture protocols each year gave good estimates of the lifespan of individual birds (a bird was consider dead when it was no longer captured or observed).

For all case studies we used 1993 to 2002 as our study period, and we used the same pedigree, which consisted of the $n_p = 3574$ individuals that were present on the 5 main study islands in this period. The pedigree spanned seven generations. For our case studies we used individual data on 1) bill depth, 2) breeding season success and 3) lifetime reproductive success. For all birds sex,

hatch year and hatch island was available.

In case study one we consider bill depth as one year old. Bill depths were approximately Gaussian distributed (see Appendix C, Figure C.1), and we have measurements for $n_d = 1025$ birds. Many individuals in the pedigree had missing data for this trait because the bird did not survive until it was one year old. For individuals that were not measured as one year old, but later in life we have used bill depths adjusted to one year old size as in Jensen et al. (2008). We standardized the data to have mean 0 and variance 1.

In case study two we considered breeding season success. If at least one of the offspring an adult bird produced in a given breeding season survived until recruitment (i.e. one year of age) we defined its breeding season a success. Otherwise the breeding season was a failure. The breeding season could be a failure either because the bird did not produce any offspring, or because all its offspring died before recruitment. The data consist of pairs of values (n_i, y_i), where n_i is the number of breeding seasons individual i had during the study period (e.g. it was alive and adult) and y_i is the number of successful breeding seasons, $y_i \leq n_i$. Individuals that died before their first breeding season (did not recruit) or that emigrated to an island not among the 5 main study islands have no data. There are $n_d = 1182$ individuals with data. Of these about 71% did not have any successful breeding seasons.

In case study three we consider data on individual lifetime reproductive success (LRS), i.e. the number of recruits the individual produced over its lifetime. Data takes the form (n_i, y_i) where n_i is identical to n_i in case study two, and y_i is the total number of recruits produced in the study period. For this trait we had data for the same $n_d = 1182$ individuals as in case study two. y_i ranged from 0 to 10, with mean 0.64. 71% produced no recruits, and about 46% of the 344 individuals that produced one or more recruits produced only one.

3 Latent Gaussian models and INLA

In this section we give a brief introduction to latent Gaussian models and how Integrated Nested Laplace approximation (INLA) can be used to make approximations for posterior marginals for these models.

In general, latent Gaussian models are hierarchical models where we assume a n_p-dimensional latent field $\boldsymbol{x}$ to be point-wise observed through $n_d \leq n_p$ data $\boldsymbol{y}$. The latent field $\boldsymbol{x}$ is assumed to have Gaussian density conditional on some hyperparameters $\boldsymbol{\theta}_1$: $\boldsymbol{x}|\boldsymbol{\theta}_1 \sim \mathcal{N}(0, \boldsymbol{Q}^{-1}(\boldsymbol{\theta}_1))$.

The data $\boldsymbol{y}$ are assumed to be conditionally independent given the latent field $\boldsymbol{x}$ and, possibly, some additional hyperparameters $\boldsymbol{\theta}_2$. The model definition is completed by assigning a prior density to the hyperparameters $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$. In addition, some linear constraints of the form $\boldsymbol{Bx} = \boldsymbol{e}$, where the matrix $\boldsymbol{B}$ has rank k, may be imposted (Rue et al., 2009).

INLA provides a recipe for computing in a fast and accurate way, approximations to marginal posterior densities for the hyperparameters $\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})$ and for the latent variables $\tilde{\pi}(x_i|\boldsymbol{y})$. Such approximations are based on a smart use of

Laplace or other related analytical approximations and of numerical integration schemes. As a by-product of the main computations INLA can also compute the Deviance Information Criteria (DIC), a measure of complexity and fit useful to compare different models. The model which receives the lowest value of DIC is considered the best model, and a difference in DIC of more than 10 definitely rule out the model with the higher DIC (Spiegelhalter et al., 2002). Moreover, it is also possible to compute posterior marginals for linear combinations of the variables in the latent field.

In order for the INLA methodology to work in a fast and efficient way, latent Gaussian models have to satisfy some additional properties. First, the latent Gaussian model $\boldsymbol{x}$, often of large dimension, admits conditional independence properties, that is, it is a Gaussian Markov random field (GMRF) with a sparse precision matrix $\boldsymbol{Q}$ (Rue and Held, 2005). The efficiency of INLA relies, in fact, on efficient algorithms for sparse matrices. Secondly, because INLA needs to integrate over the hyperparameter space, the dimension of non-Gaussian $\boldsymbol{\theta}$ should not be too large, say ≤ 14. Finally, each data point y_i depends on the latent Gaussian field only through the linear predictor $\eta_i = g(\mu_i)$ where $g(\cdot)$ is a known link function and $\mu_i = E(y_i)$, i.e. $\pi(y_i|\boldsymbol{x}, \boldsymbol{\theta}) = \pi(y_i|\eta_i, \boldsymbol{\theta})$.

INLA presents several advantages over MCMC based inference: it provides accurate results in just a fraction of the time needed by smart MCMC algorithms, and it does not require convergence diagnostics. Moreover, the `R-INLA` package (available at `www.r-inla.org`) makes inference from GRMF models using the INLA methodology easy.

4 Animal Models

In this section we show that animal models are latent Gaussian Markov random field (GMRF) models which fits into the INLA framework described in Section 3. Moreover, we describe in detail the different versions of animal models we are interested in. For a more in depth introduction to animal models, see Sorensen and Gianola (2002).

In general, animal model is a generalized linear mixed model; the observed trait y_i, $i = 1, \ldots, n_d$ belongs to an exponential family

$$y_i \sim \pi(y_i; \mu_i, \boldsymbol{\theta}_2),$$

where the expected value $\mu_i = E(Y_i)$ is linked to a linear predictor η_i through a known link function $g(\cdot)$, so that $g(\mu_i) = \eta_i$. The linear predictor η_i accounts for the effects of various covariates and the breeding value in an additive way;

$$\eta_i = \beta_0 + \boldsymbol{z}_i^T \boldsymbol{\beta} + u_i + \epsilon_i, \tag{1}$$

where β_0 is an intercept, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{n_f})$ are *fixed effects*, u_i individual i's breeding value, ϵ_i it's individual effect, and $\boldsymbol{z}_i^T$ is a known incidence vector. The *fixed effects* (in a frequentist framework) accounts for group-specific effects such as e.g. sex, year of birth and locality or sub-population. In a Bayesian framework all parameters are treated as random variables, but out of convenience we refer to $\boldsymbol{\beta}$'s as fixed effects. The breeding values are genetically linked

random effects also known as additive genetic effects. The individual effects are unstructured Gaussian random effects, often named the environmental effect in quantitative genetics.

We assign a vague Gaussian prior to $\boldsymbol{\beta}$: $\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \sigma_\beta^2 \boldsymbol{I})$, where σ_β^2 is a known (large) variance and $\boldsymbol{I}$ is the identity matrix. The individual effects are $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2 \boldsymbol{I})$. The breeding values for the population, $\boldsymbol{u} = (u_1, u_2, \dots u_{n_p})$, are assumed to have a dependency structure corresponding to the pedigree

$$\boldsymbol{u} | \boldsymbol{A}, \sigma_u^2 \sim \mathcal{N}(\mathbf{0}, \sigma_u^2 \boldsymbol{A}),$$

where $\boldsymbol{A}$ is the relationship matrix and σ_u^2 is the additive genetic variance (see e.g. Lynch and Walsh, 1998; Sorensen and Gianola, 2002). The inverse of the relationship matrix, $\boldsymbol{A^{-1}}$, is a sparse matrix due to the fact that the breeding values forms a GMRF (Steinsland and Jensen, 2010). $\boldsymbol{A^{-1}}$ can be calculated from the pedigree (Quaas, 1976). Note that there might be more individuals in the pedigree than individuals with observations, $n_d \leq n_p$, and we have assumed an indexing such that u_i corresponds to y_i.

Further, to avoid identification problems we include a common intercept and constrain all factors and the breeding values to sum to zero (see Steinsland and Jensen, 2010).

The animal model as described above is a latent GMRF model where the latent field is $\boldsymbol{x} = (\{\eta_i\}, \boldsymbol{\beta}, \boldsymbol{u})$ and the hyperparameter vector $\boldsymbol{\theta}$ includes the variances (σ_u^2, σ_ϵ^2) and, possibly, the parameters in the likelihood function. The precision matrix for the latent field $\boldsymbol{x}$ is sparse because the inverse of $\boldsymbol{A}$ is sparse. Moreover, the likelihood of each data point depends on the latent field only through the linear predictor η_i. Therefore INLA can be applied to the animal model.

In our analyses we might be interested in marginal posterior for individual breeding values, u_i, fixed effects β, the additive genetic variance σ_u^2, the individual variance σ_e^2, the heritability h^2 or to evaluate the model using DIC. The heritability is loosely speaking the proportion of the variability the genes account for in a phenotypic trait. Precise definitions of heritability are given in subsequent subsections. In addition, it might be interesting to look at linear combinations of breeding values $\sum_{i \in C} w_i u_i$, where w_i are weights, for example the mean of beeding values for different cohorts.

4.1 Animal model for Gaussian data

For many continuous traits, such as the bill depth of house sparrows, it is natural to assume a Gaussian likelihood with an identity link function, $\eta_i = \mu_i$. The animal model can then be written as: $y_i \sim \mathcal{N}(\mu_i, \sigma_e^2)$, where the linear predictor is modelled as in (1) and the variance of σ_e^2 is the variance of individual effects, often referred to as environmental variance.

A Gaussian animal model can be formulated in two alternative ways, both fitting the INLA framework. Both model formulations have their numerically advantages depending on the aim of the analysis.

Model formulation 1 (MF1): Likelihood $y_i|\eta_i \sim \mathcal{N}(\eta_i, \sigma_e^2)$ and latent field $\eta_i = \beta_0 + z_i^T\beta + u_i + \epsilon_i$, where the variance of ϵ is fixed to a small value; $\sigma_\epsilon^2 = \sigma_{small}^2$.

Model formulation 2 (MF2): Likelihood $y_i|\eta_i \sim \mathcal{N}(\eta_i, \sigma_{small}^2)$, i.e. the variance of the likelihood is fixed to a small value, and latent field $\eta_i = \beta_0 + z_i^T\beta + u_i + \epsilon_i$, where the variance of ϵ is σ_e^2.

MF1 and MF2 coincide if the same priors are used for the hyper-parameters $(\beta, \sigma_e^2, \sigma_u^2)$. For MF1 $\boldsymbol{\epsilon}$ can be omitted from the model. It is included here to be consistent with MF2. In MF2 σ_{small}^2 can be interpreted as measurement uncertainty. Both formulations are latent Gaussian fields with only two non-Gaussian parameters, namely $\boldsymbol{\theta} = (\sigma_u^2, \sigma_e^2)$.

There are two situations in which we have to be cautious which model formulation we use; when finding the posterior for the heritability h^2, and evaluating models using DIC.

In general, and in the Gaussian case, the narrow sense heritability, is defined as the proportion of the phenotypic variance which is due to additive genetic variance (Lynch and Walsh, 1998)

$$h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}. \tag{2}$$

While it is easy using the INLA algorithm to compute posterior marginals for the hyperparameters, computing functions of more than one hyperparameter becomes computationally demanding. Out of convenience, we therefore use MF2, parametrized with (σ_u^2, h^2) instead of (σ_u^2, σ_e^2). Further, (σ_u^2, h^2) is given a prior such that it corresponds to the prior of (σ_u^2, σ_e^2), and hence this is a pure reparameterization.

On the other hand, DIC is based on evaluating the likelihood, and is not invariant with respect to parametrization, (Spiegelhalter et al., 2002). Using MF2, i.e. a fixed small variance for the likelihood does not work numerically; almost all models get the same DIC to the precision given by INLA. So if DIC needs to be calculated MF1 should be used.

To summarize, when u_i, $\sum_{i\in C} w_i u_i$, β or σ_u^2 is of interest both MF1 and MF2 might be used. If σ_e^2 or DIC is the aim of the analyses MF1 has to be used, while MF2 with parametrization (σ_e^2, h^2) has to be used if h^2 is of interest. Hence we might have to fit two (INLA) models to get all estimates of interest.

4.2 Animal model for Binomial data

In case of binomial data, the animal model is defined as: $y_i \sim \text{Bin}(n_i, p_i)$ $i = 1, \ldots, n_d$, where n_i is the numer of trials and p_i is the probability of success. Moreover, we assume a logit link function, so that the linear predictor is defined as: $\eta_i = \text{logit}(p_i) = \log(\frac{p_i}{1-p_i})$. The linear predictor is then modelled as in (1). In the bivariate case $n_i = 1$. Then the variance of the non-structured random effect σ_e^2 is confounded with the link, and is not identifiable (Sorensen and Gianola, 2002) because the individual effects are already accounted for through

the link and the likelihood. Therefore we omit ϵ from the linear predictor, and use

$$\eta_i = \beta_0 + \boldsymbol{z}_i^T \boldsymbol{\beta} + u_i. \tag{3}$$

For binomial data with $n_i > 1$ a non-structured random effect could be used to account for overdispersion.

For binomial data it is not immediately obvious how to define the heritability of the trait. The most common definition is derived from the idea that there exists a latent (unobserved) continuous trait called liability l_i such that we register a success if $l_i < 0$ and a failure if $l_i > 0$ (Lerner, 1950). The definition of heritability depends also on the type of the link function and in the case of the logistic function it is

$$h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\pi^2}{3}} \tag{4}$$

were $\frac{\pi^2}{3}$ is the variance of a logistic variable (see Vazquez et al., 2009). Note that the heritability on the latent scale does not correspond to the proportion of explained variance in the phenotype, e.g. the binomial data. For a discussion on heritability for non-Gaussian traits, see Dempster and Lerner. (1950); Visscher et al. (2008).

The binomial animal model is a latent Gaussian model with only one non-Gaussian hyperparameter, $\theta = \sigma_u^2$. The heritability, as defined in (4), is a function of only one random variable, σ_u^2, and can therefore easily be calculated from σ_u^2's marginal posterior distribution.

4.3 Animal model for (zero-inflated) Poisson data

Count data are often modelled as Poisson distributed: $y_i \sim \text{Poisson}(\mu_i)$ with $\mu_i = E_i \lambda_i$, where E_i is the known exposure, e.g. the lifetime, and λ_i is the intensity, e.g. the annual reproductive success. We assume an exponential link function $\eta_i = \log(\lambda_i)$, and modell the linear predictor η as in (3).

Dataset which are almost Poisson, but have too many zero-observations, often occur. Then a zero-inflated Poisson (ZIP) distribution might be useful. ZIP models are a mixture of a Poisson distribution and a distribution with point mass one at zero. There are several versions of zero-inflated Poisson, we will use $ZIP(p, \mu_i)$ defined as: $\text{Prob}(y|\dots) = p \times 1_{[y=0]} + (1-p) \times \text{Poisson}(y; \mu_i)$, where $1_{[y=0]}$ is an indicator function and $\text{Poisson}(y; \mu_i)$ indicates the Poisson likelihood with mean μ_i, and p is the proportion of extra zeros.

Poisson and zero-inflated Poisson animal models are latent Gaussian fields with hyperparameter vectors $\boldsymbol{\theta} = \sigma_u^2$ and $\boldsymbol{\theta} = (\sigma_u^2, p)$, respectively.

In the Poisson case it has been proposed that the heritability on the log scale can be defined as (Foulley et al., 1987; Matos et al., 1997; Vazquez et al., 2009)

$$h_\eta^2 = \frac{\sigma_u^2}{\sigma_u^2 + \lambda^{-1}} \tag{5}$$

where λ is the average intensity; $\lambda = \frac{1}{n_d} \sum_{i=1}^{n_d} \lambda_i = \frac{1}{n_d} \sum_{i=1}^{n_d} \exp(\eta_i)$.

The heritability (5) is then a function of one hyper-parameter and the random variable λ which is a linear combination of functions of predictors η_i. Such

a quantity is (at least currently) not possible to compute using INLA. An approximated estimate of h^2 can be computed by using a point estimate for λ together with the marginal posterior of σ_u^2. The point estimate can either be calculated directly from data, or by plugging in point estimates for the predictors $\boldsymbol{\eta}$. With this model we calculate the heritability of the intensity, e.g. annual reproductive success. If the heritability of LRS is of interest, this can be estimated by setting the exposure $E_i = 1$ (and only using individuals that are uncensored at either end of the study period).

5 Synthetic case studies

In this section we illustrate the INLA methodology using a series of syntetic case studies for the models described in Section 4. We report here results for the Gaussian and the Binomial model. For corresponding results for the Poisson model see Appendix A, Table A.1.

To make our simulated data set as realistic as possible we do the following: first, we simulate data based on the pedigree of the house sparrow dataset with $n_p = 3574$ individuals as described in Section 2. Second, we replicate in the simulated data set the same missing data structure that we find in the house sparrow data set.

Inference is done using the `AnimalINLA` package. See Appendix B for R codes. As priors for σ_u^2 and σ_e^2 we use InvGamma$(0.5, 0.5)$.

5.1 Synthetic Gaussian case study

In our first experiment we deal with Gaussian data simulated from:

$$y_i|\mu_i, \sigma_e^2 \sim \mathcal{N}(\mu_i, \sigma_e^2) \tag{6}$$

$$\eta_i = \mu_i = \beta_0 + u_i \tag{7}$$

where $\boldsymbol{u}|\boldsymbol{A}, \sigma_u^2 \sim \mathcal{N}(0, \sigma_u^2 \boldsymbol{A}^{-1})$, and $\boldsymbol{A^{-1}}$ is computed from the house sparrow pedigree.

We simulate data for $\beta_0 = 0$ and values of σ_u^2 and σ_e^2 between 0 and 1 such that $\sigma_u^2 + \sigma_e^2 = 1$. Moreover, we assume as missing all measurements that are missing for bill depth in the house sparrow data set.

We fit the model assuming a sum to zero constraint on the breeding values, $\sum u_i = 0$. Because we in this experiment are interested in estimating the variance parameters we choose the model formulation MF1 described in Section 4.1.

Table 1 shows the estimated posterior mean together with standard deviations, and the 95% credible interval (CI) for σ_u^2 and σ_e^2. The results indicate that INLA performs well, giving posterior means quite close to the true values of σ_u^2 and σ_e^2, with small standard deviations and 95% CI that contain the true value. However, for small values for σ_u^2 (less than 0.1) there seems to be some bias in the estimate of the genetic variance. This is briefly discussed in Section 7.

Table 1: Inference from INLA for synthetic Gaussian data, simulated under model 6 with different values for σ_u^2 and σ_e^2 and $\alpha = 0$. $\hat{\sigma}_u^2$ and $\hat{\sigma^2}_e$ are the posterior means with standard deviations (sd), and 95% credible intervals (CI). ΔDIC is the difference of DIC from model (6) and (7) and a model with only an intercept, specified in model (8) and (9).

$\boldsymbol{\sigma_u^2}$	$\hat{\sigma}_u^2$ (sd)	95% CI	$\boldsymbol{\sigma_e^2}$	$\hat{\sigma}_e^2$ (sd)	95% CI	ΔDIC
0	0.09 (0.03)	(0.05,0.17)	**1**	0.88 (0.05)	(0.80,0.98)	-9.770
0.05	0.12 (0.04)	(0.06,0.21)	**0.95**	0.95(0.05)	(0.85,1.05)	-2.259
0.1	0.13 (0.04)	(0.07,0.23)	**0.9**	0.90 (0.05)	(0.80,1.01)	1.316
0.15	0.14 (0.04)	(0.07,0.23)	**0.85**	0.87 (0.05)	(0.77,0.98)	6.990
0.2	0.20 (0.06)	(0.11,0.33)	**0.8**	0.85 (0.06)	(0.74,0.97)	30.083
0.3	0.29 (0.07)	(0.18,0.45)	**0.7**	0.68 (0.06)	(0.56,0.81)	92.820
0.4	0.38 (0.07)	(0.26,0.53)	**0.6**	0.63 (0.06)	(0.51,0.76)	161.147
0.5	0.49 (0.08)	(0.36,0.66)	**0.5**	0.50 (0.06)	(0.39,0.64)	274.218
0.6	0.60 (0.07)	(0.47,0.75)	**0.4**	0.39 (0.05)	(0.29,0.50)	454.308
0.7	0.71 (0.08)	(0.56,0.90)	**0.3**	0.35 (0.06)	(0.24,0.49)	552.164
0.8	0.75 (0.08)	(0.61,0.91)	**0.2**	0.23 (0.05)	(0.14,0.35)	853.660
0.9	0.88 (0.07)	(0.75,1.02)	**0.1**	0.15 (0.04)	(0.08,0.25)	1293.086
1	0.95 (0.06)	(0.84,1.07)	**0**	0.09 (0.03)	(0.05,0.16)	1639.385

For each simulated data set we fit also a model without genetic effect, hence where the model in (6) and (7) simplifies to:

$$y_i|\mu_i, \sigma_e^2 \sim \mathcal{N}(\mu_i, \sigma_e^2) \tag{8}$$

$$\eta_i. = \mu_i = \beta_0. \tag{9}$$

The aim is to check whether, using DIC, it is possible to detect when it is important to include the genetic effect in the model. Table 1 reports the difference in DIC (ΔDIC) between model (7) and model (9). The results show that when the heritability is practically zero ($h^2 = \sigma_u^2 < 0.1$) we, correctly, choose the simpler model, while for larger values of σ_u^2 the genetic component is identified as important. The posterior marginal of σ_u^2 and σ_e^2 for INLA and MCMC are compared in Appendix C, Figure C.2.

5.2 Synthetic Binomial case study

Binomial data can be challenging to analyse, especially when the the number of trials n_i is very low (Fong et al., 2010). To analyse the performance of INLA for binomal data we have carried out different simulation studies and compared the estimates obtained with INLA with those obtaind using MCMC (`MCMCglmm`, Hadfield, 2010).

We simulate data from the model $y_i|p_i \sim \text{Bin}(n_i, p_i)$ with a logit link function $\eta_i = \text{logit}(p_i) = \alpha + u_i$. Where $\boldsymbol{u}|\boldsymbol{A}, \sigma_u^2 \sim \mathcal{N}(0, \sigma_u^2 \boldsymbol{A}^{-1})$, and $\boldsymbol{A^{-1}}$ is

computed from the house sparrow pedigree. We simulate data for $\alpha = 0$ and values of σ_u^2 such that the corresponding heritability, computed as in Equation (4), varies between 0 and 1.

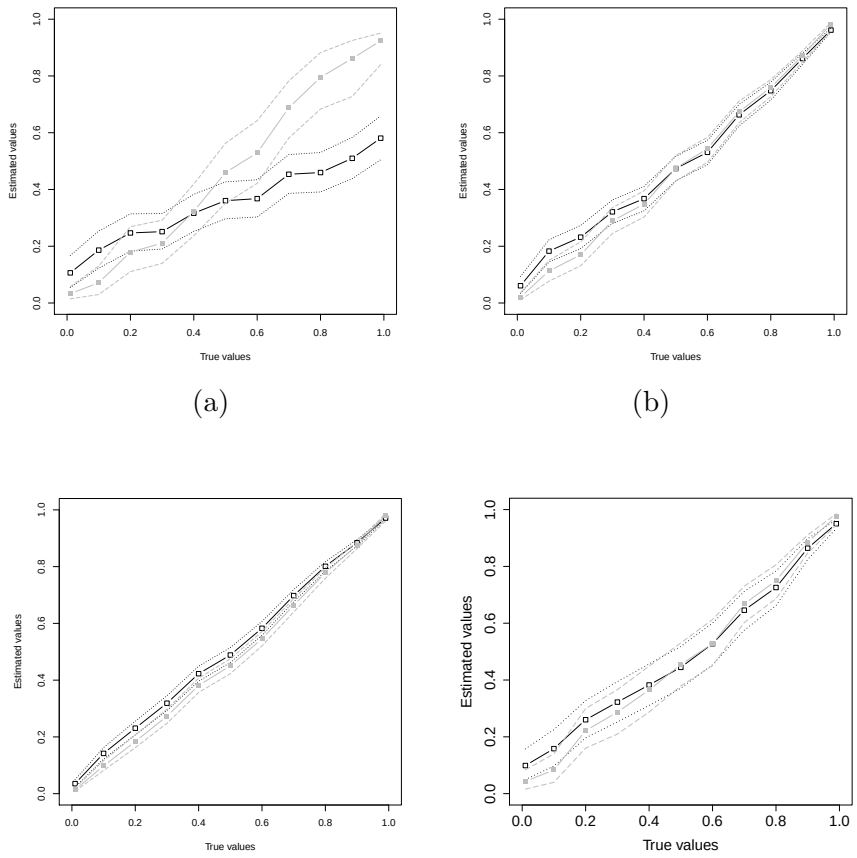
In our first experiment we let $n_i = 1$, $\forall i = 1, \ldots, n_p$, hence we have binary data for all the individuals in the pedigree. This case is, in general, particularly difficult, because with no replicates for any of the individuals the genetic variance is difficult to identify. When we look at the posterior estimate for σ_u^2, we see that the performance of INLA is quite bad (see panel a of Figure 1). The estimates for the heritability are rather biased and, in practice, it is almost impossible to distinguish bethween cases with high and low heritability of the binary trait. INLA is based on a Gaussian approximation of the log-likelihood functions which, in this case, has a very non-Gaussian behaviour. Moreover, the dependence structure induced by the house sparrow pedigree is not strong enough to allow for a correct estimation of the genetic variance.

The performance of INLA improves very fast with increasing number of trials. In our second experiment we let $n_i = 2$, $\forall i = 1, \ldots, n_p$, hence we have two trials for each individual in the pedigree. In this case, the presence of replicated measures makes it possible to estimate the genetic variance more accurately. Panel b of Figure 1 shows that the posterior means computed by INLA are very close to those computed using MCMC and close to the true value of h^2. We still see a small bias for small values of h^2 but not such that it would be problematic in a real data scenario. Even better estimates are obtained in the third experiment where the number of trials n_i changes from individual to individual in the pedigree and is randomly sampled between 1 and 9 (see panel c of Figure 1).

In the last experiment the number of trials n_i is as in the house sparrow breeding season success data set (see Section 2). Moreover, we also reproduce in the simulated data set the same missing data structure as in the real data set. In this experiment the number of trials is sampled uniformly between 1 and 9 and there are are 2392 individuals with missing data. That is, for more than 65% of the individuals in the pedigree the trait under consideration was not recorded. Results shown in Figure 1, panel d, are similar to those for the two previous cases. The estimates seem to be rather accurate with a small bias for very small values of the heritability. Moreover, results from INLA agree well with those from MCMC. In this experiment we have larger CI around the posterior mean when compared to the one in Figure 1, panel c. This is due to the presence of missing data.

6 House sparrow case studies

In this section we analyze the data introduced in Section 2 using the animal models in Section 4. To do inference we use INLA, described in Section 3. We have three case studies; bill depth, breeding season success and lifetime reproductive success (LRS). For each case study we first do model comparison using DIC to choose which fixed effects (sex, hatch year and hatch island) and random effect (additive genetic effect) to include in our model. For the best



(a) (b)

(c) (d)

Figure 1: True vs estimated heritability: posterior mean (solid line) and 95% credible intervals for INLA (black, dotted lines) and MCMC (grey, dashed lines). The number of trials is 1 in panel (a), 2 in panel (b), uniform between 1 and 9 in panel (c) and as distributed in the house sparrow data set in panel (d).

model we do further analysis according to the chosen model and the case study. This include estimating parameters, heritability and mean breeding values for each cohort. To compare INLA and MCMC, inference for some analyses for each case are done with `MCMCglmm` (Hadfield, 2010). All computation times reported are for a dual-core 2.5GHz laptop.

For all models we set $\boldsymbol{\beta} \sim N(0, \sigma^2_\beta)$, $\sigma^2_\beta = 4.5e^5$ and assume Gamma priors for $\sigma^2_u \sim \Gamma(0.5,0.5)$ and $\sigma^2_e \sim \Gamma(0.5,0.5)$ when needed. We choose the best model by starting with the full model;

$$\eta_i = \beta_0 + \beta_{\text{sex}(i)} + \beta_{\text{year}(i)} + \beta_{\text{island}(i)} + u_i, \tag{10}$$

and remove one variable at the time in a stepwise manner. In each step all nested models are examined, but only the one with lowest DIC (i.e. the best one at each step) is reported in Table 2.

6.1 Bill depth

Bill depth is a Gaussian trait and we use the animal model described in Section 4.1.

The results from the model choice procedure are presented in Table 2. We see that the full model turns out to be the best. Our further analyses for bill depth are based on this model.

We find the marginal posterior distribution for the variances; σ_u^2 has posterior mean 0.31 (sd = 0.05) and 95% credible interval (CI) (0.22,0.42). For σ_e^2 we get a posterior mean 0.46 (sd = 0.05) with 95% CI (0.39,0.56). We also calculate the marginal posterior of h^2 using MF2; mean 0.41 (sd = 0.06) with 95% CI (0.30,0.51). The posteriors for σ_u^2, σ_e^2 and h^2 are plotted in panel a and b of Figure 2. The computation time for INLA was 5 seconds for both MF1 and MF2. `MCMCglmm` gives the same estimates as `INLA`, see Appendix C, Figure C.2. For 10000 iterations `MCMCglmm` used 60 seconds (the MCMC error is still clearly visible and we would need more samples to have the same accuracy as the approximation in `INLA`).

To investigate trends in the breeding values over years we find the posterior mean breeding values for each hatch year *year* (i.e. cohort); $\sum_{i \in C_{year}} \frac{1}{n_{year}} u_i$, where n_{year} is the number of individuals with hatch year *year*, and the sum is over all these individuals. Linear combinations are easily specified in INLA, and are estimated at the same time as the other analysis.

The results shown in panel d of Figure 2 suggest that micro-evolution has occurred during the study period, with an increase in breeding values across cohorts. The mean phenotypic bill depth for each cohort also suggests a change in bill depth, but in the opposite direction (see panel c of Figure 2). We also looked at the difference between the cohorts 1993 and 2002 in the posterior mean breeding values, $\sum_{i \in C_{1993}} \frac{1}{n_{1993}} u_i$ - $\sum_{i \in C_{2002}} \frac{1}{n_{2002}} u_i$, to investigate if the difference between those cohorts was significant. We found that the difference between mean breeding values for cohorts hatching in years 1993 and 2002 was significant, with mean difference -0.090 (sd = 0.044) and 95% CI (-0.180,-0.005). The posterior marginal of the difference is given in Appendix C, Figure C.3.

Note that the estimates of linear combinations we obtain here take into account dependencies, and hence do not suffer from the same biases as when using regression on best linear unbiased predictor (BLUP) estimates obtained from REML-based analyses as discussed in Wilson et al. (2009) and Hadfield et al. (2010).

6.2 Breeding season success

Breeding season success is the number of breeding seasons that is a success, i.e. results in at least one recruit. These data are in nature binomial, and are analyzed using the animal model in Section 4.2.

Results from the model selection procedure are reported in Table 2. We find that the best model do not include linear additive genetic effects, and hence that the inherited part of breeding season success is zero or very close to zero.

However, if we use the full model to estimate σ_u^2 we get posterior mean 0.13,

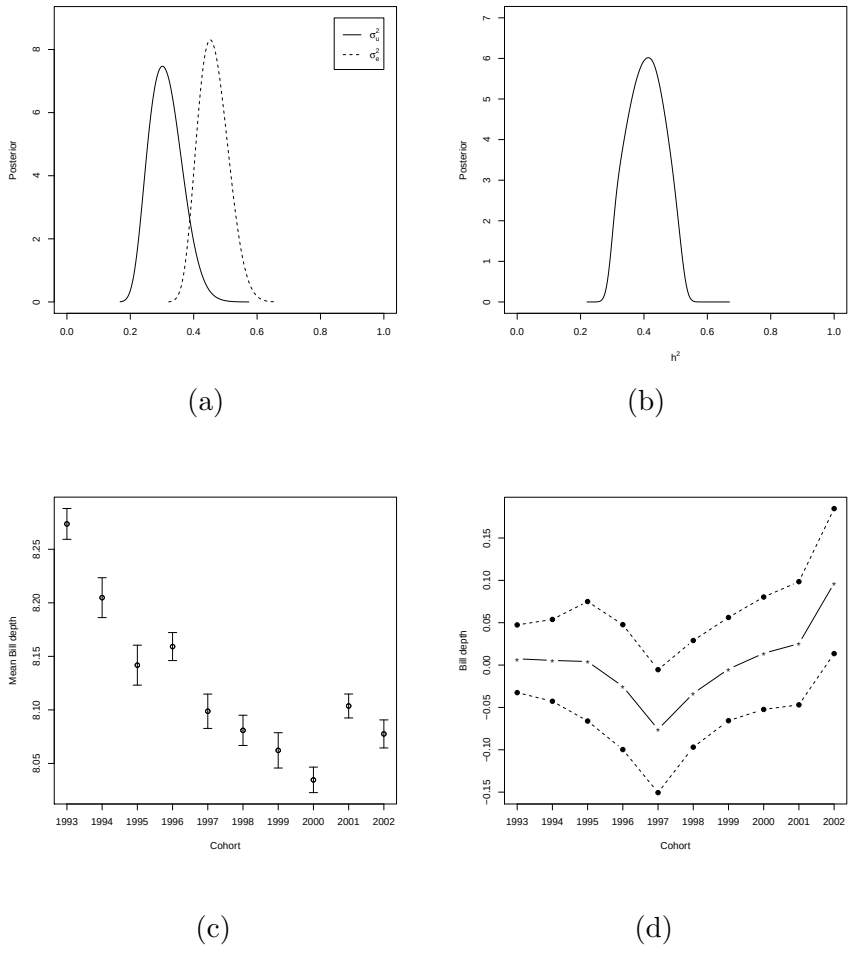


Figure 2: Parameters from animal models for the Gaussian trait bill depth in Norwegian house sparrows: posterior marginal distributions approximated by INLA. σ_u^2 (solid line) and σ_u^2 (dotted line) in panel (a), heritabililty h^2 in panel (b), mean phenotypic bill depth with 95% confidence interval for cohorts 1993-2002 in panel (c), linear combinations with posterior mean breeding values (solid line) and 95% credible interval (dashed line) for cohorts 1993-2002 in panel (d).

Table 2: Deviance information criteria (DIC) for different models explaining variance in bill depth, breeding season success and lifetime reproductive success (LRS) of Norwegian house sparrows. * indicates the best model, i.e. the model with lowest DIC value.

Bill depth∼	DIC	
sex + year + island + u	2471.359	*
year + sex + u	2471.492	
year + u	2484.178	
u	2584.277	
Breeding season success ∼	DIC	
sex + year + island + u	1718.687	
sex + year + island	1709.878	*
year + island	1710.776	
year	1713.180	
LRS ∼	DIC	
year + sex + island + u	2275.140	*
year + sex + island	2275.729	
year + sex	2283.010	
sex	2291.700	

standard deviation 0.05 and 95% credible interval (0.07,0.24). Furthermore, using (4) gives posterior heritability with mean 0.04 (sd = 0.01) and 95% CI (0.02,0.07). These estimates are similar to those from the synthetic dataset when heritability is equal or close to zero in Section 5.2 (Figure 1).

6.3 Lifetime Reproductive Success

Lifetime reproductive success for an individual is the number of recruits it produces during its lifetime. This is count data, and we analyzed this trait using the animal model in Section 4.3 with $E_i = n_i$, were n_i is the number of breeding seasons individual i has during the study period. Due to the large amount of zeros we suspect that we need a model that account for overdispersion. Therefore, we first fitted the full model with two different likelihood models; Poisson (DIC = 2421.465) and zero-inflated Poisson (DIC = 2275.140). Because zero-inflated Poisson gave lowest DIC, we proceeded with this likelihood when choosing which fixed and random effects to include in the model. Also the histogram of lifetime reproductive success divided by lifespan indicated a zero-inflated Poisson distribution (see Appendix C, Figure C.4). The model with lowest DIC is the full model, although very close to the model without additive genetic effects (Table 2). We proceed with this model in our analysis of lifetime reproductive success. Remember when modelling LRS in such a way, here controlling for lifetime, the likelihood is for LRS and the estimated heritability is actually for the annual reproductive success (intensity). Hence, the results suggests that annual reproductive success might be heritable.

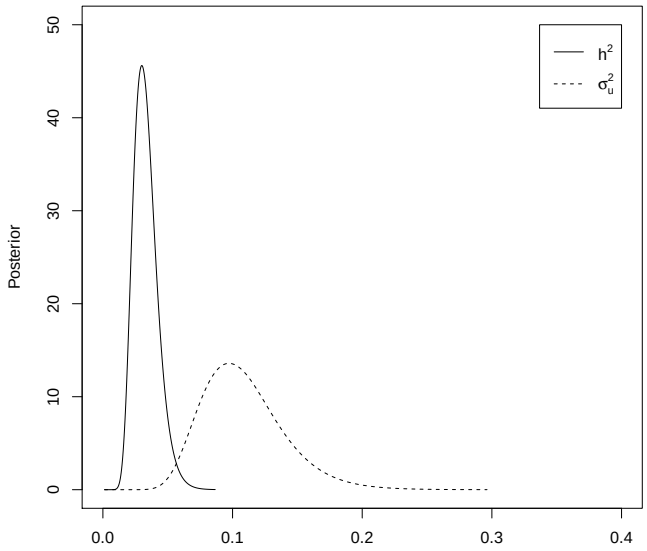


Figure 3: Posterior distribution of the heritability (h^2) and additive genetic variance (σ_u^2) of a zero-inflated Poisson distributed trait, annual reproductive success, in Norwegian house sparrows.

Accordingly, the posterior for σ_u^2 is 0.11 (sd = 0.03) with 95% CI (0.06,0.18). To obtain the posterior of h^2 defined as in (5) we plug in the point estimate $\lambda^* = \frac{\sum y_i}{\sum n_i}$. This gives a posterior mean of the heritability of 0.03 (sd = 0.01), 95 % CI (0.02,0.05). Posterior distributions for σ_u^2 and h^2 are given in Figure 3.

7 Discussion

In this section we discuss the findings of the house sparrow case study. Our study suggests that there may be an increase in mean breeding value across cohorts for bill depth (panel d of Figure 2). Accordingly, we found a significant difference between the first and last study year (1993-2002). This change indicates that a microevolutionary change has ocurred in bill depth. For the same population Ringsby et al. (2009) found evidence that bill depth is important in obtaining food (i.e. for efficiency and feeding rate). This suggests that bill depth might be under selection (see Jensen et al., 2008). When estimating linear combinations dependencies and uncertainties of breeding values have been taken into account, and we avoid the biases discussed in Wilson et al. (2009) and Hadfield et al. (2010). However, it is difficult to determine whether the observed change in breeding values is due to an evolutionary response to selection on bill depth or random genetic drift, as genetic drift may cause independent fluctuations in breeding values across generations (Hadfield et al., 2010).

The observed phenotypic mean bill depth also changed across cohorts. This change was however in the opposite direction of the change in breeding values (panel c of Figure 2). Whereas studies of other natural bird populations have found evidence for this divergence in observed and predicted evolutionary changes (Merilä et al., 2001; Postma et al., 2007), this is however not a general result, as other studies have found changes that were in the same direction (see e.g. Grant and Grant, 2002; Sheldon et al., 2003). The directions of phenotypic and genetic change may differ due to a number of reasons. For example, changes in the environment may oppose any genetic changes and thus conceal a genetic response to selection. Another possible reason is that the strength and direction of seletion fluctuate in time and space (Merilä et al., 2001). Importantly, selection can act on traits that are found to be genetically correlated with bill depth (Jensen et al., 2008), and not on bill depth itself. Former studies of these house sparrow populations indicate that both these explanations are possible (see e.g. Ringsby et al., 2002; Engen et al., 2007).

Both breeding season success and annual reproductive success are traits closely related to fitness. Fitness related traits have previously been found to be largely influenced by the environment and thus have low heritability (Merilä and Sheldon, 2000). In our study the breeding season success was not found to be heritable, and annual reproductive success had very low heritability. Consequently, our results coincide with other studies in natural populations (see Jones, 1987; Merilä and Sheldon, 2000), finding low heritability for fitness-related traits.

8 Conclusion

In this paper it is demonstrated that INLA provides a suitable methodology for doing inference for a range of animal models. In case studies we have considered models with additive genetic effects (breeding values $\boldsymbol{u}$), individual effects (environmental effects $\boldsymbol{\epsilon}$) and in addition fixed effects, all factors. These case studies required animal models with Gaussian, Binomial (with logit link), Poisson and zero-inflated Poisson (with log link) likelihoods.

Animal models might have a range of likelihoods. The R-INLA software also support different zero-inflated Gaussian and Binomial likelihoods, survival models (exponential, Weibull and Cox likelihoods), Student-T and skew-normal likelihoods (see `www.r-inla.org`). It is also straightforward to make inference with INLA for animal models extended with other additive random effects, such as maternal effects or litter effects, as well as covariates.

Furthermore, we also demonstrated that linear combinations are easily computed in INLA, and that this can give interesting insight into for instance the microevolutionary processes.

We have compared inference obtained using INLA and MCMC. The general conclusion is that INLA is a fast and accurate approximation method. However, it is less flexible than MCMC methods, and we experienced this in case study three (LRS-data, Poisson likelihood) were we were not able to calculate the heritability as defined in (5) using INLA. Though an approximated estimate

could be obtained. In the Gaussian case heritability estimates can be obtained with INLA using a tailored reparametrization.

In the synthetic case study of binary traits, we have also shown that INLA gave very biased posteriors for the additive genetic variance σ_u^2 (for the pedigree we have used). Hence, we recommend that one should not use INLA for a binary animal model unless a simulation study suggests that INLA gives correct results for the pedigree and missing data structure of the particular data set in question.

Both MCMC and INLA results showed biased estimates for small values of σ_u^2. This could be due to our choice of prior. Priors for variances are discussed in Gelman (2006). We suggest that this should be further investigated, but is outside the scope of this paper.

The R-package `AnimalINLA` has been developed for performing inference using INLA for animal models with likelihoods applied in this paper. It can be downloaded at
`www.r-inla.org`. This package includes functionality for calculating the inverse of the relationship matrix A from a pedigree. Furthermore, there are tailored functions for finding posteriors for σ_u^2, σ_ϵ^2, the heritability for Gaussian, binomial and Poisson likelihoods and linear combinations such as $\sum_{i \in C} u_i$. These functions use `R-INLA` with suitable default settings. The `R-INLA` code is also included to give a good starting point to users who wants to make modifications, e.g. other likelihoods or more random effects. Through providing easy to use software which gives results fast we hope Bayesian animal models become accessible to a wider audience of biologists and animal breeders.

Acknowledgements

The authors thank Håvard Rue for customizing the `R-INLA`, and him, Gregor Gorjanc and Andrew Finley for fruitful discussions. We also thank collegues at the Centre for Conservation Biology for help with collecting the house sparrow data. Part of the work was done while A.M.H.and I.S. visited the Statistical and Applied Mathematical Sciences Institute (SAMSI). This study was supported by Norwegian Research Council (Storforsk, Strategic University Program (SUP) in Conservation Biology, and grant no. 191847 to H.J.).

References

Blasco, A. (2001). The Bayesian controversy in animal breeding. *Journal of Animal Science* **79,** 2023–2046.

Dempfle, L. (1977). Relationship between BLUP (Best Linear Unbiased Prediction) and Bayesian method. *Annales de genetique et de selection animale* **9,** 27–32.

Dempster, E. R. and Lerner., I. M. (1950). Heritability of threshold characters. *Genetics* **35,** 212–236.

Eidsvik, J., Martino, S., and Rue, H. (2009). Approximate Bayesian inference in

Spatial Generalized Linear Mixed Models. *Scandinavian Journal of Statistics* **36,** 1–22.

Engen, S., Ringsby, T. H., Sæther, B.-E., Lande, R., Jensen, H., Lillegård, M., and Ellegren, H. (2007). Effective size of fluctuating populations with two sexes and overlapping generations. *Evolution* **61,** 1873–1885.

Fong, Y., Rue, H., and Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Biostatistics* **11,** 397–412.

Foulley, J. L., Gianola, D., and Im., S. (1987). Genetic evaluation of traits distributed as Poisson-binomial with reference to reproductive characters. *Theoretical and applied genetics* page 73:870.

Freeman, S. and Herron, J. (2004). *Evolutionary analysis.* Pearson Prentice Hall, USA.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* **1,** 515–533.

Grant, P. R. and Grant, B. R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296,** 707–711.

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R package. *Journal of Statistical Software* **33,** 1–22, http://www.jstatsoft.org/v33/i02/.

Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., and Kruuk, L. E. B. (2010). The misuse of BLUP in Ecology and Evolution. *The American Naturalist* **175,** 116–125.

Jensen, H., Steinsland, I., Ringsby, T. H., and Sæther, B.-E. (2008). Evolutionary dynamics of a sexual ornament in the house sparrow, (*Passer domesticus*), the role of indirect selection between sexes. *Evolution* **62,** 1275–1293.

Jones, J. S. (1987). The Heritability of Fitness: Bad News for 'Good Genes'? *Trends in ecology & evolution* **2,** 35–38.

Lerner, D. . (1950). Heritability of threshold characters. *Genetics* **35,** 212–236.

Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10,** 325–337.

Lynch, M. and Walsh, B. (1998). *Genetics and Analysis of Quatitative Traits.* Sinauer Associates, 5 edition.

Martino, S., Akerkar, R., and Rue, H. (2010). Approximate Bayesian Inference for Survival Models. *Scandinavian Journal of Statistics, doi: 10.1111/j.1467-9469.2010.00715.x* .

Matos, C., Thomas, D. L., Gianola, D., Perez-Enciso, M., and Young, L. D. (1997). Genetic Analysis of Discrete Reproductive Traits in Sheep Using Linear and Nonlinear Models: II. Goodness of Fit and Predictive Ability. *Journal of Animal Science* **75,** 88–94.

Merilä, J. and Sheldon, B. C. (2000). Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* **155,** 301–310.

Merilä, J., Sheldon, B. C., and Kruuk, L. E. B. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica* **112,** 199–222.

O'Hara, R. B., Cano, J. M., Ovaskainen, O., Teplitsky, C., and Alho, J. S. (2008). Bayesian approaches in evolutionary quantitative genetics. *Journal of Evolutionary Biology* **21,** 949–957.

Ovaskainen, O., Cano, J. M., and Merila, J. (2008). A Bayesian framework for comparative quantitative genetics. *Proceedings of the Royal Society B-Biological Sciences* **275,** 669–678.

Postma, E., Visser, J., and Noordwijk, A. J. V. (2007). Strong artificial selection in the wild results in predicted small evolutionary change. *Journal of Evolutionary Biology* **20,** 1823–1832.

Pärn, H., Jensen, H., Ringsby, T. H., and Sæther, B.-E. (2009). Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. *Journal of Animal Ecology* **78,** 1216–1225.

Quaas, R. (1976). Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* **32,** 949–953.

Ringsby, T. H., Berge, T., Sæther, B.-E., and Jensen, H. (2009). Reproductive success and individual variation in feeding frequency of house sparrows (*Passer domesticus*). *Journal of Ornithology* **150,** 469–481.

Ringsby, T. H., Sæther, B.-E., Tufto, J., Jensen, H., and Solberg, E. (2002). Asynchronous spatiotemporal demography of a house sparrow metapopulation in a correlated environment. *Ecology* **83,** 561–569.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*, volume 104 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian Models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B* **71,** 319–392.

Schrödle, B., Held, L., Riebler, A., and Danuser, J. (2011). Using integrated nested Laplace approximations for the evaluation of veterinary surveillance data from Switzerland: a case-study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **60,** 261–279.

Sheldon, B. C., Kruuk, L. E. B., and Merila, J. (2003). Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* **57,** 406–420.

Simm, G. (1998). *Genetic Improvement of Cattle and Sheep.* Ipswich, U.K.: Farming Press.

Sorensen, D. and Gianola, D. (2002). *Likelihood, Bayesian and MCMC Methods in Quantitative Genetics.* Springer-Verlag, New York.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society. Series B, Statistical methodology* **64,** 583–639.

Steinsland, I. and Jensen, H. (2010). Utilizing Gaussian Markov Random Field Properties of Bayesian Animal Models. *Biometrics* **66,** 763–771.

Tempelman, R. J. and Gianola, D. (1994). Assessment of a poisson animal model for embryo yield in a simulated multiple ovulation-embryo transfer scheme. *Genetics selection evolution* **26,** 263–290.

Vazquez, A. I., Gianola, D., Bates, D., Weigel, K. A., and Heringstad, B. (2009). Assessment of Poisson, logit, and linear models for genetic analysis of clinical mastitis in Norwegian Red cows. *Journal of dairy science* **92,** 739–748.

Visscher, P. M., Hill, W. G., and Wray, N. R. (2008). Heritability in the genomics era — concepts and misconception. *Nature Reviews Genetics* **9,** 255–266.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., and Nussey, D. H. (2009). An ecologist's guide to the animal model. *Journal of Animal Ecology* **79,** 13–26.

A Appendix Tables

Table A.1: Inference from INLA for synthetic Poisson data, simulated under model $y_i \mid \lambda_i \sim Pois(n_i, \lambda_i)$, $\eta_i = log(\lambda_i) = \beta_0 + u_i$ with $\beta_0 = 0$, with missing pattern as in the house sparrow Possion case study. $\hat{\sigma}_u^2$ is the posterior mean with standard deviations (sd), and 95% credible interval (CI).

$\boldsymbol{\sigma_u^2}$	$\hat{\sigma}_u^2$ (sd)	95% CI
0	0.08 (0.02)	(0.05,0.13)
0.05	0.09 (0.02)	(0.05,0.13)
0.1	0.12 (0.03)	(0.07,0.18)
0.15	0.14 (0.03)	(0.09,0.21)
0.2	0.20 (0.04)	(0.13,0.27)
0.3	0.33 (0.05)	(0.25,0.43)
0.4	0.43 (0.05)	(0.33,0.54)
0.5	0.53 (0.06)	(0.43,0.65)
0.6	0.60 (0.06)	(0.49,0.73)
0.7	0.68 (0.06)	(0.56,0.81)
0.8	0.84 (0.08)	(0.69,1.00)
0.9	0.91 (0.08)	(0.77,1.08)
1	0.99 (0.09)	(0.83,1.17)

B Appendix R codes

Simulating data with same dependency as the real pedigee, where the sparse structure matrix `Cmatrix` is obtained from $\boldsymbol{A}^{-1}$ calculated in the R package `AnimalINLA`
(`www.r-inla.org/related-projects/animalinla`).

We simulated data with different values of σ_u^2 =`var.u` and σ_e^2 =`var.e` with the function simulate.breeding.values:

```
Simulation code for breeding value:
##need the package "spam"
install.packages("spam")

inla.complete.Cmatrix <- function(C)
{
    idx = (C$i != C$j)
    return (list(i=c(C$i, C$j[idx]), j=c(C$j, C$i[idx]),
                 values=c(C$values, C$values[idx])))
}
simulate.breeding.values <- function(Cmatrix, varu, nsamples = 1)
{
    library(spam)
    prec = 1/varu
    Comp = inla.complete.Cmatrix(Cmatrix)
    S = spam(x = list(i = Comp$i, j = Comp$j, values =
                                    Comp$values))
    Q = prec * S
    breeding = rmvnorm.prec(nsamples,mu=rep(0, nrow(Q)), Q)
        breeding = as.vector(breeding)
}

##define the sparse-matrix from the relationship matrix
##computed in compute.Ainverse(), used in simulate.breeding.values()
Cmatrix = list(i= xx$Ainverse[,1],j = xx$Ainverse[,2], values =xx$Ainverse[,3])
```

Synthetic Gaussian case study (Section 5.1 in Animal models and Integrated Nested Laplace Approximations)

```
library(AnimalINLA)

##Run AnimalINLA
xx=compute.Ainverse(pedigree)

##number of individuals in the pedigree
```

```
Nbird = dim(pedigree)[1]
## choose the values of the hyperparameters
var.u = 0.6
var.e = 0.4

## simulate the breeding values and the environemental effect
breeding = simulate.breeding.values(Cmatrix, var.u)
env = rnorm(Nbird, mean = 0, sd = sqrt(var.e))

## compute the trait
trait =  breeding + env

## make the data frame
data = data.frame(y=trait,u=1:Nbird)

##Run AnimalINLA
gauss=animal.inla(response=y, genetic=c("u"),
                  Ainverse =sparseMatrix(i=xx$Ainverse[,1],
                  j=xx$Ainverse[,2],x=xx$Ainverse[,3]),
                  data=data, type.data="gaussian",
                  dic=TRUE,sigma.e=TRUE)

##hyperparameteres
gauss$summary.hyperparam
```

Synthetic Binomial case study (Section 5.2 in Animal models and Integrated Nested Laplace Approximations)

```
library(AnimalINLA)
##need the package "boot"
install.packages("boot")
library(boot)

## numbers of individuals in the pedigree
Nbird = dim(pedigree)[1]

## set the value for the hyperparameter, where beta0 is the intercept
var.u = 0.3
beta0 = 1

## set the number of trials
Ntrials = sample(1:9, 3574 , replace=T)

## simulate breeding values
breeding = simulate.breeding.values(Cmatrix, var.u)
eta =  beta0 + breeding
```

```
p = inv.logit(eta)

## simulate the trait
trait = rbinom(Nbird, Ntrials, p)

data = data.frame(y = trait,u = as.factor(1:Nbird),
                  e = as.factor(1:Nbird),
                  Ntrial = Ntrials,
                  Individual = datasetGIndividual)

##Run AnimalINLA
xx=compute.Ainverse(pedigree)

bin=animal.inla(response=y, genetic=c("u"),
                Ntrials = Ntrial,
                Ainverse =sparseMatrix(i=xx$Ainverse[,1],
                j=xx$Ainverse[,2],x=xx$Ainverse[,3]),
                data=data,type.data="binomial",
                dic=TRUE)

##hyperparameteres
bin$summary.hyperparam
```

Synthetic Poisson case study

```
library(AnimalINLA)

##number of individuals in the pedigree
Nbird = dim(pedigree)[1]

## choose the values of the hyperparameters
var.u =  0.7

##Run AnimalINLA
breeding = simulate.breeding.values(Cmatrix, var.u)

## compute the trait
eta = breeding
lambda=exp(eta)
trait=rpois(Nbird,lambda)

## make the data frame
data = data.frame(y=trait,u=1:Nbird,n=rep(1,Nbird))

##Run AnimalINLA
xx=compute.Ainverse(pedigree)
```

```
pois=animal.inla(response="y", genetic=c("u"),
                 Ainverse =sparseMatrix(i=xx$Ainverse[,1],
                 j=xx$Ainverse[,2],x=xx$Ainverse[,3]),
                 E=n,data=data,type.data="poisson",dic=TRUE)

##hyperparameters
pois$summary.hyperparam
```

C Appendix Figures

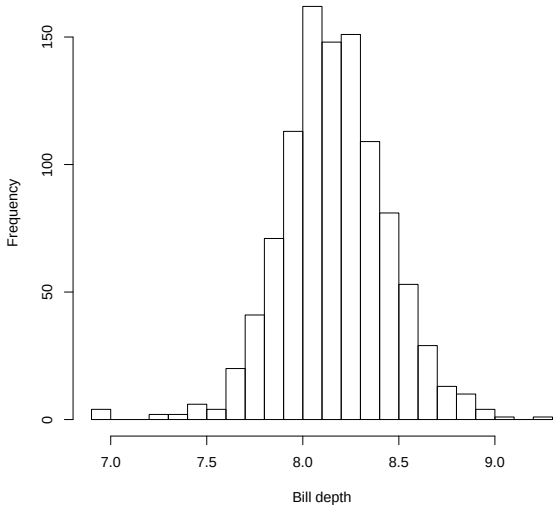
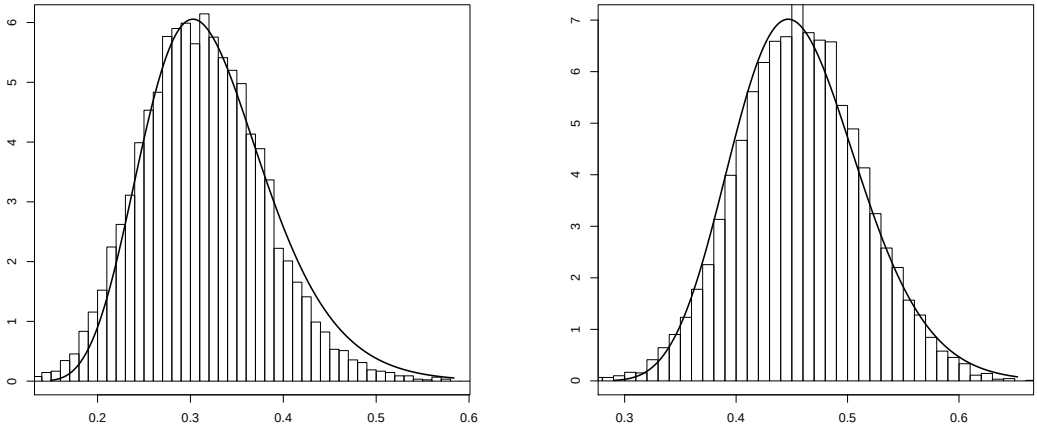


Figure C.1: Histogram showing phenotypic bill depth observations for house sparrows in northern Norway, indicating a Gaussian distribution.



(a) (b)

Figure C.2: INLA (solid line) and MCMC estimate (histogram) for the posterior marginal of σ_u^2 (panel a) and σ_e^2 (panel b) for the bill depth of house sparrows in northern Norway.

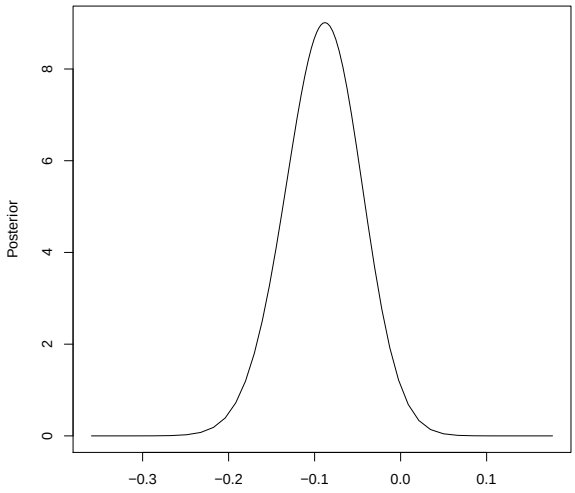


Figure C.3: Posterior of difference in mean breeding values for bill depth between cohorts 1993 and 2002 in house sparrows in northern Norway.

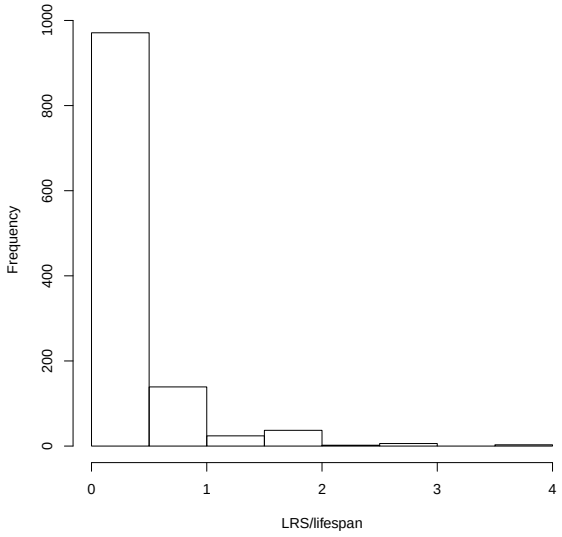


Figure C.4: Histgram showing observed lifetime reproductive sucess (LRS) relative to the lifespan (LRS/lifespan) in house sparrows in northern Norway, indicating a zero-inflated Poisson distribution.