

# Bayesian Inference of Multiple-Traits in a House Sparrow Population using the Animal Model

Ingelin Steinsland & Henrik Jensen Norwegian University of Science and Technology, Trondheim, Norway

## Why Bayesian?

• Can easily access uncertainty for all quantities of interest, e.g. for heritability and generic correlation estimates as well as response to selection.

### For the population:

 $y = X\beta + Zu + \epsilon = W\binom{u}{\beta} + \epsilon$ 

• X and Z: incidence matrices,  $W = (Z, X)^T$ .

# Results

PS: This is ongoing work, the software is still tested

### • The Gibbs sampler ran for 55000 iterations



• Uncertainty of estimated parameters are considered in the further analysis.



Hestmannøy (one of the study islands)

## Data

- Most house sparrows of six islands off the coast of Northern Norway are ringed and measuered since 1993.
- Blood samples taken  $\Rightarrow$  Genotyping of 9 microsatalites  $\Rightarrow$  Parenthood analysis  $\Rightarrow$  Pedigree
- Traits measured on adult birds; tarsus length, wing length, bill depth, bill length, body mass, total badge and visual badge.

- u: genetic effect  $u \sim N(0, \Sigma_u \otimes A)$
- A: genetic covariance matrix
  Structure of A<sup>-1</sup> from moralising the pedigree ⇒ A<sup>-1</sup> sparse.
- $\epsilon$ : environmental effect  $\epsilon \sim N(0, \Sigma_{\epsilon} \otimes I)$

### Constraints:

• For all but one fixed effect (f =hatch year and sex), for all traits t:

$$\sum_{l=1}^{Nlevel} \beta_{flt} = 0$$

• For the breeding values, for all traits:

$$\sum_{i=1}^{N} u_{it} = 0$$







- Missing data: Females do not have a badge + some other missing trait values.
- There are 3572 birds in the pedigree, of these 1004 have (some) adult traits measurements.
- For (almost) all birds with data *sex*, *hatch year* and *hatch island* are known.

# The Animal Model For bird i (i = 1, ..., 1004):

#### Of interest:

Genetic and environmental covariance matrices (Σ<sub>u</sub> and Σ<sub>ε</sub>), fixed effects (β) and/or individual breeding values (u<sub>i</sub>).
Heritability; h<sup>2</sup><sub>j</sub> = σ<sup>2</sup><sub>uj</sub>/σ<sup>2</sup><sub>ui</sub>+σ<sup>2</sup><sub>ci</sub>.

## Full conditional distributions:

β, u, y<sub>miss</sub> | y<sub>obs</sub>, Σ<sub>u</sub>, Σ<sub>ε</sub> ~ MGMRF, a multivariate Gaussian Markov Random Field with constraints.
Σ<sub>u</sub>, Σ<sub>ε</sub> | y, β, u ~ Inverted Wishart (if no constraints)

## Gibbs sampler

yellow) and environmental correlation ( $\Sigma_{\epsilon}$ , red).

- All environmental correlation estimates are positive.
- Genetic correlation estimates for traits with low heritability (body mass) and/or many missing data (badges) are uncertain.

# Summary

- Animal model of multiple traits = MGMRF model.
- Use a Gibbs sampler with only two blocks.
- Able to find uncertainty in heritability and genetic correlation
- Can handle constraints, e.g. in "fixed effects".

observed traits =fixed +genetic +environmental  $y_i = \beta_i + u_i + \epsilon_i$ 

y<sub>i</sub>: observations (traits), y<sub>i</sub> = (y<sub>tarsus</sub>, y<sub>wing</sub>,...)<sub>i</sub>
β<sub>i</sub>: "fixed effects" (sex, hatch year and hatch island), β ~ N(0, σ<sup>2</sup><sub>β</sub>I)

•  $u_i$ : genetic effects,  $u_i \sim N(0, \Sigma_u)$ 

•  $\epsilon_i$ : Environmental effects:  $\epsilon_i \sim N(0, \Sigma_{\epsilon})$ 

• Conjugate prior for the genetic and environmental covariance matrices  $\Sigma_u$  and  $\Sigma_{\epsilon}$  (inverted Wishart).

Use a Gibbs sampler with two blocks for making inference.

## Algorithm

• For each iteration

 Sample (β, u, y<sub>miss</sub>) of dimension 25000!
 Sample (Σ<sub>u</sub>, Σ<sub>ϵ</sub>) (Metropolis-Hastings step because of constraints)

Step 1 is the challenge:

• Sparse  $A^{-1} \Rightarrow$  cheaper calculations. • Complexity  $< \mathcal{O}(n^{1.5})$ , use GMRFLib.

### Further work

• Extend methodology to make inference about response to selection with uncertainty.

