



Random effect modelling of great tit nesting behaviour

William Browne (University of Bristol)

In collaboration with

Richard Pettifor (Institute of Zoology, London),

Robin McCleery and Ben Sheldon

(University of Oxford)

And Dylan Childs (University of Sheffield).

Talk Summary

- Background.
- Wytham woods great tit dataset.
- Partitioning variability in responses.
- Partitioning correlation between responses.
- Prior sensitivity.
- Including individual nestling responses.

Statistical Analysis of Bird Ecology Datasets

- Much recent work done on estimating population parameters, for example estimating population size, survival rates.
- Much statistical work carried out at Kent, Cambridge and St Andrews.
- Data from census data and ring-recovery data.
- Less statistical work on estimating relationships at the individual bird level, for example: Why do birds laying earlier lay larger clutches?
- Difficulties in taking measurements in observational studies, in particular getting measurements over time due to short lifespan of birds.

Wytham woods great tit dataset

- A longitudinal study of great tits nesting in Wytham Woods, Oxfordshire.
- 6 responses : 3 continuous & 3 binary.
- *Clutch size, lay date and mean nestling mass.*
- *Nest success, male and female survival.*
- Data: 4165 nesting attempts over a period of 34 years.
- There are 4 higher-level classifications of the data: female parent, male parent, nestbox and year.



Data background

The data structure can be summarised as follows:

Source	Number of IDs	Median #obs	Mean #obs
Year	34	104	122.5
Nestbox	968	4	4.30
Male parent	2986	1	1.39
Female parent	2944	1	1.41

Note there is very little information on each individual male and female bird but we can get some estimates of variability via a random effects model.

Univariate cross-classified random effect modelling

- For each of the 6 responses we will firstly fit a univariate model, normal responses for the continuous variables and probit regression for the binary variables. For example using notation of Browne et al. (2001) and letting response y be clutch size:

$$y_i = \beta + u_{male(i)}^{(2)} + u_{female(i)}^{(3)} + u_{nestbox(i)}^{(4)} + u_{year(i)}^{(5)} + e_i$$
$$u_{male(i)}^{(2)} \sim N(0, \sigma_{u(2)}^2), \quad u_{female(i)}^{(3)} \sim N(0, \sigma_{u(3)}^2),$$
$$u_{nestbox(i)}^{(4)} \sim N(0, \sigma_{u(4)}^2), \quad u_{year(i)}^{(5)} \sim N(0, \sigma_{u(5)}^2),$$
$$e_i \sim N(0, \sigma_e^2)$$

Estimation

- We use MCMC estimation in MLwiN and choose 'diffuse' priors for all parameters.
- We run 3 chains from different starting points for 250k iterations each (500k for binary responses) and use Gelman-Rubin diagnostic to decide burn-in length.
- We compared results with the equivalent classical models in Genstat and got broadly the same results. in Genstat.
- We fit all four higher classifications and do not consider model comparison.

Clutch Size

Parameter	Estimate (S.E.)	Percentage variance
β	8.808 (0.109)	-
$\sigma_{u(5)}^2$ (Year)	0.365 (0.101)	14.3%
$\sigma_{u(4)}^2$ (Nest box)	0.107 (0.026)	4.2%
$\sigma_{u(3)}^2$ (Male)	0.046 (0.043)	1.8%
$\sigma_{u(2)}^2$ (Female)	0.974 (0.062)	38.1%
σ_e^2 (Observation)	1.064 (0.055)	41.6%

Here we see that the average clutch size is just below 9 eggs with large variability between female birds and some variability between years. Male birds and nest boxes have less impact.

Lay Date (days after April 1st)

Parameter	Estimate (S.E.)	Percentage variance
β	29.38 (1.07)	-
$\sigma_{u(5)}^2$ (Year)	37.74 (10.08)	50.3%
$\sigma_{u(4)}^2$ (Nest box)	3.38 (0.56)	4.5%
$\sigma_{u(3)}^2$ (Male)	0.22 (0.39)	0.3%
$\sigma_{u(2)}^2$ (Female)	8.55 (1.03)	11.4%
σ_e^2 (Observation)	25.10 (1.04)	33.5%

Here we see that the mean lay date is around the end of April/beginning of May. The biggest driver of lay date is the year which is probably indicating weather differences. There is some variability due to female birds but little impact of nest box and male bird.

Nestling Mass

Parameter	Estimate (S.E.)	Percentage variance
β	18.829 (0.060)	-
$\sigma_{u(5)}^2$ (Year)	0.105 (0.032)	9.0%
$\sigma_{u(4)}^2$ (Nest box)	0.026 (0.013)	2.2%
$\sigma_{u(3)}^2$ (Male)	0.153 (0.030)	13.1%
$\sigma_{u(2)}^2$ (Female)	0.163 (0.031)	14.0%
σ_e^2 (Observation)	0.720 (0.035)	61.7%

Here the response is the average mass of the chicks in a brood. Note here lots of the variability is unexplained and both parents are equally important.

Nest Success

Parameter	Estimate (S.E.)	Percentage overdispersion
β	0.010 (0.080)	-
$\sigma_{u(5)}^2$ (Year)	0.191 (0.058)	56.0%
$\sigma_{u(4)}^2$ (Nest box)	0.025 (0.020)	7.3%
$\sigma_{u(3)}^2$ (Male)	0.065 (0.054)	19.1%
$\sigma_{u(2)}^2$ (Female)	0.060 (0.052)	17.6%

Here we define nest success as one of the ringed nestlings captured in later years. The value 0.01 for β corresponds to around a 50% success rate. Most of the variability is explained by the Binomial assumption with the bulk of the over-dispersion mainly due to yearly differences.

Male Survival

Parameter	Estimate (S.E.)	Percentage overdispersion
β	-0.428 (0.041)	-
$\sigma_{u(5)}^2$ (Year)	0.032 (0.013)	41.6%
$\sigma_{u(4)}^2$ (Nest box)	0.006 (0.006)	7.8%
$\sigma_{u(3)}^2$ (Male)	0.025 (0.023)	32.5%
$\sigma_{u(2)}^2$ (Female)	0.014 (0.017)	18.2%

Here male survival is defined as being observed breeding in later years. The average probability is 0.334 and there is very little over-dispersion with differences between years being the main factor. Note the actual response is being observed breeding in later years and so the real probability is higher!

Female survival

Parameter	Estimate (S.E.)	Percentage overdispersion
β	-0.302 (0.048)	-
$\sigma_{u(5)}^2$ (Year)	0.053 (0.018)	36.6%
$\sigma_{u(4)}^2$ (Nest box)	0.065 (0.024)	44.8%
$\sigma_{u(3)}^2$ (Male)	0.014 (0.017)	9.7%
$\sigma_{u(2)}^2$ (Female)	0.013 (0.014)	9.0%

Here female survival is defined as being observed breeding in later years. The average probability is 0.381 and again there isn't much over-dispersion with differences between nestboxes and years being the main factors.

Multivariate modelling of the great tit dataset

- We now wish to combine the six univariate models into one big model that will also account for the correlations between the responses.
- We choose a MV Normal model and use latent variables (Chib and Greenburg, 1998) for the 3 binary responses that take positive values if the response is 1 and negative values if the response is 0.
- We are then left with a 6-vector for each observation consisting of the 3 continuous responses and 3 latent variables. The latent variables are estimated as an additional step in the MCMC algorithm and for identifiability the elements of the level 1 variance matrix that correspond to their variances are constrained to equal 1.

Multivariate Model

$$y_i = \beta + u_{male(i)}^{(2)} + u_{female(i)}^{(3)} + u_{nestbox(i)}^{(4)} + u_{year(i)}^{(5)} + e_i$$

$$u_{male(i)}^{(2)} \sim MVN(0, \Omega_{u(2)}), \quad u_{female(i)}^{(3)} \sim MVN(0, \Omega_{u(3)}),$$

$$u_{nestbox(i)}^{(4)} \sim MVN(0, \Omega_{u(4)}), \quad u_{year(i)}^{(5)} \sim MVN(0, \Omega_{u(5)}),$$

$$e_i \sim MVN(0, \Omega_e)$$

Here the vector valued response is decomposed into a mean vector plus random effects for each classification.

$$p(\Omega_u^{(i)}) \sim invWishart_6[6, 6 * S_u^{(i)}],$$

$$S_u^{(i)} = \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05 \end{bmatrix}$$

Inverse Wishart priors are used for each of the classification variance matrices. The values are based on considering overall variability in each response and assuming an equal split for the 5 classifications.

Use of the multivariate model

- The multivariate model was fitted using an MCMC algorithm programmed into the MLwiN package which consists of Gibbs sampling steps for all parameters apart from the level 1 variance matrix which requires Metropolis sampling (see Browne 2006).
- The multivariate model will give variance estimates in line with the 6 univariate models.
- In addition the covariances/correlations at each level can be assessed to look at how correlations are partitioned.

Correlations in full model

	CS	LD	NM	NS	MS
LD	N, F, O -0.30	X	X	X	X
NM	F, O -0.09	F, O -0.06	X	X	X
NS	Y, F 0.20	N, F, O -0.22	O 0.16	X	X
MS	- 0.02	- -0.02	- 0.04	Y 0.07	X
FS	F, O -0.02	F, O -0.02	- 0.06	Y, F 0.11	Y, O 0.21

Note numbers give correlations in a 1 level model.

Key: Blue +ve, Red -ve: Y – year, N – nestbox, F – female, O - observation

Pairs of antagonistic covariances at different classifications

There are 3 pairs of antagonistic correlations i.e. correlations with different signs at different classifications:

CS & FS : Female 0.48 Observation -0.20

Interpretation: Birds that lay larger clutches are more likely to survive but a particular bird has less chance of surviving if it lays more eggs.

LD & FS : Female -0.67 Observation 0.11

Interpretation: Birds that lay early are more likely to survive but for a particular bird the later they lay the better!

LD & NM : Female 0.20 Observation -0.19

Interpretation: Females who generally lay late, lay heavier chicks but the later a particular bird lays the lighter its chicks.

Prior Sensitivity

Our choice of variance prior assumes a priori

- No correlation between the 6 traits.
- Variance for each trait is split equally between the 5 classifications.

We compared this approach with a more Bayesian approach by splitting the data into 2 halves:

In the first 17 years (1964-1980) there were 1,116 observations whilst in the second 17 years (1981-1997) there were 3,049 observations

We therefore used estimates from the first 17 years of the data to give a prior for the second 17 years and compared this prior with our earlier prior.

Both approaches give similar results.

MCMC efficiency for clutch size response

- The MCMC algorithm used in the univariate analysis of clutch size was a simple 10-step Gibbs sampling algorithm.
- The same Gibbs sampling algorithm can be used in both the MLwiN and WinBUGS software packages and we ran both for 50,000 iterations.
- To compare methods for each parameter we can look at the effective sample sizes (ESS) which give an estimate of how many 'independent samples we have' for each parameter as opposed to 50,000 dependent samples.
- $ESS = \# \text{ of iterations} / \kappa$, $\kappa = 1 + 2 \sum_{k=1}^{\infty} \rho(k)$

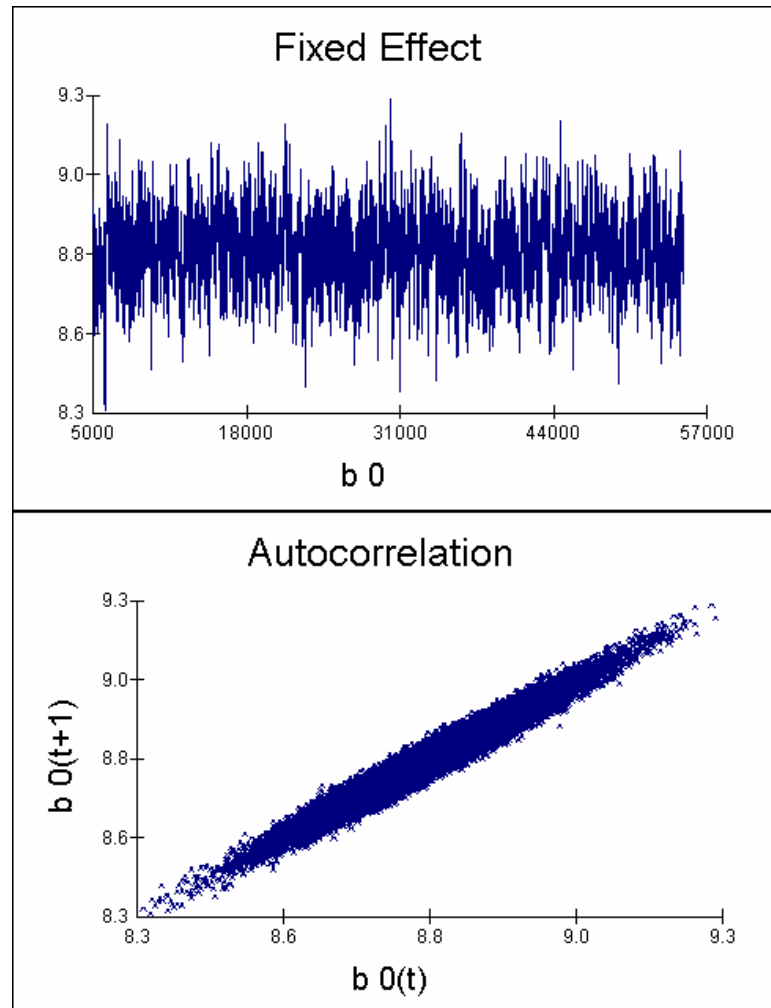
Effective Sample sizes

The effective sample sizes are similar for both packages.
Note that MLwiN is 5 times quicker!!

Parameter	MLwiN	WinBUGS
Fixed Effect	671	602
Year	30632	29604
Nestbox	833	788
Male	36	33
Female	3098	3685
Observation	110	135
Time	519s	2601s

We will now consider methods that will improve the ESS values for particular parameters. We will firstly consider the fixed effect parameter.

Trace and autocorrelation plots for fixed effect using standard Gibbs sampling algorithm



Hierarchical Centering

This method was devised by Gelfand et al. (1995) for use in nested models. Basically (where feasible) parameters are moved up the hierarchy in a model reformulation. For example:

$$y_{ij} = \beta_0 + u_j + e_{ij}, u_j \sim N(0, \sigma_u^2), e_{ij} \sim N(0, \sigma_e^2)$$

is equivalent to

$$y_{ij} = \beta_j + e_{ij}, \beta_j \sim N(\beta_0, \sigma_u^2), e_{ij} \sim N(0, \sigma_e^2)$$

The motivation here is we remove the strong negative correlation between the fixed and random effects by reformulation.

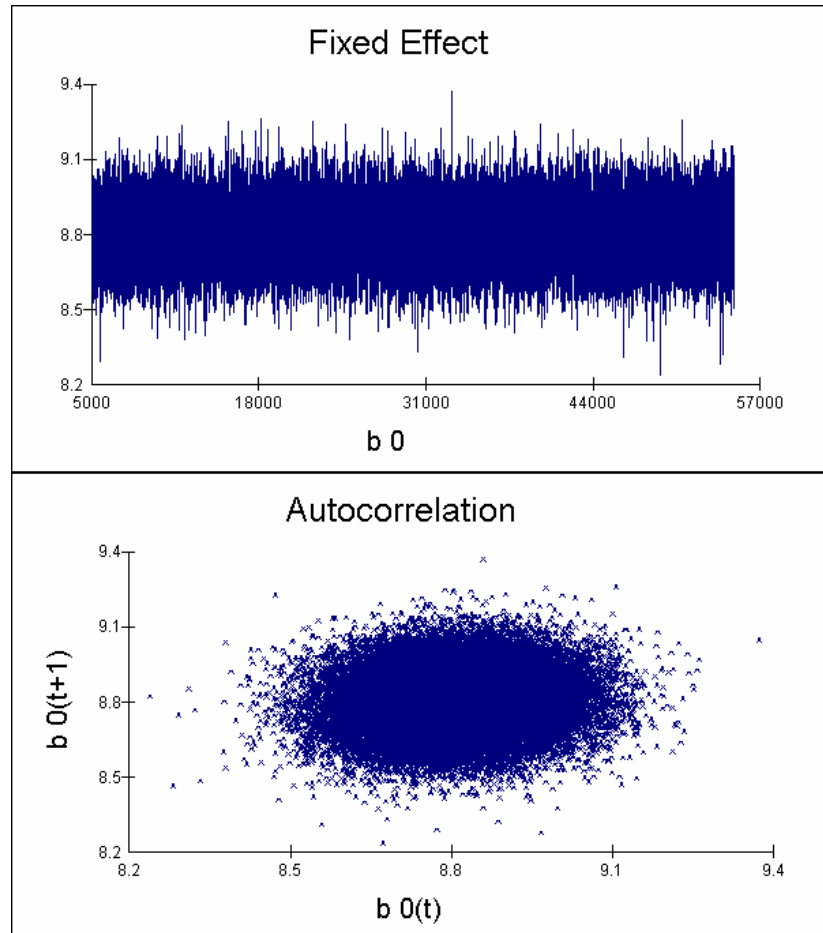
Hierarchical Centering

In our cross-classified model we have 4 possible hierarchies up which we can move parameters. We have chosen to move the fixed effect up the year hierarchy as it's variance had biggest ESS although this choice is rather arbitrary.

$$\begin{aligned}y_i &= \mathbf{u}_{female(i)}^{(2)} + \mathbf{u}_{male(i)}^{(3)} + \mathbf{u}_{nestbox(i)}^{(4)} + \beta_{year(i)}^{(5)} + e_i, \\ \mathbf{u}_{female(i)}^{(2)} &\sim N(0, \sigma_{u(2)}^2), \mathbf{u}_{male(i)}^{(3)} \sim N(0, \sigma_{u(3)}^2), \\ \mathbf{u}_{nestbox(i)}^{(4)} &\sim N(0, \sigma_{u(4)}^2), \beta_{year(i)}^{(5)} \sim N(\beta_0, \sigma_{u(5)}^2), e_i \sim N(0, \sigma_e^2), \\ \beta_0 &\propto 1, \sigma_{u(2)}^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon), \sigma_{u(3)}^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon), \\ \sigma_{u(4)}^2 &\sim \Gamma^{-1}(\varepsilon, \varepsilon), \sigma_{u(5)}^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon), \sigma_e^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon).\end{aligned}$$

The ESS for the fixed effect increases 50-fold from 602 to 35,063 while for the year level variance we have a smaller improvement from 29,604 to 34,626. Note this formulation also runs faster 1864s vs 2601s (in WinBUGS).

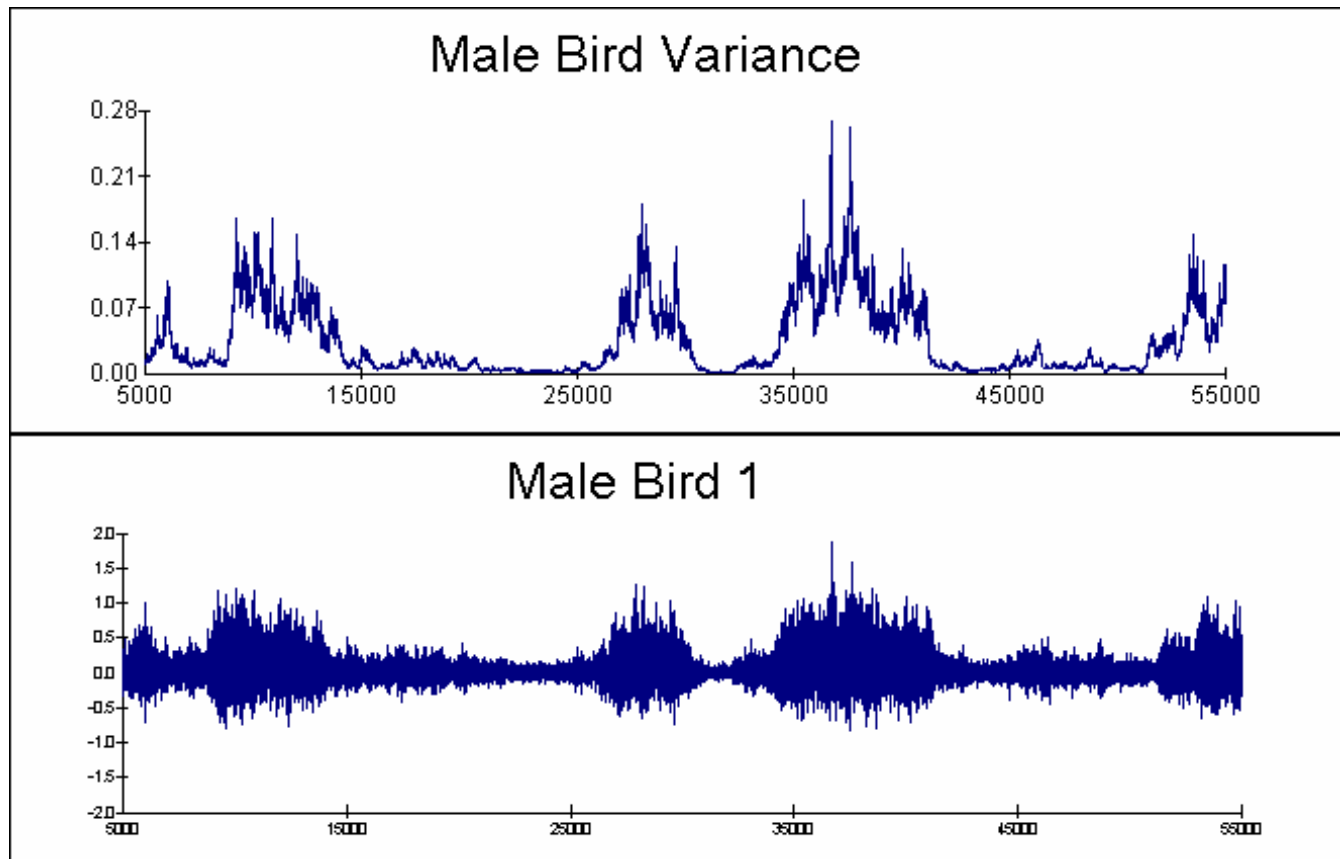
Trace and autocorrelation plots for fixed effect using hierarchical centering formulation



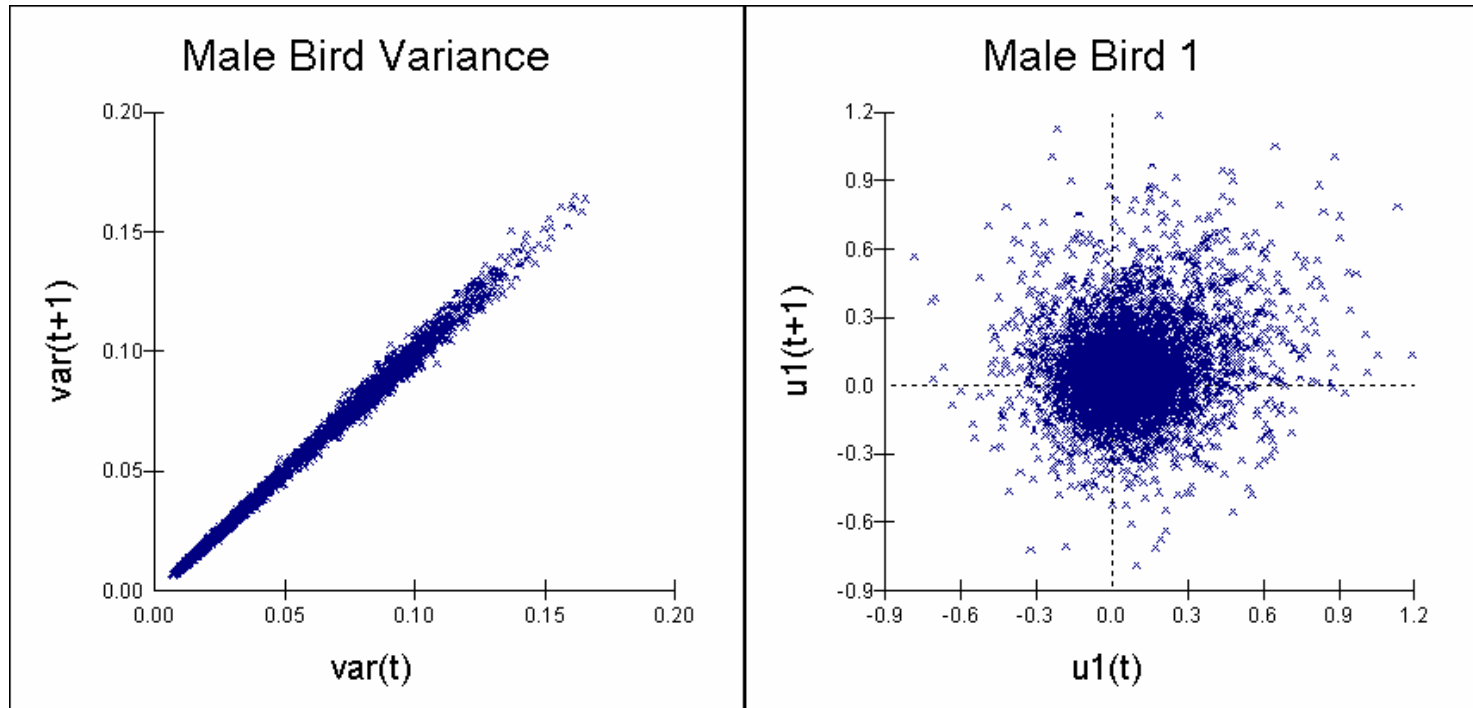
Parameter Expansion

- We next consider the variances and in particular the between-male bird variance.
- When the posterior distribution of a variance parameter has some mass near zero this can hamper the mixing of the chains for both the variance parameter and the associated random effects.
- The pictures over the page illustrate such poor mixing.
- One solution is parameter expansion (Liu et al. 1998).
- In this method we add an extra parameter to the model to improve mixing.

Trace plots for between males variance and a sample male effect using standard Gibbs sampling algorithm



Autocorrelation plot for male variance and a sample male effect using standard Gibbs sampling algorithm



Parameter Expansion

In our example we use parameter expansion for all 4 hierarchies. Note the α parameters have an impact on both the random effects and their variance.

$$y_i = \beta_0 + \alpha_2 v_{female(i)}^{(2)} + \alpha_3 v_{male(i)}^{(3)} + \alpha_4 v_{nestbox(i)}^{(4)} + \alpha_5 v_{year(i)}^{(5)} + e_i,$$
$$v_{female(i)}^{(2)} \sim N(0, \sigma_{v(2)}^2), v_{male(i)}^{(3)} \sim N(0, \sigma_{v(3)}^2),$$
$$v_{nestbox(i)}^{(4)} \sim N(0, \sigma_{v(4)}^2), v_{year(i)}^{(5)} \sim N(0, \sigma_{v(5)}^2), e_i \sim N(0, \sigma_e^2),$$
$$\beta_0 \propto 1, \alpha_k \propto 1, \sigma_{v(k)}^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon), k = 2, \dots, 5, \sigma_e^2 \sim \Gamma^{-1}(\varepsilon, \varepsilon)$$

The original parameters can be found by:

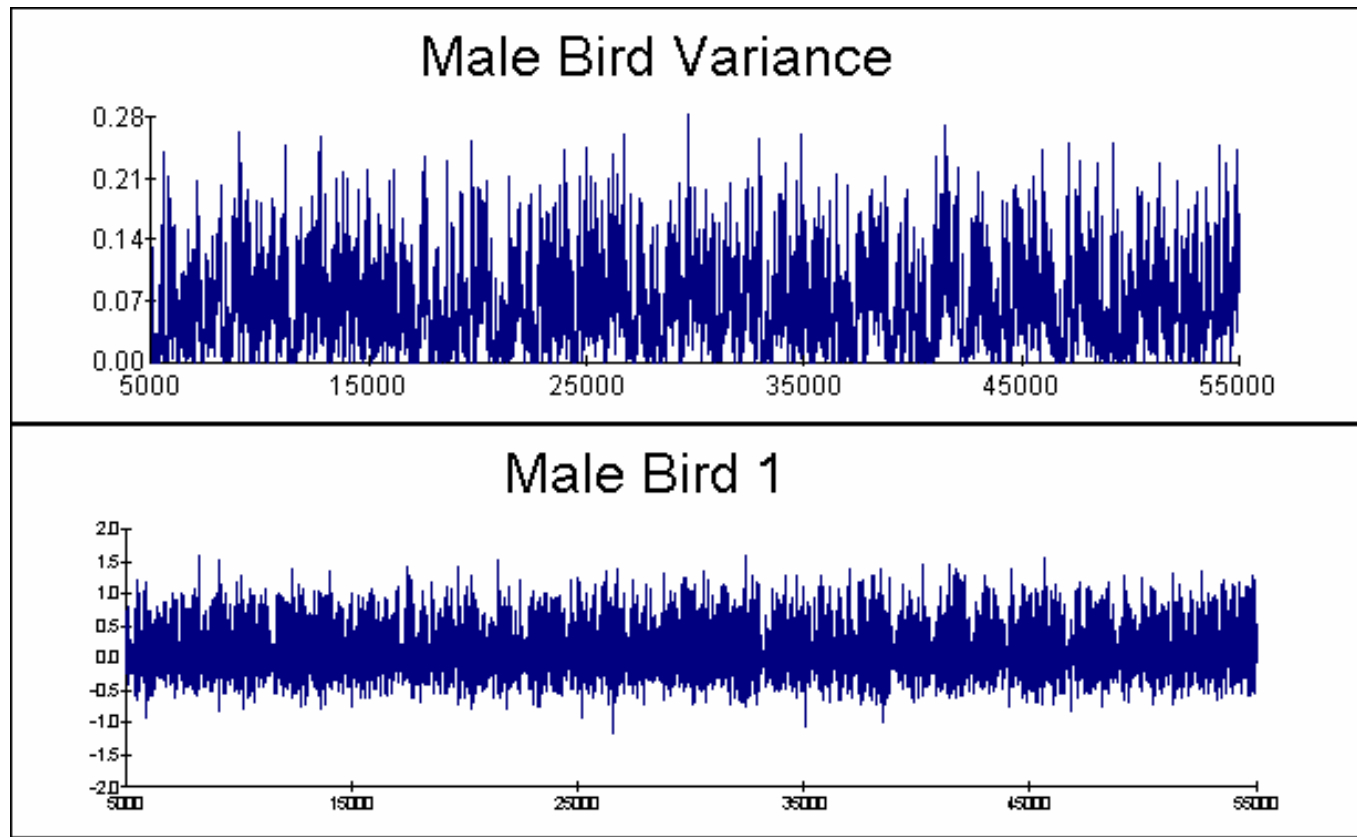
$$u_i^{(k)} = \alpha_k v_i^{(k)} \text{ and } \sigma_{u(k)}^2 = \alpha_k^2 \sigma_{v(k)}^2$$

Note the models are not identical as we now have different prior distributions for the variances.

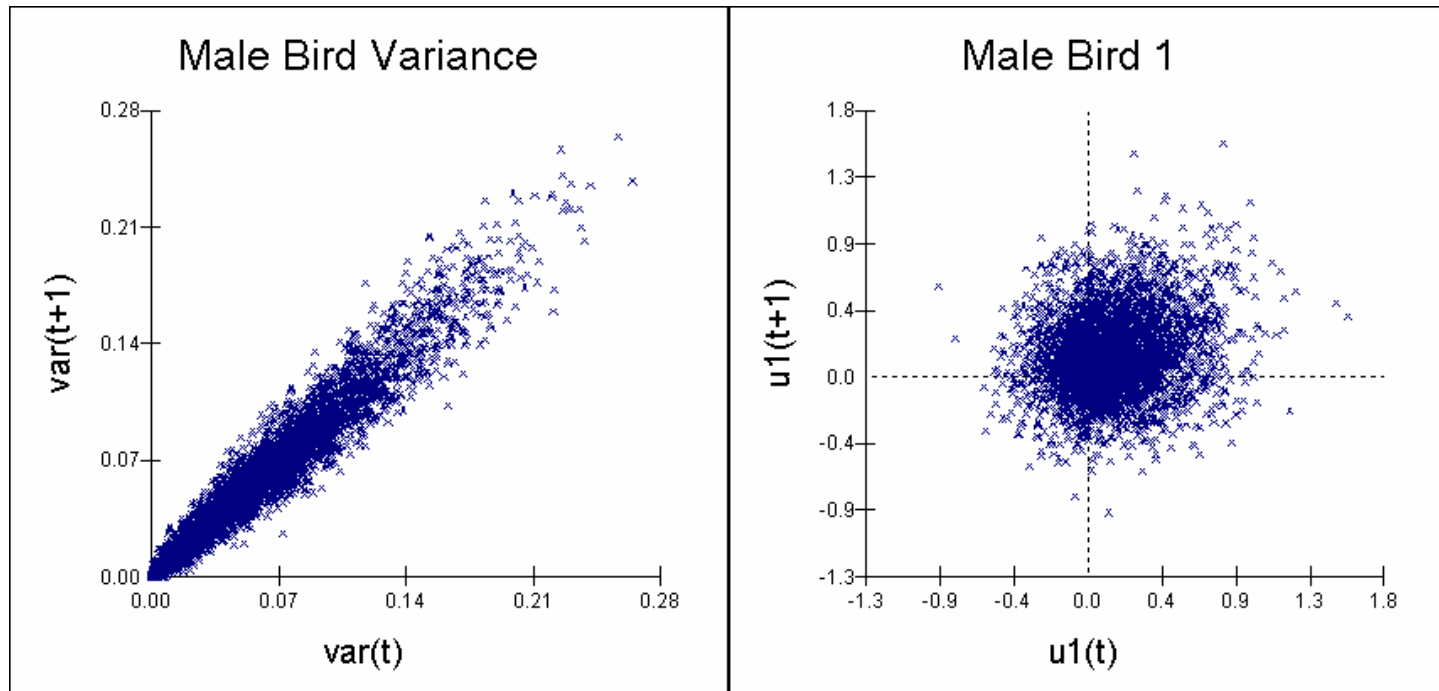
Parameter Expansion

- For the between males variance we have a 20-fold increase in ESS from 33 to 600.
- The parameter expanded model has different prior distributions for the variances although these priors are still ‘diffuse’.
- It should be noted that the point and interval estimate of the level 2 variance has changed from 0.034 (0.002,0.126) to 0.064 (0.000,0.172).
- Parameter expansion is computationally slower 3662s vs 2601s for our example.

Trace plots for between males variance and a sample male effect using parameter expansion.



Autocorrelation plot for male variance and a sample male effect using parameter expansion.



Combining the two methods

Hierarchical centering and parameter expansion can easily be combined in the same model. We performed centering on the year classification and parameter expansion on the other 3 hierarchies and got the following.

Parameter	WinBUGS originally	WinBUGS combined
Fixed Effect	602	34296
Year	29604	34817
Nestbox	788	5170
Male	33	557
Female	3685	8580
Observation	135	1431
Time	2601s	2526s

Including responses on individual chick survival

- In our earlier modelling we consider a response 'nest success' that indicates that one (or more) chicks survive to breed. This fitted nicely into the normal/probit binomial modelling framework.
- We would prefer to distinguish between nesting attempts where just one bird survives and those where lots survive.
- We can easily fit separate models for each response and include higher level correlations. To capture the lower level correlations we can instead fit responses as predictors.

Including responses on individual chick survival

- Here we consider a different subset of the Wytham Woods dataset (4864 attempts over 28 years (1976-2003) and 3,338 female birds). Note nestboxes replaced in mid 1970s with concrete boxes.
- For illustration we consider the 2 classifications: year and female, and three responses: clutch size, female survival and nestling survival. Note these are minimal sets for usage in an interesting population model.
- We will use WinBUGS to fit this model as currently MLwiN cannot fit general models with responses of different types.
- We also use the logit link instead of the probit (due to problems in WinBUGS with the probit).

Model for three responses

The following model was fitted (with the addition of non-informative priors as previously). Note that the two fixed effects, β_3 and β_4 capture the level 1 correlations (but we do not model correlation between FS & NS at level 1).

$$CS_i \sim \text{Normal}(\mu_i^{CS}, \sigma_e^2)$$

$$FS_i \sim \text{Bernoulli}(p_i^{FS})$$

$$NS \sim \text{Binomial}(CS_i, p_i^{NS})$$

$$\mu_i^{CS} = \beta_0 + u_{female(i),1}^{(2)} + u_{year(i),1}^{(3)}$$

$$\text{logit}(p_i^{FS}) = \beta_1 + \beta_3 CS_i + u_{female(i),2}^{(2)} + u_{year(i),2}^{(3)}$$

$$\text{logit}(p_i^{NS}) = \beta_2 + \beta_4 CS_i + u_{female(i),3}^{(2)} + u_{year(i),3}^{(3)}$$

$$u_{female(i)}^{(2)} \sim N(0, \Omega_{u^{(2)}}), \quad u_{year(i)}^{(3)} \sim N(0, \Omega_{u^{(3)}}),$$

Results – fixed effects

Below we fill in the fixed effects estimates for the three responses. We see clearly the ‘penalty’ terms for increasing clutch size, in particular on female survival.

$$\mu_i^{CS} = 8.733 + u_{female(i),1}^{(2)} + u_{year(i),1}^{(3)}$$

$$\text{logit}(p_i^{FS}) = 1.524 - 0.231CS_i + u_{female(i),2}^{(2)} + u_{year(i),2}^{(3)}$$

$$\text{logit}(p_i^{NS}) = -1.929 - 0.051CS_i + u_{female(i),3}^{(2)} + u_{year(i),3}^{(3)}$$

We can express the survival terms as probabilities as follows:

CS	2	7.733	8.733	9.733	15
FS	0.743	0.434	0.379	0.326	0.125
NS	0.116	0.089	0.085	0.081	0.063

Results - Covariances

Below we give the between female and between year covariance matrices with responses in the order CS, FS & NS:

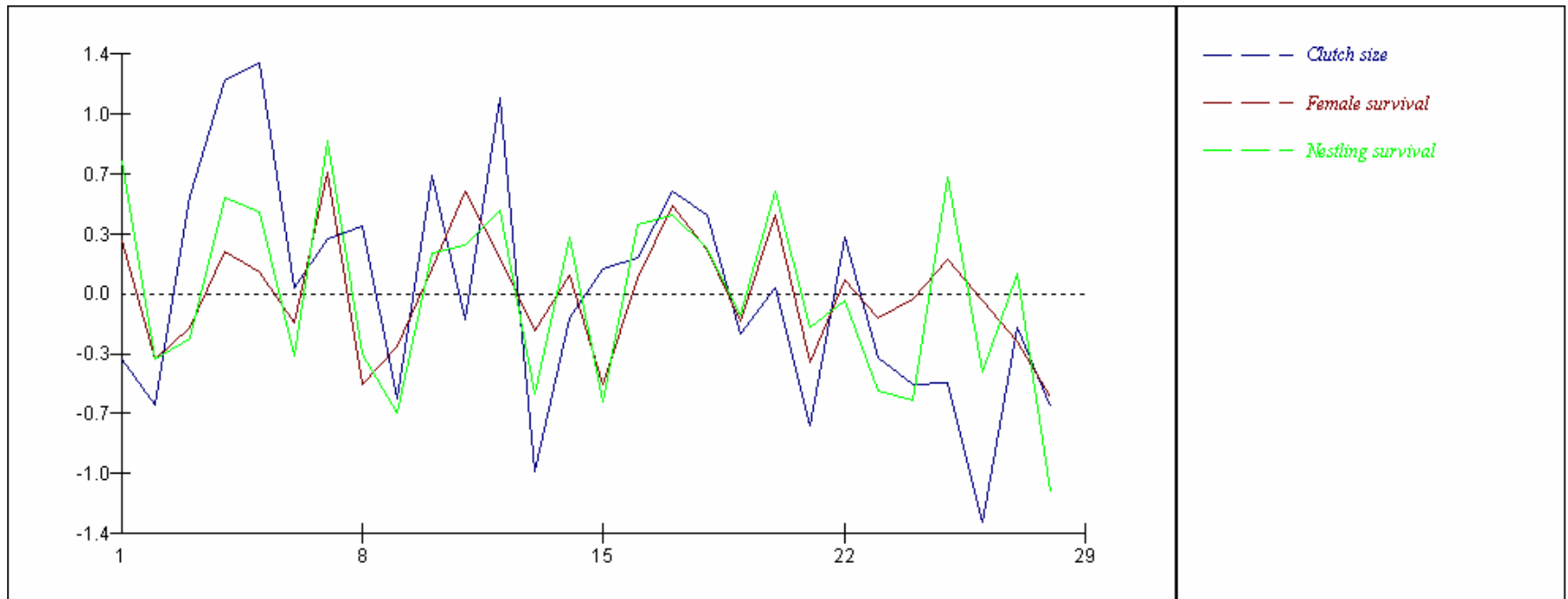
$$\Omega_{u(2)} = \begin{bmatrix} 1.149 & & \\ 0.530 & 0.357 & \\ 0.051 & 0.078 & 0.249 \end{bmatrix} \quad \Omega_{u(3)} = \begin{bmatrix} 0.583 & & \\ 0.094 & 0.162 & \\ 0.196 & 0.163 & 0.356 \end{bmatrix}$$

$$\sigma_e^2 = 1.180$$

We can see that clutch size and female survival is more influenced by female birds whilst nestling survival is more influenced by year. We see that all correlations are positive suggesting birds in better condition lay larger clutches and are more likely to survive. There are particularly strong correlations between clutch size and female survival between females and the two survival variables between years.

Results – year effects

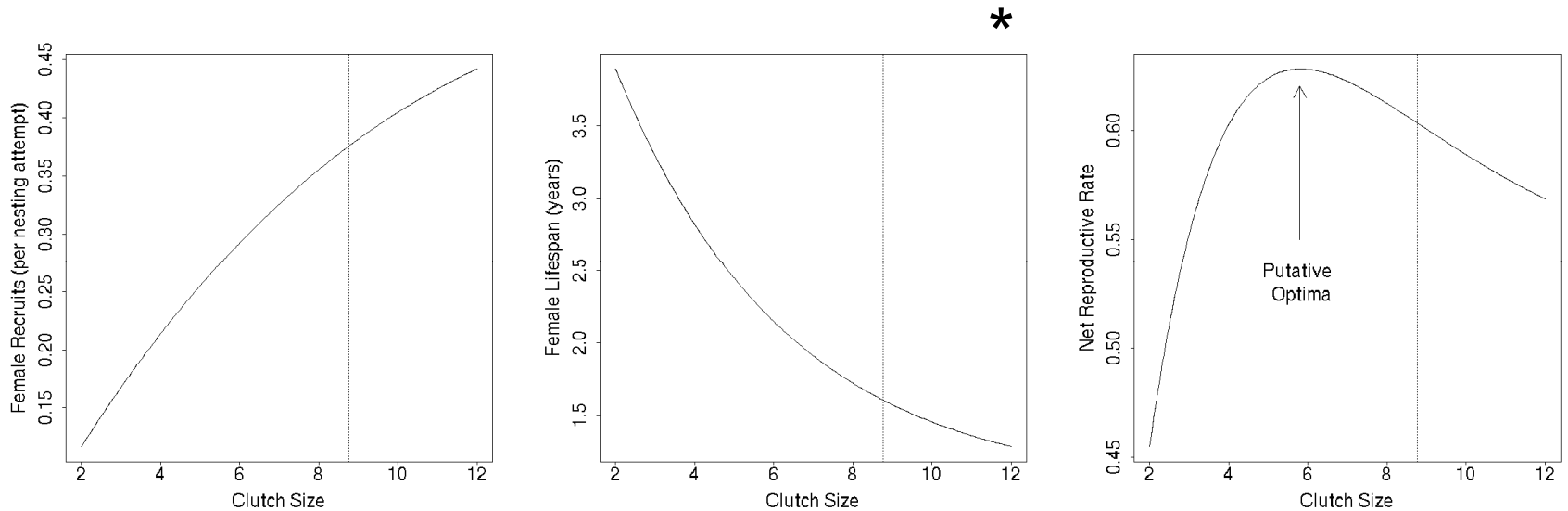
Here we see the random effects for the three responses for the 28 years. There is positive correlation between the three responses in particular between the two survivals (0.83).



Further work with Dylan Childs: Using year effects in a population model

- Population model aims to look at dynamic behaviour of population over time and look for 'optimal' strategies.
- The values of the year random effects can be used as known constants in the population model to focus the model on the time period studied or alternatively year random effects can be drawn from the variance-covariance matrix to simulate 'typical' years.
- The model can be expanded to incorporate dispersal of birds from Wytham, egg survival in the nest, fledgling survival and adult survival.

Using the estimates: Estimation of an optimal strategy



- The adult cost of reproduction (*) is very hard to detect using population level analyses
- It appears to be an important component of fitness.

Conclusions

- In this talk we have considered analysing observational bird ecology data using complex random effect models.
- We have seen how these models can be used to partition both variability and correlation between various classifications to identify interesting relationships.
- We then investigated hierarchical centering and parameter expansion for a model for one of our responses. These are both useful methods for improving mixing when using MCMC.
- We finally looked at how to incorporate individual fledgling survival in our models. This will be useful for fitting population models in the future.

References

- Browne, W.J. (2004). An illustration of the use of reparameterisation methods for improving MCMC efficiency in crossed random effect models. *Multilevel Modelling Newsletter* **16** (1): 13-25
- Browne, W.J., McCleery, R.H., Sheldon, B.C., and Pettifor, R.A. (2007). Using cross-classified multivariate mixed response models with application to the reproductive success of great tits (*Parus Major*). *To appear in Statistical Modelling*
- Browne, W.J. (2002). *MCMC Estimation in MLwiN*. London: Institute of Education, University of London
- Browne, W.J. (2006). MCMC Estimation of 'constrained' variance matrices with applications in multilevel modelling. *Computational Statistics and Data Analysis*. **50**: 1655-1677.
- Browne, W.J., Goldstein, H. and Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling* **1**: 103-124.
- Chib, S. and Greenburg, E. (1998). Analysis of multivariate probit models. *Biometrika* **85**, 347-361.
- Gelfand A.E., Sahu S.K., and Carlin B.P. (1995). Efficient Parametrizations For Normal Linear Mixed Models. *Biometrika* **82** (3): 479-488.
- Kass, R.E., Carlin, B.P., Gelman, A. and Neal, R. (1998). Markov chain Monte Carlo in practice: a roundtable discussion. *American Statistician*, **52**, 93-100.
- Liu, C., Rubin, D.B., and Wu, Y.N. (1998) Parameter expansion to accelerate EM: The PX-EM algorithm. *Biometrika* **85** (4): 755-770.