MULTILEVEL MODELLING NEWSLETTER

Centre for Multilevel Modelling
Bedford Group for Lifecourse and Statistical Studies
Institute of Education, University of London
20 Bedford Way, London WC1H 0AL, ENGLAND
Web site: http://multilevel.ioe.ac.uk/
Enquiries about newsletter to Ian Plewis
E-mail: i.plewis@ioe.ac.uk
Tel: +44 (0) 20 7612 6688
Fax: +44 (0) 20 7612 6572

Vol. 16 No. 1 **September, 2004**

Forthcoming Workshops

1-3 December 2004. A three-day introductory workshop in multilevel modelling for medical and public health researchers using *MLwiN* will take place at the Institute of Community Health Sciences, Queen Mary, University of London.

Enquiries to:

Tami Lapidot
Forensic Psychiatry Research Unit, St. Bartholomew's Hospital, William Harvey House, 61 Bartholomew Close, London EC1A 7BE, United Kingdom.
Tel: +44 (0) 20 7601 7511
Fax: +44 (0) 20 7601 7969
Email: t.lapidot@qmul.ac.uk

If you plan to run any workshops using *MLwiN*, please notify Amy Burch a.burch@ioe.ac.uk and she will advertise these workshops on the multilevel web site.

The Royal Statistical Society- Joint Meeting of the Social Statistics Section/ General Applications Section

19 October 2004. Recent Advances in Multilevel Modelling Methodology And Applications. 2.00 pm to 5.30 pm at the Royal Statistical Society ,12 Errol Street, London, EC1Y 8LX.

All are welcome and the event is free: Registration details on the Society's web page www.rss.org.uk or contact a.fielding@bham.ac.uk.

The following papers will be given:

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

A Pickles, N.Shryane, & E. Fieldhouse (University of Manchester). *Joint analysis of ranked preferences and electoral voting to identify patterns of tactical voting*. Generalised Linear Latent and Mixed Model frameworks using GLLAMM will be discussed to examine the contrast between ranked preferences and voting behavior.

A Leyland (University of Glasgow) & Ø. Næss (University of Oslo, Norway). *Correlated cross-classified multilevel models: lifecourse epidemiology in the Oslo mortality study*. Individual mortality data in which the effect of area of residence at four timepoints (censuses 1960-1990) is modelled will be investigated. Mobility of people is modelled by cross-classified effects incorporating area effects that are correlated over time.

F. Steele, C. Kallis, H. Goldstein & H. Joshi (Institute of Education, University of London). *Modelling correlated event histories: partnerships and childbearing among British women in the 1958 birth cohort*. A new method for the analysis of correlated event histories will be described using a simultaneous equation multilevel multi-state model of repeated transitions from marital and non-marital unions and childbearing within co-residential partnerships.

P. Bassett (Institute of Education, University of London). *An application of multilevel methods in examining the effects of class size upon pupil attainment in English primary schools*. In this talk an analysis of a large-scale longitudinal study investigating the effects of class size upon pupil attainment in primary schools will be discussed.

H. Goldstein (Institute of Education, University of London). *Multilevel smoothing spline models*. Generalisations of existing methods for generalised additive modelling will be examined with particular applications to repeated measures data. It will be shown how existing algorithms can be modified to incorporate such models.

J. Rasbash (University of Bristol). *Multilevel social network models and their application to family relationship data*. In family relationship studies, the data are often of the form of measured behaviour from one family member to another family member. Models that decompose family relationships into terms for actor, partner, dyad and family level effects will be described.

ESRC Research Methods Festival

This very successful event, held in Oxford on 1 – 3 July 2004, featured a number of presentations on multilevel modelling. These included papers by Kelvyn Jones, Ed Fieldhouse, Alastair Leyland and Ian Plewis in a session called 'What is multilevel modelling?' that was very well attended; and also more advanced papers by James Carpenter in the session 'Modelling complex processes', and by Fiona Steele and colleagues, and by Jon Rasbash and Tom O'Connor in the session 'Understanding family processes'.

More details about the presentations can be found at: http://www.ccsr.ac.uk/methods/festival/programme/

RC33 Sixth International Conference on Social Science Methodology

The RC33 Sixth International Conference on Social Science Methodology was held in Amsterdam on 16-20 August 2004.

The following papers were presented at the Multilevel Analysis session:

Frequentist MCMC Estimation Methods for Multilevel Logistic Regression
Carlos Coimbra and Tom A.B. Snijders, Department of Sociology, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands.
c.a.q.coimbra@ppsw.rug.nl
t.a.b.snijders@ppsw.rug.nl

The Use of Internal Pilot Studies to Derive Powerful and Cost-Efficient Designs for Studies with Nested Data
M. Moerbeek, Department of Methodology and Statistics, Utrecht University, The Netherlands.
m.moerbeek@fss.uu.nl

Performance of Likelihood-Based Estimation Methods for Multilevel Binary Regression Models
Marc Callens and Christophe Croux, Katholieke Universiteit Leuven, Belgium.
marc.callens@econ.kuleuven.ac.be

Outliers and Multilevel Models
John F. Bell and Eva Malacova, University of Cambridge Local Examinations Syndicate, UK.
bell.j@ucles.org.uk

On the Relative Efficiency of Unequal Cluster Sizes in Multilevel Intervention Studies
L. Kotova, G.J.P. van Breukelen, M.J.J.M. Candel and M.P.F. Berger, Department of Methodology and Statistics, Faculty of Health Sciences, University of Maastricht, The Netherlands.
larissa.kotova@stat.unimaas.nl

Multilevel Multiprocess Modelling of Partnership Transitions and Fertility in Britain
Fiona Steele, Constantinos Kallis, Harvey Goldstein and Heather Joshi, Bedford Group for Lifecourse and Statistical Studies, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK.
c.kallis@ioe.ac.uk

The Concept of 'Social Level' and how to Assess it
Pieter van den Eeden, Department of Social Research Methodology, Vrije Universiteit, Amsterdam, The Netherlands.
pvdeeden@inter.nl.net

The Hausman Test of Random Effects Specifications
A. Fielding, University of Birmingham, UK.
a.fielding@bham.ac.uk

The Problem of Time Dependent Explanatory Variables at the Context Level in Discrete Time Multilevel Event History Analysis.

Michael Windzio, EMPAS, University of Bremen, Germany. mwindzio@empas.uni-bremen.de

Multilevel Multiprocess Modelling of Partnership and Childbearing Event Histories

Fiona Steele, Constantinos Kallis and Harvey Goldstein
Institute of Education, University of London
f.steele@ioe.ac.uk

Introduction

The outcomes of marital and non-marital partnerships and childbearing within those partnerships are two related dynamic processes. The decision to end a partnership, or to move from cohabitation to marriage, is likely to be jointly determined with the decision to have a child with that partner. In other words, there may be factors, both observed and unobserved, which drive both processes. While previous research has examined the effects of the presence of children on partnership stability, few studies allow for the possibility that children are prior outcomes of a potentially related process. If decisions about partnerships and childbearing are jointly determined, the unobserved components of the models for each process will be correlated. Therefore indicators of the presence of children will not be independent of the residuals in the model for partnership transitions, and estimates of their effects on partnership outcomes will be biased.

In this paper, we examine the effect of the presence and age of children on partnership outcomes using a multiprocess model (Lillard, 1993), which allows for correlation between the unmeasured individual-specific determinants of partnership durations and fertility. A multilevel model is used to allow for correlation between the durations of multiple partnerships, and of intervals between children, for the same individual. Repeated events lead to a two level hierarchical structure, with events nested within individuals.

Methodology

The multiprocess model is a system of simultaneous equations for partnership transitions and childbearing. Simultaneity of the two processes comes from allowing the hazard of a partnership transition at time t to depend on prior outcomes of the childbearing process (the number and age of children born *before* time t), and allowing for correlation between unobservables affecting each process. We consider a total of three partnership transitions: marriage to separation, cohabitation to separation, and cohabitation to marriage. The hazards of these transitions are modelled jointly with the hazard of a conception, again distinguishing between marital and non-marital partnerships. Each equation defines a discrete-time hazards model. A discrete-time formulation has two main advantages. First, as with many retrospectively collected event history data, the dates of events are reported in months. It is therefore natural to specify a model that assumes measurement in discrete rather than continuous time. Second, after restructuring the data, standard multilevel methods for analysing discrete response data may be used (Steele et al., 1996). Thus complex event history models, such as the one described below, may be fitted using existing estimation procedures and software.

Model for Partnership Transitions

Marriage

A partnership is defined as a continuous period of at least one month spent living with the same partner. The unit of analysis is a partnership episode, which is defined as a continuous period of time spent in the same partnership state, marriage or (unmarried) cohabitation, with the same partner.

We denote by $h_{ij}^{PM}(t)$ the hazard of a marital separation during time interval *t* of episode *i* for individual *j*. A multilevel discrete-time event history model for marital separations may be written (omitting subscripts) as:

$$\begin{aligned} \operatorname{logit} h^{PM}(t) &= \alpha_0^M D^{PM}(t) + \\ &\alpha_1^M F(t) + \alpha_2^M X^{PM}(t) + u^{PM} \end{aligned} \qquad (1)$$

$\alpha_0^M D^{PM}(t)$ is the baseline log-hazard which is a function of marriage duration at time *t* or, for marriages immediately preceded by a period of cohabitation, partnership duration. Possible choices for the baseline log-hazard include a step function, where the duration is treated as a categorical variable, or a polynomial function. The potentially endogenous time-varying outcomes of the fertility process, which may affect both future partnership transitions and fertility, are denoted by $F(t)$, with coefficient vector α_1^M. Other covariates which affect marital dissolution are represented by $X^{PM}(t)$. Unobserved time-invariant individual-specific factors are represented by normally distributed random effects u^{PM}.

In order to estimate (1) each marriage duration, D_{ij}^{PM}, is converted to a sequence of D_{ij}^{PM} binary responses, $y_{ij}^{PM}(t)$. For *t*=1, . . ., $D_{ij}^{PM}-1$, $y_{ij}^{PM}(t)$=0; and for t=D_{ij}^{PM}, $y_{ij}^{PM}(t)$=1 if separation occurs at D_{ij}^{PM} and $y_{ij}^{PM}(t)$=0 otherwise (right-censored durations). As start and end dates of episodes were recorded to the nearest month, it is possible to have a binary response for each month. However, using discrete time intervals of one month leads to a very large dataset. We therefore grouped partnership durations (and birth intervals) into six-month intervals, with each observation weighted by the number of months for which an individual was 'at risk' of having an event.

Cohabitation

We consider two transitions from the cohabitation state: separation, and marriage to the same partner. Denote by $h_{ij}^{PC(r)}(t)$ the hazard of a transition of type r from cohabitation, in time interval t of episode i for individual j, where r=0 (no transition), 1 (separation), or 2 (marriage). Transitions from cohabitation may be modelled using a multilevel discrete-time competing risks model (Steele et al., 1996):

$$\log\left[\frac{h^{PC(r)}(t)}{h^{PC(0)}(t)}\right] = \alpha_0^{C(r)} D^{PC(r)}(t) + \alpha_1^{C(r)} F(t) + \alpha_2^{C(r)} X^{PC(r)}(t) + u^{PC(r)}, \quad r = 1,2 \qquad (2)$$

where $\alpha_0^{C(r)} D^{PC(r)}(t)$ is a function of cohabitation duration at time t, $X^{PC(r)}(t)$ are covariates that affect the hazard of a transition of type r from cohabitation, and $u^{PC(r)}$ are individual and transition-specific random effects.

To estimate (2) each cohabitation duration, D_{ij}^{PC}, is converted to a sequence of D_{ij}^{PC} multinomial responses, $y_{ij}^{PC}(t)$. The response at time t is coded 0 if still cohabiting, 1 if separation occurs, and 2 if marriage to the same partner occurs.

Equations (1) and (2) define a multilevel multistate model (Steele et al., 2004), where in the present case the states are marriage and cohabitation. To allow for unobserved individual-level characteristics that affect each type of transition, the random effects may be correlated across transitions with covariance Ω_u^P. Simultaneous estimation of (1) and (2) is achieved by pooling all episodes and defining indicator variables for marriage and cohabitation. These indicators are interacted with the explanatory variables to allow for marriage and cohabitation specific effects of partnership duration, fertility outcomes and background characteristics. The coefficients of the indicators themselves are allowed to vary randomly across women to produce the state-specific random effects.

Model For Childbearing Within Partnerships

Denote by $h_{ij}^{FM}(t)$ the hazard of a conception leading to a live birth within marriage during time interval t in partnership episode i for individual j. We denote by $h_{ij}^{FC}(t)$ the hazard of a conception within a cohabiting partnership. The model for childbearing consists of separate equations for marriage and cohabitation, which are estimated simultaneously. Both equations include as covariates prior outcomes of the childbearing process, $F(t)$, as well as background characteristics.

Marriage

A multilevel event history model for the waiting time to conception within marriage may be written (omitting subscripts):

$$\text{logit}\, h^{FM}(t) = \beta_0^M D^{FM}(t) + \beta_1^M F(t) + \beta_2^M X^{FM}(t) + u^{FM} \quad (3)$$

where $\beta_0^M D^{FM}(t)$ is a function of the partnership duration, $X^{FM}(t)$ are covariates affecting the fertility process, and u^{FM} is an individual-level random effect.

Cohabitation

The model for conceptions within cohabitation is written:

$$\text{logit}\, h^{FC}(t) = \beta_0^C D^{FC}(t) + \beta_1^C F(t) + \beta_2^C X^{FC}(t) + u^{FC} \quad (4)$$

where $X^{FC}(t)$ are covariates and u^{FC} is an individual-level random effect, which may be correlated with u^{FM} with covariance Ω_u^F.

Estimation

Equations (1), (2), (3) and (4) define a multiprocess model. These equations must be estimated simultaneously as there may be non-zero correlations between the woman-specific random effects across equations. Specifically we assume that $u = (u^{PM}, u^{PC(1)}, u^{PC(2)}, u^{FM}, u^{FC}) \sim N_5(\mathbf{0}, \Omega_u)$. Correlated random effects would arise if the unobserved characteristics that influence the timing of partnership transitions are correlated with those that affect childbearing within partnerships. Non-zero correlations between elements of $u^P = (u^{PM}, u^{PC(1)}, u^{PC(2)})$ and of $u^F = (u^{FM}, u^{FC})$ would suggest that $F(t)$, the number and/or age of children from the current or a previous partnership, is endogenous with respect to partnership transitions.

The discrete-time multiprocess event history model can be framed as a multilevel bivariate discrete response model where for each time interval t of a partnership there are two responses: 1) a binary or multinomial response for the partnership status, and 2) a binary response indicating the occurrence of a birth. The model may be estimated using existing methods for mixtures of binary and multinomial responses (Steele et al., 2004) after defining indicators for the partnership and fertility responses and interacting these with the duration variables and covariates. The results presented in this paper were obtained using Monte Carlo Markov Chain (MCMC) estimation, as implemented in *MLwiN* (Rasbash et al., 2004).

Data

The analysis uses data from female respondents in the National Child Development Study (NCDS), a longitudinal study of all those living in

Great Britain who were born in a single week in March 1958 (Shepherd, 1997). Retrospective partnership and birth histories were collected in 1981, 1991 and 2000, when the respondents were age 23, 33 and 42. One task of the current study was to link data collected at each age to form continuous partnership and birth histories from ages 16 to 42.

The explanatory variables of major interest are outcomes of the fertility process. Respondents were asked to identify the father of each child and for the date that each child left home. Thus it was possible to create time-varying counts of the number of children living with a woman, distinguishing between preschool and older children, and between children born to the current partner at time *t* and those fathered by a previous partner or a non-coresident partner. Other covariates include age at the start of the partnership, variables relating to previous partnerships, the number of years of post-compulsory education (time-varying), father's social class and the experience of parental separation during childhood.

The analysis sample contains 5142 women who had partnered before age 33; these women contribute 7032 partnerships and 9137 partnership episodes.

Results

Correlations between random effects

The estimated random effects covariance matrix obtained from the multiprocess model is shown in Table 1.

Table 1. Estimated random effects covariance matrix from the multiprocess model

	Conception within cohabitation	Conception within marriage	Marital separation	Cohabitation separation	Cohabitation to marriage
Conception within cohabitation	0.296* (0.212, 0.431)				
Conception within marriage	-0.018 (-0.041, 0.001) **-0.143**	0.050* (0.041, 0.062)			
Marital separation	0.246* (0.088, 0.417) **0.377**	-0.075* (-0.130, -0.030) **-0.278**	1.433* (0.975, 1.884)		
Cohabitation separation	0.081 (-0.057, 0.206) **0.187**	-0.026 (-0.059, 0.009) **-0.145**	0.497* (0.210, 0.741) **0.520**	0.652* (0.424, 0.928)	
Cohabitation to marriage	0.214* (0.130, 0.319) **0.591**	-0.019 (-0.047, 0.007) **-0.129**	0.237* (0.051, 0.428) **0.296**	0.095 (-0.072, 0.263) **0.178**	0.444* (0.301, 0.602)

Note: The values in each cell are the point estimate (the mean of the MCMC samples) and the 95% interval estimate (the 2.5% and 97.5% point of the distribution). In off-diagonal cells an estimate of the correlation (the mean of the correlation estimates across samples) is shown in bold. The results are based on 30,000 MCMC samples, with a burn-in of 5,000.

*Indicates that the 95% interval estimate does not contain zero

There is substantial unobserved heterogeneity in the hazards of all partnership transitions and in the hazards of conceptions within partnerships. Of most interest, however, are the covariance terms, several of which differ significantly from zero. Among partnership transitions, for example, the random effect for marital separation is positively correlated with the random effect for separation from cohabitation; this suggests that women with above average propensities of marital separation (u^{PM}>0) will tend also to have above average propensities to separate from a non-marital partnership ($u^{PC(1)}$>0).

Across processes, the random effects for marital separation and conception intervals within marriage are negatively correlated. Women with below average risks of separation, i.e. long marriages, have an above average risk of having a child with a husband. A strong positive correlation is found between the unobserved woman-specific factors affecting the hazard of converting a cohabitating partnership into marriage and those affecting the hazard of a birth within cohabitation. A possible interpretation of this correlation is that women who view cohabitation as a precursor to a more formal marital partnership (and therefore have a high probability of marrying) are likely to have a child while cohabiting, in anticipation of marriage. However, the significant positive correlation between the random effects for marital separation and births within cohabitation suggests that women with a high chance of having a child during cohabitation tend to have a high risk of separation should they marry.

Effects of prior fertility outcomes on partnership transitions

Table 2 shows estimates from two model specifications, controlling for the effects of the other covariates mentioned earlier. The first model is a single process model, where the random effects across processes are assumed to be uncorrelated. This model assumes that prior fertility outcomes are exogenous with respect to partnership transitions. The second model considered is a multiprocess model in which the correlations between u^P and u^F are estimated freely. A correlation that is significantly different from zero provides evidence that prior fertility outcomes are endogenous, in which case the estimated effects from the single process model will be biased.

Table 2. Estimated effects of preschool children with the current partner on partnership transitions

Variables	*Single process model* *Coefficient*	*(SE)*	*Multiprocess model* *Coefficient*	*(SE)*
Marital separation				
No. children (ref.=none)				
1	-0.525*	(0.067)	-0.510*	(0.067)
2+	-0.878*	(0.103)	-0.837*	(0.104)
Separation from cohabitation				
No. children				
1	-0.280*	(0.116)	-0.299*	(0.119)
2+	-0.739*	(0.258)	-0.792*	(0.265)
Cohabitation to marriage				
No. children				
1	-0.147	(0.081)	-0.230*	(0.084)
2+	-0.073	(0.158)	-0.245	(0.162)

*Indicates that the 95% interval estimate does not contain zero.

For illustration, we present only the effects of having preschool age children fathered by the current partner. (The complete set of results can be found in the full version of the paper available on request from f.steele@ioe.ac.uk.) The results from both models imply that the presence of young children reduces the risk of marital separation. The effects are slightly weaker for the multiprocess model, which can be explained by the negative correlation between the random effects for marital separation and marital fertility (Table 1). The strong negative effect obtained using a single process model is partly due to selection; women with a low risk of separation are more likely to have children within marriage. These women lower the risk of separation for women with marital children, leading to an overstatement of the negative effect of having children.

Having young children also reduces the risk of separation for a cohabiting couple. Since the random effect for separation from cohabitation is not significantly correlated with either of the random effects for fertility, we do not have sufficient evidence to reject the single process model in favour of the multiprocess model.

Based on the single process model, we would conclude that cohabiting couples who have had children together are not significantly more or less likely to marry than those who have not. However, when we move to a multiprocess model the negative effect of having one child becomes stronger and attains significance at the 5% level. This change in the estimates is due to the positive correlation between the random effects for the transition from cohabitation to marriage and births within cohabitation. On average, women with a high propensity to marry a cohabiting partner have a high propensity to have a child during cohabitation. If this form of selection is ignored, the estimated odds of marriage for women who have had children with their current (cohabiting) partner will be inflated, leading to the erroneous conclusion that having young children is not associated with marriage.

Conclusion

We have proposed a multiprocess model for the analysis of correlated event histories. By modelling jointly the processes of marital dissolution, the outcomes of cohabitation and childbearing we allow for endogeneity of the presence of children born within partnerships. While adopting a multiprocess approach leads to little change in the substantive conclusions about the effects of prior fertility outcomes on partnership dissolution, a negative effect of the presence of young children on the transition from cohabitation to marriage emerges. In addition, the multiprocess model reveals a number of interesting findings regarding correlations between the unobserved factors influencing the different processes. For example, a negative residual correlation between the hazards of marital dissolution and of a marital birth suggests that women with a high risk of dissolution tend to delay or limit childbearing within marriage.

Future research under the current project will explore partnership transitions and fertility for women from the 1970 British Cohort Study (BCS70). The experiences of this younger cohort will be compared with those of the 1958 birth cohort for ages 16-30. Questions for further research include whether the effects of the presence of children on partnership dissolution and the movement from cohabitation to marriage have changed as single parenthood and non-marital births become increasingly common.

Acknowledgements

This research was funded by the UK Economic and Research Council under the Research Methods Programme (award no. H333250044). We thank colleagues in the Centre for Longitudinal Studies at the Institute of Education for their advice on data preparation.

References

Lillard, L. (1993). Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing. *Journal of Econometrics*, **56**: 189-217.

Rasbash, J., Steele, F., Browne, W., and Prosser, B. (2004). *A User's Guide to MLwiN*, *Version 2.0*. London: Institute of Education.

Shepherd, P. (1997). *The National Child Development Study: An Introduction to the Origins of the Study and the Methods of Data Collection*. Centre for Longitudinal Studies, Institute of Education, London.

Steele, F., Diamond, I. and Wang, D. (1996). The Determinants of the Duration of Contraceptive Use in China: A Multilevel Multinomial Discrete Hazards Modelling Approach. *Demography*, **33**: 12-33.

Steele, F., Goldstein, H., and Browne, W. (2004). A General Multistate Competing Risks Model for Event History Data, with an Application to the Study of Contraceptive Use Dynamics. *Journal of Statistical Modelling*, **4**: 145-159.

An Illustration of the Use of Reparameterisation Methods for Improving MCMC Efficiency in Crossed Random Effect Models

William J. Browne
School of Mathematical Sciences, University of Nottingham
william.browne@nottingham.ac.uk

Introduction

In this article we illustrate how the performance of MCMC methods can be improved by particular reparameterisation schemes in the MCMC literature. We use as an example a four way crossed classification model fitted to the Wytham Woods great tit dataset of bird nesting attempts recently analysed in Browne et al. (2004). We describe two methods, hierarchical centering (Gelfand et al., 1995) and parameter expansion (Liu et al., 1998) and show how they can improve the efficiency of a Gibbs sampler algorithm. We then show how the methods can be combined to create a more efficient MCMC estimation algorithm.

The *MLwiN* software package (Rasbash et al., 2000), based on the IGLS algorithm (Goldstein, 1986), has introduced random effect modelling to a large number of applied researchers in many disciplines. Rasbash and Goldstein (1994) developed an extension to the IGLS algorithm that deals with cross-classified models by forming a constrained nested model formulation. This method works well with data structures that are approximately nested but has difficulties with larger datasets with many crossed classifications. Patterson and Thompson (1971) derived a restricted maximum likelihood (REML) approach for cross-classified models and an efficient implementation of this approach (Gilmour et al., 1995) is implemented in the GenStat software package.

In recent years a second estimation engine has been included in the *MLwiN* software package based on Monte Carlo Markov Chain (MCMC) estimation. MCMC algorithms are easily extended to fit cross-classified random effects models as described in Chapter 14 of Browne (2002). The MCMC engine in *MLwiN* is designed with speed in mind and uses standard Gibbs sampling and Metropolis Hastings algorithms (see Browne, 2002 for details). This sometimes leads to poor efficiency of the MCMC sampler and bad 'mixing' of the chains it produces i.e. the chains are heavily autocorrelated.

There are many algorithms that fall under the MCMC banner and there has been a lot of research on developing more efficient algorithms for specific models. The WinBUGS (Spiegelhalter et al., 2000) software package is a general purpose MCMC estimation software package which although generally slower than *MLwiN* for the equivalent model, currently has far greater flexibility in terms of model choice. In this article we review two developments in MCMC algorithm construction that can be easily implemented in WinBUGS and result in great improvements in the efficiency of the resulting MCMC sampler. Before describing each of these techniques we briefly describe our example dataset and model that will be used to compare the MCMC algorithms. We then briefly describe the two developments used in this article before looking at the effects they have on our model. We finish with some brief conclusions and discuss extensions to this work.

Wytham Woods great tit dataset

Random effect modelling can be used in many application areas and for our example we use a dataset from bird ecology. Wytham Woods in Oxfordshire is a site where a long-term individual based study of great tits has been carried out, initiated by David Lack in 1947. We consider a dataset of 4165 observations taken over a 34-year period (1964-1997). Each observation is a breeding attempt for a pair of great tits and the dataset contains six response variables for each observation. We also have for each attempt the identification of the male and female birds involved plus the year of the attempt and the nestbox. From a substantive point of view interest lies in the relative importance of the genetic and environmental effects and Browne et al. (2004) consider fitting a multivariate response cross-classified model to the dataset.

For our purposes we will consider just one of the response variables - clutch size - and examine the univariate normal response model fitted to it in Browne et al. (2004). The model can be written using the notation of Browne et al. (2001) as:

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

where y_i is the clutch size for observation *i*. The four sets of *u*'s are random effects with the superscripts identifying the respective higher levels. The subscripts are functions that for each observation identify the corresponding higher level unit. We have added diffuse priors for all unknown parameters with $\varepsilon = 10^{-3}$.

The dataset structure is described in Table 1 and here we see that for many of the random effects, in particular the male and female bird effects we have very little data to work with. In fact some of these effects are not identifiable via the data alone as we have many pairs of male and female birds who only ever mate with each other and hence their relative effects are estimated via their prior distributions. This model and dataset was chosen as the efficiency of the standard Gibbs sampling algorithm as used in Browne et al. (2004) is poor (in fact they run this model for 250k iterations to get 'reasonable' estimates).

Table 1. Summary of the potential sources of variation in the great tit dataset

Source	Number	Median observations per id	Observations per id
Year	34	104	19-250
Nestbox	968	4	1-16
Male Bird	2986	1	1-6
Female Bird	2944	1	1-6

Model (1) was run for 50,000 iterations following a burn-in of 5,000 iterations using both *MLwiN* (version 2.0) and WinBUGS (version 1.4). Note the *MLwiN* to WinBUGS interface was used to generate the WinBUGS code (see Browne, 2002, Chapter 8). The point estimates and 95% credible intervals for the fixed effect (average clutch size) and five variance estimates are given in columns 2 and 3 of Table 2. As we would expect, given both pieces of software are supposed to fit the same algorithm, we have very similar estimates. The largest differences are in the point and interval estimates of the between male variance and these may be explained by looking at the autocorrelations in the chains.

Table 2. Point estimates (means) and 95% credible interval estimates for each method

Parameter	*MLwiN* Gibbs	WinBUGS Gibbs	WinBUGS Hierarchical centering	WinBUGS Parameter Expansion	WinBUGS Both methods
β_0	8.805 (8.589,9.025)	8.810 (8.593,9.021)	8.809 (8.596,9.023)	8.806 (8.582,9.024)	8.810 (8.593,9.024)
$\sigma^2_{u(5)}$ - Year	0.365 (0.215,0.611)	0.365 (0.216,0.606)	0.365 (0.215,0.606)	0.377 (0.220,0.630)	0.365 (0.215,0.607)
$\sigma^2_{u(4)}$ - Nestbox	0.107 (0.059,0.158)	0.108 (0.060,0.161)	0.108 (0.060,0.161)	0.110 (0.060,0.165)	0.109 (0.061,0.162)
$\sigma^2_{u(3)}$ - Male	0.045 (0.001,0.166)	0.034 (0.002,0.126)	0.034 (0.002,0.126)	0.064 (0.001,0.172)	0.070 (0.001,0.178)
$\sigma^2_{u(2)}$ - Female	0.975 (0.854,1.101)	0.976 (0.858,1.097)	0.976 (0.857,1.097)	0.971 (0.853,1.094)	0.968 (0.848,1.089)
σ^2_{e} - Observation	1.064 (0.952,1.173)	1.073 (0.968,1.175)	1.073 (0.968,1.175)	1.049 (0.938,1.158)	1.046 (0.935,1.157)

Table 3 gives effective sample size (ESS) estimates for each of the parameters for each method with the standard Gibbs sampler implementations in *MLwiN* and WinBUGS given in columns 2 and 3 respectively. The ESS (Kass et al., 1998) equals the number of iterations divided by a measure of the correlation of the chain. For an independent sampler the ESS will equal the actual number of iterations. It should be noted that the differences in ESS between *MLwiN* and WinBUGS here give an indication of the size of Monte Carlo errors for this statistic. We see that the between year variance has greatest ESS whilst the between male variance has a very poor ESS of ~35 for 50,000 actual iterations. This will explain why we are observing greater variation between estimates for this parameter.

Table 3. Effective sample sizes resulting from runs of 50,000 iterations following a burn-in of 5,000 iterations

Parameter	*MLwiN* Gibbs	WinBUGS Gibbs	WinBUGS Hierarchical centering	WinBUGS Parameter Expansion	WinBUGS Both methods
β_0	671	602	35063	635	34296
$\sigma^2_{u(5)}$ - Year	30632	29604	34626	29366	34817
$\sigma^2_{u(4)}$ - Nestbox	833	788	789	4887	5170
$\sigma^2_{u(3)}$ - Male	36	33	33	600	557
$\sigma^2_{u(2)}$ - Female	3098	3685	3683	8572	8580
σ^2_e - Observation	110	135	135	1677	1431
Time	519s	2601s	1864s	3662s	2526s

In the final row of Table 3 we see the time to run for 55,000 iterations and we see that the *MLwiN* implementation is significantly faster taking roughly a fifth of the time of WinBUGS. A fair comparison measure for competing MCMC algorithms is to calculate how quickly they can produce a particular ESS or equivalently the ESS per minute and these figures are given in Table 4.

Table 4. Effective samples per minute (after burn-in) for each method

Parameter	*MLwiN* Gibbs	WinBUGS Gibbs	WinBUGS Hierarchical centering	WinBUGS Parameter Expansion	WinBUGS Both methods
β_0	85.3	15.3	1241.5	11.4	896.1
$\sigma^2_{u(5)}$ - Year	3895.4	751.2	1226.0	529.3	909.7
$\sigma^2_{u(4)}$ - Nestbox	105.9	20.0	27.9	88.1	135.1
$\sigma^2_{u(3)}$ - Male	4.6	0.8	1.2	10.8	14.6
$\sigma^2_{u(2)}$ - Female	394.0	93.5	130.4	154.5	224.2
σ^2_{e} - Observation	14.0	3.4	4.8	30.2	37.4

Here as expected we see that *MLwiN* gives roughly five times the number of samples per minute as WinBUGS. We will now describe the two techniques that we hope will improve the ESS figures and explain how they can be used in our example.

Hierarchical centering

MCMC algorithms work by aiming to create chains of independent draws from the joint posterior distribution of all unknown parameters in our statistical model. They do this by considering in turn groups of parameters and sampling from the conditional posterior distribution of each group. In the case of our model and all algorithms considered in this article each group consists of one parameter, although there exist block-updating algorithms, for example structured MCMC (Sargent et al., 2000). In our example we have many random effects at each classification each of which is a parameter, which combined with the variances and the fixed effect results in a total of 6938 parameters and hence a 6938 dimensional joint posterior distribution! The main reason that we see poor effective sample sizes is strong correlations within the joint posterior distributions of groups of the parameters and this is a motivation for block updating algorithms. The first alternative is to consider reparameterisations of the parameters that remove the correlations.

In Table 3 we may expect poor ESS for the male variance given that we have very little information about each male bird effect. However, it is also noticeable that β_0, which represents average clutch size, has reasonably small ESS (~600). We have far more information on this variable so why are we getting poor MCMC efficiency? The hierarchical centering method (Gelfand et al., 1995) was devised for nested random effect models but can also be adapted to improve the mixing of β_0.

If we consider the first line of (1) we see that β_0 is involved in the mean likelihood contribution for each observation, which consists of a sum of β_0 plus 4 random effects. There is, therefore, strong correlation between the value of β_0 and all the random effects. To remove this correlation we can consider centering with respect to one of the higher-level classifications: moving β_0 from the first line of model 1 and including it instead as the mean of one of the set of random effects. We will choose the year level as the between year variance has greatest ESS, although our choice here is rather arbitrary. The resulting model is:

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

Here we have replaced the year level residuals $u^{(5)}_{year(i)}$ (with mean 0) with the year level random effects $\beta^{(5)}_{year(i)}$ (with mean β_0). Models (1) and (2) are equivalent and are simply different parameterisations of the same model. The results of fitting (2) in WinBUGS are given in the fourth columns of Tables 2, 3 and 4. The conditional distribution of β_0 in the new parameterisation is independent of three of the four sets of residuals and is faster to compute resulting in a speeding up of the MCMC algorithm. The change in ESS for β_0 is remarkable, from ~600 to ~35,000 and the other parameter affected by the reparameterisation, $\sigma^2_{u(5)}$, also has an increase in ESS from ~30,000 to ~35,000. Figures 1 and 2 show the trace plots and lag-1 autocorrelation plots for models (1) and (2).

Figure 1. β_0 trace and autocorrelation plot for the basic Gibbs sampler

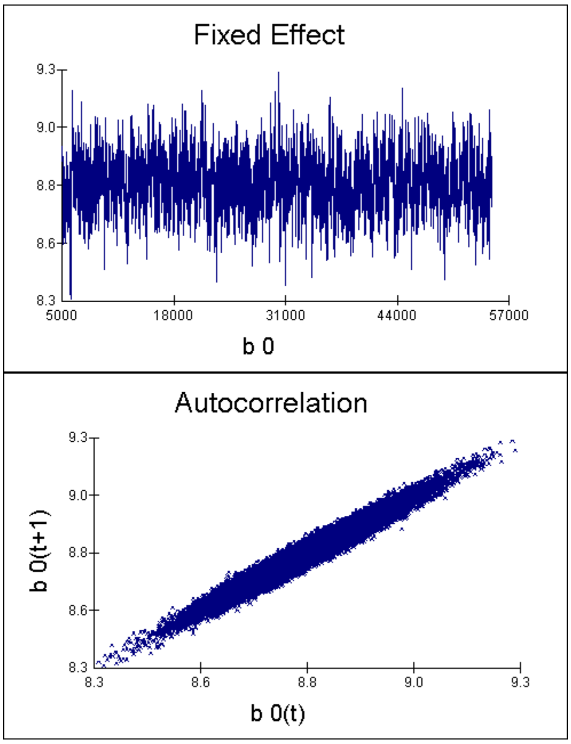
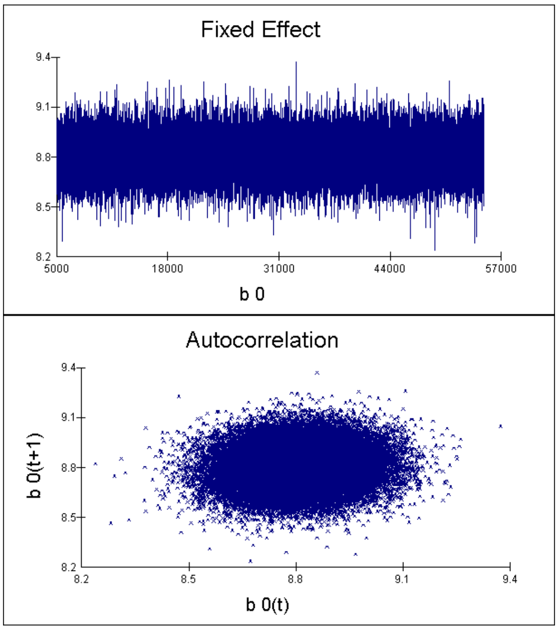


Figure 2. β_0 trace and autocorrelation plot after hierarchical centering



In Figure 1 we see the cigar shaped autocorrelation plot suggesting high autocorrelation. In Figure 2 we see that this plot is now a healthy oval shape and that the trace plot shows better mixing. Returning to Table 3 we see that none of the ESS for the other parameters is affected by hierarchical centering. In particular the ESS for the between male variance is still around 30 and to cure this we turn to our second method.

Parameter expansion

In the last section we have shown how hierarchical centering can improve the ESS for β_0 by using a reparameterisation that removes the correlation between certain parameters. Figure 3 shows the trace plots for the male bird variance and one of the male bird residuals using the standard Gibbs sampling algorithm on model (1). Here we see that the variance trace has very poor mixing and the variance gets stuck for long periods of time close to zero. The residual trace is particularly interesting as for each bird we do not have much information and hence how the trace appears is very closely linked to the variance trace. When the variance is large the residual trace covers the whole of the posterior and when the variance is small the residual trace is concentrated near zero.

Figure 4 shows lag-1 autocorrelations for the variance and residual. It is interesting that here the plot for the variance shows clearly the problem of high autocorrelation whereas the pattern for the residual trace is masked in the autocorrelation plot due to the good mixing behaviour when the variance is large. The problem we have to contend with here is the correlation between the variance and the residuals. In particular when the variance parameter becomes small, the residuals are also small and hence using univariate updating methods we struggle to escape this part of the posterior. A solution lies in the technique of parameter expansion.

Parameter expansion is a method that was originally developed by Liu et al. (1998) to speed up the EM algorithm. This method was then considered in relation to the Gibbs sampler by Liu and Wu (1999) and has been considered particularly for random effect models by Van Dyk and Meng (2001), Gelman et al. (2004) and Gelman (2004a). The method consists of augmenting the model that we wish to fit with additional parameters to form an expanded model. Some of these additional parameters are not identifiable but there exists within the model an 'embedded model' which is identifiable and is the original model we wished to fit. This means that the original parameters can be constructed from the new augmented parameter set.

In our example we wish to reduce the effect of the correlation that exists between the male residuals and the between males variance. We will also do the same for the other four classifications. This we achieve by in our model multiplying each set of residuals by an additional parameter α_i. Our model then becomes:

Figure 3. Trace plots for the male bird variance $\sigma^2_{u(3)}$ and one of the male bird random effects, $u_1^{(3)}$ for the basic Gibbs sampler formulation

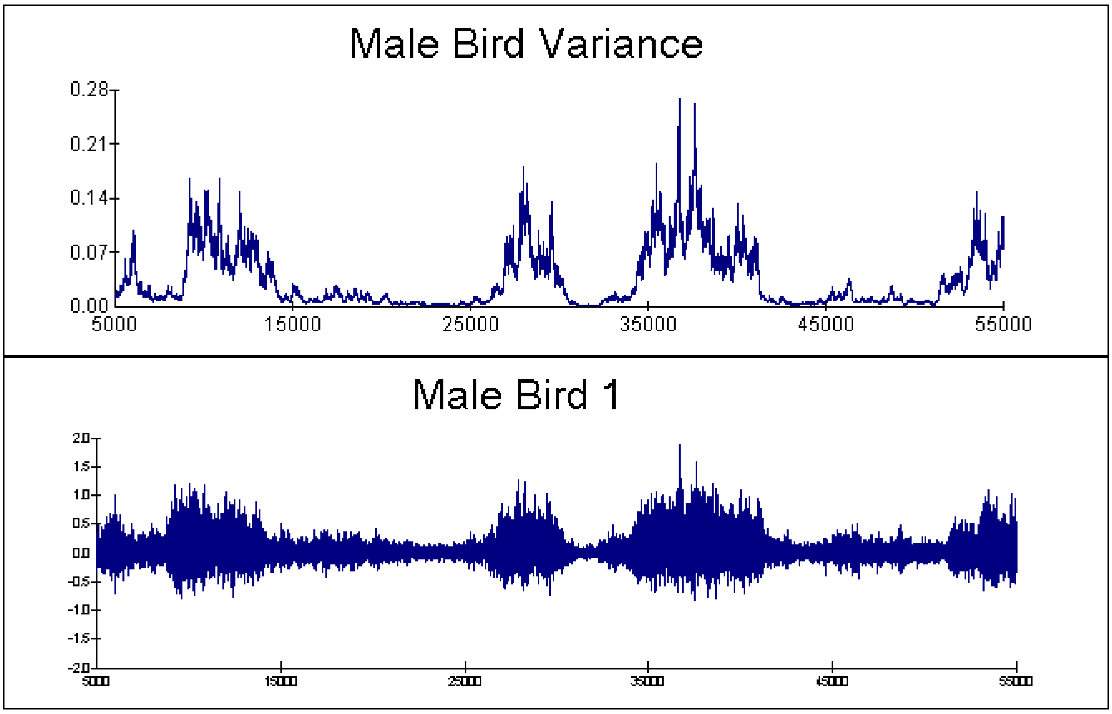
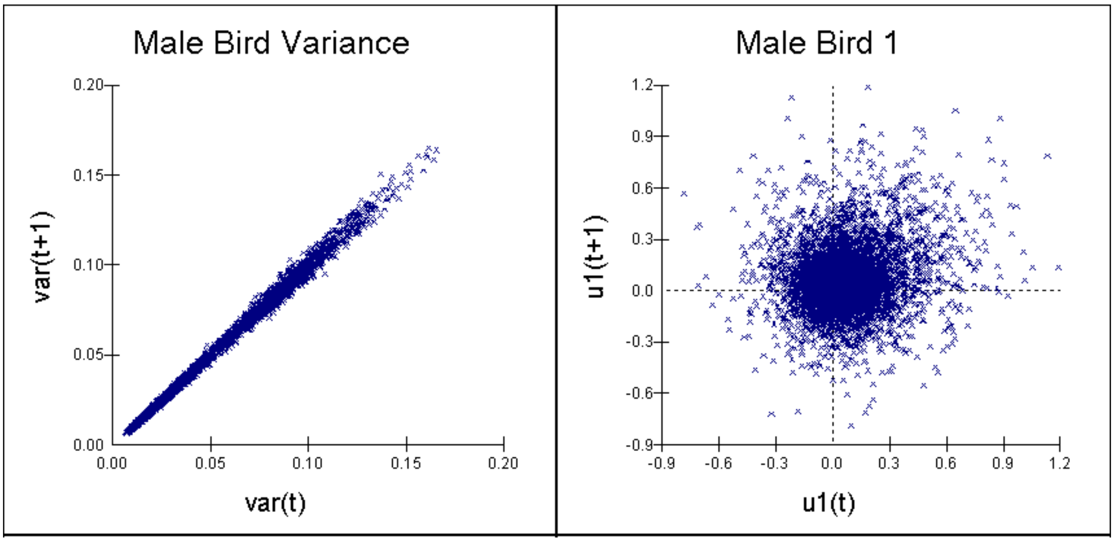


Figure 4. Lag-1 autocorrelations for the male bird variance $\sigma^2_{u(3)}$ and one of the male bird random effects, $u_1^{(3)}$ for the first 5000 stored iterations for the basic Gibbs sampler formulation



$$
\begin{aligned}
&y_i = \beta_0 + \alpha_2 v^{(2)}_{female(i)} + \alpha_3 v^{(3)}_{male(i)} + \alpha_4 v^{(4)}_{nestbox(i)} + \alpha_5 v^{(5)}_{year(i)} + e_i,\\
&v^{(2)}_{female(i)} \sim N(0,\sigma^2_{v(2)}), v^{(3)}_{male(i)} \sim N(0,\sigma^2_{v(3)}),\\
&v^{(4)}_{nestbox(i)} \sim N(0,\sigma^2_{u(4)}), v^{(5)}_{year(i)} \sim N(0,\sigma^2_{v(5)}), e_i \sim N(0,\sigma^2_e),\\
&\beta_0 \propto 1, \alpha_k \propto 1, \sigma^2_{v(k)} \sim \Gamma^{-1}(\varepsilon,\varepsilon), k = 2,..5, \sigma^2_e \sim \Gamma^{-1}(\varepsilon,\varepsilon)
\end{aligned}
\qquad (3)
$$

The original parameters can be found by $u_i^{(k)} = \alpha_k\, v_i^{(k)}$ and $\sigma^2_{u(k)} = \alpha_k^2\, \sigma^2_{v(k)}$ for individual i in classification k. When all the α parameters are one we have our original model. As the α_k parameters multiply both the variance and all the residuals they allow the sampler a quick route out of the part of the posterior near the origin. It should be noted that this model is not identical to the earlier models as the prior distribution for the (original) random effect variances is no longer the $\Gamma^{-1}(\varepsilon,\varepsilon)$ distribution (see Gelman, 2004b for details).

Model (3) was run in WinBUGS and the results are given in the fifth columns of Tables 2, 3 and 4. In Table 3 we see that all the variance parameters have improved ESS (apart from the between year variance which exhibits no change). The ESS improvement is (relatively) greatest for the between male variance with almost a 20-fold improvement. Figures 5 and 6 show trace plots and lag-1 autocorrelation plots for the between males variance and one residual using (3). Here we see great improvement in the mixing of all four parameters and although the variance trace is not perfect it is a great improvement on the trace in Figure 3.

Figure 5. Trace plots for the male bird variance $\sigma^2_{u(3)}$ and one of the male bird random effects, $u_1^{(3)}$ for the parameter-expanded Gibbs sampler formulation

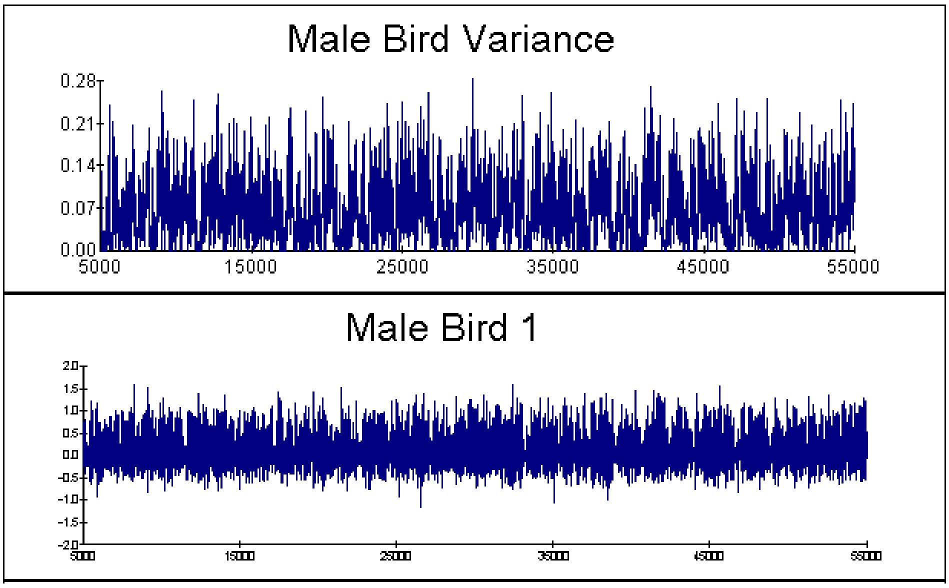
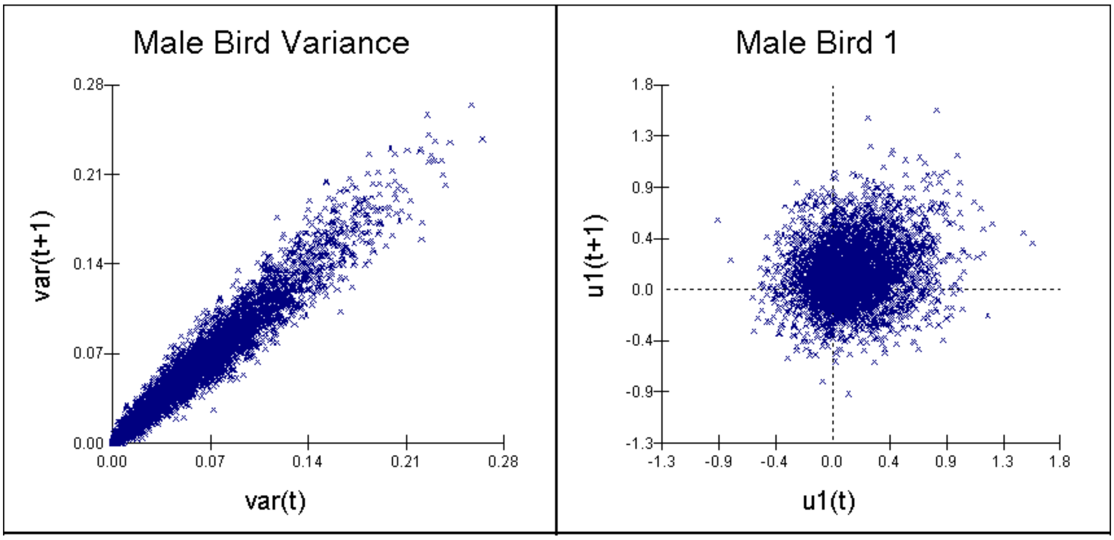


Figure 6. Lag-1 autocorrelations for the male bird variance $\sigma^2_{u(3)}$ and one of the male bird random effects, $u_1^{(3)}$ for the first 5000 stored iterations for the parameter-expanded Gibbs sampler formulation



Combining the methods

In the previous section we have considered two methods that have improved the MCMC efficiency of parts of our cross-classified model for the great tit dataset. Hierarchical centering improved the ESS for β_0 and the between years variance whilst parameter expansion improved the mixing of all the other variances. It is easy therefore to combine the two methods and this results in the following model:

$$\begin{aligned}
&y_i = \beta_0 + \alpha_2 v^{(2)}_{female(i)} + \alpha_3 v^{(3)}_{male(i)} + \alpha_4 v^{(4)}_{nestbox(i)} + \beta^{(5)}_{year(i)} + e_i, \\
&v^{(2)}_{female(i)} \sim N(0, \sigma^2_{v(2)}), v^{(3)}_{male(i)} \sim N(0, \sigma^2_{v(3)}), \\
&v^{(4)}_{nestbox(i)} \sim N(0, \sigma^2_{u(4)}), \beta^{(5)}_{year(i)} \sim N(\beta_0, \sigma^2_{v(5)}), e_i \sim N(0, \sigma^2_e), \\
&\beta_0 \propto 1, \alpha_k \propto 1, \sigma^2_{v(k)} \sim \Gamma^{-1}(\varepsilon, \varepsilon), k = 2,..5, \sigma^2_e \sim \Gamma^{-1}(\varepsilon, \varepsilon)
\end{aligned} \tag{4}$$

The results of fitting this final model in WinBUGS are given in the last column of each of Tables 2, 3 and 4. In Table 3 we see that all parameters have benefited when compared with our original formulation. Table 4 however does show something interesting. The original formulation in *MLwiN* produces larger ESS per minute for both the between-year and between-female variances than our final model in WinBUGS. This shows that work on fast implementations of algorithms is also important when aiming to improve your MCMC sampler.

Conclusions

In this paper we illustrate the use of two techniques that can improve the efficiency of MCMC sampling algorithms. Our example from bird ecology is a challenging dataset and we have shown that both hierarchical centering and parameter expansion have roles to play in improving MCMC efficiency on our example model. We have steered clear of the issue of model selection. However, in Browne et al. (2004) the DIC diagnostic of Spiegel halter et al. (2002) suggests that the model we consider with all four sets of random effects is to be preferred over any of its sub-models.

We have used WinBUGS to demonstrate both hierarchical centering and parameter expansion, as it is very easy to modify the WinBUGS code to implement these features. As noted earlier the standard Gibbs sampler in *MLwiN* still performed better in terms of ESS per minute for two of the six parameters of interest and so we intend to investigate implementing these two techniques in *MLwiN* in future research. We are also interested in comparing the performances we have achieved here with block updating techniques such as structured MCMC (Sargent et al., 2000) on this and similar examples.

References

Browne, W.J. (2002). *MCMC Estimation in MLwiN*. London: Institute of Education, University of London.

Browne, W.J., Goldstein, H. and Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling*, **1**: 103-124.

Browne, W.J., Pettifor, R.A., Sheldon, B.C., Rasbash, J. and McCleery, R.H. (2004). Using cross-classified multivariate mixed response models with application to the reproductive success of great tits (*Parus Major*). *Nottingham Statistics Research report 03-18.*

Gelfand A.E., Sahu S.K., and Carlin B.P. (1995). Efficient parametrizations for Normal linear mixed models. *Biometrika,* **82** (3): 479-488.

Gelman A. (2004a) Parameterization and Bayesian modelling. *Journal of the American Statistical Association,* **99** (466): 537-545.

Gelman, A. (2004b). Prior distributions for variance parameters in hierarchical models. Technical report, Department of Statistics, Columbia University.

Gelman, A., Huang, Z., van Dyk, D., and Boscardin, W.J. (2004). Transformed and parameter-expanded Gibbs samplers for multilevel linear and generalized linear models. Technical report, Department of Statistics, Columbia University.

Gilmour, A.R., Thompson, R. and Cullis, B. (1995). AIREML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics*, **51**: 1440-1450.

Goldstein H. (1986). Multilevel mixed linear-model analysis using iterative generalized least-squares. *Biometrika,* **73** (1): 43-56.

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.

Liu, J.S., and Wu, Y.N. (1999) Parameter expansion for data augmentation. *Journal of the American Statistical Association,* **94**: 1264-1274

Patterson, H.D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, **58**: 545-554.

Rasbash, J., Browne, W.J., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., Lewis, T. (2000). *A User's Guide to MLwiN*, Version 2.1, London: Institute of Education, University of London.

Rasbash, J. and Goldstein, H. (1994). Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational and Behavioral Statistics* **19**: 337-350.

Sargent D.J., Hodges J.S., and Carlin B.P. (2000). Structured Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics,* **9** (2): 217-234

Spiegelhalter, D.J., Thomas, A. and Best, N.G. (2000). *WinBUGS version*

1.3: user manual. Cambridge: Medical Research Council Biostatistics Unit.

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*, **64:** 583-640.

Van Dyk, D.A., and Meng, X-L. (2001) The art of data augmentation. *Journal of Computational and Graphical Statistics,* **10:** 1-50.

Review of 'Small Area Estimation'. Rao, J. N. K. (2003). New York: Wiley Series in Survey Methodology. ISBN: 0-471-41374-7 £61.95, pp. 344.

Alastair Leyland
University of Glasgow

This book provides comprehensive coverage of small area estimation. It covers the theory and methods from the most basic estimators to sophisticated model-based estimators using a variety of estimation techniques. The term "small area" in the title is actually a general term for "any sub-population for which direct estimates of adequate precision cannot be produced". This means that the book will be of interest to a broad spectrum of readers, although statisticians working in survey research will remain its primary audience.

Following a brief introductory Chapter, Chapter 2 provides background material on direct domain estimation; that is, it uses values of the variable of interest, y, taken only from sample units in that domain or area. Chapter 3 then covers traditional demographic methods for population estimation which use indirect estimators (estimators that "borrow strength" by using values of y from related areas or time periods).

Chapter 4 considers indirect domain estimation. Synthetic estimators, composite estimators and James-Stein estimators are all introduced for use with sample survey data, typically in conjunction with auxiliary population data.

The bulk of the book – the 207 pages of Chapters 5 to 10 – covers the use of models for small area estimation. Rao lists four advantages that model-based estimation provides: the use of model diagnostics to find models that fit well; the ability to attach measures of precision to each small area estimate; the extension to include non-normal and even complex (e.g. spatially structured) data; and the accurate small area inferences available through random effects models or mixed models.

Chapter 5 introduces the basic area level model and basic unit level model. Some straightforward extensions are also given including spatial models, multilevel models and generalized linear mixed models.

Chapters 6 to 8 then focus on empirical best linear unbiased prediction (EBLUP) models; Chapter 6 covers the theory, Chapter 7 the basic models and Chapter 8 extensions to the basic area level and unit level models. Chapter 6 also describes model estimation and selection using two popular software packages – SAS PROC MIXED and the S-PLUS function *lme*. The special case of block diagonal covariance structures – common in small area estimation – is given particular attention, including the use of transformation method diagnostics and influence diagnostics for the linear mixed model. Chapter 7 applies the theory relating to block diagonal covariance structures from Chapter 6 to the basic area level and unit level models of Chapter 5. The array of extensions to the EBLUP models considered in Chapter 8 include time series and cross-sectional models, spatial models and two level models.

Chapters 9 and 10 then cover empirical Bayes (EB) and hierarchical Bayes (HB) estimation methods. The main extensions of Chapter 9 include linear mixed models, disease mapping, constrained empirical Bayes and empirical linear Bayes (ELB) methods. There is also detailed coverage of the estimation of approximations to the posterior variance of the EB estimator and EB confidence intervals. Chapter 10 provides a brief introduction to Markov Chain Monte Carlo (MCMC) estimation and then covers its application to the basic area level and unit level models along with many extensions including two level models, disease mapping and exponential family models.

This largely theoretical book brings together three facets. Firstly, there is sufficient detail for anyone to apply any of the estimators to their own data. Secondly, the book contains the proofs of many theorems. Finally, methods are illustrated throughout using many varied examples. This book is not for the lazy; the reader is not led gently through the analysis of example data sets. In fact, many of the examples form a narrative as opposed to being quantitative – illustrating the data that have given rise to the use of the estimators without necessarily presenting the results. However, in the 331 references provided there are examples of applications of all of the estimators. This widens the appeal of the book and means that its audience will include those interested in the application and those interested in developing the methodology or the theory of small area estimation. This book will be a welcome addition to my bookshelf.

Some Recent Publications Using Multilevel Models

Baxter-Jones, A. D. G., Faulkner, R. A. and Whiting, S.J. (2003). Interaction between nutrition, physical activity and skeletal health. In New, S.A., and Bonjour, J.P. (Eds.), *Nutritional Aspects of Bone Health*. Royal Society of Chemistry: Cambridge.

Baxter-Jones, A., and Mirwald, R. (2004). Multilevel modelling. In: Hauspie, R.C., Cameron, N., and Molinari, L., (Eds.), *Methods in Human Growth Research*. Cambridge University Press: Cambridge.

Baxter-Jones, A. D .G., Mirwald, R. L., McKay, H. A., and Bailey, D. A. (2003). A longitudinal analysis of sex differences in bone mineral accrual in healthy 8 to 19 year old boys and girls. *Annals of Human Biology*, **30**: 160-175.

Beckman, A., Merlo, J., Lynch, J. W., Gerdtham, U. G., Lindstrom, M., and Lithman, T. (2004). Country of birth, socioeconomic position, and healthcare expenditure: a multilevel analysis of Malmo, Sweden. *Journal of Epidemiology and Community Health*, **58** (2): 145-149.

Beunen, G., Baxter-Jones, A.D.G., Mirwald, R. L., Thomis, M., Lefevre, J., Malina, R. M., and Bailey, D. A. (2002). Intra-individual allometric development of aerobic power in 8 to 16 year-old boys. *Medicine and Science in Sports and Exercise*, **34** (3): 503-510.

Ecob, R., Croudace, T. J., White, I. R., Evans, J. E., Harrison, G. L., Sharp, D., and Jones, P. B. (2004). Multilevel investigation of variation in HoNOS ratings by mental health professionals: naturalistic study of a six-month sample of consecutive referrals to NHS services. *International Journal of Methods in Psychiatric Research*, **13** (3): 152-164.

Johnell, K., Merlo, J., Lynch, J., and Blennow, G. (2004). Neighbourhood social participation and women's use of anxiolytic-hypnotic drugs: a multilevel analysis. *Journal of Epidemiology and Community Health*, **58** (1): 59-64.

Lewsey, J. D. (2004). Comparing completely and stratified randomized designs in cluster randomized trials when the stratifying factor is cluster size: a simulation study. *Statistics in Medicine*, **23**: 897-905.

Merlo, J. (2003). Multilevel analytical approaches in social epidemiology: measures of health variation compared with traditional measures of association. *Journal of Epidemiology and Community Health*, **57** (8): 550-552.

Merlo, J., Asplund, K., Lynch, J., Rastam, L., Dobson, A. (2004). Population Effects on Individual Systolic Blood Pressure: A Multilevel Analysis of the World Health Organization MONICA Project. *American Journal of Epidemiology*, **159** (12): 1168-79.

Merlo, J., Lynch, J. W., Yang, M., Lindstrom, M., Ostergren, P. O., Rasmusen, N. K., Rastam, L. (2003). Effect of neighborhood social participation on individual use of hormone replacement therapy and antihypertensive medication: a multilevel analysis. *American Journal of Epidemiology,* **157** (9): 774-783.

Moerbeek, M., van Breukelen, G. J. P., and Berger, M. P .F. (2003). A comparison of estimation methods for multilevel logistic models. *Computational Statistics*, **18** (1): 19-37.

Moerbeek, M., van Breukelen, G. J. P., and Berger, M. P. F. (2003). A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *Journal of Clinical Epidemiology*, **56** (4): 341-350.

Moerbeek, M., van Breukelen, G. J. P., Ausems, M., and Berger, M. P. F. (2003). Optimal sample sizes in experimental designs with individuals nested within clusters. *Understanding Statistics*, **2** (3): 151-175.

Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research*, **39** (1): 129-149.

Reise, S. P., and Duan, N. (Eds.) (2003). *Multilevel Modelling Methodological Advances, Issues, and Applications*. Lawrence Erlbaum Associates: Mahwah, NJ.

Thompson, A. M., Baxter-Jones, A.D.G., Mirwald, R.L., and Bailey D.A. (2003). Comparison of physical activity levels in male and female children using chronological and biological ages. *Medicine and Science in Sports and Exercise*, **35** (10): 1684-1690.

Van den Noortgate, W., and Onghena, P. (2003). A parametric bootstrap version of Hedges' homogeneity test. *Journal of Modern Applied Statistical Methods*, **2**: 73-79.

Van den Noortgate, W., and Onghena, P. (2003). Combining single-case experimental studies using hierarchical linear models. *School Psychology Quarterly*, **18**: 325-346.

Van den Noortgate, W., and Onghena, P. (2003). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers*, **35**: 1-10.

Van den Noortgate, W., and Onghena, P. (2003). Multilevel meta-analysis: A comparison with traditional meta-analytical procedures. *Educational and Psychological Measurement*, **63**: 765-790.

Please send us your new publications in multilevel modelling for inclusion in this section in future issues.

MLwiN Clinics in London

Wednesday 6 October 2004
Wednesday 3 November 2004
Wednesday 1 December 2004

at

Centre for Multilevel Modelling
11 Woburn Square, London WC1H 0NS

Contact *MLwiN* Technical Support for appointments
Tel: +44 (0) 20 7612 6688
mlwin.support@ioe.ac.uk

Future clinic dates will be announced at:
http://multilevel.ioe.ac.uk/support/clinics.html