D. J. Spiegelhalter, A. Thomas & N. G. Best (2003). *WinBUGS (Version 1.4) User Manual*. MRC Biostatistics Unit, Cambridge.

T. B. Swartz, Y. Haitovsky, A. Vexler & T. Yang (2004). Bayesian identifiability and misclassification in multinomial data. *The Canadian Journal of Statistics*, 32, 285–302.

P. C. van Deusen (2002). An EM algorithm for capture-recapture estimation. *Environmental and Ecological Statistics*, 9, 151–165.

Received 15 October 2006
Accepted 10 July 2007

Saman MUTHUKUMARANA: muthukumaranasm@yahoo.com
Carl J. SCHWARZ: cschwarz@stat.sfu.ca
Tim B. SWARTZ: tim@stat.sfu.ca
Department of Statistics and Actuarial Science
Simon Fraser University, Burnaby
British Columbia, Canada V5A 1S6

Discussion: Towards a Bayesian analysis template?

Olivier GIMENEZ

1. INTRODUCTION

I would like to congratulate Drs Muthukumarana, Schwarz and Swartz (henceforth MSS) for extending classical mark-recapture models to estimate survival in a complex situation arising from fish monitoring. Because of their model likelihood complexity, MSS used Markov chain Monte Carlo (MCMC) simulations to implement a Bayesian analysis of their data. The Bayesian framework in association with MCMC algorithms is becoming increasingly popular for fitting complex models such as models with latent structures. Two of the main reasons for this are that (i) MCMC methods are well suited to circumvent the issue of high-dimensional integrals involved in these likelihoods and, (ii) fast and powerful computers along with flexible and reliable programs are now available allowing the relatively time-realistic and easy implementation of various MCMC routines. This being said, the possibility to fit complex models comes with methodological issues that should not be overcome. In that sense, MSS have provided an impressive work that deserves to be emphasized. In this discussion, I comment on several technical points and give general considerations that were inspired by the MSS paper.

2. TECHNICAL POINTS

2.1. Model selection.

To determine whether a model specifying survival as a function of travel times was better supported by the data than without, MSS relied on BIC as it was easy to implement within WinBUGS. As acknowledged by MSS, this was one choice among many alternatives, and this is precisely the Achilles' heel of Bayesian analyses. Indeed, many procedures exist and none of them seems to be as consensual as AIC is in the statistical ecology literature (Burnham & Anderson 2002). Basically, there are two groups of methods. One produces a value for each model to be compared among a set of models (e.g., mean square predictive error: Gelfand & Ghosh 1998; DIC: Spiegelhalter, Best, Carlin & van der Linde 2002; BIC: for example, Link & Barker 2007), and the other performs automatic exploration of the model space (e.g., Gibbs variable selection: George & McCulloch 1993; reversible jump MCMC: Green 1995). Most

often, we adopt one option or another because it is convenient to calculate, or because we are familiar with it. MSS went for a method belonging to the first family of methods, in line with recent recommendations (Link & Barker 2007). Interestingly, reversible jump MCMC is now implemented in WinBUGS, and MSS could have used it to compare models. This raises the issue of which method to use. Unfortunately, I'm not aware of any comparison or any review that might give clear guidelines. Having made clear what each method does, the performances of several candidate methods could be assessed by calculating frequencies of ranking the true model as best in Monte Carlo simulations.

2.2. Identifiability and convergence issues.

We are often provided with, even in biological papers, technical details regarding convergence of the MCMC algorithms, while it wouldn't come to one's mind to mention anything about convergence in a classical analysis. This is probably because the procedures in the former case have not yet been implemented in an automatic way, and further work is needed in that direction. MSS have paid careful attention to the identifiability issue, and demonstrated that their model was not parameter redundant. Note that formal methods were developed to assess parameter redundancy of probabilistic models that could be used here too (review in Gimenez et al. 2005). Nevertheless, a cause of poor MCMC convergence is weak identifiability (rather than nonidentifiability) because it leads to large autocorrelations. Calculating the overlap between prior and posterior parameter distributions can help in diagnosing weak identifiability (Garrett & Zeger 2000; Gimenez, Morgan & Brooks 2008). Practical recommendations on checking MCMC convergence are given by experienced statisticians in Kass, Carlin, Gelman & Neal (1998) and comparisons of several available methods can be found in El Adlouni, Favre & Bobée (2006).

2.3. Goodness-of-fit testing.

Goodness–of–fit testing has received little attention in the Bayesian literature and only a few methods are available, which are reviewed by MSS: Bayesian p-values, cross-validation, and another approach developed by Box (1980). Here again, I'm not aware of any evaluation of the frequentist properties (nominal level and power) of these methods. Besides, these procedures tend to be 'omnibus', in that the alternative hypothesis is simply stated as 'the model does not fit the data at hand', without any further indication as to where to go then. Once again, the Bayesian approach may benefit from getting closer to a classical framework. Indeed, goodness-of-fit testing procedures are well developed for single (Lebreton, Burnham, Clobert & Anderson 1992) and multistate (Pradel, Wintrebert & Gimenez 2003) mark-recapture models (review in Pradel, Gimenez & Lebreton 2005). These methods rely on contingency tables that specify well-identified alternative hypotheses (e.g., a trapping effect on recapture probabilities or a memory effect on movement probabilities), which, in case of rejection, are invaluable when it comes to building a model which fits the data better.

2.4. Even more complexity?

I have two further minor suggestions that might help to improve the MSS model. First, MSS made the strong assumption that variation in survival could be fully explained by travel times variation. However, if some extra variation exists, then bias may occur in parameter estimates (Barry, Brooks, Catchpole & Morgan 2003). By using a state-space formulation of their model (Gimenez et al. 2007), MSS could have incorporated individual random effects to cope with (potential) unexplained sources of variability in survival (see Clark et al. 2005; Gimenez et al. 2006; Zheng, Ovaskainen, Saastamoinen & Hanski 2007; Royle 2008). Second, MSS wonder how their model might be extended to a two–dimensional spatial setting. We have recently

extended our nonparametric approach dealing with complex relationships between survival and covariates that was mentioned by MSS (Gimenez et al. 2006; see also Gimenez et al. 2006 in the Additional References) to cope with bivariate smoothing (Gimenez & Barbraud 2008). This methodology has also potential applications in ecology in order to estimate spatial synchrony, as well as in evolutionary biology in order to estimate fitness surfaces made of quantitative phenotypic traits.

3. GENERAL CONSIDERATIONS

3.1. Bayes or not Bayes: is that the question?

As a biostatistician, I have long adopted an eclectic and pragmatic approach and have been using either the Bayesian or the frequentist approach based on a few empirical criterions such as the time it takes to get results, the ease of programming and the nature of the biological question (is there any added value of going for a Bayesian analysis?). From a practitioner's point of view, it is worth repeating that, although it may appear obvious, both approaches are complementary, provided that one is careful in using the terminology. As a nice illustration of this statement, I would like to draw attention to the MCMC procedure recently proposed by Lele, Dennis & Lutscher (2007) which has the appealing feature of producing maximum likelihood estimates. It is fair to say that the Bayes approach is rarely used for what it is intrinsically, but rather as an excuse for implementing the MCMC machinery to cope with complex multidimensional likelihoods. Examples of incorporating prior information are still too few (see, however, Martin, Kuhnert, Mengessen & Possingham 2005; McCarthy & Masters 2005), probably due to our feebleness as referees, while every biologist would agree never to start a new data analysis without prior knowledge of the system.

3.2. Transfer to biologists.

Obviously, in a pragmatic approach, the key question is 'why should I go for a Bayesian analysis'. The answer depends obviously on the analyst and the question, but several steps may be taken to help in deciding whether jumping or not into new territory is worth the price.

- We should think more of teaching Bayesian theory in introductory statistics courses, although some colleagues still hesitate to do so. I've taught both Bayesian and frequentist theories this year in a statistical modelling course for Masters students in ecology and evolutionary biology. The discussions were stimulating, focusing mainly on the incorporation of prior information and on when to use one or the other method. I will repeat this next year as I think that this lecture has not only added an arrow to their bow, but it has also contributed to the development of their critical mind.
- Several excellent textbooks are now available which encourage self–teaching, for applied statistician readers (Gilks, Richardson & Spiegelhalter 1996; Lee 1997; Carlin & Louis 2000; Congdon 2003; Gelman, Carlin, Stern & Rubin 2003; Congdon 2006) as well as for biologist readers (Clark 2007; McCarthy 2007), and many others will surely follow. Attending Bayesian workshops is another very efficient way of learning new material, which has the non negligible advantage of keeping us stuck somewhere (often in exotic places) with limited risk of being disturbed.
- To encourage codes and data sharing, we militate with other colleagues for the creation of a statistical ecology internet platform, with a format similar to Genbank (Benson et al. 2007) in genetics, a database to which nucleotides sequences are submitted prior to publication. This web site would gather material (in particular BUGS codes) that have been used in publications, and would avert the too convenient statement 'the code is available upon request from the authors' which I have used myself too often.

- Related to that, user-friendly and reliable pieces of software are needed. WinBUGS (Spiegelhalter, Thomas & Best 2003) is very flexible (Gimenez et al. 2008), but could gain in conviviality (e.g., by improving its debugging capabilities and implementing in routine several simple analyses). The efforts to build a dialog between R and WinBUGS initiated through the R package R2WinBUGS (Sturtz, Ligges & Gelman 2005) should be continued. An alternative to WinBUGS is AD–Model Builder (Fournier 2001), which seems to be much quicker, but is neither free nor open-source, as is WinBUGS.

Overall, teaching (and research) in ecological statistics largely benefits from collaboration between statisticians and biologists.

4. CONCLUSIONS

Hierarchical analysis of data on marked animals (Clark et al. 2005; Pradel 2005; Gimenez et al. 2007; Zheng, Ovaskainen, Saastamoinen & Hanski 2007; Royle 2008) is experiencing an increasing number of applications in ecology, conservation and evolutionary biology, thanks to the Bayesian framework in conjunction with MCMC methods for its implementation. Note that even though this combination has many advantages, I do not mean to overlook other methods that are valuable to fit models with latent structures, such as particle filtering (Buckland, Newman, Thomas & Koesters 2004; Thomas, Buckland, Newman & Harwood 2005), Kalman filtering (Besbeas, Freeman, Morgan & Catchpole 2002) and Newton-type algorithms (Pradel 2005). Even better, we still need to explore other methods since, although the Bayesian framework is more than three centuries old, we have to confess that its practical implementation using MCMC simulations is not as mature as maximum likelihood analyses using standard optimization methods. As Muthukumarana, Schwarz and Swartz acknowledge, potential issues may arise at various steps of the analysis, such as model identifiability, convergence assessment, model selection and goodness-of-fit testing. We see, however, good signs of a trend towards clear guidelines on how to carry out a Bayesian analysis using MCMC algorithms, the paper by MSS being an important contribution in that direction. In that spirit, and besides the original development of new models for fish monitoring, I consider the paper by MSS as a successful attempt to produce a Bayesian analysis template for future data analyses.

ACKNOWLEDGEMENTS

Many thanks to B. J. T. Morgan who encouraged me to do my first steps in the Bayesian theory while I was his research fellow working at the university of Kent, and to Rachel Borysiewicz who checked my English with patience.

Received 7 November 2007
Accepted 7 November 2007

Olivier GIMENEZ: olivier.gimenez@cefe.cnrs.fr
Centre National de la Recherche Scientifique
Centre d'Ecologie Fonctionnelle et Evolutive –UMR 5175
1919 Route de Mende, 34293 Montpellier Cedex 5, France

Authors' response

We thank Gimenez for his commentary on our article—we agree with him on virtually all of his points.