Fitting animal survival models with temporal random effects
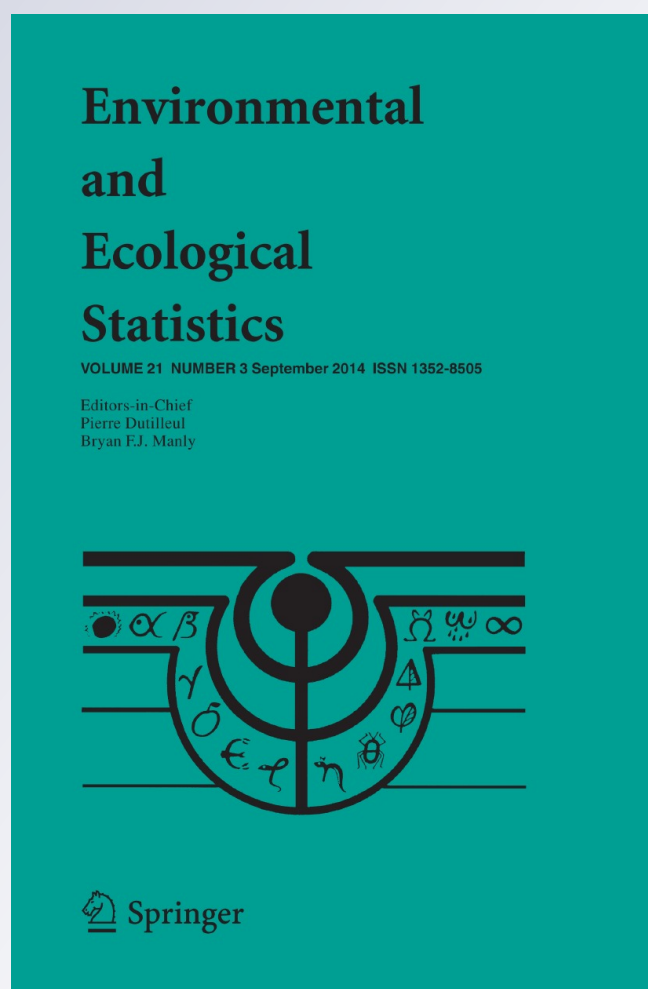
Sarah Cubaynes, Christian Lavergne & Olivier Gimenez

🦄 Springer

Springer

Fitting animal survival models with temporal random effects

Sarah Cubaynes · Christian Lavergne · Olivier Gimenez

Abstract Estimating temporal variance in animal demographic parameters is of particular importance in population biology. We implement the Schall's algorithm for incorporating temporal random effects in survival models using recovery data. Our frequentist approach is based on a formulation of band-recovery models with random effects as generalized linear mixed models and a linearization of the link function conditional on the random effects. A simulation study shows that our procedure provides unbiased and precise estimates. The method is then implemented on two case studies using recovery data on fish and birds.

Keywords Band return · Generalized linear mixed models · R software · Ring recoveries · Schall's algorithm

S. Cubaynes · O. Gimenez (✉)
Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, Campus CNRS,
1919 Route de Mende, 34293 Montpellier Cédex 5, France
e-mail: olivier.gimenez@cefe.cnrs.fr

S. Cubaynes · C. Lavergne
Institut de Mathématiques et Modélisation de Montpellier, Université Montpellier 3,
UMR 5149, Place Eugène Bataillon, 34095 Montpellier Cédex 5, France

C. Lavergne
e-mail: christian.lavergne@univ-montp3.fr

S. Cubaynes
Department of Zoology, University of Oxford,
The Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK
e-mail: sarah.cubaynes@zoo.ox.ac.uk

1 Introduction

Studying animal demography often involves dealing with a non-exhaustive monitoring of populations. For free-ranging animal populations, methods were developed to estimate demographic parameters while accounting for imperfect detection of animals (Lebreton et al. 1992). The applications and developments of these so-called capture–recapture (for alive individuals) and band-recovery models (for dead individuals) have increased exponentially over the last twenty years, and the analysis of data on marked animals has become an important field of biometry (Pollock 2000).

Of particular interest is survival, as it is a key quantity in management of species exploited via fishing or hunting to understand their population dynamics, establish quotas and forecast the effects of harvesting strategies on their demography (Williams et al. 2002). There is also a need to assess the potential interactions of anthropogenic pressure with environmental conditions on survival by considering factors such as time and weather covariates (Pollock 2002; Grosbois et al. 2008).

In this context, survival probabilities are increasingly seen as realizations of a random process rather than fixed constants (Barry et al. 2003; Burnham and White 2002). This enables one to account for unobserved environmental factors (Barry et al. 2003), model association among demographic parameters (Link and Barker 2005), assess nonlinear effects of covariates (Gimenez et al. 2006) or dependence in time-dependent survival parameters (Johnson and Hoeting 2003). Omitting random effects can lead to overestimation of the significance of climatic covariates on survival (Grosbois et al. 2008) and overoptimistic prediction intervals on future population behavior (Barry et al. 2003).

Bayesian methods have been proposed to estimate survival in models with random effects, but despite their flexibility, the use of Markov chain Monte Carlo (MCMC) simulation methods is often time-consuming and their implementation and manipulation may be difficult for non-statisticians (Gimenez 2008). Random effects have been considered in a frequentist framework (e.g., Burnham and White 2002). However, estimation via maximum likelihood remains difficult to perform since to evaluate the likelihood, one needs to integrate over the (possibly highly) multidimensional distribution of random effects, and this exercise has to be repeated for each iteration of an optimization algorithm. Ecological data being collected over several years therefore resulting in long time series, these calculations are slow if not computationally infeasible when temporal random effects need to be included.

To circumvent this issue, we propose an alternative method for incorporating temporal random effects in animal survival models using recovery data, based on the Shall's algorithm (Schall 1991). The novelty of our approach is twofold. First, we introduce a generalized linear mixed model (GLMM) formulation of mixed band-recovery models for the analysis of data on dead animals. The GLM formulation was first proposed by Cooper et al. (1998) but did not include random effects. Second, we implement the Shall's algorithm (Schall 1991) for incorporating temporal random effects in animal survival models using recovery data. The estimation of random effects and temporal variance is based on a linearization of the GLMM conditional on the random effects, made possible by the use of the Schall's algorithm. The advantages of our method include the possibility to estimate fixed and random effects in a frequentist frame-

work, which is usually faster and more accessible to biologists than the Bayesian framework.

To evaluate the performance of our proposal, we conduct a simulation study and compute bias and precision (mean square error). The method is then implemented on two case studies using recovery data on fish and birds.

2 Theory

2.1 Band-recovery models

For K sampling occasions (generally hunting or fishing seasons), $N_1, \ldots, N_K$ individuals are banded (tagged or ringed) and released back into the population (Brownie et al. 1985). The number of bands Y_{ij} recovered in time period j from dead animals originally banded in time period i is then recorded. Expected numbers of band recoveries $E(Y_{ij}) = \mu_{ij}$ can be written as functions of two probabilities: S the survival probability and f the probability that a banded animal, alive when a given cohort is banded, will be harvested and its band reported during the next time period. For example, under a model with band recovery probability f constant over time and survival probability time-dependent S_j, $j = 1, \ldots, K-1$, we have $\mu_{ij} = f$ for $j = i$ and $\mu_{ij} = f \prod_{l=i}^{j-1} S_l$ for $j > i$. Maximum likelihood estimators are traditionally obtained under the assumption of multinomial sampling for each cohort (Brownie et al. 1985). Alternatively, a product-Poisson sampling assumption can be invoked as a product-multinomial distribution can be approximated by a product-Poisson distribution when fixed sample size $N \rightarrow +\infty$ (Sandland and Cormack 1984). This equivalence allows band-recovery models to be formulated within the general framework of GLMs (Cooper et al. 1998). Conditional on N_i, the y_{ij}'s are considered as realizations of independent random variables Y_{ij}. These variables are Poisson distributed with parameter $\mu_{ij} > 0$. Using the log link function, we have:

$$\log(E(\mathbf{Y})) = \log(\boldsymbol{\mu}) = \log(\mathbf{N}) + \mathbf{X}\beta \tag{1}$$

where $\boldsymbol{\mu}$ is a vector of expected band recovery probabilities, N a vector of offsets made of the cohort sizes, X is an incidence matrix and β a vector of the model parameters on the log scale. X has one row for each recovery cell and one column for each element of the parameters vector β. The columns of X define the model being fitted with each entry of X giving the number of times the corresponding element of β occurs within the corresponding recovery cell expectation. For example, consider a band recovery dataset involving three years of releases and recoveries. To estimate the parameters of a model with constant recovery and survival probabilities, one needs $\boldsymbol{\mu} = (\log(f), \log(f) + \log(S_1), \log(f) + \log(S_1) + \log(S_2), \log(f), \log(f) + \log(S_2), \log(f))'$, $\mathbf{N} = (N_1, N_1, N_1, N_2, N_2, N_3)'$, $\boldsymbol{\beta} = (\beta_f, \beta_S)'$ where $\beta_S = \log(S_1) = \log(S_2)$ and

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 & 0 \end{pmatrix}'$$

where the prime means transpose. As in standard GLM analyses, maximum likelihood estimates of the β's are obtained via iteratively weighted least squares (IWLS; McCullagh and Nelder 1989).

2.2 Incorporation of time-varying random effects on survival

We introduce random effects on survival in band-recovery models via the usual formulation:

$$\log(E(\mathbf{Y})) = \log(\boldsymbol{\mu}) = \log(\mathbf{N}) + \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\boldsymbol{\xi} \tag{2}$$

where $\boldsymbol{\xi}$ is the vector of random effects with $\boldsymbol{\xi} \sim N(0, \mathbf{D})$ and $\mathbf{D} = \sigma^2 \mathbf{I}_{K-1}$ (with $\mathbf{I}_{K-1}$ the identity matrix with $K - 1$ rows and columns) and its associated incidence matrix $\mathbf{U}$. $\mathbf{U}$ has one row for each recovery cell and one column per recovery occasion. Each entry of $\mathbf{U}$ gives the survival probability to which a random effect applies. Note that the probability of μ_{ii} recovery cells does not involve survival, hence the corresponding row in $\mathbf{U}$ is made of 0's. For example, coming back to the previous example with three years of releases and recoveries, one would use $\xi = (\xi_1, \xi_2)'$ and

$$\mathbf{U} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}'.$$

The Poisson assumption now holds conditional on the random effects, so that we have $E(\mathbf{Y}|\boldsymbol{\xi}) = \text{var}(\mathbf{Y}|\boldsymbol{\xi}) = \mu_{ij}$. Parameter estimation in GLMMs is made difficult as random effects are not directly observed. Because the distribution of Y is known conditionally on ξ, the likelihood is obtained via integration over the multivariate distribution of the random effects. Its calculation is difficult unless simple situations are considered (e.g., Gimenez and Choquet 2010). Here, we adopted the Schall method (Schall 1991) based on a linearization of the model conditional on the random effects to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ using the Henderson's equations for mixed models (Henderson et al. 1959).

Schall's estimation approach is an iterative procedure considering at each step the GLMM conditional on the random effects. A working variable is introduced as in IWLS for GLM, then the Henderson's mixed model equations are solved for the associated linear mixed model. The key is that, at each step, this iterative procedure introduces a linearization of the model with current parameter values. This leads to the following iterative algorithm for the estimation of fixed effects and variance components in a GLMM (t denotes the iteration step):

- Step 1 : Denoting μ_ξ the conditional mean, and given the linear predictor $\eta^{[t]} = g(\mu^{[t]}/N) = X\beta^{[t]} + U\xi^{[t]}$ with g being the log function here and N the offset, we define the working variable $z^{[t]}$ as in the IWLS estimation procedure for GLM by linearizing the link function:

$$\begin{aligned} z^{[t]} &= \eta^{[t]} + (y - \mu_\xi^{[t]})g'(\mu_\xi^{[t]}) \\ &= \eta^{[t]} + (y - \mu_\xi^{[t]})/\mu_\xi^{[t]}. \end{aligned} \tag{3}$$

- Step 2 : Consider the linearized model $M^{[t]}$ for $z^{[t]}$:

$$M^{[t]} : z^{[t]} = X\beta^{[t]} + U\xi^{[t]} + \epsilon^{[t]}$$

with residual variance matrix $W^{[t]}$ defined as:

$$
\begin{aligned}
W^{[t]} &= \text{var}(\epsilon^{[t]}|\xi) \\
&= \text{var}\left(\left(Y - \mu_\xi^{[t]}\right) g'\left(\mu_\xi^{[t]}\right) |\xi\right) \\
&= \text{diag}\left\{ \text{var}\left(Y_{ij}|\xi\right) g'\left(\mu_{\xi,ij}^{[t]}\right)^2 \right\}_{i=1,\ldots,K-1,\, j=i,\ldots,K-1} \\
&= \text{diag}\left\{ 1/\left(N_i \mu_{\xi,ij}^{[t]}\right) \right\}.
\end{aligned}
$$

- Step 3 : Solve the Henderson's mixed model equations associated with $M^{[t]}$ now considered as a linear mixed model:

$$
\begin{pmatrix} X'W^{[t]-1}X & X'W^{[t]-1}U \\ U'W^{[t]-1}X & U'W^{[t]-1}U + D^{[t]-1} \end{pmatrix} \begin{pmatrix} \beta \\ \xi \end{pmatrix} = \begin{pmatrix} X'W^{[t]-1}z^{[t]} \\ U'W^{[t]-1}z^{[t]} \end{pmatrix}
$$

to obtain $\beta^{[t+1]}$ and $\xi^{[t+1]}$ the solutions of the system.
- Step 4 : Use the value $\xi^{[t+1]}$ obtained at the previous step to calculate $\sigma^{2[t+1]}$ (Searle et al. 1992) using either restricted maximum likelihood for estimation or maximum likelihood for model selection.
  (a) The restricted maximum likelihood approach gives:

$$\sigma^{2[t+1]} = \frac{\xi'^{[t+1]}\xi^{[t+1]}}{q - \text{trace}\left(C^{[t]}\right)/\sigma^{2[t]}}$$

  with C the matrix formed using the last q rows and columns of the inverse of the matrix of the Henderson system where q is the number of random effects ($q = K - 1$ here).
  (b) The maximum likelihood approach gives:

$$\sigma^{2[t+1]} = \frac{\xi'^{[t+1]}\xi^{[t+1]}}{q - \text{trace}(C^{*[t]})/\sigma^{2[t]}}$$

  with C^* the inverse of the matrix formed using the last q rows and columns of the matrix of the Henderson system.
  This value of $\sigma^{2[t+1]}$ is then used to calculate $z^{[t+1]}$, $W^{[t+1]}$ and $D^{[t+1]}$.

All steps are repeated until convergence, i.e. $|\sigma^{2[t+1]} - \sigma^{2[t]}| < \epsilon$ and $|\beta^{[t+1]} - \beta^{[t]}| < \epsilon$ where $\epsilon = 10^{-5}$ here. We therefore obtained best linear unbiased estimator for β and best linear unbiased predictor for ξ. Using the final linearized model at convergence of the Schall's estimation algorithm, standard errors can be obtained using the variance–covariance matrix of the estimators of a linear mixed model (Searle et al. 1992).

3 Simulation study

To provide empirical support for our approach, we conducted a simulation study. We considered a model with random effects on survival $\log(S_j) = \beta_S + \epsilon_j$ with $\epsilon_j \sim N(0, \sigma^2)$ for $j = 1, \ldots, K - 1$. We used $\beta_S = -0.92$ so that mean survival was 0.4 while recovery probability was set constant and equal to 0.3.

We compared two levels of temporal variance on survival (on a log scale), low with $\sigma^2 = 0.01$ and high with $\sigma^2 = 0.1$. For each of the two levels of temporal variation, we simulated 250 recovery datasets using 10 different numbers of recovery occasions, from $K = 20$ up to $K = 40$, and 10 different numbers of individuals banded each time, from $N = 250$ up to 1000. In total, we considered 100 scenarios for each value of σ^2. To each dataset, we fitted the random-effect band-recovery model using Schall's approach and computed parameter bias and mean square error (MSE). For the sake of comparison, we also fitted the same band-recovery model without random effects.

For low temporal variance ($\sigma^2 = 0.01$), bias in Schall's estimates was negligible and MSE was uniformly low for all parameters (Fig. 1).

For high temporal variance ($\sigma^2 = 0.1$), bias was also negligible and MSE decreased as the number of released individuals increased (Fig. 2).

Overall, the performance of Schall's approach was acceptable. Noteworthy, the computation of Schall's estimates was very fast, almost instantaneous.
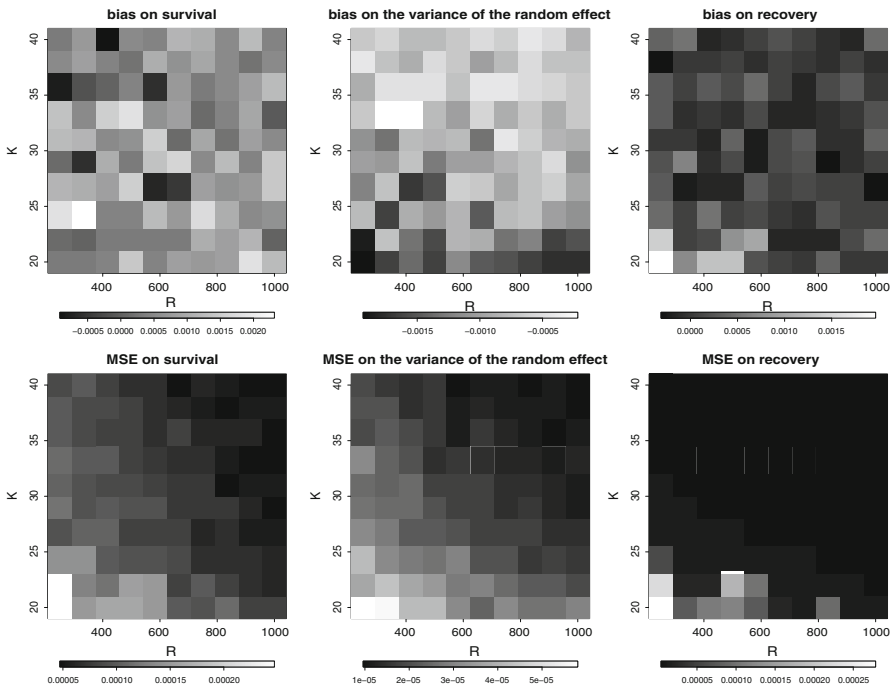
Fig. 1 Bias (*top panels*) and mean square error (MSE, *bottom panels*) for mean survival probability $\exp(\beta_S)$ (*left column*), variance of the random effect (*middle column*) and recovery probability (*right column*) as a function of the number of recovery occasions (Y-axis, K) and the number of released individuals (X-axis, R); low temporal variance $\sigma^2 = 0.01$
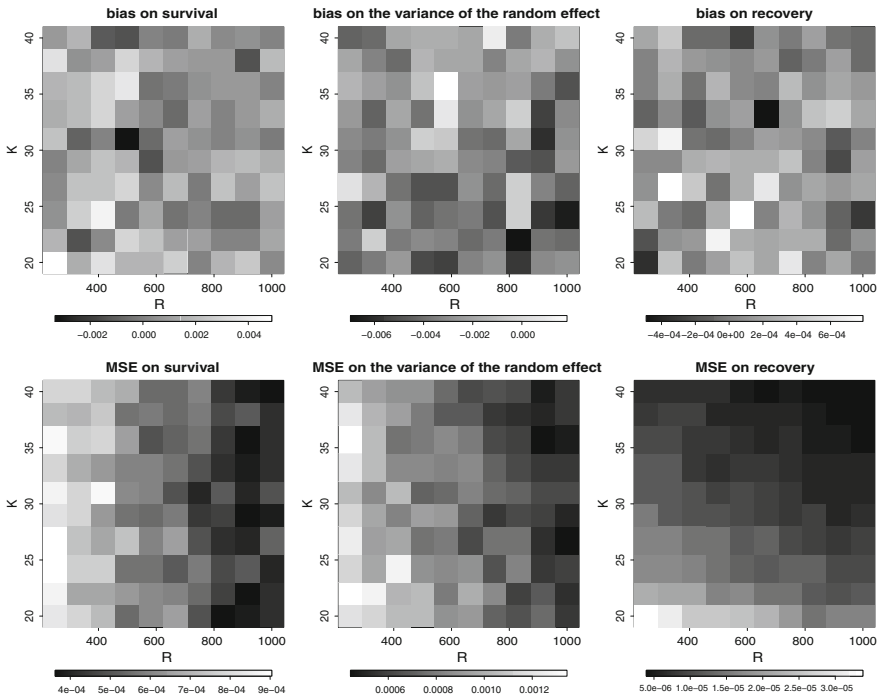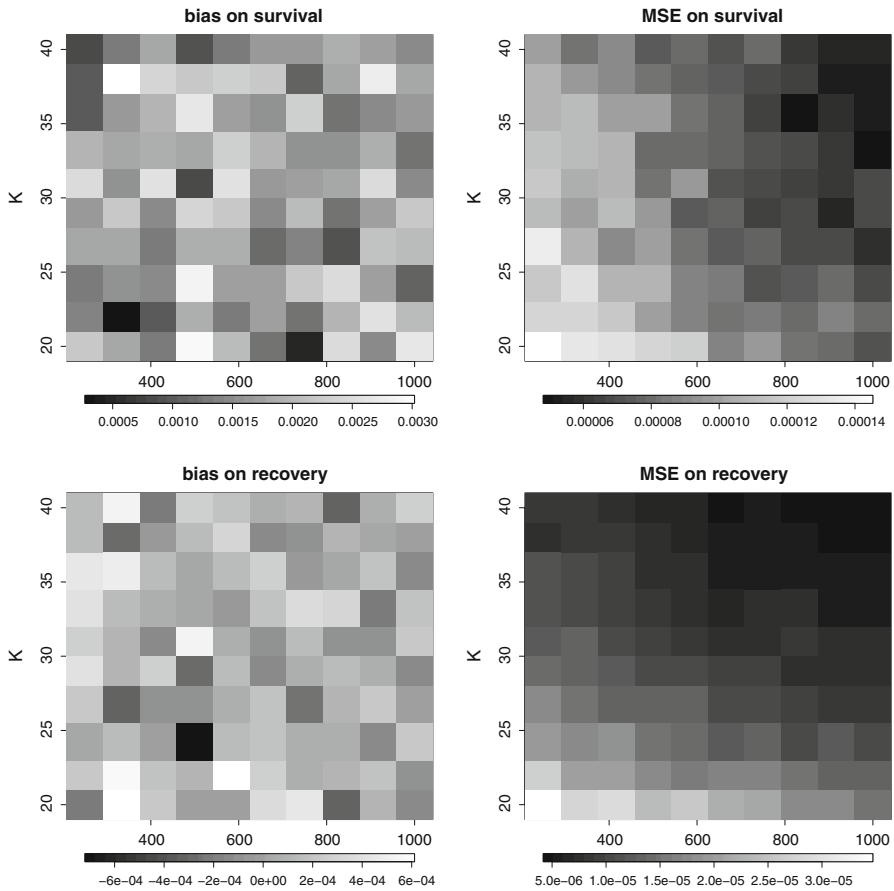
Fig. 2 Bias (*top panels*) and mean square error (MSE, *bottom panels*) for mean survival probability $\exp(\beta_S)$ (*left column*), variance of the random effect (*middle column*) and recovery probability (*right column*) as a function of the number of recovery occasions (Y-axis, K) and the number of released individuals (X-axis, R); high temporal variance $\sigma^2 = 0.1$

Last but not least, we assessed the effect of ignoring temporal variation by fitting a model without random effects. As expected, when temporal variance was low, parameter estimates were little affected meaning that bias and precision were comparable to estimates obtained from the model with random effects (Fig. 3). In contrast, when temporal variance was high, the bias in the survival probability estimate was non negligible, almost 10 times as big as the bias from the model with random effect (Fig. 4), hence confirming previous results about the risks of ignoring extra-residual temporal variation (Barry et al. 2003).

4 Applications

We applied the approach to data from two studies on fish and birds. We first considered a geographically closed population of trout in which adults were tagged during fall spawning, and subsequently reported killed by fisherman, from 1960 to 1969 (Youngs and Robson 1975). The total number of fish tagged was 7,770, with a yearly range of between 360 and 1,048. The total number of recoveries was 865. Second, we analyzed data on adult male mallards based on summer banding, and subsequent recoveries by hunters, from 1955 to 1996 (Franklin et al. 2002). The total number of birds ringed
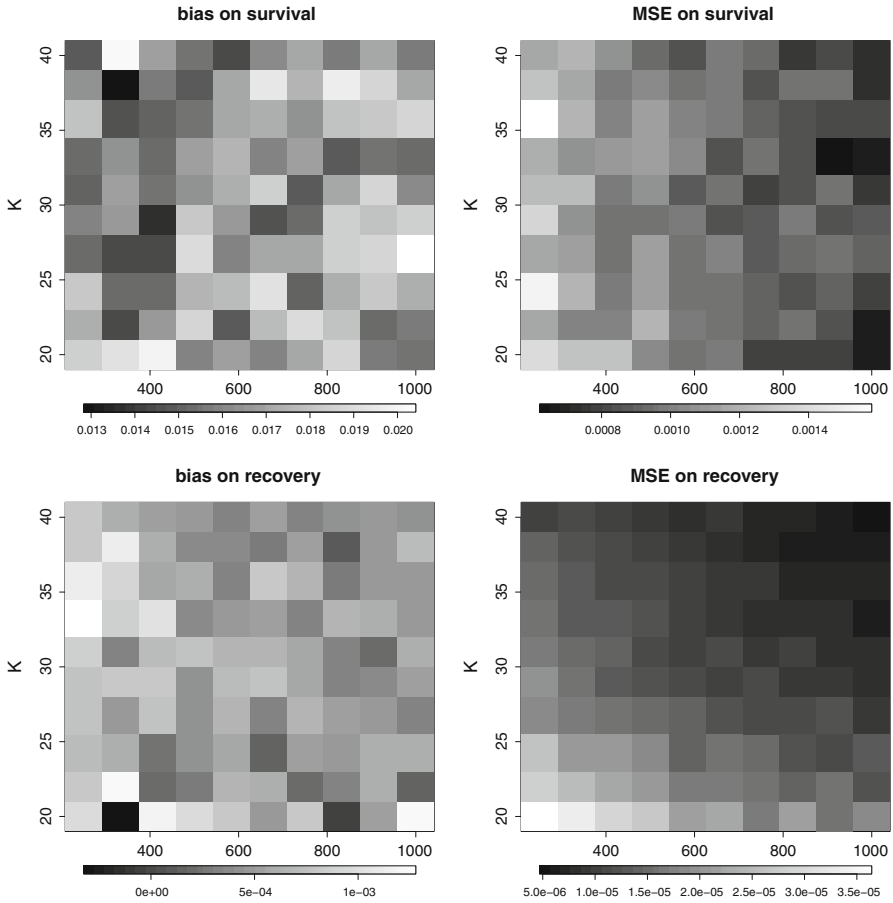
Fig. 3 Bias (*left column*) and mean square error (MSE, *right column*) for mean survival probability $\exp(\beta_S)$ (*bottom panels*) and recovery probability (*top panels*) as a function of the number of recovery occasions (Y-axis, K) and the number of released individuals (X-axis, R); low temporal variance $\sigma^2 = 0.01$

was 42,015, with a yearly minimum of 268 and a yearly maximum of 2,279. The total number of recoveries was 7,647.

Raw data and R codes to fit the model are provided in Supplementary material. The results of fitting a model with constant recovery probability and a yearly random effect on survival are provided in Table 1. Recovery probability was 7 % for mallards and 6 % for trout. Survival was 64 % for mallards, a figure very similar to that found by Franklin et al. (2002) for waterfowls using shrinkage estimators for random-effect recovery models (Burnham and White 2002). Survival was 51 % for trout, which was comparable to the results obtained by Youngs and Robson (1975) in their seminal paper on tag-recovery models.

For both mallards and trout, we found relatively low temporal variation, as previously reported for the mallards by Franklin et al. (2002). A question of paramount importance for biologists is whether temporal variation is significant. This question

Fig. 4 Bias (*left column*) and mean square error (MSE, right column) for mean survival probability $\exp(\beta_S)$ (*bottom panels*) and recovery probability (*top panels*) as a function of the number of recovery occasions (Y-axis, K) and the number of released individuals (X-axis, R); high temporal variance $\sigma^2 = 0.1$

Table 1 Parameter estimates for the random-effect band-recovery model applied to the mallard and trout datasets

	Mallard ($K = 42$)	Trout ($K = 10$)
f	0.068 (0.001)	0.057 (0.002)
S	0.640 (0.032)	0.511 (0.016)
σ^2	0.014 (0.002)	0.078 (0.031)

Schall's estimates (with standard errors) are given for recovery probability f, mean survival probability S and temporal variance σ^2

can be addressed using a model selection approach by comparing the models with and without the time random effect. To discriminate among models, we defined a modified Akaike information criterion following Lavergne et al. (2008) using the Gaussian log likelihood corresponding to the linear mixed model for the final working data, say

$z^{[f]}$, as $AIC = n \log(2\pi) + \log(|\hat{\Gamma}|) + (z^{[f]} - X\hat{\beta})'\hat{\Gamma}^{-1}(z^{[f]} - X\hat{\beta}) + 2p$ where n is the number of recovery cells, p is the number of parameters plus one for the random effect and $\hat{\Gamma} = \hat{W} + U\hat{D}U'$ where σ^2 is estimated using maximum likelihood (Step 4b of the Schall's algorithm in Sect. 2.2).

When applied to the trout dataset, we found that a model with time treated as a fixed effect on survival was better supported by the data (AIC = 124.6) than a model with a yearly random effect (AIC = 131.2) or survival constant over time (AIC = 142.6). In the mallard analysis, we found that the model with random effect on survival was best supported by the data (AIC = 4,523.2) when compared to a model with time as a fixed effect (AIC = 4,551) or no time variation on survival (AIC = 4,566.6).

From a biological point of view, even though temporal variance was statistically significant using the AIC in the mallards, its value was ecologically-speaking rather low. This result is in agreement with the hypothesis of canalization (Gaillard and Yoccoz 2003) that predicts little impact of environmental variation on adult survival of long-lived species. In contrast, we found a fixed temporal effect on survival of trout that might be related to environmental forcing (Grosbois et al. 2008) and would deserve further investigation.

5 Discussion

Estimating temporal variance in demographic rates is of particular relevance to population biology as it relates to important questions in ecology, evolution and conservation biology. Furthermore, models without random effects can overestimate the significance of covariates on demographic rates (Barry et al. 2003) as residual occasion-to-occasion variation is not accounted for.

We provide a frequentist alternative to the Bayesian approach using the Schall's algorithm to include random effects and carry out inference in survival models using recovery data. The frequentist framework being familiar to all biologists, we expect this approach will be easily adopted by them. Our method performed well in terms of estimation as bias on parameter estimates was negligible, and mean square error showed that the error surrounding these estimates was low. Because models were fitted quasi-instantaneously, our approach provides a relevant alternative to the Bayesian approach implemented with MCMC simulations methods (Barry et al. 2003), which can prove time-consuming especially when the number of recovery occasions is high as in the mallard example. We acknowledge, however, that recent developments in the Bayesian field, e.g. the integrated nested laplace approximation (Rue et al. 2009), might compete well with the Schall's algorithm. Another appealing feature of our method is its easy implementation in program R, which is familiar to most population ecologists, using the codes provided in the Supplementary material.

Incorporating covariates measured at the occasion level, like a linear trend over time or climatic conditions, may be useful to make biological sense of the temporal variation in survival, and can easily be incorporated through the X matrix in Eq. 1. Having a random effect in addition to the covariate allows modeling residual variation that is not explained by the covariate alone and formally testing for the significance

of the covariate (Grosbois et al. 2008). Last but not least, allowing for random variation in the survival rates is a natural way to accomodate overdispersion (Barry et al. 2003), a phenomenon that often occurs in data on marked indidivuals (Lebreton et al. 1992).

Throughout this paper, we have considered random effects on the survival probabilities. However, it might be useful to also include them on the recovery probabilities to conveniently account for temporal variation in the nuisance parameter. As warned by Barry et al. (2003) however, we recommend caution when using random effects on both the survival and recovery probabilities because large correlations between estimates of these parameters may occur, which makes it difficult to disentangle the relative effect of temporal variation.

This work has wider applications than just to band-recovery models. Indeed, pending that the model can be expressed as a GLM, the log-linear modelling framework can be used to estimate abundance and time trend in population size as estimated in closed population models or survival based on live recaptures for which the date of death is unknown contrary to recovery data (Cormack 1989; Rivest and Levesque 2001).

References

Barry SC, Brooks SP, Catchpole EA, Morgan BJT (2003) The analysis of ring-recovery data using random effects. Biometrics 59:54–65

Brownie C, Anderson DR, Burnham KP, Robson DS (1985) Statistical inference from band recovery data: a handbook, Second edition. U.S. Fish and Wildlife Service, Resource Publication No. 156. 305 pp

Burnham K, White G (2002) Evaluation of some random effects methodology applicable to bird ringing data. J Appl Stat 29:245–264

Cooper MC, Bonett DG, El-Saidi MA (1998) A general log-linear model for the analysis of bad recovery data. Environmetrics 9:377–391

Cormack RM (1989) Loglinear models for capture–recapture. Biometrics 45:395–413

Franklin AB, Anderson DR, Burnham KP (2002) Estimation of long-term trends and variation in avian survival probabilities using random effects models. J Appl Stat 29:267–287

Gaillard J-M, Yoccoz N (2003) Temporal variation in survival of mammals: a case of environmental canalization? Ecology 84:3294–3306

Gimenez O (2008) Discussion: towards a bayesian analysis template? Can J Stat 36:21–24

Gimenez O, Choquet R (2010) Incorporating individual heterogeneity in studies on marked animals using numerical integration: capture–recapture mixed models. Ecology 91:951–957

Gimenez O, Crainiceanu C, Barbraud C, Jenouvrier S, Morgan B (2006) Semiparametric regression in capture–recapture modelling. Biometrics 62:691–698

Grosbois V, Gimenez O, Gaillard J-M, Pradel R, Barbraud C, Clobert J, Mller A, Weimerskirch H (2008) Assessing the impact of climate variation on survival in vertebrate populations. Biol Rev 83:357–399

Henderson CR, Kempthorne O, Searle SR, VonKrosig C (1959) Estimation of environmental and genetic trends from records subject to culling. Biometrics 15:192–218

Johnson DS, Hoeting JA (2003) Autoregressive models for capture–recapture data: a Bayesian approach. Biometrics 59:341–350

Lavergne C, Martinez M-J, Trottier C (2008) Empirical model selection in generalized linear mixed effects models. Comput Stat 23:99–110

Lebreton J-D, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol Monogr 62:67–118

Link WA, Barker RJ (2005) Modeling association among demographic parameters in open population capture–recapture data. Biometrics 61:46–54

McCullagh P, Nelder JA (1989) Generalized linear models, 2nd edn. Chapman and Hall, London

Pollock KH (2000) Capture–recapture models. J Am Stat Assoc 95:293–296

Pollock KH (2002) The use of auxiliary variables in capture–recapture modelling: an overview. J Appl Stat 29:85–102

Rivest LP, Levesque T (2001) Improved log-linear model estimators of abundance in capture–recapture experiments. Can J Stat 29:555–572

Rue H, Martino S, Chopin N (2009) Approximate bayesian inference for latent gaussian models using integrated nested laplace approximations (with discussion). J R Stat Soc Series B 71:319–392

Sandland RL, Cormack RM (1984) Statistical inference for Poisson and multinomial models for capture–recapture experiments. Biometrika 71:27–33

Schall R (1991) Estimation in generalized linear models with random effects. Biometrika 78:719–727

Searle S, Casella G, McCulloch C (1992) Variance components. Wiley, New York

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press, San Diego, CA

Youngs WD, Robson D (1975) Estimating survival rate from tag returns: model tests and sample size determination. J. Fish. Res. Board Can 32:2365–2371

Sarah Cubaynes is a statistical ecologist now a post-doc at University of Oxford.

Christian Lavergne is a statistian professor at University of Montpellier 2.

Olivier Gimenez is a biostatistician senior researcher at CNRS, Montpellier.