

# *Towards built-in capture–recapture mixed models in program E-SURGE*

**R. Choquet & O. Gimenez**

**Journal of Ornithology**

ISSN 2193-7192

Volume 152

Supplement 2

J Ornithol (2012) 152:625-639

DOI 10.1007/s10336-010-0613-x



**Your article is protected by copyright and all rights are held exclusively by Dt. Ornithologen-Gesellschaft e.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

## Towards built-in capture–recapture mixed models in program E-SURGE

R. Choquet · O. Gimenez

Received: 30 September 2009/Revised: 13 October 2010/Accepted: 1 November 2010/Published online: 27 November 2010  
© Dt. Ornithologen-Gesellschaft e.V. 2010

**Abstract** We consider the first steps towards implementing capture–recapture mixed models (CR2Ms) in program E-SURGE. The main issue when estimating the parameters of mixed models is that integrals associated with the random effects distributions need to be dealt with. Rather than using a Bayesian approach with Markov chain Monte Carlo and in line with Gimenez and Choquet (Ecology 91:951–957, 2010), we show that a frequentist approach using numerical integration can be tractable when independent clusters of individuals can be identified. In this case, the maximum likelihood approach is time-efficient because the dimension of the integral for the likelihood is small. This allows us to integrate the likelihood by an efficient and appropriate quadrature method with a procedure for error control. Building on program E-SURGE (Choquet et al. in Modeling demographic processes in marked populations, volume 3 of Springer series: environmental and ecological statistics. Springer, Dunedin, 2009b), we extend the GEMACO language (Choquet in Can J Stat 36:43–57, 2008) to incorporate random effects in a large set of capture–recapture models, including multievent models (Pradel in Biometrics 61:442–447, 2005). To illustrate the flexible implementation of CR2Ms in E-SURGE, we consider two real examples, one with an individual random effect and one with group random effects. Future developments and limitations are also discussed.

**Keywords** Adaptive method · Cluster of individuals · Gauss-Hermite quadrature · GLMM · Hidden Markov chain · Normal random variable · State uncertainty

### Introduction

With the improvement of Markov Chain Monte Carlo (MCMC algorithms), the development of new software (see Lunn et al. 2000, for WinBUGS) and the availability of powerful computers, mixed models are becoming more and more popular in ecological science (Bolker et al. 2009), and particularly in capture–recapture (CR) (Wintrebert et al. 2005; Link and Barker 2005; Gimenez et al. 2009). A non-exhaustive list of advantages is:

- The possibility to explicitly decompose the response variance into several components.
- The possibility to model dependence among parameters.
- That boundary estimates are avoided for small sample size datasets.

In the CR research area, mixed models remain under utilized despite their potential. One reason is that practitioners lack sufficient statistical training to code (or pseudocode in the case of WinBUGS) Bayesian implementations of mixed effects models. In non-Bayesian applications, Burnham and Anderson (2002) and Royle and Link (2002) considered the use of shrinkage estimators for time random effects. However, the mandatory assumption of non-boundary time fixed effect estimates are often violated, making this approach difficult to implement for a general class of model. Moreover, it is not applicable to the scale of individual effect as individual fixed effects estimates are not available. In addition, there are computation challenges in

---

Communicated by M. Schaub.

---

R. Choquet (✉) · O. Gimenez  
CEFE, campus CNRS, UMR 5175, 1919 Route de Mende,  
34293 Montpellier cedex 5, France  
e-mail: remi.choquet@cefe.cnrs.fr

non-Bayesian applications, such as numerical integration. This paper addresses the latter situation.

In this paper, we propose a first step toward implementing CR mixed models (CR2Ms) in E-SURGE. We will only consider a simple case with independent and identically distributed random effects. We use the independence property to reduce the dimension of the integral associated with the marginal likelihood. Thus, because fitting problems become more and more complex, in the ideal case, we would like to consider an adaptive approach where the selected algorithm is suited to the model of interest. Developing adapted algorithms has proven in the past to be efficient in M-SURGE (Choquet et al. 2004) to deal with models with complex age structures. The situation here is even more challenging, and the road will be quite long before achieving this goal.

In line with Coull and Agresti (1999), McClintock et al. (2009) and Gimenez and Choquet (2010), we consider a numerical approximation using Gauss-Hermite quadrature to calculate the marginal likelihood which has no analytical solution in presence of random effects. Gaussian quadrature is known to work efficiently for low dimension integrals over a large class of problems (Lemuel-Diot et al. 2005; Heiss and Winschel 2008; Arndt et al. 2006). Numerical integration using Gauss-Hermite quadrature was used in Gimenez and Choquet (2010) to fit the Cormack-Jolly-Seber (CJS) model with individual random effects. We implement this approach for multi-event models (Pradel 2005) which are extensions of multistate CR models to handle state uncertainty. These models are implemented in E-SURGE (Choquet et al. 2009b). In addition, we have considered an adaptive Gauss-Hermite quadrature to get efficient algorithms for group random effects. To get built-in random effects, we also extend the language in GEMACO (Choquet 2008) to automatically build matrices of constraints so that mixed models can be specified with simple phrases.

We show how E-SURGE makes it is easy to fit CR2Ms using two applications. The first application illustrates a model accounting for a potential effect of the observation process with individual random effects applied to an European Dipper *Cinclus cinclus* dataset. The second application illustrates a model with group random effects applied to a Bank Vole *Myodes glareolus* dataset.

**Notation**

- $NI$  is the number of individuals,
- $G$  is the number of groups,
- $H$  is the dataset of capture histories,
- $I_s$  is the identity matrix of size  $s$ ,

$\mathcal{N}(\mu, \sigma^2)$  is the normal distribution with mean  $\mu$  and variance  $\sigma^2$ ,

$\Theta$  is the vector of biological parameters,  $\Theta_i$  the vector of biological parameters for individual  $i$  inside which the different biological parameters (survival, transition, capture, ...) are set in a row; see Choquet et al. (2009b) for details.

**Models under consideration**

We consider the background of multievent models (Pradel 2005) to handle state uncertainty in analyzing CR data, which are already implemented in E-SURGE. These models belong to the class of Hidden Markov Models (HMM) as the number of states is assumed finite (Cappe et al. 2005), and consequently to the class of State-Space Models (SSM) (see Murphy 2002, see Gimenez et al. 2007, for an application to CR data).

Generalized linear mixed model

The class of mixed effects models that E-SURGE may consider can be expressed in the form of generalized linear mixed models (GLMM). We consider, for both fixed and random effects, two general sets of effect:

- Set of effect 1: time, age, cohort and group effects.
- Set of effect 2: individual effect.

The general form of GLMM  $f(\Theta) = X\beta + Zb$ , with  $\beta$  the vector of fixed effects and  $b$  the vector of random effects, is computationally demanding because of the dimension of the problem with so many potential effects. Thus, we have implemented the following restricted form of GLMM by constraining separately categories 1 and 2.

$$f(\Theta_i) = X_0\beta_0 + X_i\beta_1 + \sum_{l=1}^L Z_l b_l + \sum_{l=L+1}^{L+P} Z_{l,i} b_{l,i} \quad i = 1, \dots, NI \tag{1}$$

where  $b_l \in \mathbb{R}^{s_l}$  and  $b_{l,i} \in \mathbb{R}$  are random effects given by

$$\begin{cases} b_l \sim \mathcal{N}(0, \sigma_l^2 \times I_{s_l}), l = 1, \dots, L, \\ b_{l,i} \sim \mathcal{N}(0, \sigma_l^2), l = L + 1, \dots, L + P. \end{cases} \tag{2}$$

$b_l, l = 1, \dots, L$  are random effects associated with the set of effects 1,  $s_l$  is the number of levels of the random effect  $l$  ( $s_l = G$  for a group random effects),  $b_{l,i}, l = L + 1, \dots, L + P$  are individual random effects assuming that individuals are independent. Matrices  $X_i$  are individual-specific matrices of individual covariates. They are never stored in the computer because of the memory size

needed but rather they are computed each time (see Appendix 1). In the same way, matrices  $Z_{l,i}$  contain either 0, 1 or values of individual covariates and are never stored. Because we assume that individuals are independent, then covariance matrices for each random effect are diagonal. We use this property to implement efficient algorithms.

General expression of the marginal likelihood

Assuming that individuals are independent, the likelihood for fixed effects for the entire set of capture histories is obtained as the product of the probability  $P(h_i|\beta)$  of the likelihood for each history  $h_i$

$$\prod_{h_i \in H} Pr(h_i|\beta),$$

see Pradel (2005) for details on multievent models and Choquet et al. (2009b) for the implementation.

For a GLMM like Eq. 1, we get the marginal likelihood:

$$L(\beta, \sigma) = \int \prod_{h_i \in H} Pr(h_i|\beta) \omega(\sigma, \mathbf{x}) d\mathbf{x}. \tag{3}$$

where  $\sigma$  is the vector of variances associated to the random effects and  $\omega(\sigma, \mathbf{x})$  is a product of normal densities associated to random effects  $b_l$  or  $b_{l,i}$ .

One might think that the dimension of the integral in Eq. 3 is equal to the size of the random effects and in most cases the evaluation of  $L(\beta, \sigma)$  is numerically intractable. However, we illustrate below two situations for which the computation of Eq. 3 in E-SURGE can be made by reducing the dimension of the integral.

Independent and identically distributed random effect for individual

E-SURGE can now handle mixed models with individual random effects only ( $L = 0$ ):

$$f(\Theta_i) = X_0\beta_0 + X_i\beta_1 + \sum_{l=1}^P Z_{l,i}b_{l,i} \tag{4}$$

with  $b_{l,i}$  in the form of Eq. 2. There is no limit for the number of random effects that we can build. However for  $P > 2$  the fitting step may be time consuming.

As in Gimenez and Choquet (2010), the marginal likelihood (3) is the product of the probabilities of all individual encounter histories. So Eq. 3 can be rewritten:

$$L(\beta, \sigma) = \prod_{h \in H} \int_{\mathbb{R}^P} Pr(h|\beta, \mathbf{x}) \omega(\sigma, \mathbf{x}) d\mathbf{x}. \tag{5}$$

**Example 1** Survival varying with individual covariates and random effect. The following model has been used in Gimenez et al. (2006) with a constant survival across time

but dependent on an individual covariate (body weight denoted by  $m$ ) and from an individual random effect ( $b_i$ ).

$$\text{logit}(\phi_i) = \beta_0 + \beta_1 m_i + b_i, \quad i = 1, \dots, I \tag{6}$$

where each  $b_i \sim \mathcal{N}(0, \sigma_b^2)$  is independent and identically distributed (i.i.d). In Eq. 6,  $P = 1$ , the computation of the marginal likelihood according to Eq. 5 leads to the evaluation of several integrals with a single dimension instead of a potentially large dimension according to Eq. 3.

Independent and identically distributed random effect for group

E-SURGE can also handle mixed models with group random effects only ( $P = 0$ ) like:

$$f(\Theta_i) = X_0\beta_0 + X_i\beta_1 + \sum_{l=1}^L Z_l b_l \tag{7}$$

with  $b_l$  in the form of Eq. 2. There is no limit for the number of random effects that we can build. However, for  $L > 2$ , the fitting step may be time consuming.

As in “Independent and identically distributed random effect for individual”, the marginal likelihood (3) is the product of the marginal probability for each group. We denote  $H_g$ , the set of encounter histories inside a group  $g$  considered together. So Eq. 3 can be rewritten:

$$L(\beta, \sigma) = \prod_{g=1}^G \int_{\mathbb{R}^L} \left( \prod_{h \in H_g} Pr(h|\beta, \mathbf{x}) \right) \omega(\sigma, \mathbf{x}) d\mathbf{x}. \tag{8}$$

**Example 2** We consider a basic model where recapture rates vary with a group random effect.

$$\text{logit}(p_g) = \beta_0 + b_g, \quad g = 1, \dots, G \tag{9}$$

where  $b_g \sim \mathcal{N}(0, \sigma_b^2)$ , (i.i.d). In this situation  $L = 1$  so for evaluating the marginal likelihood (8) we have only  $G$  one dimensional integrals to evaluate.

Numerical integration

Integrals in (5) and (8) have no analytical expression so require numerical approximations. After reparameterization, those integrals have the form:

$$K(\mathbf{y}) = \int_{\mathbb{R}^d} k(\mathbf{y}, \mathbf{x}) \exp^{-x^i x} d\mathbf{x} \tag{10}$$

where the dimension  $d$  depends on the number of random effects. For Eq. 5,  $d = P$ ; see also Gimenez and Choquet (2010) for the reparameterization in the case of the CJS model with an individual random effect. We will not consider in this paper models with dependence among

individuals or groups, some of them are formally described in Cam (2009). Dependences can be accommodated by normal random effects with non-diagonal covariance matrices. The marginal likelihood associated to such models can also be expressed as Eq. 10 after reparameterization; see for example Heiss and Winschel (2008). In the particular case of dependence of individuals and groups, the dimension  $d$  of the integral will depend of the structure of the dependence.

Order and accuracy

Let  $X_N$  be a set of  $N$  quadrature nodes of  $\mathbb{R}^d$ ,  $X_N = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}\}$  with  $\mathbf{x}^{(n)} \in \mathbb{R}^d$  for  $n = 1, \dots, N$ . Let  $\Omega_N$  be the set of quadrature weights associated with the nodes  $X_N$ ,  $\Omega_N = \{\omega_1, \dots, \omega_N\}$  with  $\omega_n \in \mathbb{R}$  for  $n = 1, \dots, N$ . The integrals  $K(\mathbf{y})$  can be approximated by  $I_N$  defined by

$$I_N = \sum_{n=1}^N \omega_n \times k(\mathbf{y}, \mathbf{x}^{(n)}) \tag{11}$$

$I_N$  is called a quadrature formula and  $K(\mathbf{y}) - I_N$  the error functional corresponding to  $K(\mathbf{y})$ . For relevant sets  $X_N$  and  $\Omega_N$ , the error functional decreases as  $N$  increases. Roughly, two methods of integration exist with different choices for  $\Omega_N$  and  $X_N$ . The first method is based on the random generation of  $X$  (see Rosenberg 1967; Genz and Monahan 1999 with or without equidistant properties; Zaremba 1968; Gonzalez et al. 2006). Under weak conditions, the simulated value is unbiased and  $\sqrt{N}$ -consistent by a law of large numbers and independent of the problem. The major drawback of this method relies on the difficulty in determining an appropriate value for  $N$  in order to control the error of integration  $K(\mathbf{y}) - I_N$ . The second method is based on the use of quadrature; see Arndt et al. (2006). The Laplace (Liu and Pierce 1994) method is not considered precise enough for complex CR likelihoods. Consequently, higher orders of Gauss-Hermite quadrature will be considered to obtain better precision. In this case,  $N$  increases with  $d$  and the order of the quadrature  $r$ ; thus, in some cases, we can replace the index  $N$  by the two indices  $d$  and  $r$ .

For  $d = 1$ , we directly use the Gauss-Hermite quadrature because it adapts easily to the normal density. The nodes of the quadrature  $X$  are related to the zeros of the  $r$ -th Hermite polynomial with  $N = 2 \times r - 1$ . The Gauss-Hermite quadrature with  $r$  nodes is exact for polynomials of order  $N$ . Here, as stated in Gimenez and Choquet (2010),  $k(\mathbf{y}, \mathbf{x})$  is not a polynomial but rather a composite function of a polynomial following an inverse logit function, which can be approximated as closely as long as  $N$  is large enough.

Moreover, under mild condition (if  $|K(\mathbf{y}) - I_{d,r+1}| < \kappa |K(\mathbf{y}) - I_{d,r}|$ ,  $\kappa < 1$ ), we can obtain the error estimate

associated with a polynomial of order  $r$  by considering two successive orders as:

$$|K(\mathbf{y}) - I_{d,r}| < \frac{1}{1 - \kappa} |I_{d,r+1} - I_{d,r}|.$$

One severe limitation of Gauss-Hermite quadrature is the effort growing with the dimension of the integral. Multivariate integration formulas are often constructed as a tensor product of quadrature formulas. Let  $I_{1,r}$  be a sequence of quadrature rules on  $\mathbb{R}$  then

$$I_{d,r} = \sum_{\mathbf{x}^{(1)} \in X} \dots \sum_{\mathbf{x}^{(d)} \in X} k(\mathbf{y}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d)}) \prod_{dd=1}^d \omega_{dd}(\mathbf{x}^{(dd)})$$

For dimension  $d$  and order  $r$ , a regular grid involves  $M = (2 \times r - 1)^d$  nodes. The calculation of Eq. 10 very soon becomes intractable. Thus, we try to reduce the order as much as possible with an appropriate scaling.

Centering and scaling

If the integrand in a numerical integration is not well behaved, algorithms can perform poorly. For a single dimension, the order of integration can be set higher to overcome this problem. But it is sometimes still not enough, and the error of integration can decrease very slowly. The problem in multiple dimensions is even worse because the number of nodes increases exponentially. Thus, for Eq. 8, we use centering and scaling in a similar way as Lemenuel-Diot et al. (2005) to improve the integration.

In order to base the approximation of the integral (Eq. 10) on an appropriate range of values of  $\mathbf{x}$ , two transformations are made: centering and scaling. If we assume that  $k(\mathbf{y}, \mathbf{x})$  is close to a normal distribution of mean  $\mathbf{x}^*$  and variance  $S$ ,  $k(\mathbf{y}, \mathbf{x}) \times \exp((\mathbf{x} - \mathbf{x}^*)'S(\mathbf{x} - \mathbf{x}^*))$  will be close to a constant and the approximation of the integral exact. Thus, we take  $\mathbf{x}^*$  as the mode of  $k(\mathbf{y}, \mathbf{x})$  and  $S$  as half of the inverse of the hessian matrix of the same function.

With the overall transformation  $\mathbf{x} = U^{-1}\mathbf{z} + \mathbf{x}^*$  with  $S = U^tU$  then Eq. 10 becomes

$$\begin{aligned} K(\mathbf{y}) &= \frac{1}{\det U} \int_{\mathbb{R}^d} k(\mathbf{y}, U^{-1}\mathbf{z} + \mathbf{x}^*) \exp^{-(U^{-1}\mathbf{z} + \mathbf{x}^*)'(U^{-1}\mathbf{z} + \mathbf{x}^*)} d\mathbf{z} \\ &= \frac{1}{\det U} \int_{\mathbb{R}^d} k_1(\mathbf{y}, \mathbf{z}) \exp^{-\mathbf{z}' \times \mathbf{z}} d\mathbf{z} \end{aligned} \tag{12}$$

with  $k_1(\mathbf{y}, \mathbf{z}) = k(\mathbf{y}, U^{-1}\mathbf{z} + \mathbf{x}^*) \exp^{-(U^{-1}\mathbf{z} + \mathbf{x}^*)'(U^{-1}\mathbf{z} + \mathbf{x}^*) + \mathbf{z}'\mathbf{z}}$  and  $\det(U)$  is the determinant of the matrix  $U$ .

We compute  $U$  as in Lemenuel-Diot et al. (2005). To get the inverse of  $U$  and when necessary (i.e., when  $d > 1$ ), we use an algorithm proposed by Xu and Qiao (2008) for a singular value decomposition of the symmetric matrix  $S$  (the Takagy Factorization).

Avoiding underflow and overflow

The approximation  $I_N$  is not straightforward to implement in the case of Eq. 8. In fact, the function  $k(\mathbf{y}, \mathbf{x})$  is a polynomial of high degree, with parameters belonging to  $[0,1]$ . Thus, very soon, values of the function  $k$  will be not representable in finite numerical precision, i.e.  $k(\mathbf{y}, \mathbf{x}) < 2^{-128}$ . This limitation is called underflow. Thus, we reformulate Eq. 10 to avoid underflow.

A solution to deal with underflow in Eq. 10 is to consider  $\log k(\mathbf{y}, \mathbf{x})$  so that the sum of individual contributions is involved rather than the product.

$$\begin{aligned}
 K(\mathbf{y}) &= \int_{\mathbb{R}^d} k(\mathbf{y}, \mathbf{x}) \exp^{-x^t \mathbf{x}} \mathbf{d}\mathbf{x} \\
 &= \int_{\mathbb{R}^d} \exp(\log k(\mathbf{y}, \mathbf{x})) \exp^{-x^t \mathbf{x}} \mathbf{d}\mathbf{x}
 \end{aligned}
 \tag{13}$$

However, overflow also occurs with  $\exp(\log k(\mathbf{y}, \mathbf{x}))$ . A way to deal with this new problem is to scale the quantity  $\log k(\mathbf{y}, \mathbf{x})$  by an appropriate value called  $dev^*$ .

$$\begin{aligned}
 K(\mathbf{y}) &= \exp(0.5 \times dev^*) \times \int_{\mathbb{R}^d} \exp(\log k(\mathbf{y}, \mathbf{x}) + 0.5 \times dev^*) \\
 &\times \exp^{-x^t \mathbf{x}} \mathbf{d}\mathbf{x} = \exp(0.5 \times dev^*) \times \int_{\mathbb{R}^d} k_2(\mathbf{y}, \mathbf{x}) \exp^{-x^t \mathbf{x}} \mathbf{d}\mathbf{x}
 \end{aligned}
 \tag{14}$$

where  $k_2(\mathbf{y}, \mathbf{x}) = \exp(\log k(\mathbf{y}, \mathbf{x}) + 0.5 \times dev^*)$ .

**Implementation in program E-SURGE**

Details of the global algorithm

To get the MLE, we consider a Quasi-Newton algorithm with a first order difference scheme to approximate the gradient and a relative tolerance of  $10^{-6}$ . The error made by using a first or second degree finite difference scheme does not influence the result in most cases (Appendix 2). In this paper, the results are obtained with a second order finite difference scheme although the use of the first order scheme led to the same results in both applications. For a group random effect, we consider the adaptive scaling described previously. To ensure precision of the quadrature scheme, we update the matrix  $S$  every 15 iterations of the quasi-Newton algorithm.

Description of the random effects

We extend the model definition language of the tool GEMACO used in M-SURGE and E-SURGE. We refer to

Choquet (2008) for the description of the language used for fixed effects.

Although individual covariates are permitted for a fixed effect, individual fixed effects are not allowed in E-SURGE. Therefore, we introduce a new built-in keyword factor denoted IND for individual random effects, and implement random effects for groups with the keyword RANDOM, which translate fixed effects into random effects. These additions fit naturally into E-SURGE's model specification syntax. However contrary to traditional effect like TIME, AGE, GROUP, direct addressing of levels of IND (one level corresponding to one individual) is not currently allowed. We extend also the operator + to concatenate fixed effect and random effect to generate mixed models of the form (4,7). Examples include:

The phrase "I+XIND(1)+IND" models Eq. 6.

The phrase "I+RANDOM(GROUP)" models Eq. 9.

More generally, two general forms of phrase are currently allowed:

"phrase1+phrase2.IND" for Eq. 4 and

"phrase1+RANDOM(phrase2)" for Eq. 7,

where phrase1 and phrase2 are any general phrases for fixed effects.

**Applications**

Application 1

We consider the European Dipper data used in Lebreton et al. (1992). Although affecting only the two last occasions, we found a trap-dependence effect using program U-CARE (Choquet et al. 2009a). This local effect on time may be a consequence of the fidelity of birds to the nesting site: an individual seen one year on a nesting site has a greater probability to be detected the next year on the same nesting site. This effect may be not detected the first years because of the flood years: their livelihood is closely dependant on stream as their nests are close to the water. Therefore, we considered a modified version of the standard Cormack-Jolly-Seber model incorporating a trap-dependence effect to account for this lack of fit. More precisely, we considered two detection probabilities at time  $t$   $p_t^*$ , and  $p_t$  depending on whether an individual was previously captured or not at the previous occasion  $t - 1$ .

As in Royle (2008), we tested for individual effects in both survival ( $\phi$ ) and capture ( $p^*, p$ ). Similar to Lebreton et al. (1992), the fixed effect for survival is the flood effect ( $t = 1, 4, 5, 6$  vs  $t = 2, 3$ ). With E-SURGE, the sentence for the survival TIME(1 4 5 6,2,3)+IND builds a model where

$$\begin{aligned} \text{logit}(\phi_{t=(1,4,5,6),i}) &= \beta_1 + b_{1,i}, \\ \text{logit}(\phi_{t=(2,3),i}) &= \beta_2 + b_{1,i}. \end{aligned} \tag{15}$$

where  $b_{1,i} \sim \mathcal{N}(0, \sigma_1^2)$ .

Using the formulation for trap-dependence described in Gimenez et al. (2003), the set of states used in E-SURGE is {'Alive and captured previously', 'Alive and not captured previously', 'Dead'}. The sentence for capture TIME(1:4)+FROM.TIME(5 6)+IND builds a model where

$$\begin{aligned} \text{logit}(p_{t=(1,\dots,4),i}^*) &= \beta_3 + b_{2,i}, \\ \text{logit}(p_{t=(1,\dots,4),i}) &= \beta_3 + b_{2,i}, \\ \text{logit}(p_{t=(5,6),i}^*) &= \beta_4 + b_{2,i}, \\ \text{logit}(p_{t=(5,6),i}) &= \beta_5 + b_{2,i}. \end{aligned} \tag{16}$$

where  $b_{2,i} \sim \mathcal{N}(0, \sigma_2^2)$ .

We fit this model  $\phi(\text{flood} + \text{ind})$ ,  $p(\text{partial}(m) + \text{ind})$  using  $r = 15$  for the quadrature without centering and scaling. In this example, centering and scaling failed to improve the precision of the Gauss-Hermite quadrature because the distribution of  $h(x)$  is not close enough to a gaussian. Table 1 shows that no individual effect is detected. For any of the two models  $\phi(\text{flood} + \text{ind})$ ,  $p(\text{partial}(m) + \text{ind})$  and  $\phi(\text{flood} + \text{ind})$ ,  $p(m + \text{ind})$  for the full trap-dependence, the deviance does not even change compared to  $\phi(\text{flood})$ ,  $p(\text{partial}(m))$  and  $\phi(\text{flood})$ ,  $p(m)$ . In each case,  $\sigma_1$  and  $\sigma_2$  are estimated close to zero (resp.  $5.10^{-5}$  and  $2.10^{-5}$ ; see Fig. 8). Thus, the individual effect detected in Royle (2008) can be explained by unmodeled trap-dependence. This is also supported by simulation where data with trap-dependent effect were generated and then analyzed with an individual effect, the model with individual effect was far better than the constant model (unpublished results). A step-by-step procedure for the implementation of the model in E-SURGE is given in Appendix 3.

### Application 2

Hantaviruses are the ethiological agents of several more or less severe diseases in humans. In northern, western and

central Europe, Puumala hantavirus causes Nephropathia epidemica, a mild form of haemorrhagic fever with renal syndrome (Klein and Calisherz 2007). However, for the virus, humans act only as dead-ends and its reservoir host is a small mammal, the Bank Vole, *Myodes glareolus*. To better understand the epidemics in humans, it is crucial to understand the dynamics of Puumala virus within the reservoir host populations. The data were collected to investigate the impact of Puumala virus on the demography of Bank Voles. Nine sites, typical for the optimal bank vole habitat, were monitored in the endemic zone where most human cases have occurred in recent years in Belgium. In each site, six trapping sessions were carried out from 2004 to 2006 with 100 live traps in a 1-ha grid Tersago (shed). Sites are coded as nine groups which are in fact geographic sites without movement among sites. Because the purpose here is not a full epidemiological study, we consider only two states ('Alive', 'Dead') and we construct a simple survival model with coupled random group effects on capture. The sentence for capture (decomposed in E-SURGE as first capture=FIRSTE and recapture=NEXTE) FIRSTE+NEXTE+RANDOM(NEXTE.GROUP) builds a model where

$$\text{logit}(p_{\text{age}=2,g}) = \beta_2 + b_g.$$

where  $b_g \sim \mathcal{N}(0, \sigma^2)$  and the first mathematical parameter(FIRSTE) is fixed to 1 in the biological scale ( $p_{\text{age}=1} = 1$ ). Note that this approach is equivalent to conditioning at first capture.

We fit the model using  $r = 15$  with centering and scaling. A step-by-step procedure for the implementation of the model in E-SURGE is given in Appendix 4. Results are represented in Table 2. The asymptotic LRT distribution for the test of  $H0:\sigma = 0$  against  $H1:\sigma > 0$  is a 50:50 mixture of  $\chi^2(0)$  and  $\chi^2(1)$ . Using the mixture of distributions corresponds to halving the p-value from naively using the  $\chi^2(1)$  distribution (Pinheiro and Bates 2000). So the model  $\phi(i)$ ,  $p(i)$  was rejected ( $p$  value = 0.033) and  $\sigma$  was estimated to 0.58 (see Fig. 16).

### Discussion

We have implemented effective algorithms in E-SURGE to fit models and get maximum likelihood estimates with random effects in cases where independence of individuals

**Table 1** European Dipper example revisited

Model	Np	Deviance
$\phi(\text{flood} + \text{ind})$ , $p(\text{partial}(m) + \text{ind})$	7	650.910
$\phi(\text{flood})$ , $p(\text{partial}(m) + \text{ind})$	6	650.910
$\phi(\text{flood} + \text{ind})$ , $p(\text{partial}(m))$	6	650.910
$\phi(\text{flood})$ , $p(\text{partial}(m))$	5	650.910
$\phi(\text{flood} + \text{ind})$ , $p(m + \text{ind})$	6	656.225
$\phi(\text{flood})$ , $p(m)$	4	656.225
$\phi(\text{flood})$ , $p$	3	660.102

Np Number of mathematical parameters of the model

**Table 2** Bank Vole example with a random effect on group

Model	Np	Deviance
$\phi(i)$ , $p(i + r(g))$	3	944.013
$\phi(i)$ , $p(i)$	2	947.381

$i$  is for the intercept,  $r(g)$  is for a group random effect. Np Number of mathematical parameters of the model



or groups is assumed. Instead of a MCMC approach, we consider a Gauss-Hermite scheme to cope with high-dimensional integrals. This approach has several advantages over Monte Carlo approaches. For example, we obtain a cheap measure of the error by the estimate of two marginal likelihoods. A better measure could be achieved for both methods by computing the discrepancy (Frank and Heinrich 1996), but this option comes at a price of a considerable effort of programming. Furthermore, the language inside GEMACO has been extended to allow built-in mixed models incorporating individual and group random effects with i.i.d. assumptions. Thus, practitioners without experience in MCMC modeling can access these modeling tools within the scope of a familiar, user-friendly software package.

However several challenges remain:

- First, built-in time, age and cohort random effects still need to be implemented in E-SURGE. Dimensions of the associated integrals are very high and additional algorithms must be implemented to fit these models in a reasonable amount of time. Much progress have been done the past few years with MCMC methods or EM-type algorithms (Chaubert-Pereira 2008). Our ultimate goal is to implement the most appropriate algorithm for each model.
- Second, in the present version of E-SURGE, the independence of groups needs to be specified in a menu of E-SURGE. Description or automatic detection of independence between groups remains to be implemented.
- Third, models with non-diagonal variance–covariance matrices as in Pinheiro and Bates (1996) have to be implemented. The description of dependence between individuals (which would allow fitting animal models like in Meyer 2001) or groups is the first step of this implementation.
- Fourth, besides fitting models, model selection using AIC is not trivial in presence of random effects. The main issue lies in counting the number of parameters: should it be the number of parameters plus the number of variance components or plus the number of random effects? There has been a recent attempt to cope with this issue (Vaida and Blanchard 2005) that still needs to be investigated in the context of capture–recapture mixed models.
- Fifth, the acute computation of the rank of CR2Ms remains to be done to investigate if redundant parameters are present in the models. In the current version of E-SURGE, the rank is set equal to the rank of the hessian of the likelihood.
- Last but not least, goodness-of-fit procedures are missing for CR2Ms. We hope that further progress

will be made with GOF as an inappropriate random effect can be selected even with marginal overdispersion. This is the case in application 1.

### Supplementary materials

The program E-SURGE is available from <http://www.cefe.cnrs.fr/BIOM/logiciels.htm>

**Acknowledgments** The authors warmly thank Gilbert Marzolin for sharing the Dipper data used in Application 1 and providing useful information on the Dipper, Laurent Crespin for suggesting Application 2 and Tersago Katrien for providing the Bank Vole data, and William L. Kendall and two anonymous referees for very useful comments that improved the presentation of the paper. This research was supported by a grant from the 'Jeunes Chercheuses et Jeunes Chercheurs' program of the French ANR (ANR-08-JCJC-0028-01).

### Appendix 1: Implementation of the structure of $X_i$ and $Z_{l,i}$

$X_i = A \times B_i$  where  $B_i$  is set of square matrices dependent on individuals. This structure allows us to consider model like: 1+T.XIND. In this case,  $A$  is a time dependent matrix and the matrix  $B_i$  is a  $(K - 1) \times (K - 1)$  diagonal matrix which entries are the value of the covariates for individual  $i$ .

$Z_{l,i} = C_l \times D_{l,i}$  where  $D_{l,i}$  is a square matrix depending of the individual and of the effect. This structure allows us to consider model like: 1+XIND+RAND(XIND). In this case  $L = 0, P = 1, C_1$  is a constant vector of 1 and  $D_{1,i}$  is the value of the covariates for individual  $i$ .

### Appendix 2: Influence of the finite difference scheme

We check that the tolerance ( $Tol = 10^{-6}$ ) used for the quasi-Newton algorithm on the gradient and the error ( $Err$ ) made by the approximation of the gradient using a finite difference scheme and a numerical integration are consistent, i.e. of the same order. Let  $\varepsilon$  be the computer precision, we will demonstrate that the global error  $Err$  defined by:

$$Err = K'(y_0) - \sum_{n=1}^N \omega_n \times \frac{k(y_0 + \sqrt{\varepsilon}, \mathbf{x}^{(n)}) - k(y_0, \mathbf{x}^{(n)})}{\sqrt{\varepsilon}}$$

can be decomposed as a sum of two sources of error; the error of the finite difference scheme applied to the gradient of  $k$  and the error made by the quadrature formulae applied to the gradient of  $k$ . By the Fubini theorem,

$$K'(y_0) = \int_{\mathbb{R}^d} \frac{\partial k(y, \mathbf{x})}{\partial y} \Big|_{y_0} \exp^{-x'x} \mathbf{dx}$$

The error made by approximating a derivatives by a first order finite-difference scheme is  $O(\sqrt{\varepsilon})$  (see Dennis and Schnabel 1983) so

$$Err = K'(y_0) - \sum_{n=1}^N \omega_n \times \left( \frac{\partial k(y, \mathbf{x}^{(n)})}{\partial y} \Big|_{y_0} + O(\sqrt{\varepsilon}) \right)$$

$$Err = \int_{\mathbb{R}^d} \frac{\partial k(y, \mathbf{x})}{\partial y} \Big|_{y_0} \exp^{-x^t \mathbf{x}} \mathbf{d}\mathbf{x}$$

$$- \sum_{n=1}^N \omega_n \times \frac{\partial k(y, \mathbf{x}^{(n)})}{\partial y} \Big|_{y_0} + O(\sqrt{\varepsilon})$$

For Matlab on a 32-bit Windows system,  $\varepsilon = 2.2204 \times 10^{-16}$  so as soon as the error made by the quadrature formulae is lower than  $Tol$ , the global error is lower than  $Tol$ . For a lower tolerance ( $Tol = 10^{-6}$ ) used for the quasi-Newton algorithm on the gradient then a second-order finite-difference scheme should be used.

### Appendix 3: Application 1 with E-SURGE

E-SURGE can accept capture–recapture data either in MARK or BIOMEKO format. The two types of data file are not very different: each row corresponds to a particular capture history followed by the number of individuals with that history. In MARK format, this count is followed by a semi-colon and in BIOMEKO format the data are preceded by the number of different capture histories and the number. In this study, there is only one site. We will fit here the model  $\phi$  (flood)  $p$ (partial( $m$ ) + ind) described in Application 1.

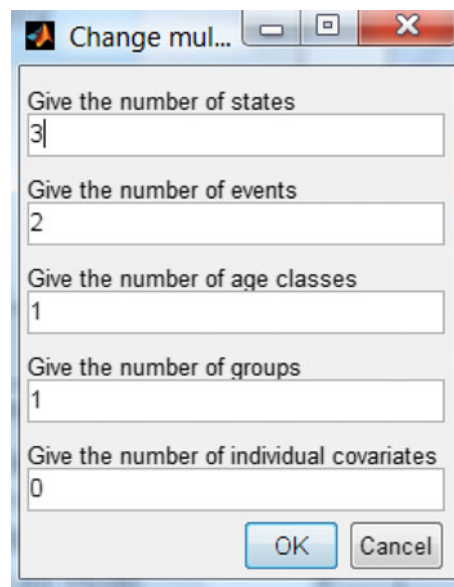
#### Starting E-SURGE

From E-SURGE, start a new session named 'result.mod'. Read in the data file and tell E-SURGE that there are no individuals covariates. Check that the numbers of capture occasions, groups and events are correct (in this case, 7 capture occasions, 1 group and 2 events): E-SURGE makes assumptions about the number of states, but these need to be modified depending on the problem you have to treat. In this case, we change the number of states to three. We will also have to set the number of age classes to one as for the present, we will not consider any age effect.

#### Fitting the model

Press the Modify button, and change the settings to specify that there a single age class, a single group and 3 states (see Fig. 1).

The main menu should show the changes. Fitting the models in E-SURGE involves four steps:



**Fig. 1** E-SURGE: the number of states is set to 3 and the number of age classes is set to one

1. The Gepat step: specifying which ones of all the potential parameters have to be estimated, which ones will be calculated as the complement to 1 (there is one such parameter per multinomial) and which ones correspond to impossible events or transitions and are fixed to zero;
2. The Gemaco step: specifying the effects (time, age...) acting on the active parameters;
3. The IVFV step: specifying initial values for the optimization procedure and/or fixed values for the active parameters;
4. The RUN step: launching the optimization procedure.

#### Specifying the pattern matrices using the GEPAT interface

There are three types of parameters used in the definition of a multi-event model (Pradel 2005):

1. the initial state probabilities;
2. the transition probabilities;
3. the event probabilities.

Each type of parameter is gathered into a row–stochastic matrix, i.e. each row corresponds to a multinomial. (Each matrix can be further decomposed into a product of several stochastic matrices allowing for example to estimate separately survival and transition parameters. However, for the current model, only 1 step is required for each type of matrix.)

To enter the GEPAT interface, click the GEPAT button at the bottom of the main window. The GEPAT interface screen for the initial state pattern matrix appears.

By default, E-SURGE lets all live states be available as initial steps. The state dead, last in the list, is impossible and does not even appear. The last live step is taken by default to be the one whose probability will be calculated indirectly, as the complement of those of the other live states. This is specified in the above window using the following general conventions for Gepat:

- a minus sign (–) indicates that a potential parameter is unavailable in the current model (impossible transition, for instance). This is equivalent to fixing it to zero but more explicit;
- any Greek letter (strike a Latin letter and E-SURGE will show its Greek counterpart) indicates a free parameter, one to be estimated directly. Note that the particular Greek letter entered is totally irrelevant to E-SURGE. In particular, the same Greek letter is used repeatedly by default within a pattern matrix (by default for initial states); this does NOT mean that the parameters are being forced to be equal;
- the asterisk (\*) indicates the parameter that is calculated indirectly, as the complement to 1 of all the other parameters on the same row. There MUST always be one and only one asterisk per row because each row corresponds to a multinomial.

Note that the order of the states is chosen by the user except for the dead state that is always positioned last. Here, the default pattern is not correct for transition and event. We need to change them.

For initial states, we implement the pattern given in Fig. 2. For transitions, we need two steps, one for survival and one for capture. So we set the number of steps to 2. For survival, we implement the pattern given in Fig. 3. For capture, we implement the pattern given in Fig. 4. For

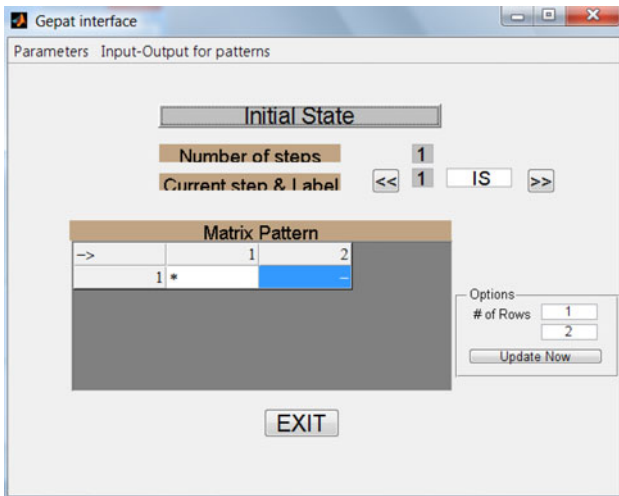


Fig. 2 E-SURGE: pattern for the initial states vector

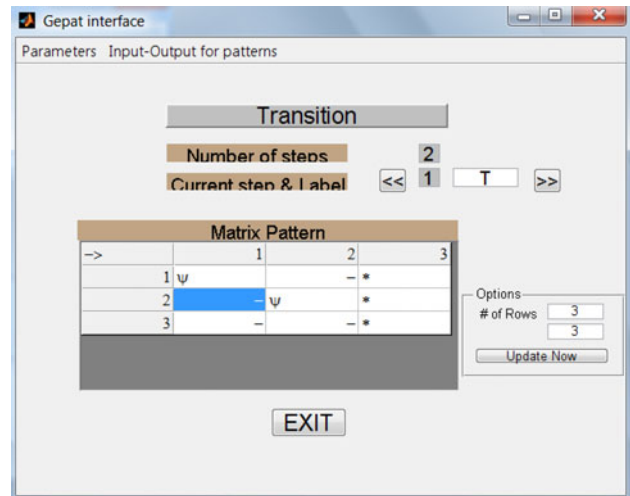


Fig. 3 E-SURGE: pattern for the survival matrix

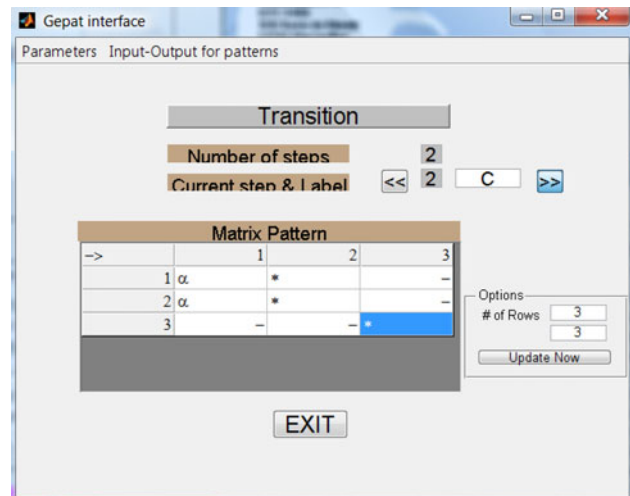


Fig. 4 E-SURGE: pattern for the capture matrix for the trap-dependent model

event, we implement the pattern given in Fig. 5. Press the “EXIT” button to return to the main window.

*Specifying the model using the GEMACO interface*

The GEMACO interface uses keywords to create a modeling sentence that indicates how parameters vary by time, over groups, over age classes, etc. At the end of the GEMACO procedure, a design matrix is created for each type of parameters. Each row of the design matrix will correspond to a parameter of the full model (all potential variability: time, age, group...) and each column corresponds to a parameter of the actual model.

The GEMACO syntax is fairly intuitive but the “sentences” you enter in the GEMACO interface must respect some priority rules that we will not develop here. We

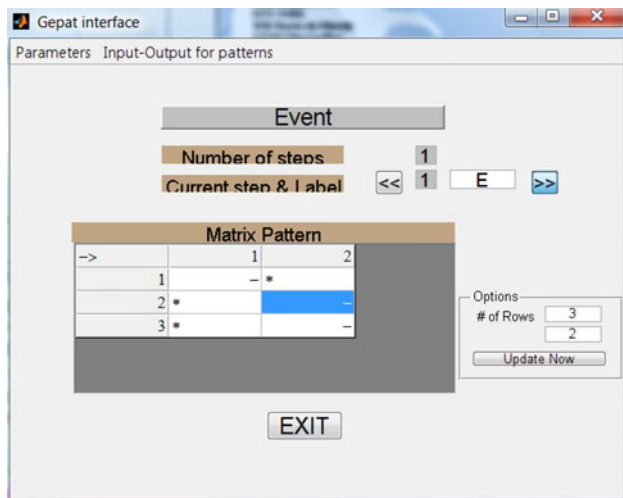


Fig. 5 E-SURGE: pattern for the event matrix

encourage the user to read the E-SURGE user manual and Choquet (2008) in which the GEMACO syntax is fully explained.

In this example, we only want to show how to use E-SURGE to fit our model.

For the trap-dependent model, the set of initial state probabilities and the set of event probabilities are empty. Click on the top “Initial State” button to go to the “Transition” screen.

To specify the model on survival, we use the phrase  $\tau(1\ 4\ 5\ 6,2\ 3)$  for the flood effect, and the new keyword *IND* for the individual random effect. As we combine all, the GEMACO sentence becomes  $\tau(1\ 4\ 5\ 6,2\ 3)+IND$ . Select the next step for transitions which corresponds to the modeling of trap dependence. We use the phrase  $\tau(1:4)+\tau(5\ 6).F+IND$ .

At this stage, to create the design matrices, we click on the Gemaco item in the top menu and select the “call GEMACO (all phrases)” submenu. At this stage, all the model structures are specified and the design matrices appear in the left window of each screen of the GEMACO interface; press the “EXIT” button to return to the main window.

*Specifying the initial and fixed values using the IVFV interface*

In E-SURGE, the user can choose the way to generate the initial values of the optimization procedure. They can be either “constant”, “randomly generated” or “equal to the estimates of a previously fitted model”. Once the type of initial values is chosen, the user can also fix the values for some parameters using the IVFV interface. Press the IVFV button to enter the interface (Fig. 6). As there is no need to specify either fixed values or initial values, click on the “EXIT” button to return to the main window.

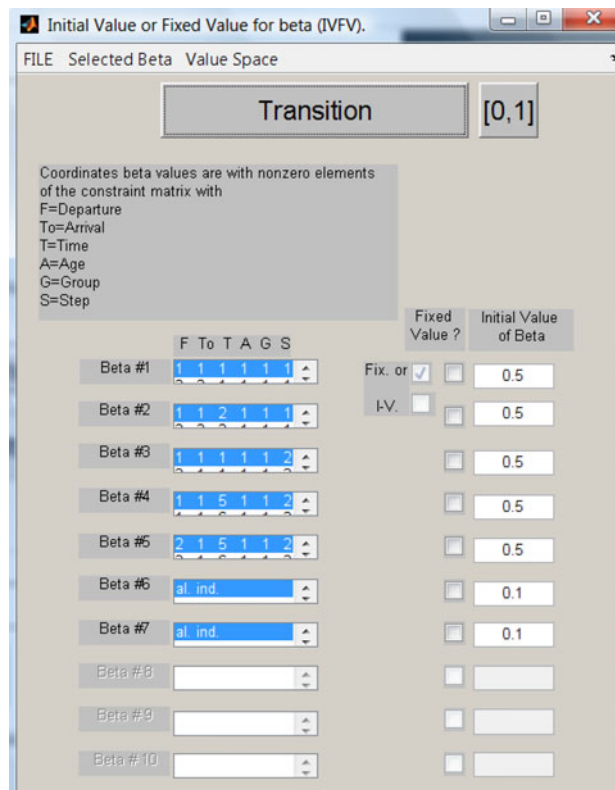


Fig. 6 E-SURGE: the initial values fixed values interface. The two last parameters are those associated to the two random effects

*Running the model*

Before running the model, we have to specify the method of integration; here, we choose the classical Gauss-Hermite method (set by default) described in the paper with 29 quadrature nodes ( $r = 15$ , set by default). We also tick the “compute C-I(Hessian)” option to get confidence intervals. The model is now ready to be fitted to the data. Press the RUN button. We observe in Fig. 7 that at the end of the fit, the estimate error made by the GH scheme to get the likelihood is lower than  $10^{-6}$  as  $-15.64 \leq -6$ .

In Fig. 8, we get the estimates for the model  $\phi$  (*flood* + *ind*)  $p(\text{partial}(m) + \text{ind})$  in the mathematical scale.

**Appendix 4: Application 2 with E-SURGE**

We will fit here the model  $\phi$  (*flood*)  $p(\text{partial}(m) + \text{ind})$  described in part 2 with E-SURGE.

**Starting E-SURGE**

From E-SURGE, start a new session and load the dataset. For this application, we only need to change the number of

**Fig. 7** E-SURGE: the Output during the RUN step shows that 20 iterations are needed for convergence and that the error due to numerical integration is small (lower than  $10^{-6}$ )

7	126	651.14	1.72186	-0.000833
8	145	650.978	4.71931	9.37e-005
9	163	650.933	1.23147	-3.58e-006
10	181	650.917	1.32543	-4.34e-005
11	200	650.911	2.33857	-2.65e-005
12	218	650.91	1.65152	-1.34e-005
13	236	650.91	0.954357	-1.12e-006
14	254	650.91	0.791328	-3.81e-007
15	272	650.91	0.923187	-1.86e-007
16	290	650.91	1	2.89e-008
17	309	650.91	1.62486	-2.57e-009
18	328	650.91	15.3682	-1.02e-009
19	346	650.91	1.86922	-1.38e-008
20	364	650.91	1.97872	8.8e-010

Optimization terminated successfully:  
 Current search direction is a descent direction, and magnitude of directional derivative in search direction less than 2\*options.TolFu  
 log10 of relative estimate error in integration: order16-15=-14.6439

**Fig. 8** E-SURGE: estimates and standard errors for the model  $\phi(\text{flood} + \text{ind})$   
 $p(\text{partial}(m) + \text{ind})$

```

-----
Beta (Mathematical parameters)
-----
      Index          Beta | Lower & Upper 95 percent CI | S.E.
-----
Beta#  1# | +0.479532164 -0.252184503 +1.211248831 +0.373324830
Beta#  2# | -0.101744542 -0.449092341 +0.245603257 +0.177218265
Beta#  3# | +1.888057726 +1.002096719 +2.774018734 +0.452020922
Beta#  4# | +2.546300713 -2.642247578 +7.734849005 +2.647218516
Beta#  5# | -0.867836745 -4.079459173 +2.343785682 +1.638582871
#####
Beta#  6# | -0.007022382 -48.820600307 +48.806555543 +24.904886696
Beta#  7# | +0.004270352 -277.068305824 +277.076846528 +141.363559274
#####
##### SE IND.R.E. ESTIMATES #####
Beta#  6# | +0.000049314 -0.685525859 +0.685624487 +0.349783252
Beta#  7# | +0.000018236 -2.366376626 +2.366413098 +1.207344317
#####
    
```

age classes to one as, for the present, we will not consider any age effect.

Fitting the model

Press the Modify button, and change the settings to specify that there is a single age class, nine groups and 2 states (see Fig. 9).

The main menu should show the changes. Fitting the models in E-SURGE involves four steps:

*Specifying the pattern matrices using the GEPAT interface*

Note that the order of the states is chosen by the user except for the dead state always positioned last. Here, the default pattern is one of the CJS model. We do not need to change them.

*Specifying the model using the GEMACO interface*

In this example, we only want to show how to use E-SURGE to fit our model.

The initial states modeling

For the CJS model, the set of initial state probabilities is empty. Click on the top “Initial State” button to go to the “Transition” screen.

To specify the model on survival, we use the phrase 1 (see Fig. 10).

Click on the top “Transition” button to go to the “Event” screen.

Because the model conditions on the first capture occasion of each individual, the only event to model at the time of the first encounter is the site of capture. It is only later on that the event ‘not encountered’ becomes possible. Thus, the event probabilities at the time of the first encounter must be treated separately. This is achieved through the use of the keywords “firste”, which stands for ‘first encounter’, and “nexte”, which stands for ‘next encounters’, respectively.

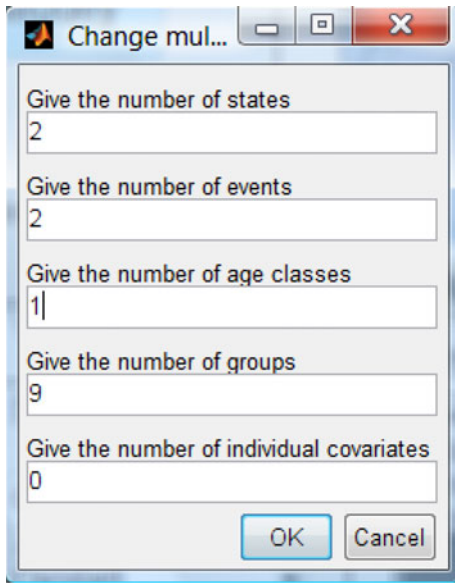


Fig. 9 E-SURGE: the number of age classes is set to one

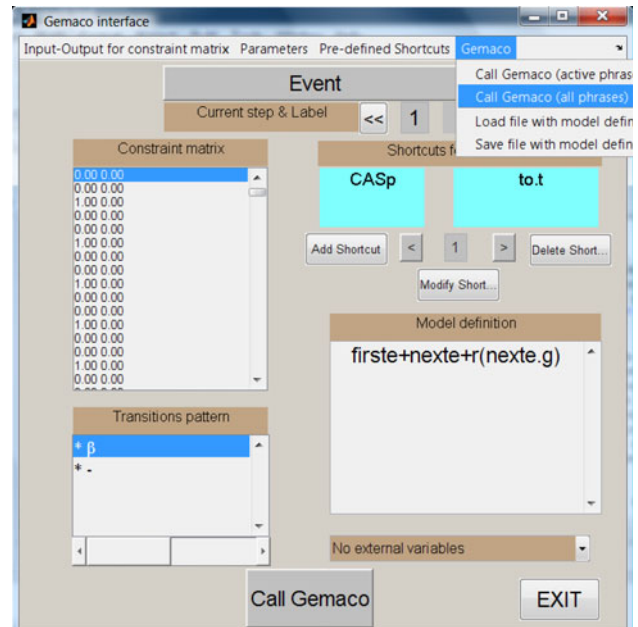


Fig. 11 E-SURGE: the sentence 'firste+nexte+r(nexte.g)' builds the model  $\phi(i+r(g))$

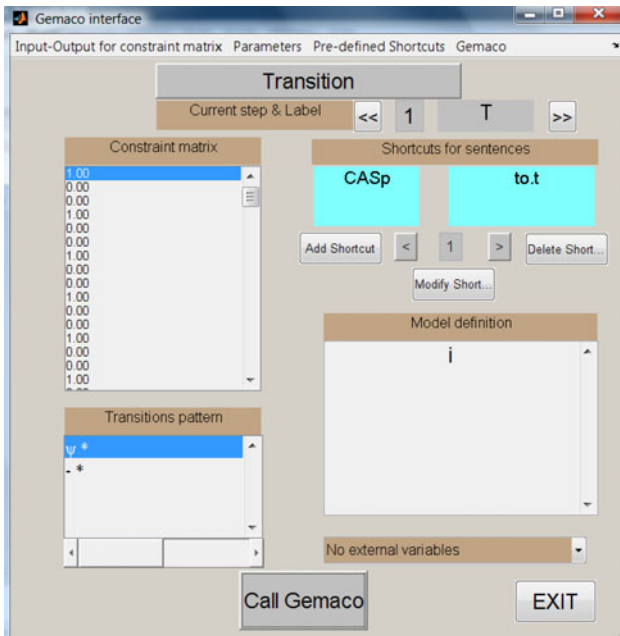


Fig. 10 E-SURGE: the sentence 'i' builds the model  $\phi(.)$

The output of Gepat is shown in the 'transitions pattern' subwindow at the bottom left. For each state (in row), there is only one active event which is the capture on the relevant site (where the stands), and the other possible event taken as the complement is 'not encountered' (first column). For instance, for state 11, the first row, the individual can be encountered on site 1 (second column, probability to be estimated) or 'not encountered' (first column, probability calculated as 1 – the other probability). The active parameters are thus just the capture probabilities. At the

time of the first encounter, the capture is certain and the capture probabilities will all be 1. At this stage, we cannot specify a fixed value, but we can specify that we need just one parameter common to all states by keeping "firste" by itself. Later, capture probability will be constant. Thus, the complete sentence is "firste+nexte+random(nexte.g)".

At this stage, to create the design matrices, we click on the Gemaco item in the top menu and select the "call GEMACO (all phrases)" submenu (see Fig. 11). All the model structures are now specified and the design matrices for fixed effect appear in the left window of each screen of the GEMACO interface; press the "EXIT" button to return to the main window.

*Specifying the initial and fixed values using the IVFV interface*

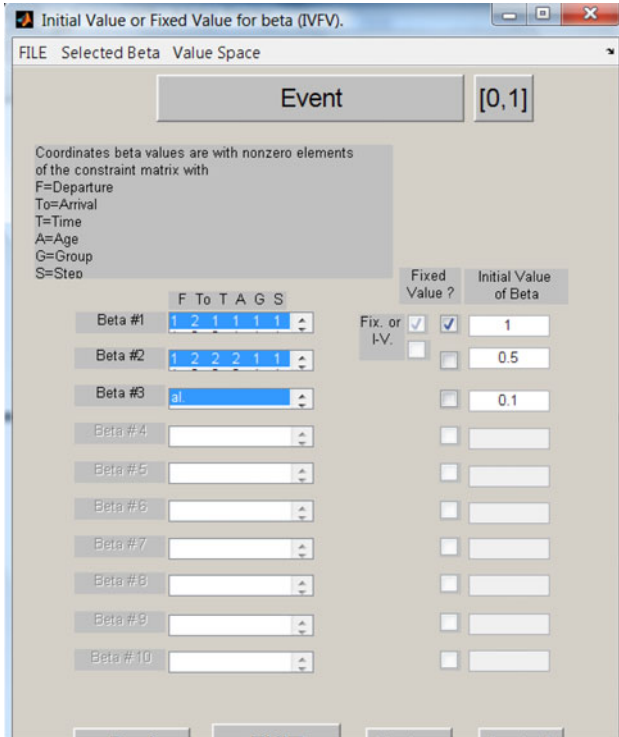
In E-SURGE, the user can choose the way to generate the initial values of the optimization procedure. They can be either "constant", "randomly generated" or "equal to the estimates of a previously fitted model". Once the type of initial values is chosen, the user can also fix the values for some parameters using the IVFV interface. Press the IVFV button to enter the interface.

*The initial states probabilities* In this case, there is no need to specify neither fixed values nor initial values. Click on the top "Initial State" button to arrive at the "Transition" screen.

*The survival-transitions probabilities* There is no need to fix values for the transition probabilities so this screen

can be left in its default state. Click on the top “Transition” button to arrive at the “Event” screen.

*The event probabilities* We can see here the different capture rate appearing in the definition of the model. The series of number indicate successively:



**Fig. 12** E-SURGE: the first mathematical parameter corresponding to 'firste' is fixed to 1 by ticking the box nearby. The third mathematical parameter is the starting value for the square root of the standard error; it must be strictly positive

- the line in the event matrix (corresponding to the state),
- the column in the event matrix (corresponding to the event),
- the capture occasion,
- the age class,
- the group,
- the step in the matrix decomposition of the event matrix (here 1).

Thus, the first parameter corresponds to the capture rate at the first capture occasion for the first age class, i.e. time of first encounter ( $A=1$ ). This is the only parameter with  $A=1$  because we have gathered all the capture rates relative to the first encounter into a single parameter. This parameter needs to be fixed to 1. We do this by entering the value 1 as “Initial Value” and ticking the box nearby (see Fig. 12). The other parameter corresponds to the following capture rates ( $A=2$ ); there is no need to fix these parameters.

After all the fixed values have been specified, press the EXIT button.

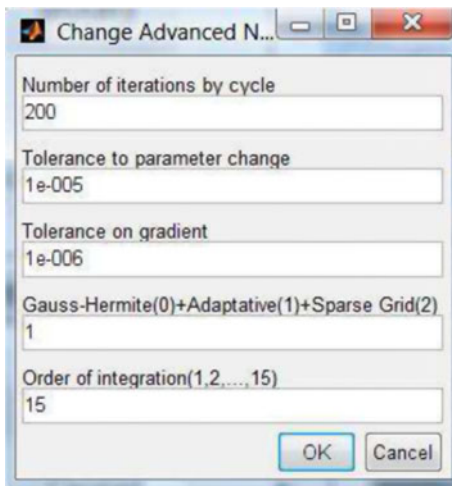
*Running the model*

Before running the model, we have to specify the method of integration. Click on the button 'Advanced Numerical Options > Modify' in the main window of E-SURGE. Here we choose the Adaptive Gauss-Hermite method described in the paper (the fourth value is set to 1) with 29 quadrature nodes (the fifth value is set to 15) (see Fig. 13).

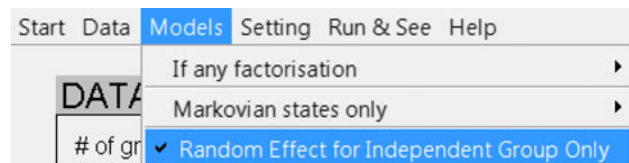
We also tick the submenu 'Random Effects for Independent Group Only' in the menu 'Models' (Fig. 14).

We also tick the 'compute C-I(Hessian)' option to get confidence intervals and an estimated of the model rank. The model is now ready to be fitted to the data. Press the RUN button. We observe in Fig. 15 that at the end of the fit, the estimate error made by the GH scheme to get the likelihood is lower than  $10^{-6}$  as  $-6.45 \leq -6$ .

In Fig. 16, we get the estimates for the model  $\phi(\cdot)$   $p(i + r(g))$  in the mathematical scale.



**Fig. 13** E-SURGE: we choose the Adaptive Gauss-Hermite method (the fourth value is set to 1) with 15 quadrature nodes (the fifth value is set to 15)



**Fig. 14** E-SURGE: to speed-up calculations and improve the precision of the integration, we set that the random group effect is i.i.d

**Fig. 15** E-SURGE: the Output during the RUN step shows that only 10 iterations are needed for convergence and that the error due to numerical integration is small (lower than  $10^{-6}$ )

Iteration	Func-count	f(x)	Step-size	Directional derivative
1	6	1159.5	0.0124954	-1.86e+005
2	16	951.553	0.00243892	2.14e+003
3	27	947.522	0.0456034	-4.09
4	41	944.493	0.895395	2.31
5	51	944.394	0.118707	0.0121
6	62	944.167	1.57906	-0.000216
7	72	944.017	2.31208	-5.68e-006
8	82	944.013	0.838651	-1.6e-005
9	92	944.013	1.04738	-2.15e-007
10	102	944.013	0.967475	5.68e-013

```

Optimization terminated successfully:
Search direction less than 2*options.TolX
|
log10 of relative estimate error in integration: order16-15=-6.4556
    
```

**Fig. 16** E-SURGE: estimates and standard errors for the model  $\phi(\cdot) p(i + r(g))$

```

-----
Beta (Mathematical parameters)
-----
      Index      Beta | Lower & Upper 95 percent CI | S.E.
-----
Beta#  1# | -1.154749301 -1.470615637 -0.838882966 +0.161156294
#####
Beta#  2# | -0.189295962 -0.927577181 +0.548985257 +0.376674091
Beta#  3# | +0.765001557 +0.388753867 +1.141249247 +0.191963107
##### SE R.E. ESTIMATES #####
Beta#  3# | +0.585227382 +0.009567244 +1.160887520 +0.293704152
#####
    
```

**References**

Arndt C, Kozlitina J, Preckel PV (2006) Efficient survey sampling of households via gaussian quadrature. *J R Stat Soc C Appl Stat* 55:355–364

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol* 24:127–135

Burnham K, Anderson D (2002) Model selection and inference: a practical information–theoretic approach, 2nd edn. Springer, New York

Cam E (2009) Contribution of capture-mark-recapture modeling to studies of evolution by natural selection. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 83–129

Cappe O, Moulines E, Ryden T (2005) Inference in Hidden Markov models. Springer, New York

Chaubert-Pereira F (2008) Combinaisons Markoviennes et semi-Markoviennes de modeles de regression. Application a la croissance d’arbres forestiers. PhD thesis

Choquet R (2008) Automatic generation of multistate capture-recapture models. *Can J Stat* 36:43–57

Choquet R, Lebreton JD, Gimenez O, Reboulet A, Pradel R (2009) U-CARE: Utilities for performing goodness of fit tests and manipulating CApture-REcapture data. *Ecography* 32:1071–1074

Choquet R, Reboulet A, Pradel R, Gimenez O, Lebreton JD (2004) M-SURGE: New software specifically designed for multistate recapture models. In: Animal biodiversity and conservation, vol 27. pp 207–215

Choquet R, Rouan L, Pradel R (2009) Program E-SURGE: a software application for fitting multievent models. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, Dunedin, pp 845–865

Coull BA, Agresti A (1999) The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics* 55:294–301

Dennis J, Schnabel R (1983) Numerical methods for unconstrained optimization and nonlinear equations. Prentice-Hall, Englewood Cliffs, NJ

Frank K, Heinrich S (1996) Computing discrepancies of smolyak quadrature rules. *J Complex* 12:287–314

Genz A, Monahan J (1999) A stochastic algorithm for high-dimensional integrals over unbounded regions with gaussian weight. *J Comput Appl Math* 112:71–81

Gimenez O, Bonner S, King R, Parker RA, Brooks SP, Jamieson L, Grosbois V, Morgan B, Thomas L (2009) Winbugs for population ecologists: Bayesian modeling using markov chain monte carlo methods. In: Thomson D, Cooch E, Conroy M (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, Dunedin, pp 883–915

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

Gimenez O, Choquet R, Lebreton JD (2003) Parameter redundancy in multistate capture-recapture models. *Biom J* 45:704–722

Gimenez O, Covas R, Brown C, Anderson M, Brown M, Lenormand T (2006) Nonparametric estimation of natural selection on a quantitative trait using mark-recapture data. *Evolution* 60:460–466



- Gimenez O, Rossi V, Choquet R, Dehais C, Doris B, Varella H, Vila J, Pradel R (2007) State space modelling of data on marked individuals. *Ecol Model* 431–438
- Gonzalez J, Tuerlinckx F, De Boeck P, Cools R (2006) Numerical integration in logistic-normal models. *Comput Stat Data Anal* 51:1535–1548
- Heiss F, Winschel V (2008) Likelihood approximation by numerical integration on sparse grids. *J Econom* 144:62–80
- Klein SL, Calisherz CH (2007) Emergence and persistence of hantaviruses. *Wildlife and emerging zoonotic diseases: the biology, circumstances and consequences of cross-species transmission*. vol 315. Springer, Berlin
- Lebreton J-D, Burnham K, Clobert J, Anderson D (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol Monogr* 62:67–118
- Lemuel-Diot A, Mallet A, Laveille C, Bruno R (2005) Estimating heterogeneity in random effects models for longitudinal data. *Biom J* 47:329–345
- Link WA, Barker RJ (2005) Modeling association among demographic parameters in analysis of open population capture-recapture data. *Biometrics* 61:46–54
- Liu Q, Pierce D (1994) A note on gauss-hermite quadrature. *Biometrika* 81:624–629
- Lunn D, Thomas A, Best N, Spiegelhalter D (2000) Winbugs a bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 10:325–337
- McClintock BT, White GC, Antolin MF, Tripp DW (2009) Estimating abundance using mark-resight when sampling is with replacement or the number of marked individuals is unknown. *Biometrics* 65:237–246
- Meyer K (2001) Estimating genetic covariance functions assuming a parametric correlation structure for environmental effects. *Genet Sel Evol* 33:557–585
- Murphy K (2002) *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis
- Pinheiro JC, Bates DM (1996) Unconstrained parametrizations for variance-covariance matrices. *Stat Comput* 6:289–296
- Pinheiro JC, Bates DM (2000) *Mixed effects models in S and S-Plus. Statistics and computing*. Springer, Berlin
- Pradel R (2005) Multievent: an extension of multistate capture-recapture models to uncertain states. *Biometrics* 61:442–447
- Rosenberg I (1967) Bernstein polynomials and monte-carlo integration. *J Numer Anal* 4:566–574
- Royle JA (2008) Modeling individual effects in the cormack-jolly-seber model: a state-space formulation. *Biometrics* 64:364–370
- Royle JA, Link WA (2002) Random effects and shrinkage estimation in capture-recapture models. *J Appl Stat* 29:329–351
- Tersago K (Unpublished). Spatio-temporal variation of puumala hantavirus in belgium: an eco-epidemiological study. PhD thesis
- Vaida F, Blanchard S (2005) Conditional akaike information for mixed-effects models. *Biometrika* 92:351–370
- Wintrebert CMA, Zwinderman AH, Cam E, Pradel R, van Houwelingen JC (2005) Joint modelling of breeding and survival in the kittiwake using frailty models. *Ecol Model* 181:203–213
- Xu W, Qiao SZ (2008) A divide-and-conquer method for the takagi factorization. *J Matrix Anal Appl* 30:142–153
- Zaremba SK (1968) Mathematical basis of monte carlo and quasi-monte carlo methods. *SIAM Rev* 10:303–314