Age–dependent capture–recapture models and unequal time intervals

A. Sanz–Aguilar, R. Pradel, G. Tavecchia

Abstract

Age–dependent capture–recapture models and unequal time intervals. Estimates of survival probabilities in natural populations can be obtained through capture–mark–recapture (CMR) models. However, when capture sessions are unevenly spaced, age–dependent models can lead to erroneous estimates of survival when individuals change age class during the time interval between two capture occasions. We propose a solution to correct for the mismatch between time intervals and age class duration in two age class models. The solution can be implemented in different ways. The first consists of adding dummy occasions to the encounter histories and fixing the corresponding recapture probabilities at zero. The second makes use of the log–link function available in some CMR software (e.g. program MARK). We used simulated and real data to show that the proposed solution delivers unbiased estimates of age–dependent survival probabilities.

Key words: Age–dependent models, Capture–mark–recapture, Missing data, Survival probability

Resumen

Modelos de captura y recaptura dependientes de la edad e intervalos de tiempo desiguales. A través de modelos de captura, marcaje y recaptura (CMR) se puede estimar la probabilidad de supervivencia en poblaciones naturales. Sin embargo, cuando las sesiones de captura están espaciadas de manera desigual, los modelos dependientes de la edad pueden producir estimaciones erróneas de la supervivencia si los individuos cambian de clase de edad durante el intervalo entre dos sesiones de captura. Proponemos una solución para corregir el desajuste entre los intervalos de tiempo y la duración de las clases de edad en modelos con dos clases de edad. La solución puede aplicarse de diferentes formas. Una consiste en añadir muestreos ficticios en las historias de captura y fijar las probabilidades de recaptura correspondientes a cero. Una segunda aplicación usa la función log–link disponible en algunos programas informáticos de CMR (p.ej., el programa MARK). Usamos datos simulados y reales para mostrar que la solución propuesta produce estimaciones no sesgadas de las probabilidades de supervivencia dependientes de la edad.

Palabras clave: Modelos dependientes de la edad, Captura, marcaje y recaptura, Datos incompletos, Probabilidad de supervivencia

Ana Sanz–Aguilar, Giacomo Tavecchia, Animal Demography and Ecology Group, Institut Mediterrani d'Estudis Avançats IMEDEA (CSIC–UIB), c/Miquel Marques 21, 07190 Esporles, Baleares, Spain.– Roger Pradel, Biostatistics and Population Biology Group, Centre d'Ecologie Fonctionelle et Evolutive, CEFE–CNRS, 1919 route de Mende, F–34293, Montpellier, France.

Corresponding author: Ana Sanz–Aguilar. E–mail: asanz@imedea.uib-csic.es

Introduction

Capture–mark–recapture methods (CMR) are widely used for diagnosis of natural populations because they can be applied to obtain robust estimates of demographic parameters accounting for imperfect detection of individuals (Lebreton et al., 1992; Williams et al., 2002; Sanz–Aguilar et al., 2016). Cormack–Jolly–Seber models for the estimates of survival probability in natural populations are based on the important assumption that animals share the same parameters regardless of their past or present history (Pradel et al., 2005). When animals are marked as young this assumption does not hold because newly marked individuals typically have a lower survival probability than already marked individuals (adults). This difference can be accommodated by including age–dependent parameters into the CMR model (Pollock, 1981; Lebreton et al., 1992). In a simple two–age–class model, one parameter, noted ϕ', would apply to the survival probability of young individuals and a second, noted ϕ', would apply to the survival of adults (see examples in Hiraldo et al., 1996; Prugnolle et al., 2003). Age–dependent parameterizations have also been considered when only adults are marked to correct for an excess of animals seen only at marking, i.e. transients (Pradel et al., 1997), for example, when tags are potentially harmful (Saraux et al., 2011) or to model an effect of breeding experience (Sanz–Aguilar et al., 2008, 2012). Age–dependent survival probabilities are parameters of interest in many ecological studies (e.g. Clobert et al., 1988; Loison et al., 1999; Tavecchia et al., 2001; Bonenfant et al., 2002; Perret et al., 2003; Catchpole et al., 2004; Sanz–Aguilar et al., 2015). However, while age–classes are equally spaced, intervals between capture–recapture occasions may not be equally spaced on the same scale, leading to erroneous estimates (see the problem in, for example, Covas et al., 2002; Zabala et al., 2011; Zuberogoitia et al., 2016). This is because individuals would change their age class within the interval between two sampling occasions rather than at the end as assumed by CMR models. We briefly introduce the problem and illustrate how it can be solved by taking advantage of the flexibility of CMR models.

The problem

Logistic, financial or weather–dependent constraints can interrupt monitoring or modify the temporal frequency of sampling occasions, leading to different time length between capture–mark–recapture sessions. Unequal time intervals, alone, do not present a major problem in CMR models (Bears et al., 2009; Cooch, 2009; Schmidt et al., 2007). Consider a study with k sampling occasions with intervals between the occasions $j = 1, 2, …, k–1$. The length of the intervals between the sampling occasions is l_j. The l value is taken as the exponent of the survival parameters expressed in some common time unit for the interval $j, j + 1$, as ϕ^{lj}. For example, the survival parameter over a unit interval ($l = 1$) would be ϕ^1, for a two–unit interval ($l = 2$) it would be ϕ^2, and so on. Values in

the vector l are commonly integers, e.g., years or months, but can also be decimal numbers; for example, the survival parameter over an eighteen–month interval can be written in terms of yearly survival as $\phi^{1.5}$ ($l = 1.5$). The freely available software for CMR analyses, such as MARK (White and Burnham, 1999), RMARK (Laake, 2013) or ESURGE (Choquet et al., 2009), allows users to specify the vector of l_j values. However, unequal time intervals pose a problem in age–dependent models because, contrarily to intervals between occasions, the age classes always retain the same length. As a consequence, an individual may 'move' through age classes during an interval of length l_j and the survival parameter can no longer be written as ϕ^{lj} because the instantaneous survival probability changes with the age classes spanned by the interval of length l_j. The mismatch between the duration of an age class and the time interval would, for instance, lead to an overestimation of the first–year survival probability if the sampling interval were greater than one year. The fundamental problem is that a given age–dependent parameter applies to only a part of the time interval. This can be solved by specifying the length of the interval for each parameter considered. We outline the solution and provide a step–by–step illustration of how this can be implemented in freely available software for CMR analysis (e.g. MARK, RMark or E–SURGE, see details in the Supplemental Information S1, S2 and S3). Note that the two implementations below are simply two practical approaches to solve the problem (see supplementary material S1, S2 and S3).

Methods

Implementation 1: adding dummy encounter occasions

When intervals are of unequal lengths, a possible solution is to add dummy encounter occasions in the encounter histories to 'fill' the temporal gaps between occasions, which means, in practice, adding columns of 0s (e.g. Grosbois and Tavecchia, 2003; Sanz–Aguilar et al., 2010): the recapture probabilities corresponding to these dummy occasions should be fixed at 0. For example, let us consider a 7–year study with five capture–mark–recapture occasions in years 1, 2, 5, 6 and 7. The interval between the second and third occasions lasts three years instead of one. The l–vector of interval lengths would be 1, 3, 1, 1. The encounter history of animals released at the beginning of the study and always seen would be '1 1 1 1 1'. When columns of '0' are added to fill the temporal gaps, the encounter history above becomes '1100111' and all six elements of the l–vector are now equal to 1. The encounter probabilities at dummy occasions (3 and 4) should be fixed at 0 (figs. S2.1 and S2.3 supplementary material S2). The survival parameter of the first age class, ϕ', now always refers to an initial one–year interval. This approach can be implemented in programs MARK and E–SURGE. Adding dummy occasions permits to manipulate the correct survival parameters, but the dummy occasions come with no

Table 1. Not separately identifiable parameters in presence of missing occasions: ϕ' juvenile survival; ϕ adult survival; p, recapture; '$_t$' time effect; '$_{cov}$' covariate effect; '' constant parameter; Np, number of separately identifiable parameters. Note that no individuals were marked during missing occasions and consequently juvenile survival parameters do not exist in the model for cohorts without released juveniles (i.e. ϕ'_3, ϕ'_4 for dataset 1 and ϕ'_3 for dataset 2). Similarly, adults were not marked on the first occasion and, consequently, ϕ_1 do not exist in the model.

Tabla 1 . Parámetros no identificables por separado en presencia de ocasiones sin muestreo: ϕ' supervivencia juvenil; ϕ supervivencia en adultos; p, recaptura; '$_t$' efecto del tiempo; '$_{cov}$' efecto de una covariable; '' parámetro constante; N, número de parámetros identificables por separado. Durante las ocasiones sin muestreo no se marcó ningún individuo y, en consecuencia, no existen parámetros de supervivencia juvenil para los grupos sin juveniles liberados en el modelo (ϕ'_3, ϕ'_4 para el conjunto de datos 1 y ϕ'_3, para el conjunto de datos 2). De igual forma, en el primer muestreo no se marcó ningún adulto y, por lo tanto, ϕ_1 no existe en el modelo.

Model	Dataset 1 Two missing occasions [3, 4] Np	Not identifiable	Dataset 2 One missing occasion [3] Np	Not identifiable
$\phi'_t \phi_t p_t$	10	$\phi'_2; \phi'_6; \phi_2; \phi_3; \phi_4; \phi_6; p_6$	13	$\phi'_2; \phi'_6; \phi_2; \phi_3; \phi_6; p_6$
$\phi'_t \phi_t p.$	8	$\phi'_2; \phi_2; \phi_3; \phi_4$	10	$\phi'_2 ; \phi_2 ; \phi_3$
$\phi'_t \phi. p_t$	9		11	
$\phi'_t \phi. p.$	6		7	
$\phi'. \phi_t p_t$	9	$\phi_3; \phi_4$	11	
$\phi'. \phi_t p.$	6	$\phi_3; \phi_4$	7	
$\phi'_{cov} \phi_{cov} p_t$	8		9	
$\phi'_{cov} \phi_{cov} p.$	5		5	
$\phi'_{cov} \phi_t p_t$	10	$\phi_3; \phi_4$	12	
$\phi'_{cov} \phi_t p.$	7	$\phi_3; \phi_4$	8	
$\phi'_t \phi_{cov} p_t$	10		12	
$\phi'_t \phi_{cov} p.$	7		8	
$\phi'_{cov} \phi. p_t$	7		8	
$\phi'_{cov} \phi. p.$	4		4	
$\phi'. \phi_{cov} p_t$	7		8	
$\phi'. \phi_{cov} p.$	4		4	
$\phi'. \phi. p_t$	6		7	
$\phi'. \phi. p.$	3		3	

additional information and identifiability issues in full time–dependent models (tables 1, 2). If the parameters that appears in the gap are unrelated to known parameters from other intervals they will not be separately identifiable, e.g. if *all* survival parameters are time– and age–dependent, the first age–class survival probabilities ϕ' at the beginning of a gap occasion and the second age–class survival probabilities ϕ that follow them are not separately identifiable (tables 1, 2). However, the approach works well as long as one of the two types of survival parameters is kept constant or modelled as a function of environmental covariates (see simulated data results below, tables 1, 2).

Implementation 2: using a log–link function

An alternative to the above implementation and especially useful when the l_j are not commensurate (or when too many dummy occasions are required) relies on the use of a logarithm transformation (this implementation is not available in program E–SURGE). The survival probability over the initial interval of a young individual can be decomposed into its initial survival as a young for a duration r with a survival probability per time unit of ϕ' followed by the survival as an adult for a duration s with a survival probability per time unit of ϕ. The survival probability over the

Table 2. Identifiable quantities (i.e. * products of parameters) in the presence of missing occasions. Only models with redundant parameters are presented; see table 1. (Notation as in table 1).

Tabla 2. Cantidades identificables (productos de parámetros) en presencia de ocasiones sin muestreo. Solo se presentan los modelos con parámetros redundantes; véase la tabla 1. (Notación como en la tabla 1).*

| | | Dataset 1 | | Dataset 2 |
| | | Two missing occasions [3, 4] | | One missing occasion [3] |
Model	Np	Identifiable products	Np	Identifiable products
$\phi'_t \phi_t p_t$	10	$(\phi'_2{*}\phi_3{*}\phi_4); (\phi_2{*}\phi_3{*}\phi_4); (\phi'_6{*}p_6); (\phi_6{*}p_6)$	13	$(\phi'_2{*}\phi_3) ; (\phi_2{*}\phi_3); (\phi'_6{*}p_6); (\phi_6{*}p_6)$
$\phi'_t \phi_t p.$	8	$(\phi'_2{*}\phi_3{*}\phi_4) ; (\phi_2{*}\phi_3{*}\phi_4)$	10	$(\phi'_2{*}\phi_3) ; (\phi_2{*}\phi_3)$
$\phi'_t \phi_t p_t$	9	$(\phi_3{*}\phi_4)$		
$\phi'. \phi_t p.$	6	$(\phi_3{*}\phi_4)$		
$\phi'_{cov} \phi_t p_t$	10	$(\phi_3{*}\phi_4)$		
$\phi'_{cov} \phi_t p.$	7	$(\phi_3{*}\phi_4)$		

whole interval of length $r + s$ is then $\phi'^r \phi^s$ (for year based sampling $r = 1$). Applying a log–link function the product $\phi'^r \phi^s$ is replaced with a linear combination of survival related quantities, $\log(\phi'^r \phi^s) = r\log(\phi') + s\log(\phi)$, to be estimated. The known quantities, r and s, can be used as covariates of the survival probability pertaining to the interval (fig. S2.2 supplementary material S2). This approach does not require changing the encounter histories contrary to implementation 1. A similar solution was used by Tavecchia et al. (2001, 2002) to estimate monthly survival of game species when marking occurred at different moments during the hunting season.

Simulated cases

To demonstrate the problem generated by unequal time interval in combination with age–dependent models, we considered a simple scenario with a model assuming two age classes and constant survival and recapture parameters. Note that this simple scenario is only for illustrative purposes. We simulated 100 datasets with five sampling occasions during a 7–years period ($k = 7$). A hundred new juvenile animals were released at each occasion. The time span elapsed between occasions was as $l = 1, 3, 1, 1$. We assumed constant yearly survival of newly marked juvenile individuals ($\phi' = 0.4$) and constant yearly survival of adult individuals ($\phi = 0.8$). We first analysed these datasets using unequal time interval (the incorrect approach) to illustrate the biases. We subsequently analysed them adding dummy columns to 'fill' the years without monitoring (implementation 1) and using the log–link implementation with r (= 1) and s (= 2) values for newly marked individuals in the second cohort to constrain the corresponding survival parameters appropriately (implementation 2). For each analysis, maximum likelihood estimates of ϕ' and ϕ were obtained using RMark (Laake, 2013). The code to simulate the data and run the analysis is provided

in supplementary material S1. Supplementary material on S2 and S3 illustrate how to implement the solution in MARK (White and Burnham, 1999) and E–SURGE, respectively (Choquet et al., 2009).

Parameter identifiability in time–dependent models

In the example above we have assumed constant survival and recapture probabilities to illustrate the problem and its solution. However, in many cases, parameters are time dependent. To show the applicability of the solution to more complex parameter structures and to study parameter identifiability, we simulated two different datasets with seven intervals ($k = 7$, supplementary material S4). Both datasets considered temporal variation on juvenile and adult survival parameters as a function of a hypothetical temporal covariate D (adult survival was modelled as $1/(1 + \exp(- (1.386 + 0.55 * D)))$ and juvenile survival as one half of adult survival at each occasion) and a constant recapture probability of 0.7. Datasets differed in the length of the period with no CMR information: dataset 1 considered a gap of two years (data from sessions 3 and 4 are missing), while dataset 2 considered a gap of one year only (data from session 3 are missing) so that $l = 1, 3, 1, 1$ and $l = 1, 2, 1, 1, 1$, respectively. To avoid identifiability problems associated with sample size (which is not within the scope of this note) a thousand new juvenile animals were released on each occasion. To explore parameter identifiability, we implemented 18 models to each simulated dataset, considering different combinations of temporal, covariate and constant effects on juvenile and adult survival probabilities, and temporal vs. constant effect on recapture probability. Datasets were analysed using program E–SURGE (Choquet et al., 2009), which provides detailed results on parameter identifiability using the explicit method proposed by Catchpole and Morgan (1997) to detect parameter redundancy (supplementary material S4).
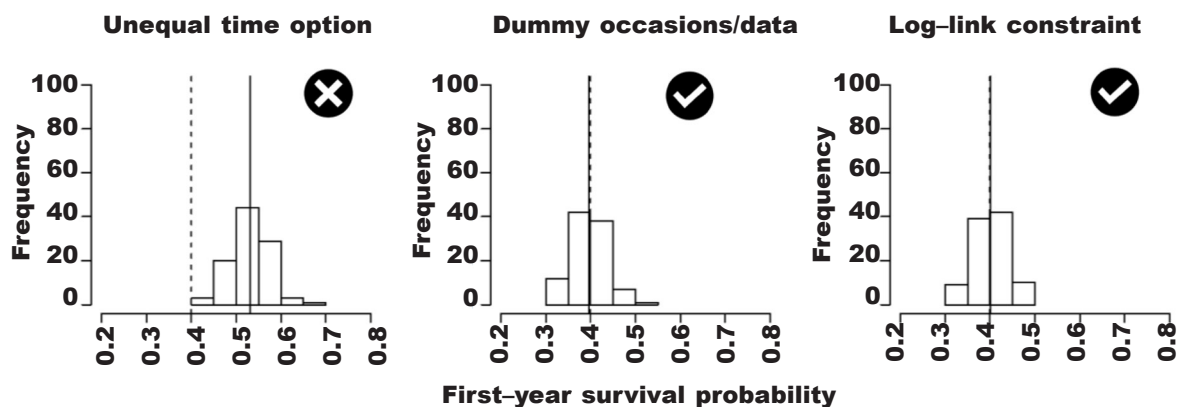
Fig. 1. Maximum likelihood estimates of first–year survival probability from 100 hypothetical datasets (see text). Vertical lines indicated the estimated average values (solid) and the true simulated value (dashed).

Fig. 1. Estimaciones por máxima verosimilitud de la probabilidad de supervivencia en el primer año de 100 conjuntos de datos hipotéticos (véase el texto para obtener información más detallada). Las líneas verticales continuas indican el promedio de los valores estimados y las discontinuas, el valor simulado verdadero.

Application to real case

We considered a dataset of capture–mark–recapture data of adult Mediterranean storm petrels (*Hydrobates pelagicus melitensis*) from Palomas Island (Eastern Spain). Birds were captured using mist–nets from 1996–2000 and from 2004–2006. The vector I_j was 1,1,1,1,4,1,1 where the '4' stands for the 3–year gap between 2000 and 2004. A full analysis of this dataset can be found in Sanz–Aguilar et al. (2010). Here we report results obtained by using the unequal time interval option and the proposed solution for comparative purpose. We only present implementation 1 as results of both implementations are equivalent (see results). The goodness of fit test of a model assuming all parameters time dependent (Pradel et al., 2005) indicated a surplus of animals seen only at marking, i.e., transients ($\phi^2_6 = 29.96$, $p < 0.05$). As a consequence, survival during the first year after marking, ϕ', was considered separately from the subsequent survival, noted ϕ, in 2–age class models (Sanz–Aguilar et al., 2010).

Results

Survival estimates in simple two–age classes constant models

Models in which we specified the unequal time interval in the *I*–vector delivered average estimates of first–year survival probability larger than the true value (average $\phi' = 0.52$; 95% CI = 0.46–0.60) instead of 0.4; fig. 1; supplementary material S1). This approach also led to a slightly underestimated recapture probability (average $p = 0.63$ (95% CI = 0.53–0.72)

instead of 0.70; supplementary material S1), while adult survival was close to the real value of 0.8 (average $\phi = 0.79$; 95% CI = 0.71–0.85). In contrast, the solution outlined above delivered the true values for all parameters regardless of the way the model was implemented ($\phi' = 0.40$, $\phi = 0.80$ and $p = 0.70$; fig. 1; supplementary material S1).

Parameter identifiability in time–dependent models

Our results indicate that models with two consecutive gaps present more problems of parameter identifiability than models with a single missing occasion (table 1). When juvenile and adult survival parameters are fully time dependent, survival parameters during the occasions without monitoring are not separately identifiable (tables 1, 2). Moreover, when the recapture probability is also time dependent, the last survival and recapture are not separately identifiable (tables 1, 2). All parameters became identifiable when juvenile and/or adult survival is constant or modelled as a function of temporal covariates with the exception of models in which more than one consecutive occasion without monitoring and adult survival was time dependent (tables 1, 2). In this case, only the adult survival parameter corresponding to the year in which the gap begins was separately identifiable (tables 1, 2).

Application to real case

As in the simulated example, when using vector *I* as exponent of survival parameters, we obtained higher estimates for first–year survival probabilities. When gap years were not properly considered, models with the unequal time interval option delivered survival estimates of $\phi' = 0.73$ and $\phi = 0.80$ (transient pro-

Table 3. AIC$_c$ values of CMR models used to estimate survival of Storm petrels captured at Palomas Island using unequal time interval only (the wrong approach) and the solution using dummy occasions (see text for details). In all models, the recapture probability was kept time dependent (see Sanz–Aguilar et al., 2010): Np, number of parameters; Ui, unequal intervals; Do, dummy occasions. (The lowest AIC$_c$ value in each approach is in bold.)

Tabla 3. Valores de AIC$_c$ de los modelos de CRM utilizados para estimar la supervivencia de los paíños europeos capturados en la Isla de las Palomas usando la opción de intervalos espaciados desigualmente (el método incorrecto) y la solución añadiendo ocasiones sin datos (véase el texto para obtener información más detallada). La probabilidad de recaptura se mantuvo variable en el tiempo en todos los modelos (véase Sanz–Aguilar et al., 2010): Np, número de parámetros; Ui, intérvalos desiguales; Do, ocasiones sin datos. (El valor de AIC$_c$ más bajo para cada método se muestra en negrita.)

Model	Np	Ui	Do
$\phi'_t \phi_t$	19	781.73	781.73
$\phi'_t \phi$	15	**777.31**	777.31
$\phi' \phi_t$	14	786.60	779.86
$\phi' \phi$	9	780.61	**772.91**

portion π = 0.087). However, when gaps were filled with dummy data and recapture probability fixed to 0 in years 2001, 2002 and 2003 the survival estimates were ϕ' = 0.40 and ϕ = 0.83 (transient proportion π = 0.518). Moreover, in this analysis, model information theory based on Akaike's Information Criterion (AIC$_c$, Burnham and Anderson, 2002) would have led to an over complex structure of survival selecting a model assuming a temporal variation in ϕ', and consequently in transient proportion π, when gap years were not properly considered (table 3).

Discussion

Obtaining unbiased estimates of demographic parameters (such as age–dependent survival) is essential for population diagnosis (Williams et al., 2002; Sanz–Aguilar et al., 2016). To achieve informed management decisions concerning biodiversity conservation, therefore, demographic parameters must be accurately estimated. Here we demonstrate that when not handled properly, the combination of unequal time intervals and age dependence in capture–recapture models can lead to erroneous estimates of survival and model selection (i.e. biological inference). To avoid the problem of unequal time intervals, sampling protocols should be properly designed. Here we showhow to partially overcome the s problem of uneven intervals in two age–classes capture–recapture models. Adding dummy columns to the encounter history can generally be used when interval lengths are commensurable with a same unit of time during which no change of age class occurs (one month, for instance). However, there might be practical limitations because in some cases this solution would lead to add a great number of dummy occasions and would 'push' estimate of survival probability close to the upper boundary value of 1. In these cases, a log–link function can be especially useful to accommodate heterogeneity in the duration of encounter occasions when uneven periods are of small duration and too many dummy occasions should otherwise be incorporated to account for the unequal time period. However, not all problems can be solved using the approaches suggested. For example, the log–link function works well in rich datasets but might cause numerical problems when data are sparse (Tavecchia et al., 2001). Moreover, by using the log–link function, the effect of temporal covariates cannot be modelled. Also, when survival parameters are fully time–dependent there are still some parameters that cannot be estimated separately, with only their products being identifiable (tables. 1, 2). The longer the period with missing information the higher the number of redundant parameters (tables 1, 2). However, by constraining survival using external covariates, most parameters become identifiable. Here we focused on two age class, single state models, but more complex models such as multistate models or models including multiple age–classes will present additional parameter identifiability problems. Despite these limitations, the solution presented here performed well in relatively simple situations and we recommend its use when age–dependent parameters are incorporated in models with uneven intervals between sampling occasions. Finally, the presence of transient animals can be accommodated in CMR models by using age–dependent models (see the real case, Pradel et al., 1997; Sanz–Aguilar et al., 2010). However, recently developed multi–event models (Pradel, 2005) allow to model transients as a specific uncertain category of individuals with known parameter values (survival probability = 0). In this case, age–dependent models are no longer necessary and the problem does not apply (e.g. Genovart et al., 2012; Santidrián Tomillo et al., 2017).

Acknowledgements

References

Bears, H., Martin, K., White, G. C., 2009. Breeding in high-elevation habitat results in shift to slower life-history strategy within a single species. *Journal of Animal Ecology*, 78: 365–375.

Bonenfant, C., Gaillard, J., Klein, F., Loison, A., 2002. Sex-and age-dependent effects of population density on life history traits of red deer *Cervus elaphus* in a temperate forest. *Ecography*, 25: 446–458.

Burnham, K. P., Anderson, D. R., 2002. Model selection and multi–model inference: a practical information–theoretic approach. Springer, New York.

Catchpole, E. A., Fan, Y., Morgan, B. J., Clutton–Brock, T. H., Coulson, T., 2004. Sexual dimorphism, survival and dispersal in red deer. *Journal of Agricultural, Biological, and Environmental Statistics*, 9: 1–26.

Catchpole, E. A., Morgan B. J., 1997. Detecting parameter redundancy. *Biometrika*, 84, 187–196.

Choquet, R., Rouan, L., Pradel, R., 2009. Program E–SURGE: a software application for fitting multievent models. In: *Modelling demographic processes in marked populations*: 845–865 (D. L. Thomson, E. G. Cooch, M. J. Conroy, Eds.), Springer series: Environmental and ecological statistics Vol. 3, New York.

Clobert, J., Perrins, C. M., McCleery, R. H., Gosler, A. G., 1988. Survival rate in the great tit Parus major in relation to sex, age, and immigration status. *Journal of Animal Ecology*, 57: 287–306.

Cooch, E., 2009. Program MARK: a gentle introduction. http://www.phidot.org/software/mark/docs/book/

Covas, R., Brown, C. R., Anderson, M. D., Brown, M. B., 2002. Stabilizing selection on body mass in the sociable weaver Philetairus socius. *Proceedings of the Royal Society London Series B Biological Sciences,* 269: 1905–1909.

Genovart, M., Pradel, R., Oro, D., 2012. Exploiting uncertain ecological fieldwork data with multi–event capture–recapture modelling: an example with bird sex assignment. *Journal of Animal Ecology*, 81: 970–977.

Grosbois, V., Tavecchia, G., 2003. Modeling dispersal with capture–recapture data: disentangling decisions of leaving and settlement. *Ecology*, 84: 1225–1236.

Hiraldo, F., Negro, J. J., Donázar, J. A., Gaona, P., 1996. A demographic model for a population of the endangered lesser kestrel in southern Spain. *Journal of Applied Ecology*, 33: 1085–1093.

Laake, J. L., 2013. *RMark: An R Interface for Analysis of Capture–Recapture Data with MARK*. Alaska Fisheries Science Center, NOAA, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Seatle.

Lebreton, J. D., Burnham, K. P., Clobert, J., Anderson, D. R., 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological monographs,* 62: 67–118.

Loison, A., Festa–Bianchet, M., Gaillard, J. M., Jorgenson, J. T., Jullien, J. M., 1999. Age-specific survival in five populations of ungulates: evidence of senescence. *Ecology*, 80: 2539–2554.

Perret, N., Pradel, R., Miaud, C., Grolet, O., Joly, P., 2003. Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology*, 72: 567–575

Pollock, K. H., 1981. Capture–recapture models allowing for age–dependent survival and capture rates. *Biometrics*, 37: 521–529.

Pradel, R., 2005. Multievent: An extension of multistate capture–recapture models to uncertain states. *Biometrics*, 61: 442–447.

Pradel, R., Gimenez, O., Lebreton, J.–D., 2005. Principles and interest of GOF tests for multistate capture–recapture models. *Animal Biodiversity and Conservation*, 28: 189–204.

Pradel, R., Hines J. E., Lebreton, J.–D., Nichols, J. D., 1997. Capture–recapture survival models taking account of transients. *Biometrics*, 53: 60–72.

Prugnolle, F., Pilard, P., Brun, L., Tavecchia, G., 2003. First–year and adult survival of the endangered Lesser Kestrel *Falco naumanni* in southern France: Local first–year survival is higher in La Crau than in Spain with a high proportion of first–year birds returning to their natal area. *Bird Study*, 50: 68–72.

Sanz–Aguilar, A., De Pablo, F., Donázar, J. A., 2015. Age–dependent survival of island vs. mainland populations of two avian scavengers: delving into migration costs. *Oecologia*, 179: 405–414.

Sanz–Aguilar, A., Igual, J. M., Oro, D., Genovart, M., Tavecchia, G., 2016. Estimating recruitment and survival in partially monitored populations. *Journal of Applied Ecology*, 53: 73–82.

Sanz–Aguilar, A., Mínguez, E., Oro, D., 2012. Is laying a large egg expensive? Female–biased cost of first reproduction in a petrel. *The Auk*, 129: 510–516.

Sanz–Aguilar, A., Tavecchia, G., Mínguez, E., Massa, B., Lo Valvo, F., Ballesteros, G. A., Barberá, G. G., Amengual, J. F., Rodriguez, A., McMinn, M., Oro, D., 2010. Recapture processes and biological inference in monitoring burrow–nesting seabirds. *Journal of Ornithology*, 151: 133–146.

Sanz–Aguilar, A., Tavecchia, G., Pradel, R., Mínguez, E., Oro, D., 2008. The cost of reproduction and experience–dependent vital rates in small petrel. *Ecology*, 89: 3195–3203.

Santidrián Tomillo, P., Robinson, N. J., Sanz–Aguilar, A., Spotila, J. R., Paladino, F. V., Tavecchia, G., 2017. High and variable mortality of leatherback turtles reveal possible anthropogenic impacts. *Ecology*, 98: 2170–2179.

Saraux, C., Le Bohec, C., Durant, J. M., Viblanc, V. A., Gauthier–Clerc, M., Beaune, D., Park, Y.–H.,

Yoccoz, N. G., Stenseth, N. C., Le Maho, Y., 2011. Reliability of flipper–banded penguins as indicators of climate change. *Nature,* 469: 203–206.

Schmidt, B. R., Schaub, M., Steinfartz, S., 2007. Apparent survival of the salamander *Salamandra salamandra* is low because of high migratory activity. *Frontiers in Zoology*, 4: 19.

Tavecchia, G., Pradel, R., Genovart, M., Oro, D., 2007. Density–dependent parameters and demographic equilibrium in open populations. *Oikos*, 116: 1481–1492.

Tavecchia, G., Pradel, R., Gossmann, F., Bastat, C., Ferrand, Y., Lebreton, J–D., 2002. Temporal variation in annual survival probability of the Eurasian woodcock *Scolopax rusticola* wintering in France. *Wildlife Biology*, 8: 21–30.

Tavecchia, G., Pradel, R., Lebreton, J–D., Johnson, A. R., Mondain–Monval, J–Y., 2001. The effect of lead exposure on survival of adult mallards in the Camargue, southern France. *Journal of Applied Ecology*, 38: 1197–1207.

White, G. C., Burnham, K. P., 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study*, 46: S120–S139.

Williams, B., Nichols, J. D., Conroy, M. J., 2002. Analysis and Management of Animal Populations: Modeling, Estimation and Decision Making. Academic Press, London, UK.

Zabala, J., Zuberogoitia, I., Martínez–Climent, J. A., Etxezarreta, J., 2011. Do long lived seabirds reduce the negative effects of acute pollution on adult survival by skipping breeding? A study with European storm petrels *(Hydrobates pelagicus)* during the "Prestige" oil–spill. *Marine Pollution Bulletin*, 62: 109–115.

Zuberogoitia, I., Zabala, J., Etxezarreta, J., Crespo, A., Burgos, G., Arizaga, J., 2016. Assessing the impact of extreme adverse weather on the biological traits of a European storm petrel colony. *Population Ecology*, 58: 303–313.

Supplementary material

Supplementary material S1. R–code for data simulation and analyses using RMARK.

Supplementary material S2. Design matrices for solution in MARK and E–SURGE.

Supplementary material S3. Implementing the solution using dummy occasions and a log–link in MARK.

Supplementary material S4. Datasets 1 and 2 and models (tables 1, 2) implemented in E–SURGE.

Supplementary material S2

The design matrices to implement the analyses in program MARK using implementation 1 (fig. S2.1) and implementation 2 (fig. S2.2), respectively. Note that encounter histories must be changed before implementing solution 2 to include the dummy occasion to fill–in years without monitoring. The files for the full analyses with *.inp, *dbf and *ftp can be found in supplementary material S3.
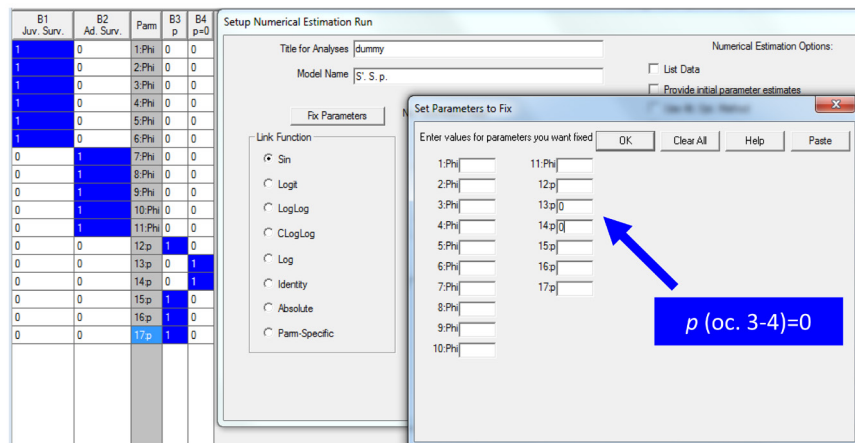

Fig. S2.1. Design matrix of a model with constant ϕ', ϕ and p in program MARK for the simulated datasets adding dummy occasions: B1 = ϕ', B2 = ϕ, B3 = p (files in supplementary material S3).

Fig. S2.1. Diseño de matriz de un modelo con ϕ' constante, ϕ y p en el programa MARK para los conjuntos de datos simulados añadiendo ocasiones sin muestreo: B1 = ϕ', B2 = ϕ, B3 = p (archivos en material suplementario S3).
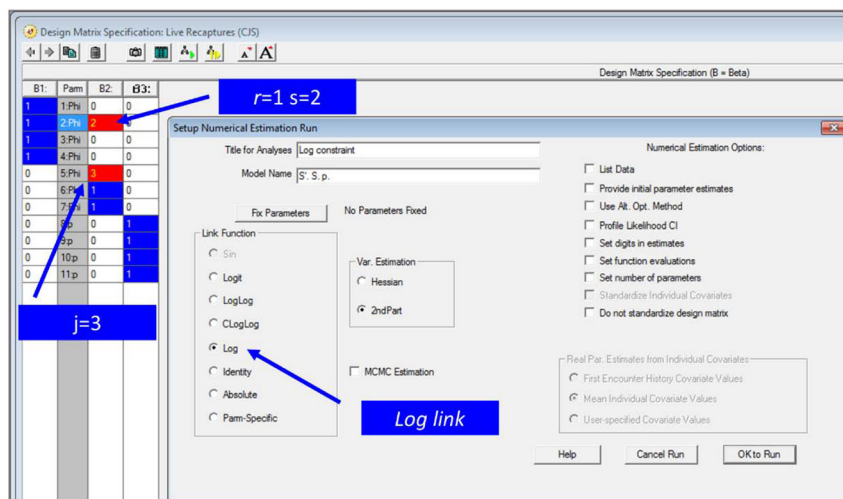

Fig. S2.2. Design matrix and run window of a model with constant ϕ', ϕ and p in program MARK: B1 = ϕ', B2 = ϕ, B3 = p (files in supplementary material S3).

Fig. S2.2. Ventana de diseño de matriz y ejecución de un modelo con ϕ' constante, ϕ y p en el programa MARK: B1 = ϕ', B2 = ϕ, B3 = p (archivos en material suplementario S3).
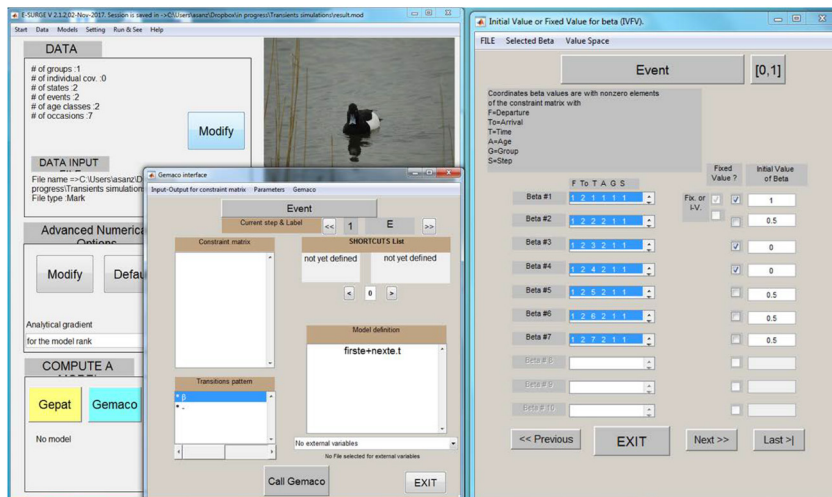
Fig. S2.3. GEMACO and IVFV steps in program E–SURGE to implement the solution to account for unequal time intervals by fixing recapture probabilities to zero.

Fig. S2.3. Pasos GEMACO y IVFV del programa E–SURGE para aplicar la solución con intervalos de tiempo desiguales fijando las probabilidades de recaptura a cero.