Methods for investigating parameter redundancy

O. Gimenez, A. Viallefont, E. A. Catchpole, R. Choquet & B. J. T. Morgan

Gimenez, O., Viallefont, A., Catchpole, E. A., Choquet, R. & Morgan, B. J. T., 2004. Methods for investigating parameter redundancy. *Animal Biodiversity and Conservation*, 27.1: 561–572.

Abstract
Methods for investigating parameter redundancy.— The quantitative study of marked individuals relies mainly on the use of meaningful biological models. Classical inference is then conducted based on the model likelihood, parameterized by parameters such as survival, recovery, transition and recapture probabilities. In classical statistics, we seek parameter estimates by maximising the likelihood. However, models are often overparameterized and, as a consequence, some parameters cannot be estimated separately. Identifying how many and which (functions of) parameters are estimable is thus crucial not only for proper model selection based upon likelihood ratio tests or information criteria but also for the interpretation of the estimates obtained. In this paper, we provide the reader with a description of the tools available to check for parameter redundancy. We aim to assist people in choosing the most appropriate method to solve their own specific problems.

Key words: Mark–recapture data, Mark–recovery data, Profile–likelihood, Analytical–numerical method, Symbolic algebra software.

Resumen
Métodos para investigar la redundancia de parámetros.— El estudio cuantitativo de individuos marcados se basa fundamentalmente en el uso de modelos biológicamente significativos. Posteriormente, la inferencia clásica se lleva a cabo a partir de la probabilidad del modelo, parametrizada mediante parámetros tales como las probabilidades de supervivencia, de recuperación, de transición y de recaptura. En la estadística clásica, intentamos obtener estimaciones de parámetros maximizando la probabilidad. Sin embargo, los modelos a menudo se parametrizan en exceso, por lo que algunos parámetros no pueden estimarse por separado. Por consiguiente, identificar qué parámetros, cuántos y qué funciones de los mismos son estimables resulta crucial, no sólo para poder efectuar una adecuada selección de modelos basada en pruebas de razón de verosimilitud o criterios de información, sino también para la interpretación de las estimaciones obtenidas. En este trabajo presentamos una descripción de las herramientas disponibles para verificar la redundancia de parámetros. Nuestro objetivo es ayudar a elegir el método más apropiado para la resolución de sus problemas específicos.

Palabras clave: Datos sobre recaptura de marcas, Datos sobre recuperación de marcas, Probabilidad del perfil, Método numérico analítico, Software de álgebra simbólica.

Olivier Gimenez, CEFE/CNRS, Equipe Biométrie et Biologie des Populations, 1919 route de Mende, 34293 Montpellier Cedex 5, France; Inst. de l'Ingénierie de l'Information de Santé, Lab. TIMC UMR CNRS 5525, Équipe TIMB, Fac. Médecine, 38706 La Tronche Cedex, France.– Anne Viallefont, Equipe de Recherche en Ingénierie des Connaissances, Univ. de Lyon 2, 69676 Bron Cedex, France.– Edward. A. Catchpole, Univ. of New South Wales at the Australian Defence Force Academy, Canberra, ACT 2600, Australia.– Rémi Choquet, CEFE/CNRS, Equipe Biométrie et Biologie des Populations, 1919 route de Mende, 38706 La Tronche Cedex 5, France.– Byron J. T. Morgan, Inst. of Mathematics, Statistics and Actuarial Science, Univ. of Kent, Canterbury CT2 7NF, England.

Corresponding author: O. Gimenez. E–mail: gimenez@cefe.cnrs-mop.fr

ISSN: 1578–665X

Introduction

Capture–mark–recapture (CMR) models in the broad sense include all the models developed to estimate demographic parameters based on data from marked animals. The initial papers are those of Cormack (1964), Jolly (1965) and Seber (1965). Over the past two decades, many improvements of the methods have been provided, that have led to both a diversification and a generalization of the tools available. We can now describe briefly the situation in this field as follows: when the marked animals are recaptured (or resighted) alive, the models used are "mark–recapture" models stricto sensu. The parameters estimated in this case are survival and capture probabilities. For a thorough review of the developments of the method since the sixties, including constancy over time, group effects, constraints on the parameters, etc. see Lebreton et al. (1992). In other cases, for example game species, the animals are not seen again during their lifetime, but their time of death is known. Brownie et al. (1990) provide a clear review of many of the models that can be used, called "mark–recovery" models. The parameters estimated are survival probabilities, and "return probabilities", i.e. the probability that the mark of a dead animal is found. See also the development in Freeman & Morgan (1992). Integrated modelling of mark–recovery and mark–recapture data is considered by several authors, see e.g. Catchpole et al. (1998).

Models have also been developed to estimate transition rates between sites or states jointly with the survival probabilities, in either of the two main situations above (Arnason, 1973; Schwarz et al., 1993; Brownie et al., 1993). All these models have a common structure and can be combined in the framework of multi–state models, as has been shown by Lebreton & Pradel (2002).

We shall here outline the common structure of many CMR models. For details concerning the one–site capture–recapture models, see Lebreton et al. (1992). Let us call Ω the set of all the different "encounter histories" that have been observed in a data set. Let q denote the number of encounter histories, ω_i (i = 1,…, q) the i–th capture history, and n_i the number of animals with this encounter history. The ω_i constitute the q cells of a multinomial model with individual probabilities π_i, where the π_i is the probability of observing the encounter history ω_i, conditionally on the time of marking and first release of the corresponding individuals. The π_i can be expressed as functions of the parameters to estimate survival and/or transition probabilities, capture and/or return probabilities, etc. The likelihood of a model L is proportional to the product of the π_i namely $\prod_i (\pi_i)^{n_i}$. Estimating the parameters of the model by the maximum likelihood method will thus consist in finding the values of the parameters that maximize L.

A model is defined to be identifiable if no two values of the parameters give the same distribution function. A particular case of non–identifiability usually occurs due to overparameterization. The likelihood of a redundant model can be expressed as a function of fewer than the original number of parameters (Catchpole & Morgan, 1997). When CMR protocols are considered, one is often faced with this form of non–identifiability and that is probably the reason why "non–identifiable" has been widely used in place of "parameter redundant" in the literature.

To fix ideas, let us consider the standard mark–recapture Cormack–Jolly–Seber (CJS) model with K capture occasions. The raw encounter histories can be fruitfully summarized in the so–called reduced *m*–array (Burnham et al., 1987, p.36) which summarises the data in the form of the number of individuals released per occasion i denoted R_i ($1 \le i \le K–1$) and the number of first recaptures given release at occasion i at the succeeding occasions j ($2 \le j \le K$) denoted m_{ij} (table 1).

For instance, throughout this paper, we will consider the well known Dipper example (Lebreton et al., 1992). During the breeding season, over a period of 7 years (1981–87), a total of 294 birds were marked and resighted. The data are summarized in table 2.

Conditioning on the releases and assuming independence among cohorts, the CJS model likelihood can therefore be easily written down as a product of multinomial probability distributions with the *m*–array cell probabilities given in table 3 (e.g. Lebreton et al., 1992).

For the CJS model, it is well known that the last survival probability and the last recapture probability cannot be estimated separately, only their product being estimable (e.g. Lebreton et al., 1992). Because ϕ_2 and p_3 only appear together in the cell probabilities, the likelihood can be rewritten in terms of ϕ_1, p_2 and $\beta = \phi_2 p_3$ as shown in table 4.

Table 1. The observed m–array for the CJS model with 3 capture occasion: Ocr. Occasions of release; Nr. Number released; FRc. First recapture occasion; R_i. Number released at occasion i; m_{ij}. Number of first recaptures at occasion j, given release at occasion i.

Tabla 1. Matriz m *observada para el modelo CJS con 3 capturas: Ocr. Lliberaciónes; Nr. Cantidad de liberados; FRc. Primera ocasion de recaptura; R_i. Número liberado en la ocasión* i*; m_{ij}. Número de primeras recapturas en la ocasión* j*, cuando la liberación se ha producido en la ocasión* i.

		FRc	
Ocr	Nr	2	3
1	R_1	m_{12}	m_{13}
2	R_2		m_{23}

Table 2. The *m*–array of the Dipper data set. In 1981, 22 birds were released among which, 11 were first recaptured in 1982, 2 in 1983, and 9 (= 22 – 11 – 2) were never observed again: Yr. Year of release, Nr. Number released; YFRr. Year of first recapture.

Tabla 2. Matriz m *para los datos del Mirlo acuático. En 1981, se libararon 22 aves de las cuales 11 fueron recapturadas por primera vez en 1982, 2 en 1983 y 9 no sa han vuelto a observar: Yr. Año de liberación; Nr. Cantidad liberada; YFRc. Año de la primera recaptura.*

		YFRc					
Yr	Nr	1982	1983	1984	1985	1986	1987
1981	22	11	2	0	0	0	0
1982	60		24	1	0	0	0
1983	78			34	2	0	0
1984	80				45	1	2
1985	88					51	0
1986	98						52

As a consequence, it has to be borne in mind that both estimates of ϕ_2 and p_3 cannot be separately discussed, except under the form of their product, otherwise one will certainly get misleading conclusions. E.g. for the Dipper data set, maximum likelihood estimates (MLEs) for ϕ_6 and p_7, both equal to 0.728, were obtained both with software MARK (White & Burnham, 1999) and M–SURGE (Choquet et al., 2003), resulting in a MLE for the product equal to 0.53. In fact, infinitely many other combinations also work e.g. ϕ_6 and p_7 equal respectively to 0.59 and 0.9. In this particular case, one should not look for a complex explanation of a decrease in the survival probability at the end of the study, but rather, consider it as an artefact due to the redundancy.

Moreover, model selection is often achieved in capture–recapture studies via the Akaïke information criterion (AIC), as recommended by Burnham & Anderson (1998). This criterion is:

$$AIC = -2 \log (L_{max}) + 2 n_p$$

where L_{max} is the maximum likelihood, and n_p is the number of estimable (functions of) parameters. To calculate this criterion in the CJS case, one has to subtract one from the total number of parameters in order to obtain the number of actually estimable parameters. Consequently, a naive computation of the number of parameters may lead to a wrong AIC–ranking of the models.

Table 3. The *m*–array cell probabilities for the CJS model with 3 capture occasions: Ocr. Occasion of release; Nr. Number released; FRc. First recapture occasion; ϕ_i. Survival probability between i and i + 1; p_j. Detection probability at occasion j.

Tabla 3. Probabilidades de las celdas de la matriz m *para el modelo CJS con 3 ocasiones de captura: Ocr. Lliberación; Nr. cantidad de liberados; FRc. Primera ocasión de recaptura;* ϕ_i. *Probabilidad de supervivencia entre* i *e* i + *1;* p_j. *Probabilidad detectada en la ocasión* j.

		FRc	
Ocr	Nr	2	3
1	R_1	$\phi_1 p_2$	$\phi_1 (1 - p_2) \phi_2 p_3$
2	R_2		$\phi_2 p_3$

Table 4. The re–parameterization of the *m*–array cell probabilities for the CJS model with 3 capture occasions: Ocr. Occasion of release; Nr. Number released; FRc. First recapture occasion; ϕ_i. Survival probability between i and i + 1; p_j. Detection probability at occasion j.

Tabla 4. Reparametrización de las probabilidades de las celdas de la matriz m *para el modelo CJS con 3 ocasiones de captura. Ocr. Liberación; Nr. Cantidad de liberados: FRc. primera oportunidad de recaptura;* ϕ_i. *Probabilidad de supervivencia entre* i *e* i + *1;* p_j. *Probabilidad detectada en la ocasión* j.

		FRc	
Ocr	Nr	2	3
1	R_1	$\phi_1 p_2$	$\phi_1 (1 - p_2) \beta$
2	R_2		β

Determining how many and which functions of the original parameters are estimable is thus crucial in model selection and in the interpretation of estimates. Two questions naturally arise, and we will focus on them in the next section:

Question 1 (Q1) – How many parameters are estimable? This number is called the rank of the model.

Question 2 (Q2) – Which parameters are estimable?

In the preceding didactic example, the conclusions could have been reached by visual inspection, trying to find the parameters which appear only together. However this approach becomes intractable for complex models.

Additionally, parameter redundancy is not always intuitive. Indeed the common belief that parameter redundancy is the result of "too many" parameters can be completely misleading. A good example occurs in modelling data from mark–recovery studies of animals banded at birth. The so–called Seber model (Seber, 1971) has a fully age–dependent survival and a constant recovery probability, with no time dependence in survival or recovery probabilities. This model is parameter redundant, and yet when extra parameters are added by allowing first–year survival to be time dependent, the model becomes non redundant (for details see Morgan & Freeman, 1989; Catchpole et al., 1996).

Another issue arises demonstrating that there exists no simple "rule of thumb" that allows us to compute the number of non redundant parameters in a model. When constraints are used on the parameters, the number of non redundant parameters may depend on the type of link function that is used. Viallefont (1995) illustrated such a situation with group dependence (see also Catchpole et al., 2002).

What precedes is an inherent model property, called intrinsic redundancy. Such a situation can be detected a priori, by methods allowing us to detect redundancy problems in the structure of a model, independently for any specific data set. It could also be detected a posteriori, i.e. after fitting the model to the data set of interest.

However, there exists a second sort of redundancy, called extrinsic redundancy, due to a particular structure of the data, usually missing or sparse data. E.g. it may happen that no individuals are detected at time i, inducing the redundancy of ϕ_{i-1} and ϕ_i, with only the survival probability between i – 1 and i + 1 being estimable. Such redundancy can only be detected "a posteriori", i.e. when the model has been fitted to the specific data set for which the problem appears.

The purpose of this paper is to provide the reader with the tools available to check for intrinsic and extrinsic parameter redundancy and to choose the most relevant method to solve their own problems. In the next section, we review the procedures available for checking for parameter redundancy, giving explanations and illustrations. Four approaches are considered and illustrated with the CJS model in conjunction with the Dipper data. We emphasize relative drawbacks and advantages and provide recommendations concerning parameter redundancy for the user of models for marked individuals.

Methods to check for parameter–redundancy

To our knowledge, there exist four different methods that can be used to detect parameter–redundancy. Table 5 presents these methods according to their ability to detect a priori intrinsic redundancy, or a posteriori both types of redundancy, and to answer Q2. The first two methods are more "intuitive", a posteriori methods, whereas the next two methods are more appropriate to a priori detect intrinsic redundancy problems.

Table 5. Summary of the conditions of use and relative advantages of the four methods proposed. For details see the text.

Tabla 5. Resumen de las condiciones de uso y ventajas relativas de los cuatro métodos propuestos. Para detalles al respecto ver el texto.

Name of the method	Detection of intrinsic redundancy	Detection of extrinsic redundancy	Necessary software	Answer to Q2
Profile likelihood	Possible on simulated data	Yes	Any CMR software for "by hand" plots; routinely implemented in M–SURGE	No
The Hessian	Possible on simulated data	Yes	Implemented in MARK, M–SURGE	No
Simulation	Simulated data with large numbers released	No	RELEASE (see also MARK) for computation of expected numbers and any CMR software for the optimization step	Partial
The formal derivative matrix	Yes	Yes	MAPLE or MATHEMATICA	Yes

Profile likelihood

This a posteriori method is based on the fact that redundancy results in a flat ridge in the likelihood, hence inducing an infinity of solutions. The redundancy can thus be shown by plotting the profile likelihood i.e. the likelihood as a function of a parameter of interest and simultaneously maximized over the other parameters (e.g. Freeman et al., 1992; Lebreton & Pradel, 2002).

This method is normally used a posteriori, i.e. on a given data set with a given model, and more specifically it is used to detect problems concerning a specific subset of parameters. One needs first to have an idea of which parameters are redundant. For example, if in two quite close models, very different estimates are found for one parameter, this may mean that it is redundant.

This method can also be used a priori on simulated data to detect intrinsic redundancy, or to distinguish between extrinsic and intrinsic redundancy problems: indeed, an intrinsic problem is model–dependent and remains so whatever the data set used, whereas an extrinsic problem is due to the structure of a specific data set, and disappears if a simulated data set with a different structure is used.

For the Dipper example, graphs of the profile deviance are shown in figure 1. Several model–fitting steps are necessary, with different fixed values at each time for the parameter under investigation. Then the different deviance values are plotted against the corresponding fixed value of the parameter. If the parameter is estimable uniquely by maximum likelihood, for example ϕ_1, one will get a higher deviance for any value of this parameter other than the MLE. On the contrary, if the parameter is not estimable, e.g. p_7, one will get a flat ridge according to the direction of this parameter. Actually, this flat ridge does not extend completely from 0 to 1. In this example, because we have to deal with probability, once the value of p_7 has been fixed between 0 and 1, the value of ϕ_6 has to be between $\beta = \phi_6 p_7$ and 1, otherwise the deviance is no longer constant, but as long as the ridge exists in the neighbourhood of the MLE, the concerned parameter must be considered redundant. Further detail on the extent of a ridge in the context of Seber's mark–recovery models is given in Catchpole et al. (1993) and Catchpole & Morgan (1994).

This method can be used "by hand" with any CMR software, while the program M–SURGE can automatically give the plot of profile deviance.

Applications can be found in Viallefont (1992), Freeman & Morgan (1992), Catchpole et al. (1993), Catchpole & Morgan (1994), Lebreton & Pradel (2002), Pradel et al. (in prep.) and Gimenez et al. (submitted).

This method is a graphical diagnostic only of the model parameter redundancy. It should not be used systematically, which would necessitate drawing as many graphs as there are parameters in the model, which would be very time–consuming.

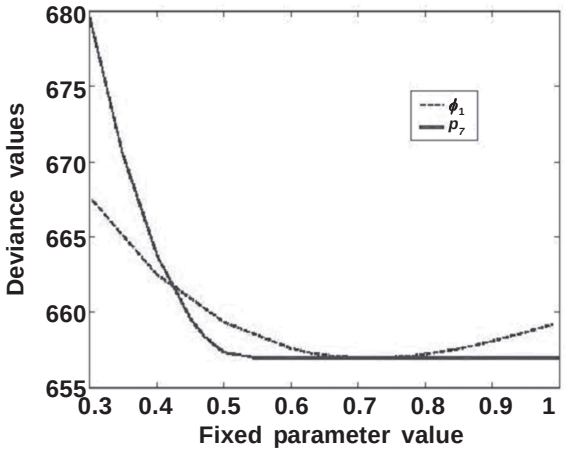

Fig. 1. Profile deviance for two parameters of the CJS model: application to the Dipper data set.

Fig. 1. Desviación del perfil para dos parámetros del modelo CJS: aplicación al conjunto de datos del Mirlo acuático.

The Hessian

This a posteriori method is based on detecting zero eigenvalues of the matrix of the second derivatives of the log–likelihood with respect to the parameters —namely the Hessian matrix— evaluated at the MLE. The model rank is computed as the number of non–zero eigenvalues (Viallefont et al., 1998) i.e. the numerical rank of the Hessian (the number of linearly independent rows). It is also possible to determine the parameters that are separately estimable by computing the eigenvectors associated with the zero eigenvalues and identifying their null co–ordinates (Reboulet et al., 1999).

For the Dipper example, the 12 x 12 Hessian matrix and the associated eigenvalues are given in table 6. The eigenvalue in bold is close to zero meaning that the model rank equals 11. Out of the 12 original parameters, only 11 are estimable, confirming what is known about the CJS model. In addition, by considering the entries in the eigenvector corresponding to the smallest eigenvalue, it is confirmed that only the last survival and capture probabilities are redundant parameters (values in bold).

This method should be used cautiously because: (1) as with any a posteriori method, it does not distinguish between intrinsic and extrinsic redundancy, thus it does not allow generalisation of the results found for one specific data set to other data sets with the same structure mode; (2) it requires the Hessian matrix, which is often done numerically through an approximation via a finite difference scheme; and (3) a perfect tuning of a zero threshold value probably does not exist (Viallefont et al., 1998).

Moreover, using this method without caution induces another problem when some parameters are estimated at a boundary value of the parameter interval. For example, if one of the survival probabilities is estimated at the value of one or zero, the numerical computation of the Hessian by finite differences may induce one eigenvalue of this matrix to be null, thus leading to counting the corresponding parameter as if it was not separately estimable, whereas it obviously is (with a value at a boundary).

Despite all these problems, this numerically tractable method has been implemented extensively in the mark–recapture software SURGE (Viallefont et al., 1998; Reboulet et al., 1999), MARK (White & Burnham, 1999) and M–SURGE (Choquet et al., 2003). We advise the reader to be very careful when using MARK, where a wrong computation of the number of independent parameters in the models can lead to an unreliable ordering of the models via AIC. When using M–SURGE, the first derivatives are analytically computed which improves the precision.

Simulation

The first step of the simulation method (also called the analytical–numerical method) generates an artificial data set, by assuming a (realistic) set of parameter values and (large) ringing numbers, and then using the model to generate expected encounter histories. In a second step, this generated data set is analyzed with the model of interest using standard software such as MARK, M–SURGE, SURVIV (White, 1982) or MSSURVIV (Hines, 1994). For large numbers of released individuals per occasion, the MLEs and standard errors produced are approximately the expected values of the param-

Table 6. Numerical diagnostics for CJS model parameter redundancy with the Dipper data set.

Tabla 6. Diagnósticos numéricos para la redundancia de parámetros en el modelo CJS con el conjunto de datos del Mirlo acuático.

Hessian matrix computed by a finite difference scheme

2.524	0.779	0.027	0.001	0.000	0.000	1.343	0.055	0.002	0.000	0.000	0.000
0.779	14.478	0.495	0.019	0.001	0.000	–0.542	1.022	0.031	0.002	0.000	0.000
0.027	0.495	18.485	0.716	0.046	0.001	–0.019	–0.468	1.166	0.075	0.004	0.001
0.001	0.019	0.716	17.067	1.100	0.028	–0.001	–0.018	–0.662	1.790	0.084	0.028
0.000	0.001	0.046	1.100	20.076	0.503	–0.000	–0.001	–0.043	–1.017	1.528	0.503
0.000	0.000	0.001	0.028	0.503	8.325	–0.000	–0.000	–0.001	–0.025	–0.485	8.325
1.343	–0.542	–0.019	–0.001	–0.000	–0.000	2.407	–0.038	–0.001	–0.000	–0.000	–0.000
0.055	1.022	–0.468	–0.018	–0.001	–0.000	–0.038	1.035	–0.029	–0.002	–0.000	–0.000
0.002	0.031	1.166	–0.662	–0.043	–0.001	–0.001	–0.029	1.965	–0.069	–0.003	–0.001
0.000	0.002	0.075	1.790	–1.017	–0.025	–0.000	–0.002	–0.069	3.008	–0.077	–0.025
0.000	0.000	0.004	0.084	1.528	–0.485	–0.000	–0.000	–0.003	–0.077	2.042	–0.485
0.000	0.000	0.001	0.028	0.503	8.325	–0.000	–0.000	–0.001	–0.025	–0.485	8.325

Hessian eigenvalues	Eigenvector associated with the null eigenvalue	
0.00000364	–3.484365e–18	ϕ_1
0.94321127	1.371273e–16	ϕ_2
1.05686038	–8.138873e–17	ϕ_3
1.85022632	2.486811e–17	ϕ_4
1.87616909	6.028106e–17	ϕ_5
2.70801343	**–7.071068e–01**	ϕ_6
3.80631847	8.958153e–19	p_2
14.56896925	–1.775061e–15	p_3
16.52430693	2.851347e–16	p_4
16.87618746	2.764039e–16	p_5
18.83059874	5.154638e–17	p_6
20.69498495	**7.071068e–01**	p_7

Table 7. Expected m–array using R_i = 10,000 for all *i* and the parameter values ϕ_1 = 0.4; ϕ_2 = 0.5; ϕ_3 = 0.6; ϕ_4 = 0.6; ϕ_5 = 0.7; ϕ_6 = 0.7; p_2 = 0.9; p_3 = 0.9; p_4 = 0.9; p_5 = 0.9; p_7 = 0.7. Values are rounded to the nearest integer: YFRc. Year of first recapture; Yr. Year of release; Nr. Number released.

Tabla 7. Matriz m esperada utilizando R_i = 10.000 para todos los i *y los valores paramétricos ϕ_1 = 0,4; ϕ_2 = 0,5; ϕ_3 = 0,6; ϕ_4 = 0,6; ϕ_5 = 0,7; ϕ_6 = 0,7; p_2 = 0,9; p_3 = 0,9; p_4 = 0,9; p_5 = 0,9; p_7 = 0,7. Los valores se redondean al número entero más próximo: YFRc. Año de la primera recaptura; Yr. Año de liberación; Nr. Cantidad liberada.*

		YFRc					
Yr	Nr	1982	1983	1984	1985	1986	1987
1981	10,000	4,900	756	38	2	1	1
1982	10,000		3,600	180	11	1	1
1983	10,000			4,500	270	16	1
1984	10,000				5,400	324	29
1985	10,000					5,400	480
1986	10,000						8,000

eter estimators and their standard errors. If a parameter estimator is unbiased to the 5th decimal place and has a coefficient of variation less than 100%, then it is declared estimable (Kendall & Nichols, 2002).

For the CJS model example for the Dipper study, we used R_i = 10,000 for all i and parameter values shown in table 7. The resulting expected m–array values are given in table 7. Using SURVIV e.g., we analyze these data fitting the CJS model. The results in table 8 suggest that all parameters are estimable except for ϕ_6 (biased estimate) and p_7 (biased estimate and large coefficient of variation).

Recent applications can be found in Schaub et al. (2004) and Kendall & Nichols (2002).

This method can be used as a completely a priori method, to investigate the intrinsic redundancy of a model, by generating the simulated data set with arbitrarily fixed values of the model parameters. Note however that the ringing numbers used in the simulation should be large enough to ensure that there are no zero cells in the *m*–array, to ensure that any redundancy found must be intrinsic. However, it is also often used a posteriori, using the point estimates obtained to generate the simulated data. The investigation of the extrinsic problems of redundancy (i.e. due to the data) cannot be assessed by this method, because it relies on the data set being very large, which requires simulated rather than actual data.

Also, the simulation method is only valid for the particular values of the parameter that are chosen to compute the expected probabilities. We recommend using several set of different values to be sure that you do not have to deal with a very particular case (a model conditionally of full rank i.e. a model of full rank, but parameter redundant for one or several values of the parameters). For the Dipper example, we tried several sets of different values for the parameters and were led to the same conclusions.

Table 8. MLEs obtained from the simulated data of table 6: P. parameters; MLEs. Maximum likelihood values; SE. Standard error; B. Bias; Cv. Coefficient of variation (%).

Tabla 8. MLE obtenidos a partir de los datos simulados de la tabla 6: P. Parametros; MLEs. Valores de probabilidad máxima; SE. Error estandar; B. Sesgo; Cv. Coeficiente de variación (%).

P	MLEs	SE	B	Cv
ϕ_1	0.700	0.078	0	0.11
ϕ_2	0.400	0.070	–0.00001	0.17
ϕ_3	0.500	0.072	–0.00000	0.14
ϕ_4	0.600	0.071	–0.00000	0.12
ϕ_5	0.600	0.070	–0.00000	0.12
ϕ_6	0.894	0.067	0.19442	0.07
p_2	0.700	0.092	0.00000	0.13
p_3	0.900	0.080	0.00002	0.09
p_4	0.900	0.073	0.00002	0.08
p_5	0.900	0.070	0.00001	0.08
p_6	0.900	0.065	0.00001	0.07
p_7	0.894	225.184	0.19442	251.76

Given that the variances computed by SURVIV should not be trusted when one or several parameters are redundant in the model (Hines, pers. com.), we greatly encourage using MARK or M–SURGE which are more reliable in computing the Hessian matrix. Of course, generating all possible encounter histories for a large fixed number of capture occasions, and then writing down their expected values under a complex model, can quickly become time–consuming and often intractable. Note however that simulations of the CJS model can be conducted with program RELEASE which can be used as a standalone application (Burnham et al., 1987) or found as part of program MARK.

Finally, it is important to notice that this method also allows us to study relevant statistical quantities, such as bias, precision of estimators, or power of likelihood–ratio and goodness–of–fit tests (Viallefont et al., 1995; Pollock et al., 1985).

The formal derivative matrix

This a priori method is based on the analytical computation of the matrix D of derivatives of the vector of the multinomial distribution cell probabilities with respect to the vector of model parameters. This method gives the answer to both Q1: the number of estimable parameters is the symbolic row rank of D, and Q2: the estimable (functions of) parameters are the formal solutions of a system of partial differential equations (PDEs). The successive steps required to perform this method are shown in table 9 and will now be detailed using the CJS model example in conjunction with the Dipper data set. Symbolic calculus software such as Maple or Mathematica can be used at each step to greatly ease the mathematical burden.

Intrinsic parameter–redundancy

The first step requires forming the vector q of original parameters and a vector m of *m*–array cell probabilities under the CJS model. In both vectors the order is arbitrary. We choose

$$\theta = (\phi_1,...,\phi_6,\ p_2,...,p_7)^T$$

and $\mu = (\phi_1 p_2, \phi_1(1 - p_2)\phi_2 p_3,...,\phi_6 p_7)^T$

Then, step 2, we calculate the symbolic derivative matrix D of $\log(\mu)$ with respect to θ. The result is: Then the answer to Q1 (step 3) is simply the symbolic rank of D which can be easily obtained, again with Maple or Mathematica. For the CJS model, we find a deficiency of 1 i.e. a rank equal to 11. The eigenvector corresponding to the zero eigenvalues is (step 4)

Table 9. Different steps required to perform the formal method.

Tabla 9. Distintos pasos requeridos para ejecutar el método formal.

Steps	Mathematical objects and notation
1. Write down the vector of log–probabilities as a function of parameters	Vector of parameters: $\theta = (\theta_1,...,\theta_p)^T$ Vector of log–probabilities: $\log(\mu_1(\theta),...,\mu_n(\theta))^T$
2. Differentiate formally log ($\mu(\theta)$) wrt the components of θ	The derivative matrix: $D = \left(\frac{\partial \log(\mu_i(\theta))}{\partial \theta_l}\right)_{\substack{1\le l\le p \\ 1\le i\le n}}$
3. Determine the number *q* of estimable parameters: if $q < p$ the model is parameter redundant then go to step 4, otherwise the model is of full rank	Symbolic rank of the derivative matrix: *r = rank (D)*
4. Write down formal solutions of:	$a_i(\theta)^T D(\theta) = 0, \quad i = 1,...,d$
5. Determine position $i_1,...,i_s$ of 0 in common to all a_s	The *s* separately non–redundant parameters: $\theta_{i_1},...,\theta_{i_s}$
6. Write down the system of partial differential equations and solve it formally to obtain the estimable functions	The system of PDEs: $\sum_{l=1}^{p} a_{l,j} \frac{\partial f}{\partial \theta_l} = 0 \quad j = 1,...,d$

$$D = \begin{pmatrix}
\frac{1}{\phi_1} & \frac{1}{\phi_1} & \frac{1}{\phi_1} & \frac{1}{\phi_1} & \frac{1}{\phi_1} & \frac{1}{\phi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & \frac{1}{\phi_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & 0 & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & \frac{1}{\phi_3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & 0 & 0 & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & 0 & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & \frac{1}{\phi_4} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{\phi_5} & \frac{1}{\phi_5} & 0 & 0 & 0 & \frac{1}{\phi_5} & \frac{1}{\phi_5} & 0 & 0 & \frac{1}{\phi_5} & \frac{1}{\phi_5} & 0 & \frac{1}{\phi_5} & \frac{1}{\phi_5} & \frac{1}{\phi_5} & \frac{1}{\phi_5} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\phi_6} & 0 & 0 & 0 & 0 & \frac{1}{\phi_6} & 0 & 0 & 0 & \frac{1}{\phi_6} & 0 & 0 & \frac{1}{\phi_6} & 0 & \frac{1}{\phi_6} & \frac{1}{\phi_6} \\
\frac{1}{p_2} & -\frac{1}{\bar{p}_2} & -\frac{1}{\bar{p}_2} & -\frac{1}{\bar{p}_2} & -\frac{1}{\bar{p}_2} & -\frac{1}{\bar{p}_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{p_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & \frac{1}{p_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & -\frac{1}{\bar{p}_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{p_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & 0 & \frac{1}{p_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & \frac{1}{p_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & -\frac{1}{\bar{p}_4} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{p_5} & -\frac{1}{\bar{p}_5} & -\frac{1}{\bar{p}_5} & 0 & 0 & \frac{1}{p_5} & -\frac{1}{\bar{p}_5} & -\frac{1}{\bar{p}_5} & 0 & \frac{1}{p_5} & -\frac{1}{\bar{p}_5} & -\frac{1}{\bar{p}_5} & \frac{1}{p_5} & -\frac{1}{\bar{p}_5} & \frac{1}{p_5} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{p_6} & -\frac{1}{\bar{p}_6} & 0 & 0 & 0 & \frac{1}{p_6} & -\frac{1}{\bar{p}_6} & 0 & 0 & \frac{1}{p_6} & -\frac{1}{\bar{p}_6} & 0 & \frac{1}{p_6} & -\frac{1}{\bar{p}_6} & \frac{1}{p_6} & -\frac{1}{\bar{p}_6} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{p_7} & 0 & 0 & 0 & 0 & \frac{1}{p_7} & 0 & 0 & 0 & \frac{1}{p_7} & 0 & 0 & \frac{1}{p_7} & 0 & \frac{1}{p_7} & \frac{1}{p_7}
\end{pmatrix}$$

where $\bar{p}_j = 1 - p_j$, $j = 2,...,7$

$$\frac{\partial f}{\partial \phi_6}\frac{\phi_6}{p_7} - \frac{\partial f}{\partial p_7} = 0$$

Since this (sole) eigenvector has zeroes in all entries except those corresponding to ϕ_6 and p_7, all parameters except those two are estimable step 5). To determine the full set of estimable functions, step 6 requires the solution of the following partial differential equation (PDE):

$$a = \left[0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; -\frac{\phi_6}{p_7} \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 1\right]^T$$

The solutions of this PDE are ϕ_1, ϕ_2, ϕ_3, ϕ_4, ϕ_5 and p_2, p_3, p_4, p_5, p_6 and the product $\phi_6 p_7$.

Extrinsic parameter–redundancy

What precedes deals only with inherent properties of the CJS model. If the behaviour of parameter estimates is needed, one has to take the data into account. By considering non–structural zeroes i.e. missing data in the *m*–array, it is easy to adjust the method for checking for extrinsic parameter redundancy. One has just to modify step 1 by forming a vector *m* of *m*–array cell probabilities incorporating only the probabilities corresponding to non–zero m_{ij} (cell probabilities corresponding to zero m_{ij} do not play any role in building the likelihood, since they are raised to the power zero).

With the Dipper data set, there is a substantial amount of missing data, so that only 11 cells do not contain missing data (see table 2); m_{12}, m_{13}, m_{23}, m_{24}, m_{34}, m_{35}, m_{45}, m_{46}, m_{47}, m_{56} and m_{67}. The expression of the derivative matrix D is simpler, consisting of just 11 of the columns of the previous D, but its rank still remains equal to 11 so that in this case the missing data do not render any extra parameters redundant.

Recent applications can be found in Schaub et al. (2004). Other examples with more details about the theory and references can be found for single–state models in Catchpole et al. (2002) with associated Maple code freely available from http://www.ma.adfa.edu.au/~eac/Redundancy/Maple and for multistate capture–recapture models in Gimenez et al. (2003), for which Maple code is freely available from ftp://ftp.cefe.cnrs–mop.fr/bio/PRM.

Discussion

The four methods all give some information about the redundant parameters in a particular model or for the case when that model is fitted to a specific data set. They should be used in preference to any "rules of thumb". If, using one of these methods, the number of non–redundant parameters in a model is known with certainty, then information criteria can safely be computed. Otherwise, it is advisable to use the value automatically computed with the Hessian method by software such as MARK or M–SURGE. It has to be noted that the *intrinsic* number of non–redundant parameters is known for many models (Lebreton et al., 1992 for CJS–type models; Gimenez et al., 2003 for multistate models; see also Viallefont, 1995; Kendall & Nichols, 2002; Schaub et al., 2004). Hence, if the Hessian method yields different results, the structure of the data has to be checked to determine whether *extrinsic* reasons induce redundancy.

Another easy way to count the number of non–redundant parameters is the simulation method. But, both the Hessian and simulation methods may be flawed by numerical issues, such as deciding what "close to zero" means. There is a slight advantage of the simulation method as it does give an immediate answer to the question: "Which parameters are not separately estimable?" (e.g. ϕ_6 and p_7 in the CJS model with the Dipper example). Concerning the Hessian method, the variance of the redundant parameters is not always very different from that of the separately estimable parameters, and some insight in the eigenvectors is thus required.

Knowing which parameters are not separately estimable only constitutes a partial answer to Q2 in the introduction of this paper. In the Dipper example, only the formal method yields the information that the product $\phi_6 p_7$ is estimable, rather than any other function of these two parameters. Hence, only the formal method gives the full answer to Q2.

To make efficient use of the profile–likelihood method, one needs to have an idea about which parameters might be redundant, in other words, one should have an idea of the "partial answer" to Q2. The profile–likelihood is then an easy–to–use method, but the results only apply to the parameters for which the profile–likelihood plot has been drawn. To obtain a complete answer to Q1 with this method, one needs to plot the profile–likelihood for all the parameters, which will be very time–consuming in many cases.

Even if a model is non–redundant, difficulties in estimation can occur because of local minima (Lebreton & Pradel, 2002) or duality phenomena (Pradel et al., in prep.) when two different estimates of the same parameters can give the same value of the smallest deviance (this is a case of non–identifiability). In such cases, a graph of the profile–deviance can add valuable information.

While building the profile deviance for a parameter, it is also quite easy to derive at the same time profile–deviance–based confidence intervals which are known to be more robust than the classical Wald confidence interval with boundary estimates (Catchpole & Morgan, 1994 for mark–recovery models; Gimenez et al., submitted for multistate mark–recapture models).

To know *a priori* the intrinsic number of estimable parameters in a model, all four methods proposed here may be used on simulated data for the profile–likelihood, the Hessian or the analytical–numerical methods: it is then advisable to generate two or more data sets to check that the result is not due to a special case in the simulated data.

The formal method that does not require any simulated or expected data is not only more reliable, but also the only one to provide a clear answer to Q2.

Moreover, having worked out a particular case (e.g. CJS with 3 capture occasions) it is often possible to extend the conclusions to larger examples with the same structure (i.e. more years of data) without having to repeat the analysis (e.g. Catchpole & Morgan, 1997, 2001; Gimenez et al., 2003). We are currently trying to extend these ideas in order to provide a taxonomy of intrinsic redundancy of many standard models.

We thus advise anyone wanting to develop new models for the analysis of marked data to use the formal method to assess the properties of the new model they develop.

Clearly, it requires the use of specific software and some knowledge of algebra. We hope to see an automatic implementation in standard CMR software in the near future. Pending this progress, we anticipate that in the near future, combining the formal method with other methods will be important. It could be the way to go towards a reliable routine computation of numbers of estimable parameters. Such an approach is adapted in program M–SURGE where the first derivatives are analytically computed. A second example is given by Choquet & Pradel (unpublished results). They developed practical rules in order to simplify the structure of derivative matrix D which makes calculation of the formal rank easier.

From a Bayesian perspective, maximum likelihood theory is equivalent to finding the mode of the joint posterior distribution of the parameters, given uniform priors. Since Bayesians usually examine posterior means, rather than modes, issues of parameter redundancy are not apparent, and might well be thought to be of no importance. If there is parameter–redundancy, then the likelihood surface is flat, however if a Bayesian approach is adopted, then the posterior can result in remarkably precise estimators. This is investigated and explained in Brooks et al. (2000) for a particular example. In their case, Brooks et al. were aware that the ridge existed, but one could envisage examples arising when that was not the case. It is therefore important to be aware of parameter redundancy. Barry et al. (2000) found that the existence of parameter redundancy can have substantial impact on posterior means and standard deviations. Carlin & Louis (1996, p. 203) recommend against the use of Markov

chain Monte Carlo methods in the presence of parameter redundancy.

Knowing that a model is full–rank may not be enough. This has been shown by Catchpole et al. (2001). It is shown there that if a full–rank model has a parameter–redundant sub–model, and is insufficiently different from that submodel, then it may perform badly in practice. Thus in practice one should think carefully about the models to be used, and make use of whatever knowledge and general results that are available (see, eg., Catchpole et al., 1996). Even though a model may be full–rank, it can still be useful to check the values of the eigenvalues. This then produces a kind of synthesis of intrinsic and extrinsic procedures.

As a conclusion, the choice between methods clearly depends on the purpose of the study. As the tools presented here are enough to tackle all sorts of problems concerning parameter redundancy, we do hope that people will use them, in order to ensure valid biological conclusions.

References

Arnason, A. N., 1973. The estimation of population size, migration rates and survival in a stratified population. *Researches on Population Ecology* 15: 1–8.

Barry, S. C., Brooks, S. P., Catchpole, E. A. & Morgan, B. J. T., 2003. The Analysis of Ring–Recovery Data Using Random Effects. *Biometrics,* 59(1): 54–65.

Brooks, S. P., Catchpole, E. A., Morgan, B. J. T. & Barry, S. C., 2000. On the Bayesian analysis of ring–recovery data. *Biometrics,* 56: 951–956.

Brooks, S. P. & Morgan, B. J. T., 1995. Optimization using simulated annealing. *The Statistician,* 44: 241–257.

Brownie, C., Anderson, D. R., Burnham, K. P. & Robson, D. S., 1985. *Statistical inference from band recovery data — a handbook,* 2nd Ed. U.S. Fish and Wildlife. Service, Washington, DC.

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H. & Hestbeck, J. B., 1993. Capture–recapture studies for multiple strata including non–Markovian transitions. *Biometrics,* 49: 1173–1187.

Burnham, K. P., 1993. A theory for combined analysis of ring recovery and recapture data. In: *Marked Individuals in the Study of Bird Populatio*: 199–213 (J.–D. Lebreton & P. M. North, Eds.). Birkhauser Verlag, Basel, Switzerland

Burnham, K. P. & Anderson, D. R., 1998. Model selection and inference, a practical Information theoretic approach. Springer, New–York.

Burnham, K. P., Anderson, D. R., White, G. C., Brownie, C. & Pollock, K. H., 1987. *Design and analysis methods for fish survival experiments based on release–recapture.* Bethesda, Maryland, American Fisheries Society.

Carlin, B. P. & Louis, T. A., 1996. *Bayes and empirical Bayes methods for data analysis*. Chapman and Hall, London.

Catchpole, E. A., Freeman, S. N. & Morgan, B. J. T., 1993. On boundary estimation in ring recovery models and the effect of adding recapture information. In: *Marked individuals in the study of Bird Populations:* 215–228 (J.–D. Lebreton & P. M. North, Eds.). Birkhauser–Verlag, Basel, Switzerland.

– 1996. Steps to parameter redundancy in age–dependent recovery models. *J. R. Statistic. Soc., B,* 58: 763–774.

Catchpole, E. A., Freeman, S. N., Morgan, B. J. T. & Harris, M. P., 1998. Integrated recovery/recapture data analysis. *Biometrics,* 54: 33–46.

Catchpole, E. A., Kgosi, P. M. & Morgan, B. J. T., 2001. On the near–singularity of models for animal recovery data. *Biometrics*, 57: 720–726.

Catchpole, E. A. & Morgan, B. J. T., 1994. Boundary estimation in ring recovery models. *J. R. Statist. Soc. B*, 56: 385–391.

– 1997. Detecting parameter redundancy. *Biometrika,* 84: 187–196.

– 2001. Deficiency of parameter–redundant models. *Biometrika,* 88: 593–598.

Catchpole, E. A., Morgan, B. J. T. & Viallefont, A., 2002. Solving problems in parameter redundancy using computer algebra. *Journal of Applied Statistics,* 29: 625–636.

Choquet, R., Reboulet, A. M., Pradel, R., Gimenez, O. & Lebreton, J.–D., 2003. *User's manual for M–SURGE 1.0.* Montpellier, France, CEFE/CNRS.

Cormack, R. M., 1964. Estimates of survival from the sighting of marked animals. *Biometrika*, 51: 429–438.

Currie, I. D., 1995. Maximum likelihood estimation and Mathematica. *Applied Statistics*, 44: 379–394.

Freeman, S. N. & Morgan, B. J. T., 1992. A modelling strategy for recovery data from birds ringed as nestlings. *Biometrics,* 48: 217–236.

Freeman, S. N., Morgan, B. J. T. & Catchpole, E. A., 1992. On the augmentation of ring–recovery data with field information. *Journal of Animal Ecology,* 61: 649–657.

Gimenez, O., Choquet, R. & Lebreton, J. D., 2003. Parameter Redundancy in Multi–state Capture–Recapture Models. *Biometrical Journal,* 45(6): 704–722.

Gimenez, O., Choquet, R., Lamor, L., Scofield, P., Fletcher, D., Lebreton, J. D. & Pradel, R. (submitted). Efficient profile likelihood confidence intervals for capture–recapture models. *Journal of Agricultural, Biological, and Environmental Statistics.*

Hines, J. E., 1994. *MSSURVIV User's Manual*. Laurel, MD, National Biological Survey: 21.

Jolly, G. M., 1965. Explicit estimates from capture–recapture data with both death and immigration – stochastic model. *Biometrika,* 52: 225–247.

Kendall, W. L. & Nichols, J. D., 2002. Estimating state–transition probabilities for unobservable states using capture–recapture/resighting data. *Ecology,* 83: 3276–3284.

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(1): 67–118.

Lebreton, J. D. & Pradel, R., 2002. Multi–state recapture models: modelling incomplete individual histories. *Journal of Applied Statistics,* 29: 353–369.

Reboulet, A. M., Viallefont, A., Pradel, R. & Lebreton, J. D., 1999. Selection of survival and recruitment models with SURGE 5.0. *Bird Study,* 46 (suppl.): 148–156.

Schaub, M., Gimenz, O., Schmidt, B. R. & Pradel, R., 2004. Estimating survival and temporary emigration in he multistate cature–recapture framework. *Ecology*, 85: 2107–2113.

Schwarz, C. J., Schweigert, J. F. & Arnason, A. N., 1993. Estimating migration rates using tag–recovery data. *Biometrics,* 49: 177–193.

Seber, G. A. F., 1965. A note on the multiple–recapture census. *Biometrika,* 52: 249–259.

– 1971. Estimating age–specific survival rates from bird–band returns when the reporting rate is constant. *Biometrika,* 78: 917–919.

Viallefont, A., 1992. Robustesse et flexibilité des modélisations en capture–recapture. Application à l'oie des neiges (*Anser caerulescens caerulescens*). Mémoire de DEA, Ecole Nationale Supérieure d'Agronomie de Montpellier, France.

– 1995. Robustesse et flexibilité des analyses démographiques par capture–recapture. De l'estimation de la survie à la détection de compromis évolutifs. Unpublished Ph. D. Thesis. Univ. of Montpellier II.

Viallefont, A., Lebreton, J.–D., Reboulet, A.–M., & Gory, G., 1998. Parameter identifiability and model selection in capture–recapture models: a numerical approach. *Biometrical Journal,* 40: 1–13.

White, G. C., 1983. Numerical estimation of survival rates from band–recovery and biotelemetry data. *Journal of Wildlife Management,* 47: 716–728.

White, G. C. & Burnham, K. P., 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study,* 46 (suppl.): 120–139.