The Canadian Journal of Statistics
Vol. 35, No. 2, 2007, Pages ???-???
La revue canadienne de statistique

Estimation of sex-specific survival with uncertainty in sex assessment

Roger PRADEL, Lory MAURIN-BERNIER, Olivier GIMENEZ, Meritxell GENOVART, Rémi CHOQUET and Daniel ORO

Key words and phrases: capture-recapture; label switching; multievent models; parameter redundancy.
MSC 2000: Primary 92D25; secondary 60J20.

Abstract: In monomorphic species, determination of sex from behavior is prone to errors. The authors develop capture-recapture survival models that account for uncertainty in the assessment of sex. They examine parameter redundancy for four basic models with constant or time-dependent survival and encounter probabilities. They further develop a more refined and more appropriate model for an Audouin's gull data set where four distinct behavioral clues have been used. They examine how useful it is to incorporate the least reliable of the clues and the genetic determination of sex available for only a handful of individuals. They finally discuss the implications of their findings for the design of field studies.

Title in French: Estimation d'une survie séparée pour chaque sexe quand l'identification du sexe est incertaine.

Résumé : Dans les espèces monomorphes, la détermination du sexe à partir du comportement est sujette à erreurs. Les auteurs développent des modèles de capture-recapture qui prennent en compte l'incertitude dans la détermination du sexe. Ils examinent la redondance en paramètres de quatre modèles de base où les probabilités de capture et de survie sont constantes ou dépendantes du temps. Ils développent ensuite un modèle plus raffiné et plus adapté à des données portant sur le goéland d'Audouin où quatre critères comportementaux distincts ont été utilisés. Ils examinent l'utilité d'incorporer le critère le moins fiable et la détermination génétique du sexe disponible pour une poignée d'individus. Ils discutent finalement les implications de leurs découvertes pour la planification des études de terrain.

1. INTRODUCTION

Sex differences in survival, dispersal or movement may have important consequences in demography, mating and parental investment patterns (Breitwisch 1989; Gowaty 1993), even for species where the two sexes have very similar body size and appearance. Thus, a modern study in ecology cannot ignore a priori this factor. However, naming the sex of an individual may be particularly tricky for monomorphic species observed in the field (i.e. species in which males and females are structurally identical). Often the field biologist will rely on behavior to distinguish males from females. Some behavioral clues like position during copulation are thought to be quite reliable but other clues like the relative body size in a pair not so much and yet the latter type are easier to gather. With this kind of approach, it is thus not rare that an animal that has once been recorded as a male is later referred to as a female. Confidence in the correct determination of sex will progressively increase with repeated and consistent observations and, eventually, it may be decided that there is no reasonable doubt about the true sex of a particular individual. However, this does not hold for animals seen only a few times and for those with a contradictory record of given sex. A natural tendency for studying nonetheless sex-specific parameters is then to consider the sole individuals for which the sex has been determined with reasonable certainty. However, this approach has shortcomings, especially when the aim is to study survival. First, it may represent a huge waste of data. For instance, in the field study that has motivated this paper, approximately 80% of the individuals had never been sexed and would thus be discarded. Second, to be sexed with reasonable certainty an individual must often have gone through several observations and hence have survived meanwhile. Then, estimation of survival based on the subsample of known-sex individuals will inevitably be positively biased. Another possibility is to segregate the animals into three groups: males, females and unknown-sex individuals. The survival of the unknown-sex individuals can then be constrained to be a weighted average of the survivals of males and females (Oro & Pradel 2000) with the weights reflecting the proportions of males and females among the unknown. However, these weights are not easy to determine as the relative proportions will depend on the sex-ratio but also on the relative catchability of males and females and on the ability to identify the sex, which may differ between males and females. Thus, while this second method is preferable, it is still imperfect and anyway does not make full use of the information available.

We know of one example where sex uncertainty was directly incorporated in a survival analysis (Conroy, Senar, Hines & Domenech 1999). This study dealt with serins *Serinus serinus*, a species which is monomorphic only at the juvenile stage. A biometrical measure (wing length) was used to predict the sex of the captured juveniles and an initial state 'predicted male' or 'predicted female' was then assigned to each bird on this basis. If they were ever seen again, their true sex was then determined and they were assigned the new state 'known male' or 'known female'. This procedure allowed the estimation of transition probabilities between the 'predicted sexes' and the 'known sexes' and thus the incorporation of the never recaptured juveniles in the analysis. In this way, survival could be estimated without bias. This model was cast within the framework of multistate capture-recapture (hereafter CR) models (Arnason 1972, 1973; Schwarz, Schweigert & Arnason 1993). Unfortunately, this approach does not work for species which remain monomorphic all their life and this is why we had to develop specific models. We present here a general solution to the problem of sex uncertainty by introducing explicitly probabilities of errors. *In essence, as compared to previous approaches, instead of summarizing the data previous to the analysis, we model the very process of sex identification.* The models we build in this way generalize the mixture models of Pledger, Pollock & Norris (2003) by allowing the incorporation of elements of information about the classes (here the two sexes). They are also more general than the specific models for unknown sex of Nichols, Kendall, Hines & Spendelow (2004) because these latter do not allow for mistakes. They actually belong to the class of CR models defined by Pradel (2005) to deal with uncertainty in the determination of a state of interest. This is the particular case where the state under study (sex) is static (Pradel 2007).

This work was motivated by the demographic study of an Audouin's gull *Larus audouinii* colony at the Ebro Delta (Spain) (e.g., Oro & Pradel 2000; Cam, Oro, Pradel & Jimenez 2004) where the birds are sexed upon observation according to four criteria (see Oro, Pradel & Lebreton 1999 for details): position during copulation, begging food during courtship, courtship feeding and relative body size (the males being on average slightly heavier than the females, see Oro 1998; Genovart, Oro & Bonhomme 2003). The information can be coded as three basic events: "the animal is judged to be a male", "the animal is judged to be a female", "no judgment is made". As this scheme should be relevant to many studies, we start (section 2) by developing a set of four general models that extend the Cormack-Jolly-Seber CR model and constant-parameter versions of it (Cormack 1964; Jolly 1965, 1982; Seber 1965) to uncertainty in sex assignment. These models incorporate probabilities of judgment and probabilities of error and can be further developed if needed as will be seen in Section 3. They are described in subsection 2.1. Because of the additional parameters involved, there is a legitimate concern that not all parameters may be identifiable. In subsection 2.2, we set off to study redundancy by the method of Catchpole, Morgan and Freeman (1997, 1998) (review by Gimenez, Viallefont, Choquet, Catchpole & Morgan 2004). However, the models of subsection 2.1 do not account for likely differences in probability of error with each of the four criteria used in the Audouin's gull study, nor do they incorporate such peculiarity of the study as a trend over time in the attempt to identify the sex of the birds. In Section 3, we construct a more refined and more specialized model, suited to our data. At the same time, we assess the effect of dealing or not dealing with different pieces of information: is it useful to gather the least reliable clue? Should genetic sexing be developed? The last section is a discussion of the interest of this kind of models and of the implications of our findings in terms of field work.

2. FOUR BASIC MODELS

2.1 Data, assumptions and parameters

The typical data are presented in Table 1. They consist of a series of encounter histories followed by a frequency number. For instance, the first line in Table 1 refers to 41 individuals that were encountered for the first time at occasion 6. Their sex could not be determined at that time (code 3) nor at occasion 7 when they were last encountered. All of the birds were originally marked as chicks on the nest and the data we are using are their resightings as live adults. Thus the protocol is the same for the first encounter and afterwards.

h	$\#\{h\}$
0000033000	41
0000112133	4
0001011000	2
0033000300	8
0330003313	4
0102220203	1
1001101100	1

Table 1: Excerpt from the Audouin's gull data coded according to three criteria: 1=judged male; 2=judged female; 3=no judgment made. $\#\{h\}$ is the number of individuals having encounter history h. The data set has 4025 individuals spread among 917 encounter histories over 10 years of study.

We make the following assumptions:

1. Each individual has a unique mark that is neither lost nor misread.
2. Individuals within each sex category are identical.
3. Encounter and survival events are independent between animals and between samples.
4. The true sex of every single individual is a priori unknown.

Because we are studying survival, we condition each time on the first encounter event. With the above assumptions, the encounter history of an individual first caught at occasion i thus follows a multinomial distribution. Hence, the likelihood is a product multinomial likelihood. To write the probability of encounter history h, we apply the law of total probability:

$$P(h) = P(f)P(h|f) + P(m)P(h|m),$$

where $P(f)$ and $P(m)$ are the sex proportions among the unmarked encountered, and $P(h|f)$ and $P(h|m)$ are the conditional probabilities of h for a female and a male respectively. To go further, we need the following parameters:

1. Traditional CR parameters
    - ϕ_m, ϕ_f sex-specific survival rates
    - p_m, p_f sex-specific encounter rates
2. New parameters
    - μ, proportion of males in the population
    - e_m, e_f sex-specific probabilities to judge the sex of an individual
    - x_m, x_f sex-specific probabilities to make the correct judgment.

It is now possible to write out $P(m)$ as $\frac{p_m\mu}{p_m\mu+p_f(1-\mu)}$ ($P(f) = 1 - P(m)$). As for the conditional probabilities of h, let us write them for the example encounter history $h = (1\ 3\ 2)$. If this relates to a male, it has been sexed the first time (probability e_m) correctly (probability x_m). It has then survived to occasion 2 (probability ϕ_m), when it was encountered (probability p_m), but not sexed (probability $1 - e_m$). It survived again (ϕ_m), was encountered at occasion 3 (p_m), and sexed (e_m) incorrectly ($1 - x_m$). All put together, we obtain $P(h|m) = e_m x_m \phi_m p_m (1-e_m) \phi_m p_m e_m (1-x_m)$. In a similar way, we would get $P(h|f) = e_f(1-x_f)\phi_f p_f(1-e_f)\phi_f p_f e_f x_f$. The complete probability for this encounter history is thus

$$\begin{aligned} P(h) &= \frac{p_m\mu}{p_m\mu + p_f(1-\mu)} e_m x_m \phi_m p_m (1-e_m)\phi_m p_m e_m (1-x_m) \\ &\quad + \frac{p_f(1-\mu)}{p_m\mu + p_f(1-\mu)} e_f(1-x_f)\phi_f p_f(1-e_f)\phi_f p_f e_f x_f. \end{aligned} \tag{1}$$

The probability of each individual encounter history will be obtained in this way and hence the likelihood of the entire data set which, assuming independence of fates, is the product of them. With $\#\{h\}$ denoting the number of animals with encounter history h, the likelihood can be written in compact form:

$$L = \prod_h P(h)^{\#\{h\}}.$$

Maximum Likelihood Estimators (MLE) are derived by maximizing L as a function of the different parameters. These estimators are known to have excellent statistical properties such as being asymptotically unbiased and this is the approach to parameter estimation retained in leading CR software. However, because the current models do not fall in the category of existing models, we could not take advantage of this feature and wrote our own maximizing programs in MATLAB (Hanselman & Littlefield 2000). The code is available upon demand from the first author. Alternatively, the model can be fitted as a multi-event model using a free program called E-SURGE (Choquet, Rouan & Pradel 2007) downloadable from http://www.cefe.cnrs.fr/biom/logiciels.htm.

2.2 Parameter redundancy

The aim of this subsection is to study the impact of the introduction of sex uncertainty on parameter redundancy. As one can easily realize by browsing through the extensive CR literature, the number of traditional CR models in use is virtually unlimited as new models can be created by changing the kind of effects considered on each type of parameters. For instance, parameters may be constant over time or time-dependent; they may vary by sex, age... It is thus impossible to tell in advance which parameters are or are not redundant in each possible model. However, it is useful to have some general guidance as to where problems are likely to occur. In traditional CR, the fully time-dependent Cormack-Jolly-Seber model (CJS) plays this role. In this model, it is known that survival over the last time interval is confounded with encounter probability at the last occasion.

This problem disappears if survival or encounter parameters are hold constant. The introduction of sex uncertainty may well add new general redundancy problems. The minimal generalization of the CJS model consists in having the survival and encounter parameters, the 'traditional' parameters of subsection 2.1, still time-dependent and in considering that all of the 'new' parameters are constant over time. In this subsection, we examine whether this model is parameter-redundant. We also examine the three models derived by holding survival or encounter probabilities or both constant over time. In the following, we denote the four models considered by their sole variable part, i.e. (ϕ_{t*s}, p_{t*s}), (ϕ_{t*s}, p_s), (ϕ_s, p_{t*s}), (ϕ_s, p_s) where s stands for sex and t for time. Superscripts are used to represent capture occasions; thus p_m^1 will denote the male encounter probability at occasion 1.

Catchpole, Morgan and Freeman (1997, 1998) have developed a formal method (hereafter the CMF method) for studying parameter redundancy in models belonging to the exponential family of probability distributions. This method can be applied to the multinomial distribution of animals over the observable encounter histories. It indicates which parameters are directly identifiable and provides identifiable functions of the redundant parameters (see Catchpole and Morgan (1997) and Catchpole, Morgan and Freeman (1998) for details). All the calculations being formal, we carried them out with MAPLE (Gimenez, Choquet & Lebreton 2003). Because the number of capture histories $4^k - 1$ increases rapidly with the number of time steps k, we could apply the procedure only for $k = 3$, but this is sufficient as the results obtained are easily demonstrated to hold in general.

A first conclusion (see Table 2) is that improving the realism of the models by adding new constant nuisance parameters to account for sex uncertainty does not restrict the ability to estimate the parameters of interest, i.e. the survival probabilities. Indeed, the only restriction we observe in the estimation of survival relates to the fully time-dependent model and this restriction is the same that was already present in the corresponding fully time-dependent CJS model. Actually, the new model improves over the CJS model applied to each sex separately in allowing the estimation of the ratio of survivals of males and females at the last occasion. We believe that this result is due to the assumption of a constant sex-ratio in the population. The other case of parameter redundancy concerns the initial encounter probabilities, p_m^1 and p_f^1 (absent from the CJS model). When encounter probabilities are time-dependent (models (ϕ_{t*s}, p_{t*s}) and (ϕ_s, p_{t*s})), only the ratio $\frac{p_f^1}{p_m^1}$ i.e. the relative catchability of males and females at the initial occasion is estimable. This is easily understood from the likelihood. The initial encounter probabilities appear in it only in the terms $P_t(m)$ and $P_t(f)$, and there only through their ratio. For instance, $P_t(m)$ can be rewritten

$$P_t(m) = \frac{p_m^t \mu}{p_m^t \mu + p_f^t (1-\mu)} = \frac{\mu}{\mu + \frac{p_f^t}{p_m^t}(1-\mu)}.$$

2.3 Dual solutions

A redundant parameter is also one that can take all of a continuous range of values while the likelihood remains at its maximum. For a given data set, this can be seen by drawing the profile likelihood curve, i.e. the curve of the maximum value that the likelihood can assume for each value of the parameter under scrutiny: for a redundant parameter, the profile likelihood presents a characteristically flat area at its top (see Gimenez et al. 2004). The shape of the profile likelihood is interesting more generally. For instance, when the profile likelihood decreases rapidly away from the optimal value of the parameter, this parameter is estimable with great precision. Away from its maximum, the profile likelihood may exhibit local maxima. If one of these happens to be exactly on the same level as the 'absolute' maximum, then this is an instance of non-identifiability without redundancy: two entirely different values of the parameter maximize the likelihood equally well

Table 2: Identification of estimable quantities in 4 models incorporating probabilities of error. Superscripts are for time steps. Only 3 occasions were considered. ϕ^2 and p^3 are thus the last survival and last encounter parameters respectively.

Model	#par	Model rank	non estimable parameters	estimable functions
(ϕ_s, p_s)	9	9	none	—
(ϕ_{t*s}, p_s)	11	11	none	—
(ϕ_s, p_{t*s})	13	12	p_m^1, p_f^1	$\frac{p_f^1}{p_m^1}$
(ϕ_{t*s}, p_{t*s})	15	13	$p_m^1, p_m^3, p_f^1, p_f^3, \phi_m^2, \phi_f^2$	$\frac{p_f^1}{p_m^1}, \frac{\phi_f^2}{\phi_m^2}, \phi_m^2 p_m^3, \phi_f^2 p_f^3$

Table 3: Maximum-likelihood estimates for each of the four models. For time-dependent parameters, the median is given. Each model has also a dual solution with probabilities of errors >50%.

Model	encounter probability (p)		survival (ϕ)		probability of judgment (e)		probability to be right (x)		proportion of males (μ)
	m	f	m	f	m	f	m	f	(m)
(ϕ_s, p_s)	0.61	0.72	0.88	0.95	0.02	0.79	0.53	0.56	0.86
(ϕ_{t*s}, p_s)	0.60	0.72	0.89	0.97	0.02	0.79	0.53	0.56	0.87
(ϕ_s, p_{t*s})	0.61	0.74	0.87	0.95	0.02	0.79	0.53	0.56	0.88
(ϕ_{t*s}, p_{t*s})	0.61	0.76	0.86	0.97	0.02	0.79	0.53	0.56	0.88

but not the values in-between. This situation cannot be detected by the CMF method. Thus, we examined the profile likelihood curves of the model with traditional parameters hold constant (ϕ_s, p_s) applied to the Audouin's gull data already mentioned (Table 1). The parameters were maintained within range by logit transformations. For practical reasons, instead of drawing the profile likelihoods, we drew the profile deviances ($D = -2 \log L$) and hence observed the minima rather than the maxima (Fig. 1).

A striking feature of model (ϕ_s, p_s) is the existence for each scalar parameter of two distinct values where the deviance is minimized. These values organize into two sets that represent two alternative solutions. This result is in fact data-independent and holds as well for the three other models of this section: (ϕ_{t*s}, p_{t*s}), (ϕ_{t*s}, p_s) and (ϕ_s, p_{t*s}) (see the Appendix for a theoretical demonstration). The deep reason for this fundamental duality comes down to this: given that the true sex of any individual is never known, it is equally acceptable, from the model point of view, to consider that an animal estimated many times to be, say, a male is indeed a male and that most judgments were correct, or that it is a female and most judgments were wrong. Yet, one solution produces probabilities of error above 50%, which is probably not acceptable from the practitioner's point of view. More precisely, one solution is derived from the other by exchanging the survival, encounter and judgment probabilities of males and females, replacing x_m with $1 - x_f$ and x_f with $1 - x_m$ and finally reversing the sex-ratio (see Appendix). As a consequence, the profile deviance curves of sex-specific parameters are the same for males and females. In the case of the Audouin's gull data, they present two close minima (Fig. 1). The profile deviance curve of the proportion of males is symmetrical around 0.5 with two minima, which happen to be very distant with our data, at 0.14 and 0.86, and, incidentally, utterly unrealistic from a biological point of view.

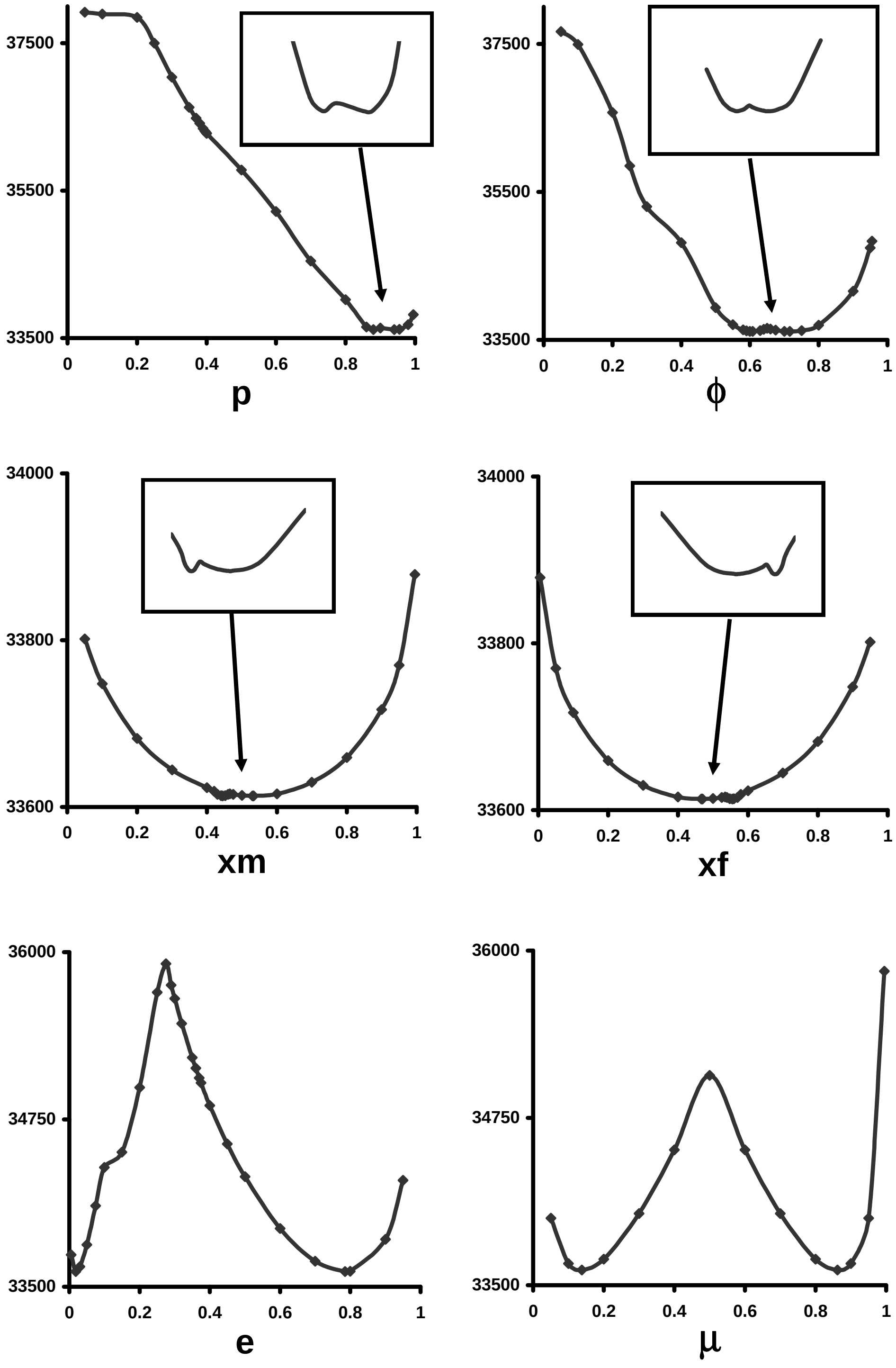

Figure 1: Profile-deviance curves of the parameters of model (ϕ_s, p_s) applied to the Audouin's gull data set. For each parameter, the deviance (y-axis) presents two distinct minima. The curves for p, ϕ and e are the same for males and females.

Table 4: Codes used to take into account the unequal reliability of each of the 4 criteria used to assess the sex of Audouin's gulls upon observation at the Ebro delta colony (Spain).

judged from	to be		not judged
	m	f	
copulation	1	5	9
begging food	2	6	
courtship feeding	3	7	
body size	4	8	

3. A MORE REALISTIC MODEL FOR THE AUDOUIN'S GULL DATA

The four models introduced in the previous section, when fitted to the Audouin's gull data, produce very unreasonable if similar estimates: very high proportions of males that are almost never judged and probabilities of error close to 50% (Table 3). However, these models do not incorporate several known important features of the study. In this section, we illustrate the flexibility of our approach by building a more realistic model incorporating our knowledge of the biology and of the way the field work had been conducted.

To start, there was no good reason to believe that the error attached to each of the 4 criteria used (copulation, begging food, courtship feeding and relative body size) was the same. Copulation for instance was suspected to be the most reliable and relative body size, the most error-prone. We thus decided to distinguish the different criteria and recoded the data accordingly (see Table 4). With the notation of Table 4, a possible reexpression for the capture history h=(1,3,2) discussed in Section 2.1 is h=(3,9,6). This would correspond to a bird been classified as a male on its first capture based on its courtship feeding display, not classified on his second capture, and classified as a female on his third capture according to its begging for food behaviour. When an individual was observed several times during the same season, priority was given to the criteria in the following order : copulation, courtship feeding, begging food and relative body size. All along the study, sex determination has been a secondary activity but it has been conducted on stringent criteria. This is reflected in the very few obvious mistakes (sexing is consistent over time for a given individual), and in the roughly 80% of never-sexed individuals in the data set. Yet, during the course of the study, sexing has gained ground; especially the criterion 'relative body size', initially used very sparingly, has become more common by the end of the study. These features were incorporated in a model denoted (ϕ_s, p_t, μ, e_T, $m4_T$, $m1$, $m2$, $x1$, $x2$, $x3$, $x4$) with the following characteristics:

- p_t, encounter probability: time but not sex-dependent.
- ϕ_s, survival rate: possibly sex-dependent but constant over time.
- μ, proportion of males in the population: held constant.
- e_T, probability to attempt to judge the sex upon encounter: time-dependent with a logit-linear trend over time (T). Note that because birds are always judged in pairs there is no sex-dependency.
- $m4_T$, frequency of the criterion 'relative body size' ($m4$) among the different criteria: time-dependent with a logit-linear trend over time.
- $m1$, frequency of use of the criterion 'copulation' among the behavioural criteria (i.e. excluding 'body size'): held constant.
- $m2$ and $m3$: same as $m1$ for 'begging food' and 'courtship feeding' respectively (one of $m1$, $m2$ or $m3$ is redundant).

Table 5: Results of model (ϕ_s, p_t, μ, e_T, $m4_T$, $m1$, $m2$, $x1$, $x2$, $x3$, $x4$) applied to the Audouin's gull data set with different amount of information. Analysis A uses all 4 criteria and the genetic determination of the sex of 24 birds; Analyses B and C do not use the genetic determination of sex; Analysis C does not use the judgments based on body size. Standard errors are given between parentheses. With less information (from left to right), precision decreases. For analyses B and C, we have retained the dual solution that yields proportions of error <50%.

Analysis	A	B	C
proportion of males	0.53 (0.027)	0.53 (0.029)	0.55 (0.034)
female survival	0.91 (0.013)	0.91 (0.014)	0.93 (0.024)
male survival	0.86 (0.013)	0.86 (0.014)	0.84 (0.025)
error copulation	0.06 (0.041)	0.06 (0.041)	0.08 (0.055)
error begging food	0.05 (0.029)	0.06 (0.031)	0.09 (0.077)
error courtship feeding	0.00 (0.161)	0.00 (0.155)	0.00 (0.286)
error body size	0.11 (0.064)	0.09 (0.074)	NA

- xi, probability to be right when using criterion i ($i = 1, \dots 4$): held constant.

This model has 23 parameters including 1 intercept and 1 slope for e and 1 intercept and 1 slope for $m4$ and 10 capture probabilities.

In addition, a limited number of birds, 24, had been sexed genetically (Genovart, Oro & Bonhomme 2003; Genovart, Oro, Ruiz, Griffiths, Monaghan & Nager 2003). We fitted the previous model using (analysis A) and not using (analysis B) this limited information. When we used it, $P(m)$ in the capture history of the genetically sexed individuals was set to 1 or 0 as appropriate. On the other hand, we considered the effects of not using the presumably less reliable criterion, 'body size'. To do that, the corresponding observations were recoded as a 'no judgment' observation. When doing this latter analysis (analysis C), the genetic determination of sex was not used. Thus, we have a gradient of decreasing amount of information from analysis A to analysis C (Table 5).

All three analyses yield estimates in agreement with what is known of the biology of the species. For instance, survival estimates are very close to those estimated previously by CR on the study site (Oro, Pradel & Lebreton 1999; Cam, Oro, Pradel & Jimenez 2004). We also note that, as anticipated, body size is the least reliable clue but the copulation criterion does not come out as the obvious best method. The results with and without genetic sexing are very similar (first 2 columns of Table 5). However, the known sex of only 24 birds suffices to break the tie between the dual solutions: while analyses B and C still have two solutions (only the reasonable one is presented in Table 5), analysis A has only one, as can be seen on the profile deviance curves relative to the proportion of males (Fig. 2). In other, not shown models, where sex-ratio is initially not identifiable, the additional information brought about by the 24 birds renders this and other parameters identifiable. As for the criterion 'body size', even though mistakes are made on average once in every 10 judgments and only 15.70% of the judgments made use of this criterion, dropping it results in a disproportionate loss of precision (compare the standard errors in the last two columns of Table 5).

4. DISCUSSION

A main concern when we started building these models was the risk of parameter redundancy. The models would have been useless if no parameter of interest was estimable. This fear had been voiced in particular by Nichols *et al.* (2004) when they examined the situation where the sex is

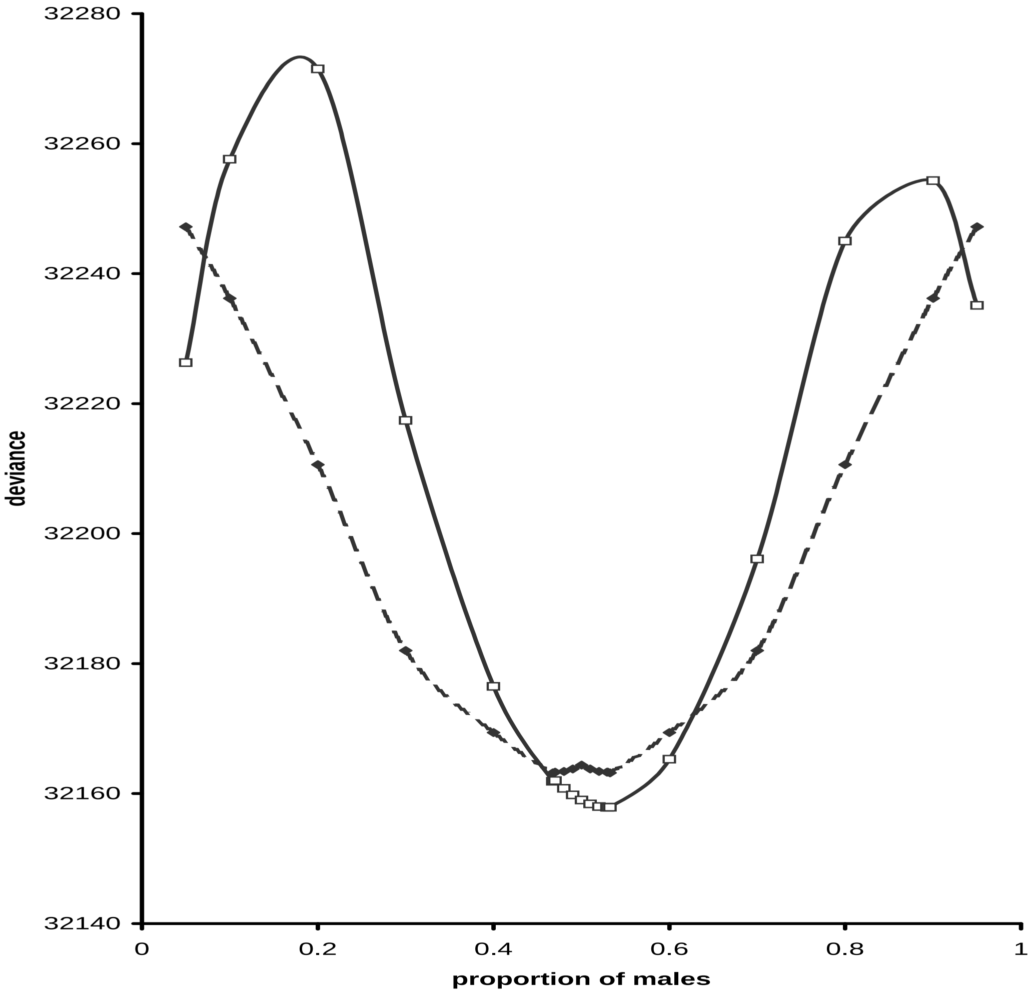

Figure 2: Improvement in the shape of the deviance brought about by a limited amount of external information shown on the profile deviance of model (ϕ_s, p_t, μ, e_T, $m4_T$, $m1$, $m2$, $x1$, $x2$, $x3$, $x4$) applied to the Audouin's gull data set and relative to the proportion of males μ. The genetic sexing of 24 individuals renders the profile deviance steeper at its bottom where it retains just one minimum (Analysis A; continuous line) as compared to when the genetic information is not used (Analysis B; dotted line). There is a narrow ridge between the 2 dual solutions in Analysis B.

not always assessed but, when it is, this is done without error. Our results show on the contrary that it is quite possible to run models incorporating probabilities of error and nonetheless gain knowledge about such important biological parameters as survival rates. Although these models have in general two mathematical solutions due to a label switching problem (Redner & Walker 1984, see Appendix), one is so unreasonable that it should not be difficult to sort out which one is to be retained.

A more serious difficulty is numerical. These models, and more generally the multievent models (Pradel 2005) to which they belong—but also the now classical multistate models—present local minima to which the optimization algorithm occasionally converges. The profile deviance curves may help diagnose the problem (Gimenez, Choquet, Lamor, Scofield, Fletcher, Lebreton & Pradel 2005) and improved algorithms may be sought (see, in particular, section 6.2 in Choquet, Reboulet, Pradel, Gimenez & Lebreton 2005), but it is also possible to guard against local minima by augmenting the information available. Often, such information already exists, but could not be exploited in the traditional CR models. One of the major aims of the newly developed models was indeed to alleviate and valorize the field work by comprehending and exploiting a wider range of data.

In this paper, we have examined two ways of gathering additional information: using less efficient clues and sexing genetically some animals. Genetic sexing is of course paramount for the determination of sex, but for practical reasons it cannot usually be performed on every single individual. In our example, genetically sexing very few birds proved sufficient to greatly improve the shape of the deviance and hence the efficiency of the optimization algorithm. We also observed that it eliminated the wrong one of the two dual solutions. Finally, genetic sexing occasionally rendered identifiable new parameters in parameter-redundant models. It should also be noticed that genetic sexing is only one approach to specifying the a priori probability to be a male, $P(m)$, for some individuals. Other, not necessarily $0-1$ predictors of sex frequently available in old studies, like biometrical measures, can be incorporated in the model in the same way ($P(m)$ can be set to some intermediate value between 0 and 1). This approach of incorporating external information has been used by Fujiwara and Caswell (2002) although in a way that, past the first encounter, may not be correct: the conditioning in their stage-assignment matrix, $P(stage|measure)$, is the reverse of what is needed, $P(measure|stage)$, in the calculation of the likelihood (see Pradel 2005 for a formal demonstration). As for the use of less efficient clues, our example proved that they may be valuable beyond what seems at first sight attainable based on their frequency of use and efficiency (percentage of error). We thus encourage people to gather such clues especially if they are not time consuming (Redman, Lewis, Griffiths, Wanless & Hamer 2002). After three or four visits, it will always be possible to assess the quality of the prospective clues as well as the improvement brought about by such and such identification method in terms of gained precision. If different methods for sex identification are in competition, it is then possible to decide objectively, from the informed assessment of their relative costs and benefits, which ones to favor and which ones to discard.

APPENDIX

The aim of this appendix is to establish the existence of two distinct dual solutions to the maximization of the likelihood of model (ϕ_s, p_s) presented in subsection 2.1, as well as some related properties of its profile deviance curves. Why this happens is rooted in the law of total probabilities applied to an arbitrary encounter history h.

$$P(h) = P(f)P(h|f) + P(m)P(h|m)$$

In this formula, the conditional probability $P(h|f)$ is a function of the sole parameters ϕ_f, p_f, e_f, x_f. This function depends on the particular encounter history h. Let us write

$$P(h|f) = g_h(\phi_f, p_f, e_f, x_f). \tag{2}$$

The probability of the same history conditional on the animal being this time a male is obtained by using the male parameters instead of the female ones. However, when a judgment was right for a female, it is wrong for a male (and vice-versa) so that x_f should be replaced with $1 - x_m$ (see the example of section 2.1 leading to equation 1 if this is not immediately clear). Thus,

$$P(h|m) = g_h(\phi_m, p_m, e_m, 1 - x_m). \tag{3}$$

The unconditional probabilities $P(m)$ and $P(f)$ are respectively

$$\frac{p_m\mu}{p_m\mu + p_f(1-\mu)} \text{ and } \frac{p_f(1-\mu)}{p_m\mu + p_f(1-\mu)}. \tag{4}$$

The transformation $\Theta\colon [0,1]^9 \mapsto [0,1]^9$

$$(\theta_1, \ldots \theta_9) \mapsto (\theta_2, \theta_1, \theta_4, \theta_3, 1-\theta_5, \theta_7, \theta_6, 1-\theta_9, 1-\theta_8)$$

plays a key role. We first establish

Lemma 1. *The probability of any encounter history, seen as a function of the parameter vector* $(\phi_m, \phi_f, p_m, p_f, \mu, e_m, e_f, x_m, x_f)$*, is invariant under* Θ*.*

$$\forall h, P(h) \circ \Theta = P(h)$$

.

Proof. When the parameters are changed in the following way:

$$(\phi_m, \phi_f, p_m, p_f, \mu, e_m, e_f, x_m, x_f) \overset{\Theta}{\longmapsto} (\phi_f, \phi_m, p_f, p_m, 1-\mu, e_f, e_m, 1-x_f, 1-x_m),$$

$$\begin{aligned}
P(f) \text{ becomes } & \frac{p_m\mu}{p_m\mu + p_f(1-\mu)} = P(m); && \text{[from (4)]}\\
P(m) \text{ becomes } & \frac{p_f(1-\mu)}{p_m\mu + p_f(1-\mu)} = P(f); && \text{[from (4)]}\\
P(h|f) \text{ becomes } & g_h(\phi_m, p_m, e_m, 1-x_m) = P(h|m); && \text{[from (2) and (3)]}\\
P(h|m) \text{ becomes } & g_h(\phi_f, p_f, e_f, x_f) = P(h|f); && \text{[from (2) and (3)]}
\end{aligned}$$

so that, eventually, $P(h)$ itself is unchanged. A consequence of this lemma is that the likelihood and the deviance are equally unchanged under Θ. Consequently, if the parameter vector $(\hat{\phi}_m, \hat{\phi}_f, \hat{p}_m, \hat{p}_f, \hat{\mu}, \hat{e}_m, \hat{e}_f, \hat{x}_m, \hat{x}_f)$ maximizes the likelihood, its transform by Θ, $(\hat{\phi}_f, \hat{\phi}_m, \hat{p}_f, \hat{p}_m, 1-\hat{\mu}, \hat{e}_f, \hat{e}_m, 1-\hat{x}_f, 1-\hat{x}_m)$, maximizes it too. Hence, the theorem:

Theorem 1. *The likelihood of model* (ϕ_s, p_s) *is maximized at two generally distinct points one of which is the transform of the other by the unipotent mapping* Θ*.*

We now examine some ensuing properties of the profile deviance curves. Let D be the deviance of model (ϕ_s, p_s). $D\colon [0,1]^9 \mapsto [0, +\infty]$, is a function of the 9 probability parameters: $\phi_m, \phi_f, p_m, p_f, \mu, e_m, e_f, x_m, x_f$, which has the property that $D \circ \Theta = D$. If

$$E_i^a = \left\{\theta \in [0,1]^9, \theta_i = a\right\},$$

the profile deviance function for parameter i is

$$P_i(a) \overset{\text{def}}{=} \min_{E_i^a} D = \min_{E_i^a} (D \circ \Theta) = \min_{\Theta\left(E_i^a\right)} D.$$

Given that:

$$\begin{aligned}
&\Theta(E_1^a) = E_2^a, && \Theta(E_2^a) = E_1^a,\\
&\Theta(E_3^a) = E_4^a, && \Theta(E_4^a) = E_3^a,\\
&\Theta(E_5^a) = E_5^{1-a}, && \\
&\Theta(E_6^a) = E_7^a, && \Theta(E_7^a) = E_6^a,\\
&\Theta(E_8^a) = E_9^{1-a}, && \Theta(E_9^a) = E_8^{1-a},
\end{aligned}$$

then,

$$\begin{aligned}
P_1(a) &= \min_{E_2^a} D = P_2(a),\\
P_3(a) &= \min_{E_4^a} D = P_4(a),\\
P_5(a) &= \min_{E_5^{1-a}} D = P_5(1-a),\\
P_6(a) &= \min_{E_7^a} D = P_7(a),\\
P_8(a) &= \min_{E_9^{1-a}} D = P_9(1-a).
\end{aligned}$$

Thus, the profile deviances of the survival (parameters 1 and 2), encounter (parameters 3 and 4) and judgment probabilities (parameters 6 and 7) are the same for both sexes. The profile deviance of the probability of error for the males (parameter 8) is the symmetrical with respect to a=0.5 of that for the females (parameter 9), and the profile deviance curve of the proportion of males (parameter 5) is symmetrical with respect to a=0.5.

REFERENCES

A. N. Arnason (1972). Parameter estimates from mark-recapture experiments on two populations subject to migration and death. *Researches on Population Ecology*, 13, 97–113.

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

R. Breitwisch (1989). Mortality patterns, sex ratios, and parental investment in monogamous birds. *Current Ornithology*, 6, 1–50.

E. Cam, D. Oro, R. Pradel & J. Jimenez (2004). Assessment of hypotheses about dispersal in a long-lived seabird using multistate capture-recapture models. *Journal of Animal Ecology*, 73, 723–736.

E. A. Catchpole, B. J. T. Morgan & S. Freeman (1998). Estimation in parameter-redundant models. *Biometrika*, 85, 462–468.

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

R. Choquet, A.-M. Reboulet, R. Pradel, O. Gimenez & J.-D. Lebreton (2005). *M-SURGE 1.8 User's manual.* Mimeographed document, CEFE, Montpellier.

R. Choquet, L. Rouan & R. Pradel (2007). Program E-SURGE: a software application for fitting multievent models. *Environmental and Ecological Statistics*, to appear.

M. J. Conroy, J. C. Senar, J. E. Hines & J. Domenech (1999). Development and application of a mark-recapture model incorporating predicted sex and transitory behaviour. *Bird Study*, 46, 62–73.

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

M. Fujiwara & H. Caswell (2002). Estimating population projection matrices from multi-stage mark-recapture data. *Ecology*, 83, 3257–3265.

M. Genovart, D. Oro & F. Bonhomme (2003). Genetic and morphological differentiation between the two largest breeding colonies of audouin's gull larus audouinii. *Ibis*, 145, 448–456.

M. Genovart, D. Oro, X. Ruiz, R. Griffiths, P. Monaghan & R. G. Nager (2003). Seasonal changes in brood sex composition in audouin's gulls. *Condor*, 105, 783–790.

O. Gimenez, R. Choquet & J.-D. Lebreton (2003). Parameter redundancy in multistate capture-recapture models. *Biometrical Journal*, 45, 704–722.

O. Gimenez, R. Choquet, L. Lamor, P. Scofield, D. Fletcher, J.-D. Lebreton & R. Pradel (2005). Efficient Profile-Likelihood Confidence Intervals for Capture-Recapture Models. *Journal of Agricultural, Biological and Environmental Sciences*, 10, 184–196.

O. Gimenez, A. Viallefont, R. Choquet, E. A. Catchpole & B. J. T. Morgan (2004). Methods for investigating parameter redundancy. *Animal Biodiversity and Conservation*, 27, 561–572.

P. A. Gowaty (1993). Differential dispersal, local resource competition, and sex ratio variation in birds. *American Naturalist*, 141, 263–280.

D. Hanselman & B. R. Littlefield (2000). *Mastering MATLAB 6.* Prentice Hall, Englewood Cliffs.

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

G. M. Jolly (1982). Mark-recapture models with parameters constant in time. *Biometrics*, 38, 301–321.

J. D. Nichols, W. L. Kendall, J. E. Hines & J. A. Spendelow (2004). Estimation of sex-specific survival from capture-recapture data when sex is not always known. *Ecology*, 85, 3192–3201.

D. Oro (1998). Audouin's gull account. in M. Ogilvie, editor *The birds of western palearctic*, 47–61. Oxford University Press, Oxford.

D. Oro & R. Pradel (2000). Determinants of local recruitment in a growing colony of audouin's gull. *Journal of Animal Ecology*, 69, 119–132.

D. Oro, R. Pradel & J.-D. Lebreton (1999). Food availability and nest predation influence life history traits in audouin's gull, larus audouinii. *Oecologia*, 118, 438–445.

S. Pledger, K. H. Pollock & J. L. Norris (2003). Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics*, 59, 786-794.

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

R. Pradel (2007). The stakes of capture-recapture models with state uncertainty. *Environmental and Ecological Statistics*, to appear.

K. K. Redman, S. Lewis, R. Griffiths, S. Wanless & K. C. Hamer (2002). Sexing northern gannets from DNA, morphology and behavior. *Waterbirds*, 25, 230–234.

R. A. Redner & H. F. Walker (1984). Mixture densities, maximum-likelihood and the EM algorithm. *SIAM Review*, 26, 195–237.

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

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

Received ???
Accepted ???

Roger PRADEL: `roger.pradel@cefe.cnrs.fr`
CEFE
Montpellier
France, 34293 Cedex 5

Lory MAURIN-BERNIER: `john4lmb@wanadoo.fr`
CEFE
Montpellier
France, 34293 Cedex 5

Olivier GIMENEZ: `olivier.gimenez@cefe.cnrs.fr`
CEFE
Montpellier
France, 34293 Cedex 5

Meritxell GENOVART: `m.genovart@uib.es`
IMEDEA
Esporles
Spain, 07190

Rémi CHOQUET: `remi.choquet@cefe.cnrs.fr`
CEFE
Montpellier
France, 34293 Cedex 5

Daniel ORO: `d.oro@uib.es`
IMEDEA
Esporles
Spain, 07190