

## FW663 -- Laboratory Exercise

### Program MARK: Joint Live Recapture and Dead Recovery Data and Pradel Model

Today's exercise explores parameter estimation using both live recaptures and dead recoveries. We will be estimating survival from both dead animal recoveries and live animal recaptures, and estimating the rate of population increase with the Pradel model. The objective of this exercise is to expose you to some real-world analyses with these 2 models, and to suggest to you how useful these models are for authentic problems. We have provided data sets from goldeneye ducks banded in Germany by Richard Barker and, closer to home, from a catch/effort monitoring program begun in 1991 for pikeminnow (formerly Colorado squawfish) supplied by Dr. Kevin Bestgen. Hence, the data are messy, and the original studies did not have in mind the type of analysis we will demonstrate here.

#### Joint Live Recapture and Dead Recovery Data

The basic model (Burnham 1993) is a combination of the Cormack-Jolly-Seber (CJS) model and the band recovery models of Brownie et al. (1985). The parameter set is  $S$  for survival,  $r$  for the probability of recovery of an individual that is dead,  $p$  for the recapture probability of an individual that is alive, and  $F$  for the probability that an individual remains on the study area (fidelity) and hence is available for recapture. Keep in mind the distinction between  $\phi$  and  $S$ :  $\phi = SF$ , and  $E = 1 - F$ , where  $E$  is emigration. Thus, the joint recovery and recapture model provides unbiased estimates of  $S$  assuming that  $r$  is constant across space regardless of where the animal dies. If  $r$  varies geographically, i.e., if  $r$  declines as animals emigrate from the study area, then the estimates of  $S$  will still be biased.

Barker (1997, 1999) extended Burnham's model to incorporate live resightings between capture occasions. The problem that led to his extension was a brown trout study, where fish were marked on the spawning areas, but a catch-and-release sport fishery resulted in significant numbers of live recaptures between marking occasions because anglers record the tag number before releasing the fish. Dead recoveries result from fish put in the creel. Barker's extension added 3 parameters to Burnham's model, and changed the interpretation of  $F$ . Parameters of Barker's model are:

- $S(i)$  = the probability an animal alive at  $i$  is alive at  $i + 1$ ,
  - $p(i)$  = the probability an animal at risk of capture at  $i$  is captured at  $i$ ,
  - $r(i)$  = the probability an animal that dies in  $i, i + 1$  is found dead and the tag reported,
  - $R(i)$  = the probability an animal that survives from  $i$  to  $i + 1$  is resighted (alive) some time between  $i$  and  $i + 1$ ,
  - $R'(i)$  = the probability an animal that dies in the  $i$  to  $i + 1$  interval without being found dead is resighted alive in the  $i$  to  $i + 1$  interval before it died,
  - $F(i)$  = the probability an animal at risk of capture at  $i$  is at risk of capture at  $i + 1$ , and
  - $F'(i)$  = the probability an animal not at risk of capture at  $i$  is at risk of capture at  $i + 1$ .
- (NB – differs from the definition in Barker (1997))

The resighting parameterization used in MARK differs from that described by Barker (1997).

An advantage of the parameterization used by MARK (Barker and White 2001) is that it enforces certain internal constraints that arise because the joint probability  $\Pr(A \text{ and } B)$  should always be less than or equal to the unconditional probabilities  $\Pr(A)$  and  $\Pr(B)$ . For example, the MARK parameterization ensures that the probability an animal is resighted alive in  $i, i + 1$  and survives from  $i$  to  $i + 1$  is less than the probability it is resighted alive in  $i, i + 1$ . It also ensures that  $\Pr(\text{resighted alive and dies in } i, i + 1 \text{ without being reported}) < \Pr(\text{dies in } i, i + 1 \text{ without being reported})$ . These internal constraints are not enforced by the other parameterizations.

### Movement

Between sampling/capture sessions, the model permits animals to leave the study area for one or more occasions, and then return. If an animal is in the study area then it is considered “at risk of capture”. If it leaves the study area then it is considered “not at risk of capture”. Animals that are at risk of capture at time  $i$ , leave the study area with probability  $(1 - F(i))$ . Thus  $F(i)$  has the same interpretation as in Burnham’s (1993) live-dead model as the fidelity to the study area.

Animals not at risk of capture are permitted to return to the study area with probability  $F'(i)$ . In Barker (1997)  $F'(i)$  was the probability that an animal out of the study area at  $i$  remained out of the study area at  $i + 1$ , but the definition has been changed in the interest of having a parameterization in common with the robust design model. Under this parameterization there are 3 types of emigration:

Random	$F'(i) = F(i)$ ,	i.e., the probability of being on the study area is identical regardless of whether the animal was on or off the study area on the previous occasion; an equivalent constraint turns out to be $F(i) = 1$ and $F'(i)$ not estimated because no animal ever leaves the study area;
Permanent	$F'(i) = 0$ ,	i.e., once an animal leaves the study area, it remains away because the probability of return is 0;
Markov	no constraint.,	i.e., the probability of being on the study area is a function of the animal’s status on the previous occasion.

A complication is that in the live resightings model the parameters  $F(i) = F'(i)$  are confounded with the capture probability  $p(i + 1)$ . By making the constraint  $F(i) = 1$  in MARK the random emigration model is fitted, but now the interpretation of  $p(i)$  is the joint probability that an animal is at risk of capture and is caught,  $F(i - 1)p(i)$ .

Under Markov emigration there tends to be serious confounding of movement and capture probabilities. In a model with time-dependent capture probabilities, it is usually necessary to constrain  $F(i) = F(\cdot)$  and  $F'(i) = F'(\cdot)$  for all  $i$ . Even then, the Markov emigration model may perform poorly. In practice the parameters  $F$  and  $F'$  are usually estimable only if the movement model is markedly different to the random emigration model, that is, if there is a large difference between  $F(i)$  and  $F'(i)$ . It is recommended that analysis begins with the random or permanent emigration model. If the Markov emigration model leads to significant improvement,

it should show in lack of fit for the other two models.

To illustrate the meaning of the emigration parameters, suppose the animal is captured during the first sampling session, not captured during the second sampling session, and then captured during the third sampling session. One of several encounter histories that would demonstrate this scenario would be:

100010

The probability of observing this encounter history can be broken into 4 factors:

- P1 = Pr(animal survives from time 1 to time 3 | released at 1),  
 P2 = Pr(animal is not resighted between 1 and 3 | released at 1 and survives to 3),  
 P3 = Pr(animal is not captured at 2 but is captured at 3 | released at 1 and survives from 1 to 3 without being resighted), and  
 P4 = Pr(encounter history after sampling period 3 | events up to sampling period 3)

For describing movement, the relevant factor is P3. An animal captured at time 1 is known to be at risk of capture at time 1. Because it was captured at time 3 we also know it was at risk of capture at time 3. There are two possible histories that underlie this observed history:

- The animal was at risk of capture at time 2 and was not captured, but was captured at time 3
- The animal left the study area between time 1 and 2 but then returned and was captured.

Because we do not know which one actually occurred we instead find the probability that it was either of the two, which is:

$$P3 = \{(1 - F_1)F'_2 + F_1(1 - p_2)F_2\}p_3$$

The complicated term in the brackets represents the probability that the animal was not captured during the second sampling session but is at risk of capture at time 3. The first product within the brackets  $(1 - F_1)F'_2$  is the joint probability that the animal emigrated between the first and second sampling sessions (with probability  $1 - F_1$ ) and then immigrated back onto the study area during the interval between the second and third sampling sessions (with probability  $F'_2$ ). However, a second possibility exists for why the animal was not captured -- it could have remained on the study area and not been captured. The term  $F_1$  represents the probability that it remained on the study area between time 1 and 2 and the term  $(1 - p_2)$  is the probability that it was not captured at time 2. The final term  $F_2$  represents the probability that the animal remained on the study area so that it was available for capture during the third sampling session.

Encounter histories for this model are coded as LDLDL... because animals can be encountered in this model as either alive or dead during the interval between capture occasions, 2 different codes are required in the encounter histories to provide information. A 1 in the D

portion of an encounter history means that the animal was reported dead during the interval. A 2 in the D portion of an encounter history means that the animal was reported alive during the interval. A 1 in the L portion of an encounter history means that the animal was alive on the study area during a capture occasion.

The following are valid encounter histories for a 5-occasion example:

1010101002

Animal was captured on the first occasion, and recaptured again on the 2nd, 3rd, and 4th occasions. It was not captured on the 5th occasion, but was seen alive during the last interval.

0000120100

Animal was captured on the 3rd occasion, and seen alive during the 3rd interval. It was reported dead during the 4th interval.

Note that there can be multiple occasions with a 1 in the L columns, and multiple occasions with a 2 in the D columns, but only one D column can have a 1.

### Data Analysis

The data we will be analyzing in this exercise concerns survival of common goldeneye ducks (*Bucephala clangula*), thanks to Richard Barker. Female common goldeneyes were trapped in nest boxes near Preetz, in the Schleswig-Holstein District of Germany between 1986 and 2000, giving 15 banding occasions. Recaptures of nesting females were obtained during nest box searches and live resightings of banded birds were reported by members of the public. To make fullest use of the available information, these need to be analyzed using a joint live-recapture/live-resighting model. The general model for these data is  $\{S(t) p(t) r(t) R(t) R'(t) F(t) F'(t)\}$ , in which all parameters are time specific. However, the data do not include any resightings or dead recoveries for 2000, so  $r(15)$ ,  $R(15)$ , and  $R'(15)$  should all be set to zero, even in models that include  $r(\cdot)$ ,  $R(\cdot)$ , or  $R'(\cdot)$  terms. Further, the last survival rate,  $S(15)$  is not estimable, as only the product  $[1 - S(15)]r(15)$  is estimable. Also, if the model include both time-specific survival and capture probabilities,  $S(t) p(t)$ , then only the product  $S(14)p(15)$  is estimable.

Only in very good data sets can time-specific  $F$  and  $F'$  parameters be estimated – I've never seen any data that good!

MARK files are available as J:\CLASSES\FW663\EXERCISE.11\GOLDENEYES.DBF and J:\CLASSES\FW663\EXERCISE.11\GOLDENEYES.FPT.

Model selection results are shown in the following table. All of these models have the constraint  $r(15) = R(15) = R'(15) = 0$  imposed. The easy way to do this is to set the last column

in the PIM for each of these parameters to a value of 1, so that you fix parameter 1 to a value of zero, and it remains fixed for all your models. Study the PIMs in the MARK file provided with this exercise to understand how this trick was implemented.

Model	AICc	Delta AICc	AICc Weights	Model Likelihood	Num Par	Deviance
{S(t) p(t) r(t) R(t) R'(t) F(.) F'(.)}	1628.752	0	0.72068	1	67	876.204
{S(.) p(.) r(t) R(t) R'(t) F(.) F'(.)}	1632.54	3.788	0.10844	0.1505	46	931.676
{S(.) p(t) r(t) R(t) R'(t) F(.) F'(.)}	1632.854	4.102	0.09268	0.1286	59	900.488
{S(t) p(.) r(t) R(t) R'(t) F(.) F'(.)}	1633.194	4.442	0.07819	0.1085	57	905.777
{S(.) p(.) r(t) R(.) R'(.) F(.) F'(.)}	1705.735	76.983	0	0	20	1063.481
{S(t) p(.) r(.) R(.) R'(.) F(.) F'(.)}	1711.203	82.451	0	0	20	1068.949
{S(.) p(t) r(.) R(.) R'(.) F(.) F'(.)}	1712.374	83.622	0	0	20	1070.12
{S(.) p(.) r(.) R(.) R'(.) F(.) F'(.)}	1713.645	84.893	0	0	7	1098.698

Some questions to answer or think about for these data:

1. Is there evidence of emigration from this study? If so, is the emigration random, permanent, or Markovian?
2. How much do you suppose the resightings improve the estimates of survival? Hint: take a look at the encounter histories file to get some idea of the number of 2's that appear.
3. How much do you suppose the dead recoveries improve the estimates of survival? Hint: in the Results Browser, use the Output | Input Data Summary menu choice to compare the numbers of dead recoveries to the numbers of live captures.
4. Can incorporating dead recoveries and/or live resightings into a Cormack-Jolly-Seber model ever hurt the precision of your estimates of survival compared to just CJS estimates? See Barker and White (2001) for some discussion of this issue.

**Pradel Model**

Pradel (1996) developed a model to estimate the proportion of the population that was previously in the population. Thus, this model, labeled 'Pradel Recruitment Only' in MARK, estimates recruitment to the population. The parameters of this model are the seniority probability,  $\gamma$  (probability that an animal present at time  $i$  was already present at time  $i - 1$ ), and recapture probability  $r$ . Only LLLL encounter histories are required for this model.

Pradel (1996) extended his recruitment only model to include apparent survival ( $\phi$ ). In MARK, this model is labeled 'Pradel Survival and Rec.'. Parameters of the model are apparent survival ( $\phi$ ), recapture probability ( $p$ ), and seniority probability ( $\gamma$ ).  $\gamma$  is defined as the

probability that an animal at time  $i$  was in the population at time  $i - 1$ . Only LLLL encounter histories are required for this model.

Pradel (1996) also parameterized his model with both recruitment and apparent survival to have the parameters apparent survival ( $\phi$ ), recapture probability ( $p$ ), and rate of population change ( $\lambda$ ), where  $\lambda$  is population size at time  $i + 1$  divided by population size at time  $i$ , or  $N_{i+1}/N_i$ . Likewise, only LLLL encounter histories are required for this model.

A additional extension to the Pradel (1996) models has been parameterized with both recruitment and apparent survival to have the parameters apparent survival ( $\phi$ ), recapture probability ( $p$ ), and fecundity rate ( $f$ ), which is defined as the number of adults at time  $i$  per adult at time  $i - 1$ . Only LLLL encounter histories are required for this model.

The relationships between the parameters of the 3 models derived from Pradel (1996) are as follows. Pradel's (1996) original paper defined the parameter  $\gamma_i$  as the probability that an animal at time  $i$  had not entered the population between time  $i$  and  $i - 1$ . In terms of Jolly's original model,  $\gamma_{i+1} = 1 - B_i/N_{i+1}$ .  $\lambda$  is the rate of change of the population, so  $\lambda_i = N_{i+1}/N_i$ .  $f$  is the fecundity rate of the population, so that  $f_i$  is the number of new animals in the population at time  $i$  per animal in the population at time  $i - 1$ , or  $N_{i+1} = N_i f_i + N_i \phi_i$ .

The following table provides the relationships between these 3 parameters.

$$\gamma_{i+1} = \phi_i / (f_i + \phi_i)$$

$$\gamma_{i+1} = \phi_i / \lambda_i$$

$$\lambda_i = \phi_i / \gamma_{i+1}$$

$$\lambda_i = f_i + \phi_i$$

$$f_i = \lambda_i - \phi_i$$

$$f_i = \phi_i (1 - \gamma_{i+1}) / \gamma_{i+1}$$

Some other useful relationships are:

$$N_i \phi_i = \gamma_{i+1} N_{i+1}$$

$$N_i \phi_i + B_i = N_{i+1}$$

$$\phi_i + \frac{B_i}{N_i} = \frac{N_{i+1}}{N_i} = \lambda_i$$

$$f_i = \frac{B_i}{N_i}$$

### Data Analysis

The data we will be analyzing in this exercise concerns survival of pikeminnows in the Upper Colorado River system. Five to 25 mile reaches of river were electrofished during a 2-day period each spring beginning in 1991. Fish were marked with passive induced transponder (PIT) tags. There were 5 reaches sampled on the Green River, 3 on the Colorado River, 3 on the Yampa River, and 2 on the White River. These 13 reaches were considered groups in the following analysis to allow evaluating differences between the river systems.

The pikeminnow data will also be used for this analysis, with the intent of estimating the rate of change of the population ( $\lambda$ ). The Pradel model requires that the sampling area remain constant. To understand the importance of this assumption, think about what happens to  $\lambda$  when the sampling area is increased. As the area expands, more and more new animals are captured, making  $\lambda$  increase as if new animals were being recruited to the population. The primary objective of the analysis was to examine the trend in  $\lambda$  over time, and to make recommendations for developing a better monitoring program for the pikeminnow population in the Upper Colorado River Basin.

MARK files are available as J:\CLASSES\FW663\EXERCISE.11\PIKE4.DBF and J:\CLASSES\FW663\EXERCISE.11\PIKE4.FPT.

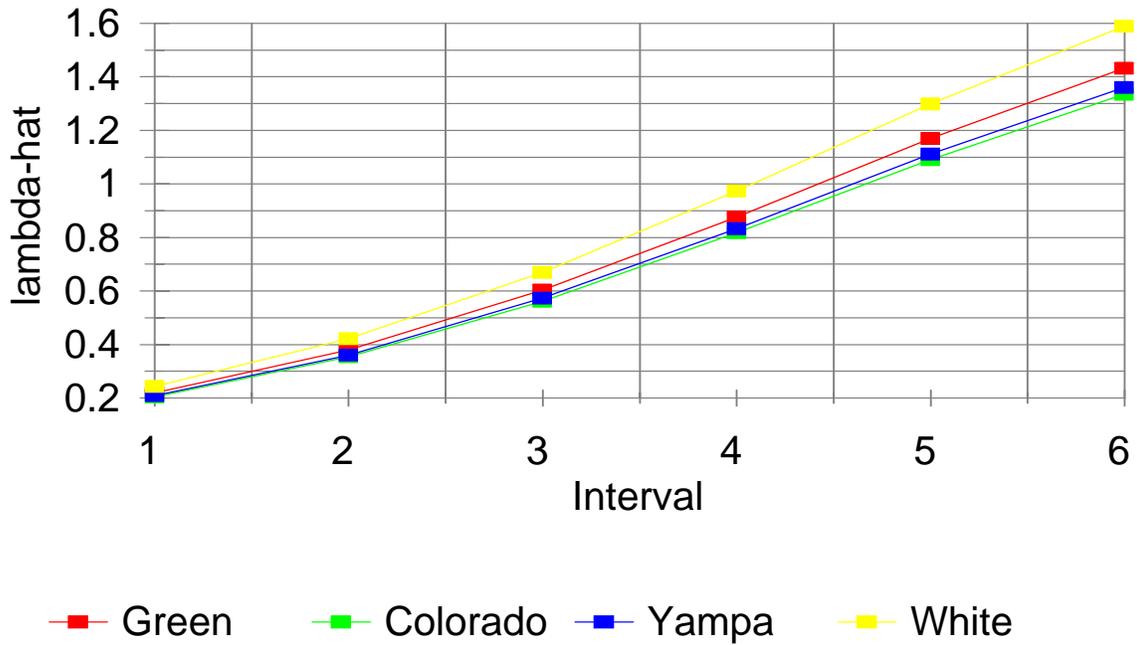
### Results

The best Pradel model had 73% of the AIC weight, and other models close to the top were flaky because of inestimable parameters, so we'll use the top AICc model for the remaining analyses. The model had survival constant across all reaches and times, and both capture probability and  $\lambda$  as a function of river and a continuous time effect plus continuous time effect squared:  $\{\Phi(\cdot) p(\text{river}+T+T^2) \lambda(\text{river}+T+T^2)\}$ . We tried some other variations of this model, but it was pretty solidly at the top, as shown by the weight.

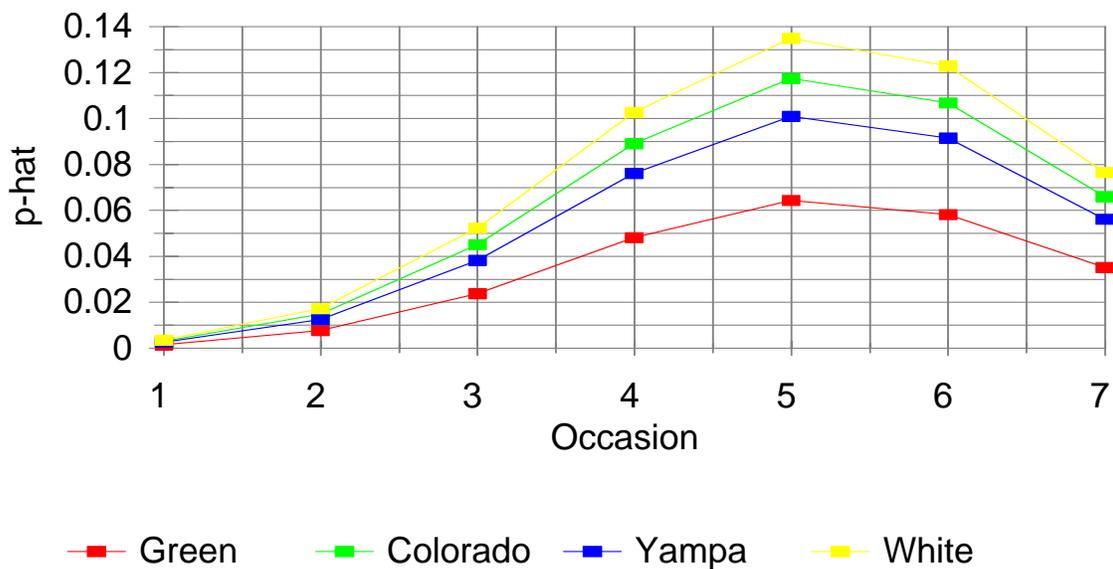
Model	AICc	Delta AICc	AICc Weights	Num. Par	Deviance
{Phi(.) p(river+T+T2) lambda(river+T+T2)}	4440.736	0	0.72901	13	4414.379
{Phi(.) p(river+T+T2) lambda(river+T)}	4444.822	4.086	0.09451	12	4420.517
{Phi(.) p(river+t) lambda(river+T)}	4445.218	4.482	0.07753	16	4412.683
{Phi(.) p(river+t) lambda(river+t)}	4445.608	4.872	0.0638	18	4408.934
{Phi(river) p(river+t) lambda(river+t)}	4448.948	8.212	0.01201	22	4403.947
{Phi(.) p(t) lambda(river+T)}	4450.339	9.603	0.00599	13	4423.983
{Phi(.) p(river+t) lambda(T)}	4450.682	9.946	0.00505	13	4424.325
{Phi(.) p(river) lambda(river+T)}	4451.35	10.614	0.00361	10	4431.135
{Phi(.) p(river+t) lambda(t)}	4451.812	11.076	0.00287	16	4419.277
{Phi(.) p(river+T) lambda(river+T)}	4452.31	11.574	0.00224	11	4430.052
{Phi(river) p(river+t) lambda(river)}	4453.117	12.381	0.00149	18	4416.443
{Phi(river) p(river) lambda(river)}	4454.735	13.999	0.00067	12	4430.429
{Phi(river+t) p(river+t) lambda(river)}	4455.668	14.932	0.00042	22	4410.667
{Phi(river+t) p(river+t) lambda(river+t)}	4456.129	15.393	0.00033	26	4402.735
{Phi(river) p(river) lambda(river+t)}	4456.643	15.907	0.00026	17	4422.041
{Phi(.) p(river+t) lambda(.)}	4457.417	16.681	0.00017	12	4433.111
{Phi(river+t) p(river) lambda(river+t)}	4460.828	20.092	3E-05	22	4415.827
{Phi(.) p(.) lambda(.)}	4461.367	20.631	2E-05	3	4455.344
{Phi(g*t) p(g*t) Lambda(g*t) PIM coding}	4676.656	235.92	0	221	4194.523

The main conclusion from this model is that  $\lambda$  has been increasing with time – see the following graph, and current values are  $>1$ . A value of  $\lambda = 1$  means the population size is not changing;  $\lambda > 1$  means the population is increasing, and  $\lambda < 1$  means the population is declining. Thinking about trends in  $\lambda$  is a little tricky, because  $\lambda$  is a rate to start with. The main conclusion from the best model is that the population is increasing, and current values of  $\lambda$  are  $>1$ . Likely the reason that lambda is estimated so high is because there must have been a pretty large increase in the population.

# Rate of Population Change



# Capture Probabilities



**Topics for Discussion**

1. Would you always expect to find  $R > R'$  in Barker's model?
2. Suppose that only females are being modeled in a Pradel model with the  $f$  parameterization. What is the interpretation of the estimates of  $f$  from the model, i.e, suppose that each female has 2 female offspring each year. Would 2 times the juvenile survival rate equal the estimate of  $f$  obtained from the marked animal data? Why or why not?
3. What recommendations would you make to Kevin Bestgen to improve the pikeminnow data collection to develop a sensitive monitoring system? What happens if the number of reaches sampled is increased to improve the power of the monitoring scheme to detect changes? What happens if the capture probability ( $p$ ) is increased by more intensive shocking in the same river reaches? What happens if each reach is sampled more extensively, i.e., backwaters that had been previously ignored will now be carefully electrofished?

**Literature Cited**

- Barker, R. J. 1997. Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics* 53:666-677.
- Barker, R. J. 1999. Joint analysis of mark-recapture, resighting and ring-recovery data with age-dependence and marking-effect. *Bird Study* 46 Supplement:82-91.
- Barker, R. J., and G. C. White. 2001. Joint analysis of live and dead encounters of marked animals. Pages 361-367 in R. Field, R. J. Warren, H. Okarma, and P. R. Sievert, editors. *Wildlife, land, and people: priorities for the 21<sup>st</sup> century*. Proceedings of the Second International Wildlife Management Congress. The Wildlife Society, Bethesda, Maryland, USA.
- Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1985. *Statistical inference from band recovery data - a handbook*, 2nd ed. U.S. Fish and Wildlife Service Research Publication Number 131, Washington, D.C. 305pp.
- Burnham, K. P. 1993. A theory for combined analysis of ring recovery and recapture data. Pages 199-213 in J.-D. Lebreton and P. M. North, editors. *Marked individuals in the study of bird population*. Birkhauser Verlag, Basel, Switzerland.
- Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* 52:703-709.