**Incorporating Dependence in Tag Loss Estimation**

**Introduction**

Developing methods of marking animals that minimize or eliminate mark loss are important because mark-recapture estimators assume recaptures of marked individuals are always detected (Seber 1982). However, current methods of marking many different species indicate that problems with mark retention encompass small mammals (Fokidis et al. 2006), large terrestrial mammals (Fosgate et al. 2006), aquatic mammals (Bradshaw et al. 2000), fish (Cowen and Schwartz 2006), and reptiles (Rivalan et al. 2005). If the assumption that no marks are lost is violated then estimates of population parameters (e.g., survival and abundance) will be biased (Arnason and Mills 1981, Diefenbach and Alt 1998) because previously marked animals that lose all marks are treated as being part of the unmarked population in closed population models and as immigrants in open models. Furthermore, even if an estimator is robust to violation of this assumption, the loss of marks will reduce the precision of parameter estimates because of the apparent reduction in sample size (Rotella and Hines 2005).

Using natural markings to track individual animals can have low error rates (Stevick et al. 2001), but not all species can be monitored in this manner. Consequently, methods of estimating tag loss likely will be necessary until new technologies are developed that eliminate mark loss. Historically, animals have been double-marked and the status of marks upon recapture (none or one mark missing) have been used to estimate mark loss under the assumption that each mark is lost independently of the other mark (Beverton and Holt 1957, Seber 1982). This independence assumption is required because the recapture of a marked animal that lost both marks is not observable.

In recent studies where the opportunity has occurred to observe loss of both tags via the use of a third permanent mark, however, the independence assumption has been shown to be invalid (Siniff and Ralls, 1991; Diefenbach and Alt, 1998; Bradshaw, Barker and Davis, 2000; Rivalan et al., 2005). These studies undermine the credibility of tag loss evaluations for situations in which loss of both marks is not observable (Pistorius et al., 2000). Consequently, it is important to incorporate appropriate models of mark loss to reduce bias in estimates of population parameters, even if modeling tag loss results in reduced precision (e.g., Pollock et al. 2007).

Unfortunately, not all species can be marked with a third permanent mark to detect the recapture of individuals with both marks missing to assess the assumption of independence of each mark loss event. This situation has some similarities to mark-recapture distance sampling (Laake, 1999; Borchers et al., 2006; Laake, Dawson and Hone, 2008) in which objects missed by both observers cannot be included in the sample because they are not observed (Borchers, 1999). However, <<appropriate Borchers/Laake citation here>> used ancillary distance data to weaken the independence assumption such that independence was assumed only for observations that occurred close to the transect line and improve abundance estimates. If the independence assumption can be similarly weakened in situations where both marks are lost in mark-recaptures studies, it may be possible to evaluate loss of marks for species where only double-marked individuals are feasible.

We develop a model of mark loss that explicitly models dependence in loss of marks and then we consider situations where we can weaken the independence assumption by incorporating dependence. We apply the model to tagged California sea lions (*Zalophus californianus*) and black bear (*Ursus americanus*) which were both double tagged and given a separate permanent mark. We compare the results using an independence model and dependence model with the double tag data excluding the double tag loss observations to the results using the data including the known double tag loss.

**Tag Loss Models**

We first consider a simple situation which has been well described in the literature (Seber 1982) and then we will build from this foundation. Two tags are applied to a sample of animals, they are released and at some time later a sample of animals are observed having one (n1= n01+ n10) or two (n2= n11) tags present. In this simple example, double tag loss () is not observable. Let **S** = (*S*1, *S*2) represent the status vector of the 2 tags where *Si* is 1 if the *ith* tag is present and 0 if the *ith* tag is absent (lost) (note: this is counter to the convention of Diefenbach et al. (1998), Bradshaw et al. (2000) and Rivalan et al. (2005) who use 1 to denote a tag loss event). Usually the tagged animals can only be identified if one or more tags are retained which means those that lost both tags (**S** = (0,0)) are not observable. Initially, we assume both tags have the same probability of tag loss, *p*, and we assume independence of tag fates.

Table 1. Joint () and marginal (*p*, *q*) probabilities for double tag status **S** = (*S*1, *S*2) with an assumed independence structure. The *n00* cell is shaded gray because it is not observed.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tag 2 Status (S2)** | |  |
| **Tag 1 Status (S1)** | *Present (S2=1)* | *Absent (S2=0)* | *Marginal* |
| *Present (S1=1)* |  |  |  |
| *Absent (S1=0)* |  |  |  |
| *Marginal* |  |  |  |

Based on the independence assumption, the probability that an animal retains both tags is (1-p)2, the probability of retaining only one tag is 2p(1-p) and the probability of losing both tags is p2. The data and probability structure can be represented by a 2x2 contingency table for the status (present (retained)/absent (lost)) for each tag (Table 1). The marginal probabilities of tag loss and tag retention are *p* and *q*=1-*p* and the joint probabilities are represented as for each of the cells in the table. Due to the assumed independence the conditional probability of tag loss or retention of a tag given the status of the other tag is the same as the marginal probability. For example, the probability that tag 1 is lost given that tag 2 is retained is.

The log-likelihood function (excluding the constant) for *p* given the observed data (*n*1, *n*2) is conditioned on the observations which excludes

The maximum likelihood estimator (MLE) for tag loss probability is . Now if we allow different probabilities for each tag (e.g., different tag types or tag orientation), then the log-likelihood function is:

and the MLEs are and which are equivalent to the capture probability estimators in a two occasion capture-recapture experiment for a closed population (i.e., Lincoln-Petersen). Each of those estimators can be viewed as a conditional probability because they measure the probability that tag *i* was lost given that tag 3-*i* was retained. If independence holds, the conditional and marginal probabilities are the same.

We can also express the tag loss probabilities using a logit link (Bradshaw et al. 2000) which becomes helpful to incorporate covariates and to express dependence in tag fates. With the logit link, the natural logarithm (loge) of the odds (loge(p/(1-p)) is some linear function of the parameters. Odds are simply the ratio of the probabilities of the event occurring and not occurring. Allowing different tag loss probabilities for each tag, we can re-express the probabilities in Table 1 such that the odds of losing the *ith* tag (*Si*=0) is (Table 2). The model has the same structure and there is a 1-1 relationship between and the parameters

Table 2. Joint () and marginal () probabilities for double tag status **S** = (*S*1, *S*2), where the odds (of losing the *ith* tag is and . The *n00* cell is shaded gray because it is not observed.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tag 2 Status (S2)** | |  |
| **Tag 1 Status (S1)** | *Present (S2=1)* | *Absent (S2=0)* | *Marginal* |
| *Present (S1=1)* |  |  |  |
| *Absent (S1=0)* |  |  |  |
| *Marginal* |  |  |  |

Now we expand the set of observed data to include those animals that lost both tags () and we allow for possible dependence in the fates of the tags. In particular, we specify different odds of losing a tag that are dependent on the status of the other tag (Table 3). The odds of losing the *ith* tag (*Si*=0) given the other tag is present (*S3-i* =1) is and the odds of losing the *ith* tag (*Si*=0) given other tag is absent (*S3-i=*0) is where α determines the amount of dependence. These dependent odds are ratios of the joint probabilities (e.g., and not the marginal probabilities. Also, because of the dependence the conditional and marginal probabilities differ.

Table 3. Joint () and marginal (*pi*, *qi*) probabilities for double tag status **S** = (*S*1, *S*2), where the odds of losing the *ith* tag (*Si*=0) given the other tag is present (*S3-i* =1) is , odds of losing the *ith* tag (*Si*=0) given other tag is absent (*S3-i=*0) is , and . The *n00* cell is observed in this model.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tag 2 Status (S2)** | |  |
| **Tag 1 Status (S1)** | *Present (S2=1)* | *Absent (S2=0)* | *Marginal* |
| *Present (S1=1)* |  |  |  |
| *Absent (S1=0)* |  |  |  |
| *Marginal* |  |  |  |

We define to be the conditional probability that the *ith* tag is lost given the other tag status is *s3-i*:

Thus, whereas, . The conditional probability of tag retention is and Each joint probability can be specified by the product of a marginal and a conditional probability: and

Diefenbach et al (1998) and Rivalan et al. (2005) specified the joint probabilities in Table 3 in a slightly different manner with , and , where p is the marginal tag loss probability and *p\** is the conditional probability of losing a tag given the other tag was absent. Although not stated, their conditional probability of losing a tag given the presence of the other tag would be . In both papers the authors used separate functional forms for *p* and *p\** and specified the independence model by using the same model for both (*p*=*p\**). Their approach is viable but we believe it is preferable to have a model with a parameter that controls dependence and with independence specified simply as α=0. As we show later this enables flexibility in modeling and incorporating dependence. Our model was motivated by Bradshaw et al. (2000) which used logistic regression and specified dependence in tag loss via interactions and measured effects in terms of odds multipliers. We note that they also provide a closed form estimator for the model in Table 3 with no covariates and α=0.

As mentioned previously the probability structure for tag loss is quite similar to capture-recapture (mark-recapture) for two occasions with a closed population which has been used with two observers to measure detection probability in visual surveys. When detection probability is measured solely with the mark-recapture data, it is necessary to assume independence between the detections by the two observers because those missed by both observers () are obviously not included in the sample (Borchers, 1999). Recently, the independence assumption was weakened (Borchers et al., 2006; Laake, 1999; Laake, Dawson and Hone, 2008)(Laake and Borchers 2004) in the combined mark-recapture and distance sampling by including a dependence measure δ(*x*) which was estimated as the discrepancy between the detection probability at distance *x* measured by the mark-recapture (double observer) data (based on independence) and the distance sampling data. If δ(*x*)=1 then independence at all distances is achieved. Because detection probability at *x*=0 cannot be measured from the distance sampling data, the independence assumption for the mark-recapture data was required for *x*=0 (δ(*0*)=1) but not for the other distances.

The dependence structure we have defined for tag loss can be expressed in terms of the δ dependence of Borchers et al. (2006). Under the independence model (Table 2 excluding ), the probability that an animal would retain at least one tag is:

Likewise for the dependence model (Table 3):

The dependence of Borchers et al. (2006) is a ratio that measures the distortion between the joint probabilities from the independence model (Table 2) and the dependence model (Table 3) which can be expressed as:

The same relationship is obtained from and from the ratio of any of the joint probabilities other than for the (0,0) event which is not used in the independence model. The dependence measure can also be expressed in terms of covariance (Borchers 1999) which in this case can be expressed as:

In general there will likely be positive dependence in tag loss which means α>0 and δ>1 but negative dependence (α<0) is possible with a lower bound of δ>1-.

The joint probabilities can be rewritten in terms of δ as: and or as and . The latter form makes it obvious that once you exclude and condition on the observed data, the δ will cancel from the rescaled joint probabilities which will only be functions of the . This is also obvious by examining Table 3 and noting that the joint probabilities for the observed set of data ( would only be functions of after conditioning on the exclusion of . The same result was shown by Borchers et al. (2006) for mark-recapture distance sampling (mrds) but in that case δ could be estimated from the observed distances. For tag loss estimation, there is no equivalent ancillary data; however, there are various ways that dependence can be incorporated to improve tag loss estimation.

The most obvious and best way to incorporate dependence is to mark a subset of the animals with a permanent mark so for a subset of the data is observable. For example, all captured black bears in Pennsylvania were given 2 ear tags but some of the bears were also given a permanent unique lip tattoo (Diefenbach et al. 1998) which could be used to identify bears that had lost both ear tags. The likelihood for the tagged only ( would be:

and for the tagged and marked: (:

The combined likelihood for the data could improve on the precision of the tag loss estimates while adjusting for dependence measured from the permanently marked bears. In the absence of data on permanently marked animals, a Bayesian alternative could be used in which a prior distribution for α or δ could be used with the likelihood for the tagged only bears. Preferably the prior would be data-derived from a similar species and situation.

Next we consider incorporating dependence for situations in which all of the animals are double-tagged and we cannot observe animals that lose both tags. On the surface this appears to be impossible because there is no ancillary data as with mrds. However, we can replace the role of ancillary data with models that restrict the dependence and still provide estimable parameters. For example, consider California sea lions which have a tag applied to both fore-flippers as pups. Initially, tag loss may be due to manufacturing or application defects which may be independent between tags; however, as the animal grows, the expansion of the fore-flipper may put pressure on the tag causing the tag to split or may cause tissue damage that allows the tag to fall out. Because growth is symmetric, if one tag is lost the other is also more likely to be lost. Thus, we could propose a model in which α=0 for the first year but not subsequent years. As long as tag loss parameters are age-invariant then any dependence beyond the first year can be estimated. Conceptually, this is similar to having a subset of animals with a permanent mark but instead we have a subset of animals for which it is reasonable to assume independence holds.

**Literature Cited**

Borchers, D. L., Laake, J. L., Southwell, C., and Paxton, C. G. M. (2006). Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62**, 372-378.

Laake, J. (1999). Distance sampling with independent observers: Reducing bias from heterogeneity by weakening the conditional independence assumption. In *Marine Mammal Survey and Assessment Methods*, 137-148. Rotterdam: A A BALKEMA.

Laake, J., Dawson, M. J., and Hone, J. (2008). Visibility bias in aerial survey: mark-recapture, line-transect or both? *Wildlife Research* **35**, 299-309.

Borchers, D. L. (1999). Composite mark-recapture line transect surveys. In *Marine Mammal Survey and Assessment Methods*, 115-126. Rotterdam: A A BALKEMA.

Borchers, D. L., Laake, J. L., Southwell, C., and Paxton, C. G. M. (2006). Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62**, 372-378.

Bradshaw, C. J. A., Barker, R. J., and Davis, L. S. (2000). Modeling tag loss in New Zealand fur seal pups. *Journal of Agricultural Biological and Environmental Statistics* **5**, 475-485.

Diefenbach, D. R., and Alt, G. L. (1998). Modeling and evaluation of ear tag loss in black bears. *Journal of Wildlife Management* **62**, 1292-1300.

Pistorius, P. A., Bester, M. N., Kirkman, S. P., and Boveng, P. L. (2000). Evaluation of age- and sex-dependent rates of tag loss in southern elephant seals. *Journal of Wildlife Management* **64**, 373-380.

Rivalan, P., Godfrey, M. H., Prevot-Julliard, A. C., and Girondot, M. (2005). Maximum likelihood estimates of tag loss in leatherback sea turtles. *Journal of Wildlife Management* **69**, 540-548.

Siniff, D. B., and Ralls, K. (1991). Reproduction, Survival and Tag Loss in California Sea Otters. *Marine Mammal Science* **7**, 211-229.

Alt, G. L., C. R. McClaughlin, and K. H. Pollock. 1985. Ear tag loss by black bears in Pennsylvania. Journal of Wildlife Management 49:316-320.

Arnason, A. N., and K. H. Mills. 1981. Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. Canadian Journal of Fisheries and Aquatic Sciences 38:1077-1095.

Bradshaw, C. J. A., Barker, R. J., and Davis, L. S. (2000). Modeling tag loss in New Zealand fur seal pups. Journal of Agricultural, Biological, and Environmental Statistics 5:475–485.

Cowen, L., and C. J. Schwartz. 2006. The Jolly-Seber model with tag loss. Biometrics 62:699-705.

Fabrizio MC, Nichols JD, Hines JE, Swanson BL, Schram ST. 1999. Modeling data from double-tagging experiments to estimate heterogeneous rates of tag shedding in lake trout (*Salvelinus namaycush*). Canadian Journal of Fisheries and Aquatic Sciences 56:1409-1419.

Fosgate GT, Adesiyun AA, Hird DW. 2006. Ear-tag retention and identification methods for extensively managed water buffalo (*Bubalus bubalis*) in Trinidad. Preventive Veterinary Medicine 73:287-296.

Pollock KH (Pollock, Kenneth H.), Yoshizaki J (Yoshizaki, Jun), Fabrizio MC (Fabrizio, Mary C.), Schram ST (Schram, Stephen T.). 2007. Factors affecting survival rates of a recovering lake trout population estimated by mark-recapture in Lake Superior, 1969-1996. Transactions of the American Fisheries Society 136:185-194.

Rotella, J. J., and J. E. Hines. 2005. Effects of tag loss on direct estimates of population growth rate. Ecology 86:821-827.

Stevick PT, Palsboll PJ, Smith TD, Bravington MV, Hammond PS. 2001. Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. Canadian Journal of Fisheries and Aquatic Sciences 58:1861-1870.