

## Forum

# Analysis of survival probabilities of female common toads, *Bufo bufo*

Benedikt R. Schmidt<sup>1</sup>, Bradley R. Anholt<sup>2</sup>

<sup>1</sup> Zoologisches Institut, Universität Basel, Rheinsprung 9, CH-4051 Basel, Switzerland  
e-mail: schmidtb@ubaclu.unibas.ch

<sup>2</sup> Department of Biology, University of Victoria, PO Box 3020 STN CSC, Victoria, BC, V8W 3N5, Canada  
e-mail: banholt@uvic.ca

Recent advances in the statistical theory of mark-recapture analysis have markedly enhanced our ability to estimate survival probabilities in natural populations (Burnham et al., 1987; Pollock et al., 1990; Lebreton et al., 1992). Here, we report an analysis of a data set drawn from a published study of female common toads, *Bufo bufo* (Kuhn, 1994a). In this otherwise statistically well analysed paper, capture-mark-recapture histories of thousands of toads were not analysed statistically but inferences were drawn nonetheless. This resulted in a serious underestimation of survival probabilities. We wish to correct these survival estimates and advocate the use of robust statistics for drawing inferences. We will also show that such an analysis can reveal unexpected patterns when using drift fences around amphibian breeding sites.

In the analysis and interpretation of capture-mark-recapture data it is necessary to distinguish between survival and recapture probabilities and to estimate both, especially if one wishes to understand which factors or traits affect survival (Lebreton et al., 1992; Martin et al., 1995). The survival probability is the probability that an individual survives from time  $t_i$  to  $t_{i+1}$ , usually from one year to the next, depending on the biological question asked. The recapture probability is the probability that an individual is recaptured, given that it is alive (see Lebreton et al., 1992, 1993). Often, return probabilities (= *survival probability*  $\times$  *recapture probability*) rather than survival probabilities are calculated and analysed (e.g., Kapfberger, 1984; Elmberg, 1990; Seidel, 1993; Jehle et al., 1995; see Lebreton et al., 1993 and Martin et al., 1995 for discussion). Return probabilities systematically underestimate survival probabilities (Martin et al., 1995). Additionally, probabilities of recapture and survival may vary among sexes, age classes or other biological groups (e.g., van Gelder and Rijdsdijk, 1987; Baron et al., 1996; Flatt et al., 1997). Thus, return probabilities may be seriously biased and therefore an analysis of return probabilities is not meaningful (Martin et al., 1995). An analysis of life histories and population dynamics based on such estimates may be simply wrong (Martin et al., 1995; van Gelder and Rijdsdijk, 1987, and references therein). Because population dynamics are the result of the balance between mortality, recruitment and migration, an understanding and estimation of survival is one of the keys to the understanding of population biology, whether for fundamental (e.g., Shine and Charnov, 1992; Forsman, 1995; Baron et al., 1996) or applied goals (e.g., Crouse et al., 1987; Doak et al., 1994; Halley et al., 1996).

### *Methods and results of Kuhn (1994a)*

Kuhn (1994a) enclosed three breeding sites of the common toad, *Bufo bufo*, using drift fences. Fences were erected in early March (well before the breeding season of this explosive breeder) and removed in May when no more toads were captured for several nights. Toads were collected at least once a day. They were marked (most were marked individually), measured and released on the same day. Using these methods, Kuhn (1994a) compiled an impressive capture-mark-recapture data set of thousands of toads. Kuhn (1994a) only presents data of female toads in his paper.

Because *Bufo bufo* shows high breeding site fidelity (Heusser, 1960, 1969) and because Kuhn (1994a) assumed that drift fences would sample the populations without bias, Kuhn (1994a) assumed that survival probabilities can be estimated directly from return probabilities. Kuhn (1994a) stated that a statistical analysis of survival probabilities would require a correction factor to account for the fact that some of the toads that were marked in the first year of his study were not reproducing for the first time. Kuhn (1994a) did not discuss survival probabilities explicitly. He wrote that only 5 to 25% of the toads breed more than once, implicitly stating that survival probabilities are 5 to 25%. He was aware that he was analysing return probabilities (rather than survival probabilities) when he analysed which factors (such as age, size or reproductive effort) affected them.

Based on the difference between the number of toads arriving at the pond and the number of toads leaving the pond, Kuhn (1994a) estimated that the mortality during the breeding season was roughly 15-40%, depending on the population and year. These estimates are a component of annual mortality and set an upper limit to survival probabilities from year to year. These estimates do not contradict our analyses.

### *The statistical analysis of survival in capture-mark-recapture studies*

Here, we only present a short description of capture-mark-recapture (CMR) statistics. Lebreton et al. (1992) and Cooch et al. (1997) present more detailed introductions to CMR statistics, Burnham et al. (1987) and Pollock et al. (1990) provide comprehensive accounts.

Table 1 illustrates the recapture process and the notation of parameters in a study. It illustrates a study where a cohort of individuals is marked at time  $t_1$  and recaptured at times  $t_2$  and  $t_3$ . We use population  $B$  of Kuhn (1994a) as an example. For simplicity, we ignore toads that were first captured at times  $t_2$  and  $t_3$ . The recapture tree could easily be expanded to include these animals (see table 4 of Lebreton et al., 1992). A "1" represents capture, "0" indicates that an individual was not captured. A capture history of the type "111" indicates that an individual was marked at  $t_1$  and recaptured at both  $t_2$  and  $t_3$ . Typically, not all individuals that are alive are recaptured during all recapture sessions: they have capture histories of the type "101". This has two implications: 1) one should incorporate the fact that not all marked individuals that are alive are recaptured into the estimation of survival probabilities and 2) more than one recapture session after marking is needed to estimate survival probabilities. One recapture session allows only an estimate of the type "survival was at least 20%" (if 20 out of 100 marked individuals are recaptured). Therefore, the analysis of survival probabilities or the analysis of CMR data sets requires another parameter: the recapture probability, denoted  $p_i$ .

To analyse the recapture tree in table 1, i.e. to analyse the probability of a capture history, we need the probabilities of survival  $\phi_1$  (from  $t_1$  to  $t_2$ ) and  $\phi_2$  (from  $t_2$  to  $t_3$ ) and the probabilities of recapture at times  $t_2$  and  $t_3$ ,  $p_2$  and  $p_3$ . For example, the capture history "111" requires that an individual survived from time  $t_1$  to  $t_2$  and  $t_3$  and that it was recaptured both at times  $t_2$  and  $t_3$ . According to probability calculus, the probability of this recapture history is  $(\phi_1 p_2 \phi_2 p_3)$ . The observed number of realizations of every capture history is the basis for the estimation of the unknown parameters in the equations under the heading "expected numbers at time  $t_3$ " in table 1. The mathematics underlying the estimation are described in e.g. Burnham et al. (1987) and Lebreton et al. (1992). The probability  $p_1$  does not appear in the equations and can therefore not be estimated. The parameters of the final recapture session cannot be estimated because they appear in the model only as a product  $(\phi_2 p_3)$  that we denote  $\beta_3$ . With two recapture sessions, only  $\phi_1$  and  $p_2$  can be separately estimated.

The estimation of the survival and recapture probabilities is done in the framework of the Cormack-Jolly-Seber (CJS) model. This model makes four assumptions (Lebreton et al., 1992; Cooch et al., 1997): 1) every marked individual at time  $t_i$  has the same probability of recapture  $p_i$ , 2) every marked individual in the population immediately after time  $t_i$  has the same probability of surviving to time  $t_{i+1}$  (in other words:

**Table 1.** Recapture tree for two occasions of recapture of a single cohort of marked individuals. Upper branches mean capture (1), and lower branches mean no capture (0). Observed numbers are given as an example and are taken from Kuhn (1994a; his fig. 4), but for simplicity only individuals captured in 1989 (time 1) are included in the figure.  $N$  = number marked and released in 1989,  $\phi_i$  = survival probability,  $p_i$  = recapture probability,  $\beta_3 = \phi_2 p_3$ .

| Initial number marked and released at time 1 | Captured, $X_{11}$ , not captured, $X_{10}$ , at time 2 | Capture histories at time 3 | Expected number at time 3 (values are based on our estimates in table 3) | Observed number at time 3                               |     |
|----------------------------------------------|---------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------|---------------------------------------------------------|-----|
| $N = 276$                                    | $X_{11}$                                                | /                           | $N \phi_1 p_2 \beta_3 = 3.8$                                             | 6                                                       |     |
|                                              |                                                         | $\cap$                      | $X_{110}$                                                                | $N \phi_1 p_2 (1 - \beta_3) = 50.3$                     | 48  |
|                                              | $X_{10}$                                                | $\cap$                      | $X_{101}$                                                                | $N \phi_1 (1 - p_2) \beta_3 = 7.0$                      | 7   |
|                                              |                                                         | /                           | $X_{100}$                                                                | $N [1 - \phi_1 p_2 - \phi_1 (1 - p_2) \beta_3] = 214.9$ | 215 |
| Parameter                                    |                                                         | Recapture occasion          |                                                                          |                                                         |     |
|                                              |                                                         | $t_1$                       | $t_2$                                                                    | $t_3$                                                   |     |
| Survival probabilities                       |                                                         | $\phi_1 \rightarrow$        |                                                                          | $\phi_2 \rightarrow$                                    |     |
| Recapture probabilities                      |                                                         | $p_1$                       | $p_2$                                                                    | $p_3$                                                   |     |

individuals marked at time  $t_{i-1}$  have the same probability of surviving from  $t_i$  to  $t_{i+1}$  as do individuals marked at time  $t_i$ , 3) marks are not lost or missed, and 4) all recapture sessions are instantaneous, relative to the interval between time  $t_i$  and  $t_{i+1}$ , and each release is made immediately after the recapture session (see Cooch et al., 1997, for further discussion of the assumptions). Assumptions 1 and 2 are generally hardest to meet. However, if the assumptions of the simple CJS model are not met, additional parameters can be estimated from the data to account for these deviations. Deviations from the CJS model as revealed by the various tests that together form Test 2 and Test 3 of RELEASE provide direction for the appropriate model structure and suggest biological hypotheses (e.g., unequal recapture probabilities may hint at a biologically interesting phenomenon; Lebreton et al., 1992; Cooch et al., 1997). A CJS model with time-dependent survival and recapture probabilities needs to be fitted to the data to estimate variance inflation factors to maintain conservative Type I errors and to see whether the data meet the assumptions of binomial error distributions (Lebreton et al., 1992; Cooch et al., 1997). The expected result is that the residual deviance is approximately equal to the residual degrees of freedom. Residual values more than twice the sample size suggest some sort of heterogeneity among the animals not accounted for by the model such as different sub-populations being treated as a homogeneous unit.

After testing the assumptions, a fully saturated CJS model is fitted to the data in which both survival and recapture parameters vary among recapture sessions (denoted  $\phi_t$  and  $p_t$ , respectively). More generally, a model is fit to the data in which all parameters are included that are to be tested or are suggested by the tests of the assumptions. This initial model then serves as a null model based on which the significance of parameters of interest can be tested. Significance test are done by model simplification. Model simplification proceeds by comparing the residual deviances of two models, one of which is a special case of the second. For example, if the more general model has a separate parameter for each recapture event and the simpler model has the same parameter for every recapture event, then the simpler model is a special case of the more general. The simpler model can never explain the data as well as the more complex model, but the additional parameters may not explain a significant amount of the variation. Statistical significance is calculated by examining the increase

in deviance when the model is simplified. This is a likelihood ratio test. The expected increase in deviance is asymptotically  $\chi^2$  distributed with degrees of freedom equal to the number of parameters dropped from the model. A strict adherence to a  $P = 0.05$  level for rejecting the null hypothesis of no effect does not always produce the best model (Burnham et al., 1995). Akaike's Information Criterion ( $AIC = \text{residual deviance} + [2 \times \text{number of parameters}]$ ) provides a more reliable decision criterion for choosing to retain parameters in the model. The model with the lowest AIC value is the best compromise between bias from including poorly supported parameters in the model and residual error. AIC also provides a criterion for comparing two models that are not nested subsets of each other. No adjustment for multiple testing is needed when using AIC (Cooch et al., 1997). Model selection strategy is described in great detail in Burnham et al. (1987), Pollock et al. (1990), Lebreton et al. (1992) and Cooch et al. (1997).

Finally, survival and recapture probabilities are estimated using the model with the lowest AIC. Estimates of the probabilities are obtained using maximum likelihood methods (Lebreton et al., 1992). Estimates and their 95% profile confidence limits are presented. Note that groups (e.g. marking cohorts) may have overlapping confidence limits. Nevertheless, there may be a statistically significant difference as outlined in the section on model choice.

The analysis of survival probabilities is only the most basic use of CMR statistics. It is also possible to test the effects of age, sex, reproductive effort, color morph, and other biological or environmental parameters on survival and recapture probabilities (see Lebreton et al., 1992, for details). Furthermore, it is possible to study recruitment and population growth (Pradel, 1996), breeding probabilities (Clobert et al., 1994; Nichols et al., 1994), costs of reproduction (Nichols et al., 1994), natural selection (Kingsolver and Smith, 1995), temporary emigration from a population (Kendall et al., 1997), and to estimate transition probabilities for stage-based population matrices (Nichols et al., 1992).

#### *Analysis of the capture-mark-recapture data set of Kuhn (1994a)*

Recapture histories of adult female *Bufo bufo* were extracted from figures 4, 5 and 6 of Kuhn (1994a) and analysed using the program SURGE (Pradel and Lebreton, 1993; Cooch et al., 1997). The three populations were analysed separately because: 1) There are no a priori hypotheses about differences among the three populations that we wish to test. 2) One population was studied for three years while the others were studied for four years. 3) Analyses showed the data to be rather differently structured among populations. 4) Sample sizes are large, so power of statistical tests is relatively high.

For each population the models tested included both time and marking cohort effects (denoted with subscripts  $t$  and  $c$ , respectively) on the recapture probability in any year and the probability of survival (denoted  $p$  and  $\phi$ , respectively) from one year to the next.  $c$  denotes the number of years since marking. Marking cohorts are correlated with age after maturity.

**Population A.** Between 1989 and 1991, 2385 adult females were marked; 329 animals were recaptured at least once. The combined Tests 2 and 3 of program RELEASE showed that the data did not fit the CJS model with time dependent survival ( $\phi_t$ ) and recapture probabilities ( $p_t$ ) ( $\chi^2 = 32.6$ ; d.f. = 4;  $P < 0.001$ ). Different survival probabilities for animals of different ages, or lower survival probabilities immediately after marking are candidates for this lack of fit (Cooch et al., 1997).

The fully saturated model for survival ( $\phi_{c*t}$ ; separate survival probability parameters for each cohort by year combination) with separate recapture probabilities for each year ( $p_t$ ) (model 1 in table 2) was slightly more variable than we might expect from the pure binomial probabilities of survival and recapture ( $\chi^2 = 2507$ ; d.f. = 2378;  $P = 0.03$ ). However, the deviance is only 5% higher than expected and does not affect any conclusions we draw from the data.

The probability of survival did not vary between years in this population (Model 2 vs. Model 1;  $\chi^2 = 2.8$ ; d.f. = 2;  $P = 0.25$ ). Reducing the number of estimated parameters for survival to two with one parameter for survival the winter immediately after marking and the second for all subsequent winters, significantly increased the deviance (Model 3 vs. Model 2;  $\chi^2 = 9.7$ ; d.f. = 1;  $P = 0.002$ ). Having the initial cohort last two years produced an even worse fit. Thus, three cohort parameters were retained. Removing the time dependence for the probability of recapture also significantly increased the deviance (Model 4 vs. Model 2;  $\chi^2 = 40.7$ ; d.f. = 2;  $P < 0.0001$ ). However, this is almost entirely due to the final year having a dramatically lower

**Table 2.** Mark-recapture analysis of Population A. Table a) lists deviance, number of parameters (np), and Akaike’s Information Criterion (AIC) for models of different complexity. Survival ( $\phi$ ) is indexed by years since marking ( $c$ ) and year ( $t$ ). Recapture probability ( $p$ ) is indexed by year ( $t$ ). The numbers following the indexes indicate the number of years over which each parameter applies. Table b) shows parameter estimates for the minimal model (highlighted in bold face in a) with their 95% profile confidence limits.

| a)           | Model                                | np       | Deviance        | AIC             |
|--------------|--------------------------------------|----------|-----------------|-----------------|
| 1)           | $\phi_{c*t}, p_t$                    | 9        | 2507.0          | 2523.0          |
| 2)           | $\phi_c, p_t$                        | 6        | 2509.8          | 2521.8          |
| 3)           | $\phi_{c(1,2)}, p_t$                 | 5        | 2519.5          | 2529.5          |
| 4)           | $\phi_c, p$                          | 4        | 2550.5          | 2558.5          |
| 5)           | <b><math>\phi_c, p_t(2,1)</math></b> | <b>5</b> | <b>2509.8</b>   | <b>2519.8</b>   |
| b) Parameter | Year(s)                              | Estimate | Lower<br>95% CL | Upper<br>95% CL |
| Survival     | 1st year                             | 31.4%    | 26.0%           | 37.4%           |
| Survival     | 2nd year                             | 73.6%    | 58.7%           | 84.5%           |
| Survival     | 3rd year                             | 26.3%    | 13.6%           | 44.7%           |
| Recapture    | 1990 & 1991                          | 38.8%    | 31.4%           | 46.8%           |
| Recapture    | 1992                                 | 17.9%    | 13.1%           | 24.0%           |

recapture probability than the first two years. Using one parameter for recapture in 1990 and 1991, and a second parameter for 1992 had almost no effect on model deviance (Model 5 vs. Model 2;  $\chi^2 = 0.5$ ; d.f. = 1;  $P = 0.48$ ). This model also had the lowest AIC value and, therefore, the minimum adequate model is one with three cohort parameters for survival and two time parameters for recapture. The maximum likelihood estimate of survival of cohorts was only 31.4% the year after marking, increased to more than 70% in the second year and fell again to 26% in the third year (table 2). The model could be further reduced by having one parameter for the first and third years, but it is hard to imagine a reasonable biological explanation for the similar survival probabilities.

*Population B.* There was no marking effort at this population in 1992, but in the previous two years 422 animals were marked, of which 69 were recaptured at least once. Program RELEASE revealed no significant deviations from the CJS model. This suggests that there are no advantages to pursuing a model for separate marking cohorts. Indeed, the deviance of the model ( $\phi_t, p_t$ ) (model 1 in table 3) was very near to what we would expect from the pure binomial probabilities of survival and recapture ( $\chi^2 = 436$ ; d.f. = 419;  $P = 0.27$ ).

This population illustrates one of the problems of having only two recapture sessions. The probability of seeing an animal again must be time dependent in 1991, but because we do not know if an animal that was not seen died or was not recaptured, this time-dependence cannot be resolved. In the full model, the survival and recapture parameters cannot be estimated separately. Their product can be accurately estimated but they cannot be estimated independently (Lebreton et al., 1992). Thus, removing the time dependence of either survival or recapture does not affect the number of parameters being estimated and the deviance of the model. Dropping the time dependence of both parameters has a strong effect (Model 4 vs. Model 1, 2, or 3;  $\chi^2 = 16.1$ ; d.f. = 1;  $P < 0.0001$ ). We have assumed constant recapture probabilities for the estimation of survival probabilities.

*Population C.* Between 1989 and 1991, 1277 adult female toads were marked and 273 animals were recaptured at least once. There were no significant deviations from the CJS model as tested by program RELEASE although the combined Tests 2 and 3 were near the nominal rejection level ( $P = 0.056$ ). This suggests the possibility that there may once again be marking cohort effects in the survival probabilities as seen in population A. The full model ( $\phi_{c*t}, p_t$ ) (model 1 in table 4) was somewhat more variable than expected with binomial probabilities ( $\chi^2 = 1647$ ; d.f. = 1270;  $P < 0.001$ ) but the variance inflation factor is considerably less than 2 so the estimates should be unbiased. We have not adjusted probability estimates because there are no borderline cases of significance.

**Table 3.** Mark-recapture analysis of Population B. Table a) lists deviance, number of parameters (np), and Akaike's Information Criterion (AIC) for models of different complexity. Both survival ( $\phi$ ) and recapture probability ( $p$ ) are indexed by year ( $t$ ). The reasons for choosing Model 2 as the minimal model are clarified in the text. Table b) lists parameter estimates for the minimal model (highlighted in bold face in a)) with their 95% profile confidence limits.

| a)           | Model                         | np       | Deviance        | AIC             |
|--------------|-------------------------------|----------|-----------------|-----------------|
| 1)           | $\phi_t, p_t$                 | 3        | 436.5           | 442.5           |
| 2)           | <b><math>\phi_t, p</math></b> | <b>3</b> | <b>436.5</b>    | <b>442.5</b>    |
| 3)           | $\phi, p_t$                   | 3        | 436.5           | 442.5           |
| 4)           | $\phi, p$                     | 2        | 452.6           | 456.6           |
| b) Parameter | Year(s)                       | Estimate | Lower<br>95% CL | Upper<br>95% CL |
| Survival     | 1989 to 1990                  | 55.8%    | 25.7%           | 82.2%           |
| Survival     | 1990 to 1991                  | 20.0%    | 11.3%           | 32.8%           |
| Recapture    | 1990 & 1991                   | 35.1%    | 17.7%           | 57.6%           |

**Table 4.** Mark-recapture analysis of Population C. Table a) lists deviance, number of parameters (np), and Akaike's Information Criterion (AIC) for models of different complexity. Survival ( $\phi$ ) is indexed by years since marking ( $c$ ) and year ( $t$ ). Recapture probability ( $p$ ) is indexed by year ( $t$ ). The numbers following the indexes indicate the number of years over which each parameter applies. Table b) lists parameter estimates for the minimal model (highlighted in bold face in a)) with their 95% profile confidence limits.

| a)           | Model                                  | np       | Deviance        | AIC             |
|--------------|----------------------------------------|----------|-----------------|-----------------|
| 1)           | $\phi_{c*t}, p_t$                      | 9        | 1647.9          | 1665.9          |
| 2)           | $\phi_c, p_t$                          | 6        | 1654.2          | 1666.2          |
| 3)           | $\phi_c, p$                            | 4        | 1691.2          | 1699.2          |
| 4)           | $\phi_{c(1,2)}, p_t$                   | 5        | 1658.9          | 1668.9          |
| 5)           | <b><math>\phi_c, p_{t(1,2)}</math></b> | <b>5</b> | <b>1654.2</b>   | <b>1664.2</b>   |
| b) Parameter | Year(s)                                | Estimate | Lower<br>95% CL | Upper<br>95% CL |
| Survival     | 1st year                               | 37.2%    | 30.6%           | 44.5%           |
| Survival     | 2nd year                               | 50.4%    | 35.8%           | 64.9%           |
| Survival     | 3rd year                               | 20.2%    | 9.4%            | 38.1%           |
| Recapture    | 1990                                   | 67.7%    | 54.7%           | 78.5%           |
| Recapture    | 1991 & 1992                            | 33.5%    | 25.2%           | 42.9%           |

Dropping the time component of survival did not significantly increase the deviance (Model 2 vs. Model 1;  $\chi^2 = 6.3$ ; d.f. = 3;  $P = 0.1$ ). However, the recapture probability did vary substantially between years (Model 3 vs. Model 2;  $\chi^2 = 47.0$ ; d.f. = 2;  $P < 0.0001$ ). This was due, almost exclusively, to higher recapture probabilities in 1990. Using only one parameter instead of two separate parameters for recapture in 1991 and 1992 had no effect on the deviance (Model 5 vs. Model 2;  $\chi^2 = 0.05$ ; d.f. = 1;  $P = 0.8$ ). Reducing the number of age classes to model survival significantly increased the deviance (Model 4 vs. Model 2;  $\chi^2 = 4.7$ ; d.f. = 1;  $P = 0.03$ ). The minimal model has three cohorts for survival and they show the same humped pattern as in Population A: low survival in the first year, increasing in the second, and declining again in the third (table 4).

*The analysis of breeding patterns*

The recapture histories can also be used to analyse the breeding pattern of the species under study. We follow the methods described by Gittins (1983), Gill (1985, 1987) and Nichols et al. (1987) to answer the question whether female *Bufo bufo* breed annually or biannually, i.e. more generally, whether they skip opportunities for reproduction. The analysis tests whether recapture histories of the type “1010” or “101” occur more frequently than expected. This is analogous to flipping a coin four times. With a fair coin we expect that heads will appear, on average, twice. Occasionally, heads and tails will alternate. If we do many such coin-flipping experiments we can ask whether alternation of heads and tails occurs more (or less) often than expected by chance.

Only toads captured in the first year of the study (1989) are included in this analysis. Survival and recapture probabilities used in this analysis were calculated from the entire data set and then applied to the cohort marked in 1989, a subset of the data. Expected and observed breeding patterns match very well (table 5). There is no statistical support for significant deviations of observed from expected breeding patterns (population A:  $\chi^2 = 6.76$ , d.f. = 7,  $P = 0.4540$ ; population B:  $\chi^2 = 1.38$ , d.f. = 3,  $P = 0.7105$ ; population C:  $\chi^2 = 7.71$ , d.f. = 7,  $P = 0.2586$  [table-wide  $\chi^2$  values]).

**Table 5.** Observed and expected frequencies of breeding patterns (1 = captured, 0 = not captured) in female *Bufo bufo* from populations A (table a), B (table b), and C (table c) of Kuhn (1994a) during the breeding seasons 1989-1992 in populations A and C and breeding seasons 1989-1991 in population B. Only females first captured in 1989 are included in this analysis. N = number of toads captured in 1989 (year  $i = 1$ ),  $\phi_i$  = probability of survival from year  $i$  to year  $i + 1$ ,  $p_i$  = probability of recapture in year  $i$ . The models were taken from Lebreton et al. (1992). The capture histories 101 and 1010 (highlighted in bold face) are predicted to be significantly larger than the expected values if toads breed at two-year intervals.

| a) Breeding pattern | Model                                                              | Expected    | Observed  |
|---------------------|--------------------------------------------------------------------|-------------|-----------|
| 1111                | $N \phi_1 p_2 \phi_2 p_3 \phi_3 p_4$                               | 1.8         | 5         |
| 1011                | $N \phi_1 (1 - p_2) \phi_2 p_3 \phi_3 p_4$                         | 2.9         | 3         |
| 1101                | $N \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 p_4$                         | 2.9         | 3         |
| 1001                | $N \phi_1 (1 - p_2) \phi_2 (1 - p_3) \phi_3 p_4$                   | 4.6         | 1         |
| <b>1010</b>         | <b><math>N \phi_1 (1 - p_2) \phi_2 p_3 (1 - \phi_3 p_4)</math></b> | <b>58.5</b> | <b>60</b> |
| 1110                | $N \phi_1 p_2 \phi_2 p_3 (1 - \phi_3 p_4)$                         | 37.1        | 31        |
| 1100                | $N \phi_1 p_2 [1 - \phi_2 p_3 - \phi_2 (1 - p_3) \phi_3 p_4]$      | 94.4        | 95        |
| 1000                | $N [1 - (\text{the 7 terms above})]$                               | 915.9       | 920       |
| Total               |                                                                    | 1118.1      | 1118      |
| b) Breeding pattern | Model                                                              | Expected    | Observed  |
| 111                 | $N \phi_1 p_2 \phi_2 p_3$                                          | 3.8         | 6         |
| 110                 | $N \phi_1 p_2 (1 - \phi_2 p_3)$                                    | 50.3        | 48        |
| <b>101</b>          | <b><math>N \phi_1 (1 - p_2) \phi_2 p_3</math></b>                  | <b>7.0</b>  | <b>7</b>  |
| 100                 | $N [1 - \phi_1 p_2 - \phi_1 (1 - p_2) \phi_2 p_3]$                 | 214.9       | 215       |
| Total               |                                                                    | 276.0       | 276       |
| c) Breeding pattern | Model                                                              | Expected    | Observed  |
| 1111                | $N \phi_1 p_2 \phi_2 p_3 \phi_3 p_4$                               | 1.8         | 4         |
| 1011                | $N \phi_1 (1 - p_2) \phi_2 p_3 \phi_3 p_4$                         | 0.8         | 2         |
| 1101                | $N \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 p_4$                         | 3.5         | 2         |
| 1001                | $N \phi_1 (1 - p_2) \phi_2 (1 - p_3) \phi_3 p_4$                   | 1.7         | 0         |
| <b>1010</b>         | <b><math>N \phi_1 (1 - p_2) \phi_2 p_3 (1 - \phi_3 p_4)</math></b> | <b>11.5</b> | <b>15</b> |
| 1110                | $N \phi_1 p_2 \phi_2 p_3 (1 - \phi_3 p_4)$                         | 24.1        | 23        |
| 1100                | $N \phi_1 p_2 [1 - \phi_2 p_3 - \phi_2 (1 - p_3) \phi_3 p_4]$      | 124.0       | 129       |
| 1000                | $N [1 - (\text{the 7 terms above})]$                               | 442.5       | 434       |
| Total               |                                                                    | 609.9       | 609       |

### Discussion

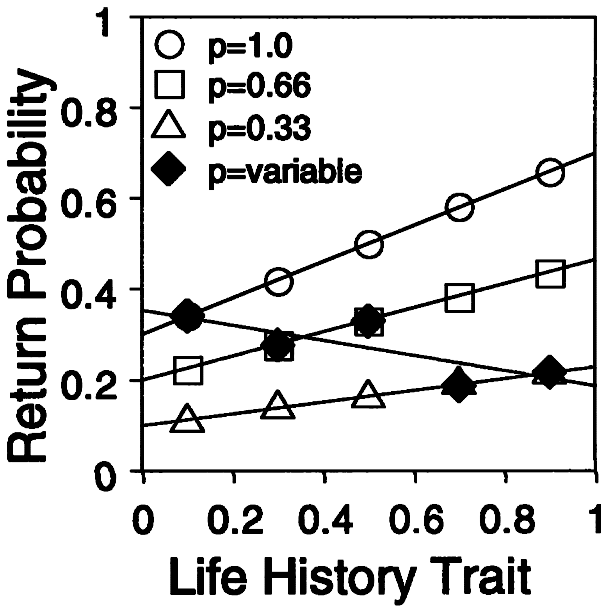
*Survival probabilities of female common toads.* Our analysis of Kuhn's (1994a) data showed that female common toads have an annual survival rate of up to 74% which is substantially higher than the 5-25% suggested by Kuhn (1994a). Survival probabilities varied greatly among marking cohorts (tables 2, 4). In populations A and C (but not in population B) overwinter survival was highest the second winter after marking (A: 73.6%, C: 50.4%) and considerably lower the winter immediately following marking (A: 31.4%, C: 37.2%) or three years after marking (A: 26.3%, B: 20.2%). Kuhn (1994a) noted that return probabilities were higher for females that had reproduced for the second time (during the study) than for females that had reproduced for the first time. Kuhn (1994a) suggested that this may be due to differences in quality among females. Tests 2 and 3 of program RELEASE showed that animals marked at different times did not have the same probability of being resighted. Different survival probabilities for animals of different ages, or lower survival probabilities immediately after marking are candidates for this variation in resighting probabilities. We cannot rule out other factors, but toe clipping, the marking technique used by Kuhn (1994a), is one possible explanation for low survival probabilities the first year after marking. However, even if toe clipping contributed to decreased survival probabilities the year after marking, it cannot explain why survival probabilities drop again three years after marking. Time since marking is correlated with age. Thus, highest survival in the second winter after marking may reflect lower survival of younger and older animals.

*The importance of recapture probabilities.* Our analysis of recapture probabilities revealed three unexpected results: recapture probabilities are generally low, variable among years, and they drop in later years of the study (tables 2, 4). These results are not what one would expect from a capture-mark-recapture study when using a drift fence. A drift fence should allow an almost complete census of the population of a non-climbing species. However, many authors (e.g., Gill, 1985, 1987; Nichols et al., 1987; Dodd, 1991; Thonke et al., 1994; Arntzen et al., 1995; Semlitsch et al., 1996) have documented or assumed that a drift fence could not sample the entire population. Here, we wish to stress that recapture probabilities may not only be  $< 1$ , but strongly so and variable among years.

Even though Kuhn (1994a) used the same method in all years, there are many possible reasons for low and variable recapture probabilities (see also Dodd and Scott, 1994): a) trespassing (Gill, 1985, 1987), b) a variable quality of the fence rebuilt every year (Arntzen et al., 1995), c) unequal catchability of e.g. size classes among years (van Gelder and Rijdsdijk, 1987), d) avoidance of the fence, e) lack of breeding pond fidelity (Schlupp and Podloucky, 1994), f) migration to the pond in autumn (toads begin to migrate in autumn but apparently they do not enter the pond in autumn; Sinsch, 1988), g) predation on toads at the fence in later years (Reading, 1989), h) regeneration of clipped toes (Kuhn, 1994b), i) irregular breeding intervals (Gill, 1985) or a lower frequency of breeding later in the study, or j) effects of weather conditions on breeding migrations and the size of the breeding population (Semlitsch et al., 1996). Some of these possibilities (e.g., a) assume that the fence has a low capture efficiency whereas others apply even if the fence captures every breeding individual (e.g., i). Testing and reporting the efficiency of the drift fence (note that the recapture probability is not a measure of the quality of the drift fence) may exclude some listed explanations for the low recapture probabilities. The low, variable and declining recapture probabilities suggest new interpretations for some patterns reported by Kuhn (1994a).

Kuhn (1994a) could not find factors such as age, size or reproductive effort that influenced return probabilities. A hypothetical example illustrates how low and variable recapture probabilities can mask real patterns in the data (fig. 1; see also Martin et al., 1995). First, consider a relationship between survival and a life history trait (e.g., body size, age, reproductive effort) which is described by a regression with slope = 0.4 and intercept of 0.3. If the recapture probability ( $p$ ) = 1.0 then the return probability can be used to describe this relationship (open circles in fig. 1). If the recapture probability is  $< 1.0$  then the slope of the regression is smaller. The relationship becomes undetectable when it is weak or recapture probabilities are low (fig. 1). If the recapture probability is not independent of the life-history character, the relationship between return probabilities and the life-history character can be opposite in sign from that of survival and the life-history character. For example, if younger animals have higher recapture probabilities, return probabilities may decline with age even though survival probability increases with age. Similar effects of the confounding of recapture and survival probabilities in return probabilities can occur for comparisons among years, sexes, habitats, or





**Figure 1.** Relationship between a life-history trait and return probabilities for a range of recapture probabilities ( $p$ ). Lower recapture probabilities lower the slope of the relationship reducing statistical power. Heterogeneous recapture probabilities can lead to slopes of the wrong sign.

any other biologically relevant comparison. Separate estimates of recapture and survival probabilities for each group overcome this confounding and allow hypothesis testing for differences among groups.

Kuhn (1994a) reported that two of his three populations were declining during the course of the study. We have shown that recapture probabilities were significantly lower in later years of the study. This results in fewer and fewer toads being seen and, as a consequence, an apparent decline in population size. Notably, population C, which was not thought to be declining, is the population with the highest recapture probabilities. The decline may not be real but an effect of recapture probabilities that are lower in later years of the study. Again, the interpretation of the pattern observed depends on the causation of the low recapture probabilities. If every toad arriving at the fence was captured the decline in recapture probabilities might reflect reduced frequency of breeding later in the study. This is a possibility that could only be detected by separating the survival probabilities from the recapture probabilities using CMR statistics.

Kuhn (1994a) suggested that a proportion of toads in the populations he studied did not reproduce every year but at intervals of 2 or 3 years. There are many examples of amphibians skipping opportunities for reproduction (Bull and Shine, 1979) including common toads (see Kuhn, 1994a). However, it must be tested whether the data really exhibit the hypothesized pattern (Gill, 1985). A capture history of the type captured-not captured-captured (101) must occur more often than expected by chance. There is no evidence in the capture histories that female toads were less likely to breed the year following a year when they did breed (table 5). It may be that the toads did not reproduce every year. If all or most (say > 95%) toads arriving at the drift fence were captured then the low recapture probabilities support the view that a large fraction of the population did not breed every year. Although the toads may not breed every year, the capture histories suggest that the probability of breeding is independent of whether they bred in the previous year.

*Conclusion*

Statistical models are a description of reality. One data set can be described by more than one statistical model. Some models, though, are better than others and we can and should search for the better ones (as outlined by

McCullagh and Nelder, 1983; Lebreton et al., 1992; Hilborn and Mangel, 1997). A general lack of statistical analysis can yield misleading results and consequently misleading conclusions. The methodology and software to analyse capture-mark-recapture data is available and well developed (e.g., Lebreton et al., 1992). A powerful analysis will not only help to avoid erroneous conclusions, it may also uncover patterns in the data that could not be seen by simple inspection. Declining recapture probabilities and cohort-dependent survival probabilities were uncovered by rigorous analysis of a remarkable data set. These results pose interesting questions. For example, the lower recapture probabilities later in the study suggest a reduced frequency of breeding. Why the toads might be breeding less frequently is an unanswered question with potentially profound conservation implications.

**Acknowledgements.** We thank J. Kuhn, G. Gollmann, U. Sinsch, and anonymous referees for discussion and comments on the manuscript, and E. Cooch for comments on the analysis. An overview of software that can be used for population analysis, including SURGE, and a guide to mark-recapture analysis using SURGE (Cooch et al., 1997) is available at <http://mendel.mbb.sfu.ca/cmr/> (address was working 9 August 1998).

## References

- Arntzen, J.W., Oldham, R.S., Latham, D.M. (1995): Cost effective drift fences for toads and newts. *Amphibia-Reptilia* **16**: 137-145.
- Baron, J.-P., Ferrière, R., Clobert, J., Saint Girons, H. (1996): Stratégie démographique de *Vipera ursinii ursinii* au Mont Ventoux (France). *C.R. Acad. Sci. Paris, Sci. Vie* **319**: 57-69.
- Bull, J.J., Shine, R. (1979): Iteroparous animals that skip opportunities for reproduction. *Am. Nat.* **114**: 296-303.
- 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. *Am. Fish. Soc. Monogr.* **5**.
- Burnham, K.P., White, G.C., Anderson, D.R. (1995). Model selection strategy in the analysis of capture-recapture data. *Biometrics* **51**: 888-898.
- Clobert, J., Lebreton, J.-D., Allainé, D., Gaillard, J.M. (1994): The estimation of age-specific breeding probabilities from recaptures or resightings in vertebrate populations: II. Longitudinal models. *Biometrics* **50**: 375-387.
- Cooch, E.G., Pradel, R., Nur, N. (1997): A Practical Guide to Mark-Recapture Analysis Using SURGE — 2nd Edition. Montpellier, Centre d'Ecologie Fonctionnelle et Evolutive — CNRS.
- Crouse, D.T., Crowder, L.B., Caswell, H. (1987): A stage-based population model for loggerhead sea turtles and implications for conservation. *Ecology* **68**: 1412-1423.
- Doak, D., Kareiva, P., Klepetka, B. (1994): Modelling population viability for the desert tortoise in the western Mojave desert. *Ecol. Appl.* **4**: 446-460.
- Dodd, C.K., Jr. (1991): Drift fence associated sampling bias of amphibians at a Florida sandhills temporary pond. *J. Herpetol.* **25**: 296-301.
- Dodd, C.K., Jr., Scott, D.E. (1994): Drift fences encircling breeding sites. In: *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*, p. 125-130. Heyer, W.R., Donnelly, M.A., McDiarmid, R.W., Hayek, L.-A.C., Foster, M.A., Eds., Washington, Smithsonian Institution Press.
- Elmberg, J. (1990): Long-term survival, length of breeding season, and operational sex ratio in a boreal population of common frogs, *Rana temporaria* L. *Can. J. Zool.* **68**: 121-127.
- Flatt, T., Dummermuth, S., Anholt, B.R. (1997): Mark-recapture estimates of survival in populations of the Asp viper, *Vipera aspis aspis*. *J. Herpetol.* **31**: 558-564.
- Forsman, A. (1995): Opposing fitness consequences of colour pattern in male and female snakes. *J. Evol. Biol.* **8**: 53-70.
- Gelder, J.J. van, Rijdsdijk, G. (1987): Unequal catchability of male *Bufo bufo* within breeding populations. *Hol. Ecol.* **10**: 90-94.

- Gill, D.E. (1985): Interpreting breeding patterns from census data: a solution to the Husting dilemma. *Ecology* **66**: 344-354.
- Gill, D.E. (1987): Reply to Nichols et al. *Ecology* **68**: 217-220.
- Gittins, S.P. (1983): Population dynamics of the common toad (*Bufo bufo*) at a lake in mid-Wales. *J. Anim. Ecol.* **52**: 981-988.
- Halley, J.M., Oldham, R.S., Arntzen, J.W. (1996): Predicting the persistence of amphibian populations with the help of a spatial model. *J. Appl. Ecol.* **33**: 455-470.
- Heusser, H. (1960): Über die Beziehungen der Erdkröte (*Bufo bufo* L.) zu ihrem Laichplatz. II. *Behaviour* **16**: 93-109.
- Heusser, H. (1969): Die Lebensweise der Erdkröte, *Bufo bufo* (L.); Das Orientierungsproblem. *Rev. Suisse Zool.* **76**: 443-518.
- Hilborn, R., Mangel, M. (1997): *The Ecological Detective: Confronting Models with Data*. Princeton, Princeton University Press.
- Jehle, R., Hödl, W., Thonke, A. (1995): Structure and dynamics of central European amphibian populations: a comparison between *Triturus dobrogicus* (Amphibia, Urodela) and *Pelobates fuscus* (Amphibia, Anura). *Austr. J. Ecol.* **20**: 362-366.
- Kapfberger, D. (1984): Untersuchungen zu Populationsaufbau, Wachstum und Ortsbeziehungen der Gelbbauchunke, *Bombina variegata variegata* (Linnaeus, 1758). *Zool. Anz.* **212**: 105-116.
- Kendall, W.L., Nichols, J.D., Hines, J.E. (1997): Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* **78**: 563-578.
- Kingsolver, J.G., Smith, S.G. (1995): Estimating selection on quantitative traits using capture-recapture data. *Evolution* **49**: 384-388.
- Kuhn, J. (1994a): Lebensgeschichte und Demographie von Erdkrötenweibchen *Bufo bufo bufo* (L.). *Z. Feldherpetol.* **1**: 3-87.
- Kuhn, J. (1994b): Methoden der Anuren-Markierung für Freilandstudien: Übersicht — Knie-Ringetiketten — Erfahrungen mit der Phalangenamputation. *Z. Feldherpetol.* **1**: 177-192.
- 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. *Ecol. Monogr.* **62**: 67-118.
- Lebreton, J.-D., Pradel, R., Clobert, J. (1993): The statistical analysis of survival in animal populations. *Trends Ecol. Evol.* **8**: 91-95.
- Martin, T.E., Clobert, J., Anderson, D.R. (1995): Return rates in studies of life history evolution: are biases large? *J. Appl. Stat.* **22**: 863-875.
- McCullagh, P., Nelder, J.A. (1983): *Generalized Linear Models*. New York, Chapman and Hall.
- Nichols, J.D., Hepp, G.R., Pollock, K.H., Hines, J.E. (1987): The Husting dilemma: a methodological note. *Ecology* **68**: 213-217.
- Nichols, J.D., Hines, J.E., Pollock, K.H., Hinz, R.L., Link, W.A. (1994): Estimating breeding proportions and testing hypotheses about costs of reproduction with capture-recapture data. *Ecology* **75**: 2052-2065.
- Nichols, J.D., Sauer, J.R., Pollock, K.H., Hestbeck, J.B. (1992): Estimating transition probabilities for stage-based population projection matrices using capture-recapture data. *Ecology* **73**: 306-312.
- Pollock, K.H., Nichols, J.D., Brownie, C., Hines, J.E. (1990): Statistical inference for capture-recapture experiments. *Wildl. Monogr.* **107**: 1-97.
- Pradel, R. (1996): Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* **52**: 703-709.
- Pradel, R., Lebreton, J.-D. (1993): *User's Manual for Program SURGE 4.2*. Montpellier, Centre d'Ecologie Fonctionnelle et Evolutive — CNRS.
- Reading, C.J. (1989): Opportunistic predation of common toads *Bufo bufo* at a drift fence in southern England. In: *Amphibians and Roads*, p. 105-111. Langton, T.E.S., Ed., Sheffield, ACO Polymer Products.
- Schlupp, I., Podlousky, R. (1994): Changes in breeding site fidelity: a combined study of conservation and behaviour in the common toad *Bufo bufo*. *Biol. Cons.* **69**: 285-291.
- Seidel, B. (1993): Bericht aus einer seit 1984 laufenden Studie über eine Gelbbauchunkenpopulation *Bombina variegata*: Ein Diskussionsansatz für feldherpetologische Studien. *Salamandra* **29**: 6-15.

- Semlitsch, R.D., Scott, D.E., Pechmann, J.H.K., Gibbons, J.W. (1996): Structure and dynamics of an amphibian community: evidence from a 16-year study of a natural pond. In: Long-term Studies of Vertebrate Communities, p. 217-248. Cody, M.L., Smallwood, J.A., Eds., San Diego, Academic Press.
- Shine, R., Charnov, E.L. (1992): Patterns of survival, growth, and maturation in snakes and lizards. *Am. Nat.* **139**: 1257-1269.
- Sinsch, U. (1988): Seasonal changes in the migratory behaviour of the toad *Bufo bufo*: direction and magnitude of movements. *Oecologia* **76**: 390-398.
- Thonke, A., Jehle, R., Hödl, W. (1994): Structure, dynamics and phenology of a population of the Danube warty newt (*Triturus dobrogicus*) on the Danube island near Vienna — a preliminary report. *Abh. Ber. Mus. Naturk. Magdeburg* **17**: 127-133.

*Received: October 3, 1997. Accepted: January 11, 1998.*

**Editorial note:** Upon acceptance of this manuscript Dr. Joachim Kuhn was invited to comment on this analysis of his results.