

# **Statistical Methods for a Study of the Maud Island Frog (*Leiopelma pakeka*)**

**Shirley Pledger<sup>1</sup> and Ben Bell<sup>2</sup>**

**May 2008**

<sup>1</sup>School of Mathematics, Statistics and Computer Science,

<sup>2</sup> School of Biological Sciences,

Victoria University of Wellington,

P.O. Box 600, Wellington 6140, New Zealand.

Email: <sup>1</sup>shirley.pledger@mcs.vuw.ac.nz

Email: <sup>2</sup>ben.bell@vuw.ac.nz



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Robust Design Temporary Emigration Model</b>	<b>1</b>
<b>3</b>	<b>Comparing Groups</b>	<b>3</b>
<b>4</b>	<b>Parameters Constant Through Time</b>	<b>3</b>
<b>5</b>	<b>Weather as a Covariate</b>	<b>4</b>
<b>6</b>	<b>Behavioural Response Models</b>	<b>4</b>
<b>7</b>	<b>Combined Models</b>	<b>6</b>
<b>8</b>	<b>Jolly-Seber Model</b>	<b>7</b>
<b>9</b>	<b>Closed Population Models</b>	<b>7</b>
<b>10</b>	<b>Testing the Time Scale of Temporary Emigration</b>	<b>8</b>
<b>11</b>	<b>Impact of Toe-clipping on Return Rate</b>	<b>9</b>
<b>12</b>	<b>Summary</b>	<b>10</b>



# 1 Introduction

This study on Maud Island in the Marlborough Sounds, New Zealand, extends over 25 years. The second author visited the island at least annually, on each trip spending several nights capturing, marking, measuring, releasing and recapturing the Maud Island frog *Leiopelma pakeka* on two  $12 \times 12$  m grids. The frogs are small (12–51 mm long), nocturnal and long-lived (up to 35 years). They occupy retreat sites under rocks and logs, emerging occasionally to feed (particularly on moist or humid nights).

As part of the analysis, estimates of population density and survival rates are required, with a check for changes over time. Comparisons of these parameters between grids and between sexes are also needed, together with estimates of the proportion underground (and hence unavailable for capture) at any one time.

The influence of the marking method (toe clipping) on survival and capture probabilities is also to be appraised.

The capture-recapture models include robust design models with temporary emigration (Sections 2 to 7), simple design Jolly-Seber models (Section 8) and closed-population models (Section 9). The other statistical methods are a customised permutation test (Section 10) and generalised linear models (Section 11).

## 2 The Robust Design Temporary Emigration Model

Open population temporary emigration capture-recapture models are the most appropriate for this study and these animals. This capture-recapture study has Pollock's robust design (Pollock 1982, Kendall *et al.* 1995), in which the population is open between the primary samples (the trips to the island), but is assumed to be closed during the secondary samples (the nights within each trip). The assumption of a closed population within each trip is reasonable for such a long-lived species, where births and deaths are rare on a short time scale. Migration is also thought to be extremely low with this species.

One important advantage of the robust design is that it allows for the fitting of temporary emigration models (Kendall *et al.* 1997). Another advantage of the robust design is the possibility of modelling and estimating short-term (within-trip) behavioural response, such as trap-shyness after handling and marking. We propose a new model in Section 6 which allows for two different types of trap-shyness.

Justification for the assumption of individual homogeneity, the evaluation of duration underground, and estimating the possible effects of toe clipping all require special methods, given later.

Our **central model**, detailed in Kendall *et al.* (1997), is based on  $M_t^t$  in Kendall *et al.* (1995), which allows the probability of capture to vary over time, both within and between trips. Extra parameters are included to allow for temporary emigration below ground. If retreat under rocks is occurring, on each trip there will be a proportion of frogs unavailable for capture. This leads to a discrepancy between population size estimates based on open populations, and those based on a series of closed populations. The open population estimates are higher, as frogs captured both before and after a particular trip are deemed to be present during that trip, even if underground, while the closed population model for that trip is estimating only the number currently on the surface and available for at least one capture that trip. This discrepancy in abundance estimates is used to estimate the proportion  $\gamma_j$  temporarily out of the (surface) study area during trip  $j$ . We assume random temporary emigration, with each animal's probability of being in retreat during trip  $j$  being independent of its retreat probability at other trips, a reasonable assumption because of the long time between trips. The model for a study with  $K$  trips has the following basic parameters:

- $N_j$  = population size, including those underground, at trip  $j$  ( $j = 1, 2, \dots, K$ )
- $\phi_j$  = survival rate from trip  $j$  to  $j + 1$  ( $j = 1, 2, \dots, K - 1$ )
- $\gamma_j$  = probability of temporary emigration (underground or elsewhere) for the duration of trip  $j$  ( $j = 2, 3, \dots, K - 1$ )
- $p_{j\ell}^*$  = probability of capture on night  $\ell$  of trip  $j$  for a frog on the surface.

Derived parameters include:

- $p_j$  = probability of at least one capture during trip  $j$  for an animal alive at  $j$   
 =  $(1 - \gamma_j)(1 - \prod(1 - p_{j\ell}^*))$  where the product  $\prod$  is taken over the nights  $\ell$  in trip  $j$
- $V_j$  = surface population on trip  $j$  (number available for at least one capture)  
 =  $(1 - \gamma_j)N_j$ .

Model fitting is done by maximum likelihood, with the equations given in Kendall *et al.* 1997. This means that comparisons by likelihood ratio tests are feasible, as well as model comparison by Akaike's Information Criterion (AIC, see e.g. Burnham and Anderson 2002).

This model assumes that temporary emigration occurs on a time scale of at least a few days underground, rather than a short phase of one or two days underground alternating with one or two days above ground. A permutation test to verify this time scale is given in Section 10.

There is a possibility that the probability of capture on any given night is largely explained by the weather conditions, with emergence more likely on moist or rainy nights. Models using weather as a covariate are in Section 5.

Our central model also assumes homogeneity of capture probability among individual animals. To justify this assumption, we used the within-trip data and closed population capture-recapture models, as detailed in Section 9.

Although we have started with a central model which excludes behavioural effects (capture shyness or capture happiness), later models check for and model such effects, including the introduction of a new model with two types of behavioural response (Section 6).

### 3 Comparing Groups

There were four basic data sets, from a cross-classification by grid and apparent sex, and four combined data sets (all apparent females, all apparent males, all grid 1 frogs and all grid 2 frogs). Different groups were compared using joint likelihoods. Survival, temporary emigration and capture parameters may differ by group as well as by time; this model is labelled  $\{\phi(tg), \gamma(tg), p^*(tg)\}$ . The label  $\phi(tg)$ , for example, specifies that each group  $g$  has its own vector of survival probabilities  $\phi_{jg}$  over time (between trips  $j$  and  $j + 1$ ). If the groups have similar survival probabilities, there is only one vector  $\phi_j$  applicable to all groups, and the model is labelled  $\phi(t)$ . Similar labelling is used for  $p$  and  $\gamma$ .

In comparing two groups, there are eight possible models, depending on whether survival, capture and/or temporary emigration match or differ between the groups. The eight models are shown in Table 1.

For the male/female groups, and then for the grouping by grid, all eight models were fitted, and AIC was used to select the best model. Likelihood ratio  $\chi^2$  tests could be used instead.

### 4 Parameters Constant Through Time

For a single group, any of the parameters could be constrained to be constant through time, replacing annual  $\phi(t)$  by annual  $\phi(c)$ ,  $\gamma(t)$  by  $\gamma(c)$ , and/or  $p^*(t)$  by  $p^*(c)$ . (Annual  $\phi$  is  $\phi_j^{365/d_j}$ , with  $d_j$  = number of days between trips  $j$  and  $j + 1$ .) These simplified models may be compared with the central model for any of the

Model	Survival	Temporary Emigration	Capture
a.	$\phi(t)$	$\gamma(t)$	$p^*(t)$
b.	$\phi(t)$	$\gamma(t)$	$p^*(tg)$
c.	$\phi(tg)$	$\gamma(t)$	$p^*(t)$
d.	$\phi(t)$	$\gamma(tg)$	$p^*(t)$
e.	$\phi(tg)$	$\gamma(t)$	$p^*(tg)$
f.	$\phi(t)$	$\gamma(tg)$	$p^*(tg)$
g.	$\phi(tg)$	$\gamma(tg)$	$p^*(t)$
h.	$\phi(tg)$	$\gamma(tg)$	$p^*(tg)$

Table 1: Possible models for groups of data, where  $\phi$ ,  $\gamma$  and  $p^*$  all vary over time, but may or may not vary by group.

data sets, to test if parameters are constant through time. Either AIC or likelihood ratio tests may be used for the comparisons. We fitted these models for each of the four basic data sets.

## 5 Weather as a Covariate

To see if  $p^*$  is mainly dependent on the weather (rainfall and moisture), categories were used to give just three distinct values of  $p^*$ . The weather covariate replaces the nightly time variation, greatly reducing the number of parameters. This model, labelled  $\{\phi(t), \gamma(t), p^*(w)\}$ , had three different time scales tried for weather  $w$ : rainfall during the preceding day ( $wd$ ), rainfall during the night of sampling ( $wn$ ), and a combined measure of rain either before or during sampling ( $wdn$ ).

A further model  $\{\phi(t), \gamma(w), p^*(t)\}$  was used to test if the temporary emigration rate could be explained by weather during the trip. For this model,  $\gamma$  took one of three values,  $\gamma_n$  for no rain that trip,  $\gamma_s$  for some rain and  $\gamma_h$  for heavy rain.

These weather models were fitted to each of the four basic data sets.

## 6 Behavioural Response Models

Because of the potential behavioural response to marking and handling, the central model was varied to allow for different short-term responses to the first capture per trip for a frog on the surface. Only temporary trap response can be modelled and



tested, as a permanent behavioural effect is confounded with survival.

We used three different models for  $p^*$ , the nightly probability of capture for surface animals.

**1. One type of response for all animals, labelled  $p^*(tb)$ :**

This is the behavioural model of Kendall *et al.* (1997), in which every animal, whether or not previously marked, responds to a first within-trip capture during trip  $j$  by moving to a new capture probability for the remainder of the nights during that trip. For this short-term behavioural response, there are two series of capture probabilities. The  $p^*$  series has different values each night, and measures the probability of capture for an animal not yet caught on this trip. The  $c^*$  series, used for within-trip recapture, has a different value for each night except the first night of each trip, where it is not relevant. Recently caught animals move onto the  $c^*$  series for the remainder of the trip, where  $c_{j\ell}^* < p_{j\ell}^*$  represents trap-shyness, a short-term avoidance of capture. We assume that all animals revert to the  $p^*$  series at the start of the next trip.

The closed population model with time and behavioural effects needs some extra assumptions to be fitted. We use the model in Pledger (2000), which is linear on the logit scale:

$$\log\left(\frac{p_{j\ell}^*}{1-p_{j\ell}^*}\right) = \alpha_{j\ell} \quad \text{and} \quad \log\left(\frac{c_{j\ell}^*}{1-c_{j\ell}^*}\right) = \alpha_{j\ell} + \beta,$$

where  $\beta$  is an adjustment factor. A negative estimate of  $\beta$  suggests trap-shyness, and the hypotheses  $H_0: \beta = 0$  versus  $H_A: \beta \neq 0$  (for trap response) or  $H_A: \beta < 0$  (for trap-shyness) may be tested to determine if there is a significant behavioural response. Alternatively, the model including  $\beta$  may be compared with the simpler model excluding  $\beta$ , using AIC to decide if the inclusion of  $\beta$  is justified.

**2. A trap response from newly-marked animals only, labelled  $p^*(tb1)$ :**

There are again two series,  $p_{j\ell}^*$  and  $c_{j\ell}^*$ , with the linear logistic connection, as in Case (1). However, in this model, only the newly toeclipped frogs move to the  $c$  series. This model assumes that only the initial capture, with the associated toe-clip, induces a behavioural response, and that previously toe-clipped animals are not substantially affected by the first within-trip capture on later trips.

In this case a test of  $H_0: \beta = 0$  versus  $H_A: \beta \neq 0$  determines if frogs respond to the first capture overall.

### 3. Two different levels of trap response, labelled $p^*(tb2)$ :

We also constructed a new model, with two different trap responses, depending on whether or not the frog had been captured and toe-clipped before the current trip. We now have two different adjustment parameters,  $\beta$  for the trap response of newly-marked animals, as in Case (2), and a different adjustment  $\delta$  for animals first seen on a previous trip. Animals marked before trip  $j$  start trip  $j$  with capture probability  $p^*$ , but after their first capture in trip  $j$ , move to a  $d^*$  series for the remainder of the trip, with linear logistic equation

$$\log\left(\frac{d_{j\ell}^*}{1 - d_{j\ell}^*}\right) = \alpha_{j\ell} + \delta.$$

The new parameter  $\delta$  measures the capture response of this group. Estimates of  $\beta < \delta < 0$  would support a stronger capture-shyness from the overall first capture with marking. If  $\beta$  and  $\delta$  are both negative, a test of  $H_0: \beta = \delta$  versus  $H_A: \beta < \delta$  determines if there is a more severe response to the initial capture, with its associated toe-clipping.

For each animal group, all three models may be fitted, and compared by AIC or likelihood ratio tests. We fitted these three models to each of our four basic data sets.

## 7 Combined Models

The behavioural models may be combined with other variations, to build and explore appropriate combinations of effects. For example,  $\{\phi(c), \gamma(w), p^*(tb2)\}$  represents the model with constant annual survival, with weather as a covariate for temporary emigration, and with two types of behavioural response in the nightly capture probabilities for frogs on the surface.

The model variations in Sections 4 to 7 were all fitted on each of the four basic groups. Since they are likelihood-based, these varied models could also be fitted on combined data sets with more than one group, using joint likelihoods. In this way, tests could be constructed to decide if any groups have similar survival, temporary emigration or capture probabilities, using any of these variations to the central model. For example, we could check if apparent males and females show a similar strength of trap response after the initial toe-clip ( $H_0: \beta_F = \beta_M$ ).

## 8 Jolly-Seber Model

After detailed investigations using the robust design models, we obtained an overview of the populations using the Jolly-Seber model (Jolly 1965, Seber 1965) for abundance ( $N_j$ ) and survival ( $\phi_j$ ) estimates over time. These match some of the parameters in the central model (Section 2). Within-trip information is pooled, and the abundance estimates are for all frogs, above or below ground.

This simplification represents just a condensation of the data, giving estimates of  $N_j$  and  $\phi_j$  which closely match those in the robust design models, but without the detailed option of modelling or estimating temporary emigration. The capture probability is now per trip, so estimation of weather or behavioural effects on nightly capture probability is no longer possible.

The advantage of the simpler model is that estimates and standard errors for  $N_j$  and  $\phi_j$  are more easily obtained, and model fitting is faster and more stable. The traditional Jolly-Seber model gives these estimates and standard errors (Seber 1973).

Tests for trends in population size are also possible, by using the fully likelihood-based version of the Jolly-Seber model proposed by Schwarz and Arnason (1996). We may construct the likelihoods to compare models with fluctuating  $N_j$  (the traditional Jolly-Seber, with no constraints on the  $N_j$  estimates) with models showing either a linear trend in  $N_j$  over time or constant  $N_j$  over time. Our methods of constructing these models are similar to those suggested in Schwarz (2001).

Any data set may have AIC comparisons, or likelihood ratio tests, among the three models, with fluctuating, linear or constant  $N_j$ .

The trend line or constant  $N$  may be fitted to just a subset of the data if desired. We chose to fit these models to just the last 15 years, as well as to the whole 25 year study. The 15-year models had fluctuating  $N_j$  for the first ten years, followed by a linear trend or constant  $N_j$ . This gave information on whether a new trend had set in over more recent years.

## 9 Closed Population Models

The within-trip closed population information enables us to check if time effects, behavioural effects or individual heterogeneity are affecting the probability of capture.

Otis *et al.* (1978) and Norris & Pollock (1996) give the development of likelihood-based closed population models in which the probability of capture may vary by time (different nights, Model  $M_t$ ), by capture response (capture shyness or capture happiness, Model  $M_b$ ) or by individual heterogeneity (intrinsic differences between animals, Model  $M_h$ ). Combined effects may also be modelled (Norris & Pollock 1996; Pledger 2000), with models  $M_{tb}$ ,  $M_{th}$ ,  $M_{bh}$  and  $M_{tbh}$ .

Since these models are all based on likelihoods, they may be compared by likelihood ratio tests or AIC, to select the most appropriate model for the population.

For each grid and for each closed-population model, we calculated the AIC from the joint likelihoods over all the trips (the sum of the separate AIC values from each trip). The models were compared to see whether the pooled information gave evidence of heterogeneity of capture, which would invalidate the temporary emigration models.

The closed population model  $M_t$  is also used in the development of a permutation test for the time scale of the temporary emigration behaviour, as detailed in the next section.

## 10 Testing the Time Scale of Temporary Emigration

The time scale of retreat-emergence behaviour was unknown. If individuals tend to spend a few days or longer underground before re-emergence, the temporary emigration models are appropriate. However the models are not sensible if the cycle is short, for example with alternate nights above and below ground. This would invalidate the temporary emigration model, which assumes the emigration lasts for the whole trip. We devised a permutation test to check whether short cycles were occurring.

In the case of short cycles, on pairs successive nights (within trips), there will be an excess of (0,1) or (1,0) values in the capture history for the individual, exceeding the number which are expected to occur with random retreat and emergence. Fewer changes than random would suggest a longer cycle time than random. We used a test statistic  $C$  = total number of changes between capture and non-capture on successive nights, totalled over the observed animals and all the trips.

We use permutations to generate the null distribution of  $C$ . If within one trip there is no drive towards a cycle which is shorter than random, the observed individuals could have their sequence of captures and non-captures in any order. Hence a

permutation of the columns of the within-trip capture matrix would give the total number of changes (0,1) or (1,0) in the absence of short-term cyclic behaviour. This permutation preserves the number of frogs seen, and the number of captures per frog - all that has occurred is a disruption of any time patterns.

### Algorithm

For a test of  $H_0$ : no short-term cyclic behaviour, versus  $H_A$ : there is short-term cycling, this algorithm is used:

1. Calculate  $C$ , the total number of changes between capture and non-capture on successive nights, for the data. Call this value  $C_{\text{data}}$ .
2. Choose a large  $R$ , number of replications.
3. For each replication:
  - (a) Within each trip, do a new random permutation of the columns of the capture matrix.
  - (b) Calculate the new total number of changes on successive nights (over all animals and all trips in the pseudo-data set).
  - (c) Save the number in a vector of  $R$  values of  $C$ , called (say)  $C_{\text{pseudo}}$ .
4. Find where  $C_{\text{data}}$  lies on the empirical distribution of pseudo- $C$  values. The proportion of  $C_{\text{pseudo}}$  values which are less than  $C_{\text{data}}$  is the p-value for the test.

See Manly (1997) for the general theory of permutation (randomization) tests.

## 11 Impact of Toe-clipping on Return Rate

The possible impact of toe-clipping on survival is of concern, and is difficult to evaluate directly when this has been the only method of marking available.

In this study, the numbers of toes clipped tended to increase over the course of the study as toe combinations involving fewer toes were used up. In retrospect, a mixture of numbers of toes clipped every year of the study might have been advisable, to avoid confounding with long-term survival.

In our analysis, we allowed for the frogs with more toes clipped having fewer opportunities for recapture by controlling for release cohort (trip when first captured). We used logistic regressions (generalized linear models with binomial data) to model return rate (probability of at least one recapture) with four covariates:

1. Release cohort (RC), to allow for the possible confounding of number of toes clipped with time remaining for opportunity for recapture (related to long-term survival estimates).
2. Location on the sampling site (LOC), whether central or peripheral (a factor at two levels). This allows for the lower recapture probability for peripheral animals, whose home ranges may be mainly off the grid.
3. The number of digits clipped (DIGITS), a discrete covariate, where we wish to test for a trend of decreased return rate with more digits clipped.
4. Snout-vent length in mm (SVL), a continuous covariate, included because return rates may vary by the size of frog.

In (3), two alternative toe-clip covariates were also appraised - number of feet with at least one toe removed (FEET), and number of digits clipped on the front feet alone (FRONT), since prey handling may be important for survival.

We used sequential analysis of deviance tables to evaluate the different predictors via  $\chi^2$  tests. One-sided Wald t-tests were used for the hypotheses of lower return rates with more toe-clips, with frogs located peripherally on the sites, and with smaller (younger) frogs more likely to disperse (pers. obs.).

## 12 Summary

The statistical methods in this report are all used in an article by Bell and Pledger (2008), in review. They provide a comprehensive analysis of a very large data set.

We note that the length of the study, over 25 years, has been invaluable in providing enough information to appraise and estimate the many effects on demographic parameters for these populations. With such a long-lived species, shorter studies could not provide these estimates. Long-term funding is essential for effective studies of such populations.

## References

- Burnham, K. P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag (2ed).

- Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration - stochastic model. *Biometrika* **52** 225–247.
- Kendall, W. L., Pollock, K. H. and Brownie, C. (1995). A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics* **51** 293–308.
- 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.
- Manly, B. F. J. (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology*. London: Chapman & Hall (2ed).
- Norris, J. L. III & Pollock, K. H. (1996). Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* **52** 639–649.
- Otis, D. L., Burnham, K. P., White, G. C. & Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62** 135 pages.
- Pledger, S. (2000). Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics* **56** 434–442.
- Pollock, K. H. (1982). A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management* **46** 757–760.
- Schwarz, C. J. (2001). The Jolly-Seber model: more than just abundance. *Journal of Agricultural, Biological and Environmental Statistics* **6** 195–205.
- Schwarz, C. J. & Arnason, A. N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* **52** 860–873.
- Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika* **52** 249–259
- Seber, G. A. F. (1973). *The Estimation of Animal Abundance and related Parameters* London: Griffiths.