

# Gary C. White

Professor Department of Fish, Wildlife and Conservation Biology



## Occupancy Estimation Robust Design

[Contents – Index](#)

### Occupancy Estimation Robust Design

The robust design occupancy estimation model (McKenzie et al. 2003) provides a method to estimate the rate (epsilon) at which occupied sites (plots) go extinct, and the rate (gamma) at which unoccupied sites are occupied. The parameters of the models implemented in MARK are psi (proportion of sites occupied), epsilon (probability of an occupied site becoming unoccupied), gamma (probability of an unoccupied site becoming occupied), and p (detection probability on a visit to the site). Three implementations of the model are present in MARK. The default, which is the parameterization used when the “Robust Design Occupancy Estimation” button is clicked in initiating a new analysis, is {psi(1), epsilon(t), gamma(t), p(session, t)}. The time intervals specified when the data are first read into MARK determine the number of intervals where epsilon applies, and the number of primary sessions that psi applies. The time intervals are specified the same as for the robust design data type.

The default parameterization {psi(1), epsilon(t), gamma(t), p(session, t)} generates estimates of the time-specific psi(t), lambda(t) and lambda'(t) parameters as derived parameters (see page 200 of the McKenzie et al. occupancy book). The lambda(t) estimates are ratios of consecutive occupancy rates, and are equivalent to the rate of change in occupancy. The lambda'(t) estimates are the log-odds ratio of consecutive occupancy rates (and are not technically defined for psi = 0 or psi = 1. The main advantage of this parameterization is that gamma and epsilon can be constrained to the [0, 1]

interval with the link function, and not affect convergence properties of the optimization. The  $\{\psi(t) \epsilon(t) p(\text{session}, t)\}$  produces derived parameters of  $\gamma(t)$  and  $\lambda(t)$ , the ratios of consecutive occupancy rates. However,  $\gamma$  is constrained to the  $[0, 1]$  interval through the penalty function approach, and for some problems, numerical convergence may be problematic. The  $\{\psi(t) \gamma(t) p(\text{session}, t)\}$  produces derived parameters of  $\epsilon(t)$  and  $\lambda(t)$ , the ratios of consecutive occupancy rates. Here,  $\epsilon$  is constrained to the  $[0, 1]$  interval through the penalty function approach, and for some problems, numerical convergence may be problematic.

To build models for the  $\{\psi(t) \epsilon(t) p(\text{session}, t)\}$  and  $\{\psi(t) \gamma(t) p(\text{session}, t)\}$  parameterizations, you change the data type from the PIM main menu. All three of these models have the same likelihood, so AICc values are comparable between them. In addition, the Pledger mixture models have been added for all 3 of the robust design occupancy models, and are also available with the change data type menu choice from the PIM main menu. All 6 of these parameterizations produce estimates of  $\lambda(t)$  and  $\lambda'(t)$ .

The usual occupancy model can also be used with robust design data by treating the primary sessions as different attribute groups, and  $\psi$  estimated for each group. However, this approach to the analysis would not provide estimates of the extinction ( $\epsilon$ ) and recolonization ( $\gamma$ ) rates.

One model often of interest is the random extinction and colonization model, where the probability of a site being occupied at time  $t+1$  is the same regardless of whether or not the site was occupied at time  $t$ . You can obtain estimates for this model in MARK by clever coding of the design matrix. Suppose the  $\epsilon$  of interest is parameter 2 in the PIM, and the  $\gamma$  of interest is parameter 3 in the PIM. Use a common beta parameter for both  $\epsilon$  and  $\gamma$ , i.e., there will be a single column that is modeled by both rows 2 and 3 of the design matrix. If you specify a logit link for both  $\epsilon$  and  $\gamma$ , and code  $\epsilon$  as -1 in the design matrix and  $\gamma$  as +1, the result is the model with  $1 - \epsilon = \gamma$ . The design matrix looks like:

Columns: B1 B2 B3

Row 1: 1 0 0

Row 2: 0 -1 0 /\* This is epsilon \*/

Row 3: 0 1 0 /\* This is gamma \*/

You also get the same results if you use a sin link instead of a logit link, because both link functions are symmetric about 0.5.

For those of you that want to be more complicated, you can also code  $\epsilon$  with -1 and use a cloglog link, and  $\gamma$  with 1 and use a loglog link, and get the

same result as you would with the logit link. Likewise, you can reverse the 2 link functions, or reverse the -1 and 1, and still get the same model, but with different interpretations of the estimate of beta.

The -1 logit link approach also works with covariates, either temporal, group, or individual covariates. Assume the following PIMs:

Epsilon PIM

2 3 4

3 4

4

Gamma PIM

5 6 7

6 7

7

Further assume that you have the time-varying covariates with values 11, 12, and 13 for the 3 time intervals. The design matrix then looks like:

Row 1: 1 0 0

Row 2: 0 -1 -11

Row 3: 0 -1 -12

Row 4 0 -1 -13

Row 5 0 1 11

Row 6: 0 1 12

Row 7: 0 1 13

Similarly, if you have an individual covariate named "covariate", you have to use the design matrix function product to specify a negative value for the covariate:

Row 1: 1 0 0

Row 2: 0 -1 product(-1,covariate)

Row 3: 0 -1 product(-1,covariate)

Row 4 0 -1 product(-1,covariate)

Row 5 0 1 covariate

Row 6: 0 1 covariate

Row 7: 0 1 covariate

In all of these cases, the estimate of  $\gamma = 1 - \epsilon$ .

An extension of the occupancy robust design model is the data type with multiple states, i.e., more than just simple occupied/not occupied. See [Occupancy Estimation Multiple States Robust Design](#) for the corresponding model.

## Individual Heterogeneity for p

Two additional sets of models have been incorporated in MARK to handle individual heterogeneity of p. Pledger mixture models (Pledger 2000), and normally distributed random errors on the logit scale (McClintock and White 2009, Gimenez and Choquet 2010). Both of these approaches would typically require more than 2 secondary occasions in a primary occasion to be able to detect individual heterogeneity. However, one advantage of the robust design is that you might look at individual heterogeneity across multiple primary occasions.

## Derived Parameters

Three derived parameters are provided. First is occupancy (psi) through time. Second is lambda, the ratio of consecutive psi estimates. Finally is lambda' (lambda prime), that is i.e.,  $[\psi(t)/(1 - \psi(t))] / [\psi(t-1)/(1 - \psi(t-1))] = [\psi(t) * (1 - \psi(t-1))] / [\psi(t-1) * \psi(1 - \psi(t))]$  (MacKenzie et al 2006, page 200).

---

© 2014 Warner College of Natural Resources  
[Apply to CSU](#) | [Equal Opportunity](#) | [Disclaimer](#) | [Privacy](#) | [Search](#)  
[Contact Webmaster](#)