

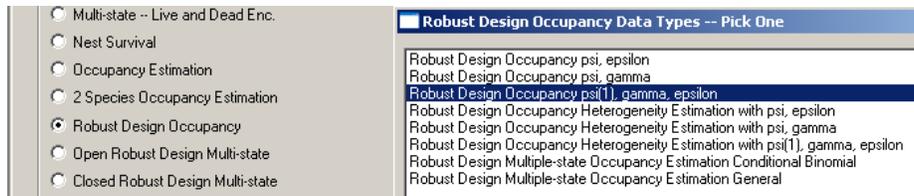
WILD 502  
 Lab 11 – Occupancy models for single-species, multi-season



Today’s lab will be based on analysis of northern spotted owl data presented in MacKenzie et al. (2003). The modeling you’ll do follows what is presented on pages 209-212 of the book titled, “Occupancy Estimation and Modeling” by MacKenzie et al. (2006). The emphasis is on how you can examine changes in occupancy and equilibrium issues with this model type. The dataset consists of detection/non-detection of northern spotted owls on multiple sites in multiple seasons (first year = 1997). We will model rates of (1) occupancy, (2) detection, (3) colonization, and (4) extinction. Here, we’ll use the parameters  $\psi$  for occupancy rate,  $p$  for detection rate,  $\gamma$  for colonization rate (going from un-occupied to occupied from 1 year to the next), and  $\epsilon$  for the local extinction rate (from 1 year to the next). Given this,  $(1 - \epsilon)$  is the rate at which occupied sites remain occupied, and  $(1 - \gamma)$  is the rate at which un-occupied sites remain un-occupied. Given the background you have in robust design, you should be able to see that the probability that a site is occupied at time  $t+1$  may or may not depend on its occupancy status in time  $t$ , i.e., be either 1<sup>st</sup>-order Markovian ( $\gamma \neq (1 - \epsilon)$ ) or random ( $\gamma = (1 - \epsilon)$ ).

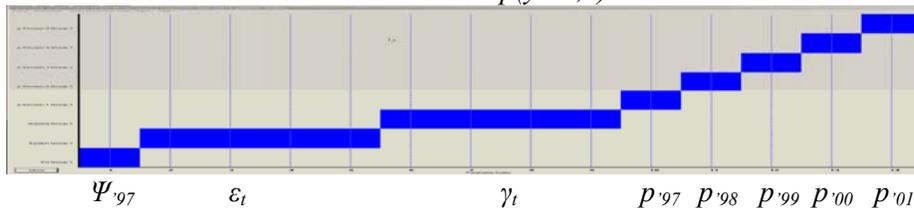
		t	
		Occupied	Not occupied
t+1	Occupied	$(1 - \epsilon)$ Local persistence	$\gamma$ Colonization
	Not occupied	$\epsilon$ Local extinction	$(1 - \gamma)$

Set up a new analysis in MARK using “Robust Design Occupancy” as the Data Type. When you do this, a 2<sup>nd</sup> window opens where you can pick the parameterization that you want to use. We’ll start with a parameterization that works with initial occupancy ( $\psi_{1997}$ ),  $\gamma_t$ , and  $\epsilon_t$  (as well as  $p_{session, t}$ ). In this parameterization, estimates of  $\psi_{1998}, \psi_{1999}, \dots, \psi_{2000}$ , and the annual rate of change in occupancy rates ( $\lambda_t = \frac{\psi_{t+1}}{\psi_t}$ ) are derived quantities that MARK calculates from ( $\psi_{1997}$ ),  $\gamma_t$ , and  $\epsilon_t$  as you’ll see.



You need to tell MARK that there are 40 occasions (5 primary sessions with 8 secondary occasions per primary occasion). There is 1 group and no covariates for this problem.

Look at the PIM chart for the default model. You’ll see that there is a single  $\psi$  (it’s for 1997, i.e., the 1<sup>st</sup> year of the study) and time-varying values for all other parameters. The authors found evidence that  $p$  varied by year but was constant across secondary occasions within each year. Thus, you can set the PIM for  $p$  within each session to be constant: we’ll call this “ $p(\text{year}, .)$ ” and use it in all of our models.



We will evaluate hypotheses represented by 6 different models.

- Model 1 = no change in occupancy = 1.  $\psi(1997), p(\text{year}, .)$ , which hypothesizes that the occupancy status doesn't change among years. You can achieve this by (a) setting the values in the PIMs for  $\epsilon$  and  $\gamma$  as constant and (b) fixing  $\hat{\gamma}$ , &  $\hat{\epsilon} = 0$  when you run the model (do this on the 'Setup Numerical Estimation Run' window by using the 'Fix Parameters' button).
- Model 2 = random changes in occupancy = 2.  $\psi(1997), \epsilon = (1 - \gamma), p(\text{year}, .)$ . This may not apply well to this species, but it is a model of interest if you think that the probability that a breeding pair will occupy a site in the next year is independent of whether the site is occupied or not in the current year (such as might be expected if site fidelity is low or site choices are made at random over the set of sites being studied). To get at this hypothesis, set  $\epsilon = (1 - \gamma)$  in each year. In Model 2, hold colonization and extinction rates constant through time. In Model 3, you'll allow colonization and extinction rates to vary across years. By comparing results for Models 2 & 3, we can evaluate whether colonization and extinction appear to be time-varying or not.

- It may not be obvious at first how you can achieve the desired constraint ( $\epsilon = (1 - \gamma)$ ). But, it turns out that your logistic regression background will help you see how the trick works.

First, consider the Design Matrix you'll use for Model 2.

B1:	B2:	B3:	Param	B4:	B5:	B6:	B7:
1	0	0	1:Psi	0	0	0	0
0	-1	0	2:Epsilon	0	0	0	0
0	1	0	3:Gamma	0	0	0	0
0	0	1	4:p Session 1	0	0	0	0
0	0	0	5:p Session 2	1	0	0	0
0	0	0	6:p Session 3	0	1	0	0
0	0	0	7:p Session 4	0	0	1	0
0	0	0	8:p Session 5	0	0	0	1

- To see how this works, consider the following.

- $\frac{\exp(-3)}{1 + \exp(-3)} = 0.047$

- $\frac{\exp(3)}{1 + \exp(3)} = 0.953$

- This trick works with covariates as well – consider the following.

-1	product(-1,PctOldGrowth)	2:Epsilon
1	PctOldGrowth	3:Gamma

- $\frac{\exp(-0.135 - 0.4 \times 4)}{1 + \exp(-0.135 - 0.4 \times 4)} = 0.15$

- $\frac{\exp(0.135 + 0.4 \times 4)}{1 + \exp(0.135 + 0.4 \times 4)} = 0.85$

- And, this trick also works with the sine link function because, like the logit link function, it too is symmetric about 0.5 (you can look at the MARK help file for further information).

- Model 3 = random changes in occupancy = 3.  $\psi(1997), \gamma(\text{year})\{\epsilon = (1 - \gamma)\}, p(\text{year}, .)$ . One way to evaluate equilibrium is by checking whether colonization and extinction rates vary in time, i.e., compare Models 2 & 3. (Another way, as nicely discussed by MacKenzie et al. in their book is to check whether or not occupancy rates are constant through time; the 2 approaches are not equivalent

and both can be evaluated). If local extinction and colonization rates are constant, the population's occupancy rate may be at, or headed for, an equilibrium level. To achieve Model 3, (a) make the PIMs for  $\varepsilon$  and  $\gamma$  to be time-varying and (b) use the following Design Matrix.

B1:	B2:	B3:	B4:	B5:	Parm	B6:	B7:	B8:	B9:	B10:
1	0	0	0	0	1:Psi	0	0	0	0	0
0	-1	0	0	0	2:Epsilon	0	0	0	0	0
0	0	-1	0	0	3:Epsilon	0	0	0	0	0
0	0	0	-1	0	4:Epsilon	0	0	0	0	0
0	0	0	0	-1	5:Epsilon	0	0	0	0	0
0	1	0	0	0	6:Gamma	0	0	0	0	0
0	0	1	0	0	7:Gamma	0	0	0	0	0
0	0	0	1	0	8:Gamma	0	0	0	0	0
0	0	0	0	1	9:Gamma	0	0	0	0	0
0	0	0	0	0	10:p Session 1	1	0	0	0	0
0	0	0	0	0	11:p Session 2	0	1	0	0	0
0	0	0	0	0	12:p Session 3	0	0	1	0	0
0	0	0	0	0	13:p Session 4	0	0	0	1	0
0	0	0	0	0	14:p Session 5	0	0	0	0	1

- Model 4 = 4.  $\psi(1997), \gamma(\cdot), \varepsilon(\cdot), p(\text{year}, \cdot)$  - this model does not set  $\varepsilon = (1 - \gamma)$  and thus the probability of being occupied at time  $t+1$  is allowed to depend on the occupancy state at time  $t$ , i.e., it can follow a 1<sup>st</sup>-order Markovian process. Think about this for a bit: if extinction rate ( $\varepsilon$ ) is different from  $(1 - \gamma)$  (which is the rate at which empty patches stay empty), (or conversely, if  $(1 - \varepsilon)$  is different from  $\gamma$ ), then the probability that a site will change states depends on what state it is in currently. Model 4 holds the local extinction and colonization rates constant while allowing them to be different from one another. To set this model up, you can just set each PIM to be constant and run the model using an Identity Matrix.

		t	
		Occupied	Not occupied
t+1	Occupied	$(1 - \varepsilon)$	$\gamma$
	Not occupied	$\varepsilon$	$(1 - \gamma)$

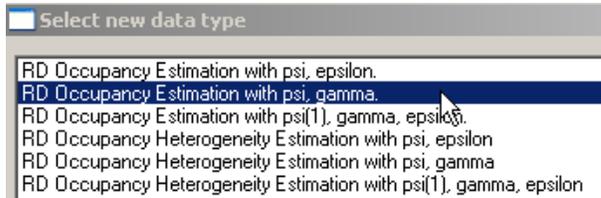
		t	
		Occupied	Not occupied
t+1	Occupied	0.8	0.6
	Not occupied	0.2	0.4

		t	
		Occupied	Not occupied
t+1	Occupied	0.8	0.8
	Not occupied	0.2	0.2

- Model 5 = 5.  $\psi(1997), \gamma(\text{year}), \varepsilon(\text{year}), p(\text{year}, \cdot)$  - this model allows local extinction and colonization rates to be different from one another and to vary over time, which is a non-equilibrium version of Model 4. This means that Model 4 can be seen as the equilibrium version of this pair.
- Model 6 = 6.  $\psi(\cdot), \gamma(\cdot), p(\text{year}, \cdot)$ , which like Model 1 and unlike Models 2-5 suggests that occupancy is constant over years. This model allows colonization to occur but holds it constant through the years. For occupancy rate to be constant, the local extinction rate must also be constant. If this is true, then local extinction rate ( $\varepsilon$ ) is defined completely by occupancy rate ( $\psi$ ) and

colonization rate ( $\gamma$ ), where  $\varepsilon = \gamma \cdot \frac{(1-\psi)}{\psi}$ . Think about why this must be so. If occupancy rate doesn't change over time and colonization occurs, then any colonization events must be offset by extinctions such that  $\varepsilon \cdot \psi = \gamma \cdot (1-\psi)$ , which can be re-written as  $\varepsilon = \gamma \cdot \frac{(1-\psi)}{\psi}$ .

Because  $\varepsilon$  is a derived parameter in this model, we drop it from the model name to make that point clearer. To run this model, we need to use a different parameterization for the data that allows us to model  $\psi_t$  directly and to estimate  $\varepsilon_t$  as a derived parameter. To do this, click on the PIM menu, choose "Change Data Type", and select the 2<sup>nd</sup> option.



Even though we are changing the parameterization, the AICc values are still comparable for Model 6 and the preceding models as the parameterizations are based on the same likelihood.

To run Model 6, you can make each of the PIMs have constant values and run the model using an identity matrix.

The modeling should provide you with the following results.

Model	K	-2log(L)
1. psi1, p(yr, .)	6	1542.564
2. psi1, gamma(.), eps(=1-gamma), p(yr, .)	7	1429.532
3. psi1, gamma(yr), eps(=1-gamma), p(yr, .)	10	1429.320
4. psi1, gamma(.), eps(.), p(yr, .)	8	1337.523
5. psi1, gamma(yr), eps(yr), p(yr, .)	14	1327.639
6. psi(.), gamma(.), p(yr, .)	7	1337.951

### Assignment

1. How many total sites were used in the study?
2. Were all sites were visited on all occasions or not? If they were not, how often does it appear that most sites were visited (just a ballpark estimate from looking at the data for a bit).
3. What are your annual estimates of  $p$  per occasion and  $p^*$  based on the typical number of visits made per site per year? Based on those estimates, does it appear that one can count on having  $p^* = 1.0$  if one does 5 surveys in any year, most years, all years? Explain.

4. Provide your estimates of  $\psi$  (SE) for each year for each model.

Model	$\psi_{1997}$ (SE)	$\psi_{1998}$ (SE)	$\psi_{1999}$ (SE)	$\psi_{2000}$ (SE)	$\psi_{2001}$ (SE)
1					
2					
3					
4					
5					
6					

5. Do the results indicate that naïve estimates of occupancy would be badly biased in this species? Explain.
6. Using the estimates of  $\psi_{1997}$ ,  $\gamma_t$ , and  $\varepsilon_t$  produced by model 5, show how to calculate the derived estimates of occupancy rate for 1998 through 2001.
7. Based on the results, does it appear that colonization and extinction are important processes for this species? Explain.
8. In the top model, what is the extinction rate and how is it derived? Based on the results for the top model, how many sites are expected to change occupancy status each year? Show your work.
9. In one paragraph, please describe what you learned about the 6 hypotheses described by the 6 competing models based on this dataset.
10. Provide a brief example of a scenario where this modeling approach would be useful for a problem of interest to you.