Single Season Occupancy Case Study #1 – Estimating Occupancy of Mahoenui giant weta in NZ.

Project Description and Context

In this module, we shall fit models that involve covariates for both occupancy and detection probabilities to data that have been collected on the Mahoenui giant weta in a scientific reserve by the Department of Conservation in the King Country district of New Zealand. 72 3m-radius circular plots were surveyed between 3-5 times over a 5-day period for weta. Each plot was assessed as to the level of browsing by feral goats and each survey was conducted by 1 of 3 observers. Although the investigation was a pilot study designed to inform future surveys, it permitted inference about the relationship between goat browsing and weta occupancy. Browsing produces dense foliage and was hypothesized to lead to reduced predation and higher weta occupancy. See MacKenzie et al. (2006) for more details. The data are included in the sample data folder that is installed along with PRESENCE in the Excel spreadsheet **Weta_pg116.xls**. This file consists of 5 sheets containing the detection-nondetection data (**detection_histories**), whether a plot was browsed or unbrowsed (**site_covar**), and which observer conducted which survey (**Obs1**, **Obs2** and **Obs3**). In this example the number of surveys is not constant for all plots, hence the detection-nondetection data include missing observations that are indicated with a "-". All of the covariates are dummy variables that =1 if the covariate is of the value indicated by the covariate name, and =0 otherwise. Note that continuous valued covariates are also possible.

References:

MacKenzie, D.I., Nichols, J. D., Royle, J.A., Pollock, K.H., Bailey, L.L., and Hines, J.E. 2006. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurence. Elsevier, Inc. 324p.

Exercise Objectives

- Learn how to create and run occupancy models where occupancy and/or detection is a function of site and/or survey specific covariates
- Learn to import and analyze temporal data
- Continue to increase comfort level and familiarity with all aspects of analysis in PRESENCE from data exploration to model selection, data filter assignment and interpretation of results

Presence spreadsheet data file: ***Weta_pg116.xls***

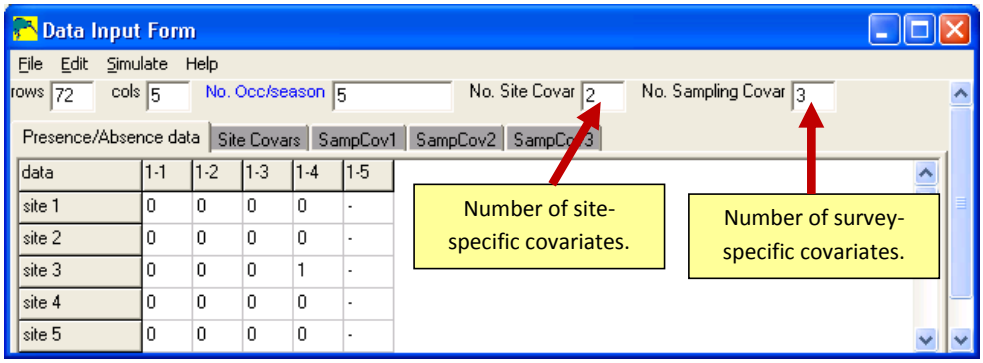
Formatted: Space After: 10 pt, Line spacing: Multiple 1.15 li, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Single Season Occupancy Case Study #1 – Estimating Occupancy of Mahoenui giant weta in NZ.

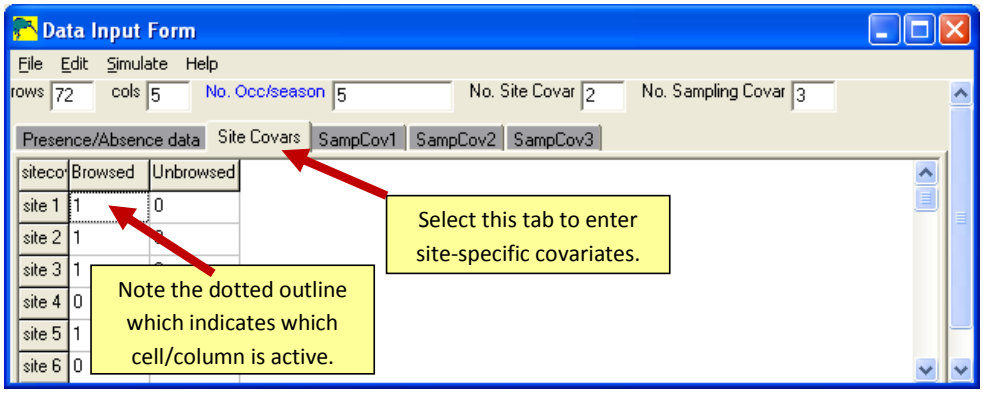
Single Season Occupancy Case Study #1 – Estimating Occupancy of Mahoenui giant weta in NZ.

INSTRUCTIONS

Step 1 – Data Import: Begin PRESENCE, start a new project and open the data input form. Copy and paste the detection-nondetection data from the spreadsheet into PRESENCE in the same manner as for the previous example, changing the number of surveys per season to 5. Next we have 2 site-specific covariates (whether a site was browsed or unbrowsed by the goats) and 3 sampling occasion, or survey-specific, covariates we want included in our data file. Therefore, change the number of site-specific covariates to 2 and the number of sampling covariates to 3. Note that when you do so additional tabs appear for sheets on which you enter your data. As the site-specific covariates must remain constant for the entire season, each one only requires 1 column hence all site-specific covariates are entered on a single sheet. The sampling occasion covariates can vary with each survey; hence the covariate will have the same dimensions as the detection-nondetection data. Therefore, 1 sheet is required by each sampling covariate.



To input the site-specific covariates, in your spreadsheet program select the sheet labelled **site_covar**, highlight the range of the covariate data including the covariate names (i.e., cells A1:B73), copy the selected cells, return to PRESENCE's data input form and select the tab labelled **Site Covar**. Click on the top-left grid cell and make sure the cell border becomes a dotted line (not a flashing cursor), then from the edit menu, select **Edit>Past>Paste w/covnames**. This paste option will not only paste the covariate values into the grid cells, but will also paste in the covariate names.



Then enter the sampling (or survey-specific) covariate that indicates whether observer 1 conducted the survey, return to your spreadsheet software and select the tab labelled **Obs1**, select and copy the values (cells A1:E72), return to the data input form and select the tab labelled **SampCov1**. Make the top-left cell active and paste the values by selecting **Edit>Paste>Paste values**. Next rename the covariate by selecting **Edit>Rename covariate** and in the text box insert **Obs1** and hit 'OK'. Repeat these steps for the covariates **Obs2** and **Obs3**.

Once you have entered all 3 observer covariates, save the data file using an appropriate name (here I have used 'weta_116') then close the data input form which will return you to the **Enter Specifications for PRESENCE Analysis** window. Select the data file you have just created, add a title for the project then click 'OK'. After a couple of seconds a blank results browser should appear. Remember, if you do not see the results browser, you have not successfully set up your project file.

Step 2 – Data Exploration:

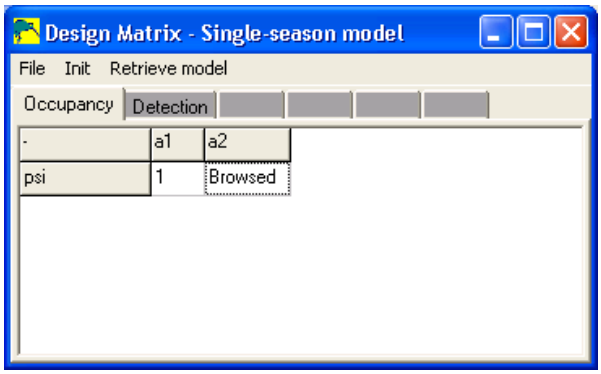
Bring back the Data window (**View>Data**) and examine which cells contain missing values for the Presence/Absence data, and covariate data (Obs1, Obs2, Obs3).

Note that the observer covariates contain missing values, and that these exactly correspond with the missing values in the detection-nondetection data. Covariate values are allowed to be missing, but only if the detection-nondetection data for the corresponding survey of the sampling unit is also missing. The assumption is required by most statistical methods. Missing values for site-specific covariates are not allowed.

Step 3 – Running a simple model : First, let's fit a simple model where the probability of occupancy is the same for all plots and the probability of detection is the same in all surveys. In our earlier notation we could call this model **psi(.),p(.)**. The design matrix for occupancy will be just a '1' in the single grid cell, and for detection probability we have five rows (for the 5 possible survey occasions) and require a single column of '1's. This is actually the default model

that appears where you select to run a custom, single-season model, hence after the design matrix appears you just need to hit 'Ok to Run' then confirm the results to add them to the results browser.

Step 4 – Analyze the influence of covariates on occupancy and detection: Next we shall fit a model where the probability of occupancy will vary among plots according to the level of browsing, and detection probability will be allowed to vary among days. We could call this model **psi(Browsed),p(Day)**. We shall first fit the model, then work back through it explaining what we're actually doing. So start the analysis by selecting **Run>Analysis:single-season** and clicking on the **Custom** radio button if it is not already selected. This will bring up the design matrices window and set the design matrix for occupancy to look like the following.



This will require you to add a column to the design matrix, then enter a covariate name. To add a column, you simply right-click anywhere on the design matrix which opens a pop-up menu and select **Add Col** (the 5th item in the menu) which will create an empty column on the right-hand side of the design matrix. You can then type the name of the covariate into the grid cell, but if you do so you must get the name 100% correct (and note that PRESENCE is case sensitive). Alternatively, if you open the **Init** menu you will notice that all of the available covariates are listed there, prefixed with a '*'. After activating column 2 (so the grid cell in the second column has the dotted outline), select **Init>*Browsed** and the covariate name **Browsed** will be inserted into the correct column.

Once the design matrices are set up, return to the SNER window (without closing the design matrix window), rename the model then hit 'OK to Run' and confirm the results when prompted.

So what have we just done? Let's start with the simple one; the design matrix for detection probability. Recall that we have already used this design matrix in the previous example, and that to read the design matrices we move along the row, we sum the terms produced by

multiplying the values in the grid cell with the corresponding beta parameters for each column. So here we could write the set of equations:

$$\text{logit}(p_{i1}) = \mathbf{1} \text{ X } b_1 + \mathbf{0} \text{ X } b_2 + \mathbf{0} \text{ X } b_3 + \mathbf{0} \text{ X } b_4 + \mathbf{0} \text{ X } b_5 = b_1$$

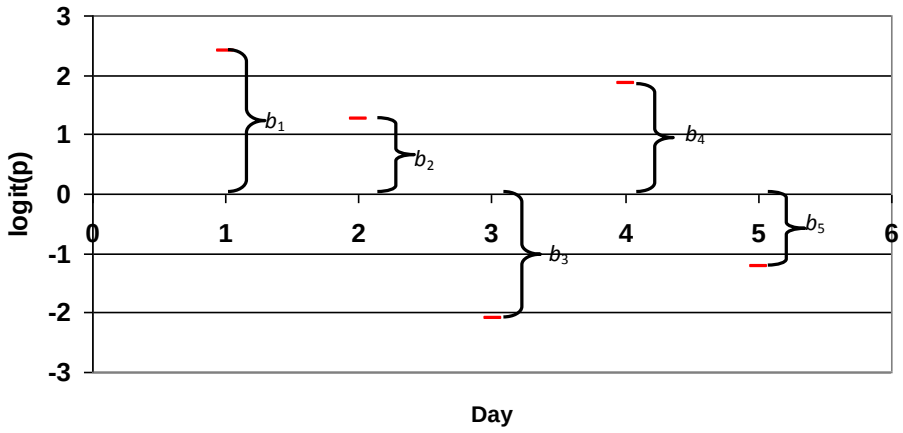
$$\text{logit}(p_{i2}) = \mathbf{0} \text{ X } b_1 + \mathbf{1} \text{ X } b_2 + \mathbf{0} \text{ X } b_3 + \mathbf{0} \text{ X } b_4 + \mathbf{0} \text{ X } b_5 = b_2$$

$$\text{logit}(p_{i3}) = \mathbf{0} \text{ X } b_1 + \mathbf{0} \text{ X } b_2 + \mathbf{1} \text{ X } b_3 + \mathbf{0} \text{ X } b_4 + \mathbf{0} \text{ X } b_5 = b_3$$

$$\text{logit}(p_{i4}) = \mathbf{0} \text{ X } b_1 + \mathbf{0} \text{ X } b_2 + \mathbf{0} \text{ X } b_3 + \mathbf{1} \text{ X } b_4 + \mathbf{0} \text{ X } b_5 = b_4$$

$$\text{logit}(p_{i5}) = \mathbf{0} \text{ X } b_1 + \mathbf{0} \text{ X } b_2 + \mathbf{0} \text{ X } b_3 + \mathbf{0} \text{ X } b_4 + \mathbf{1} \text{ X } b_5 = b_5$$

Or we could represent it graphically as:



So there are no constraints on how detection probability might vary among days; each day is free to have a unique value that differsfrom the others, and the logits of these values are b_1-b_5. That is, this design matrix represents the model that allows detection probability to be day-specific, with all plots having the same probability on each day.

Next, let's consider the design matrix for the occupancy probability, which is a representation of the following equation.

$$\text{logit}(\psi_i) = \mathbf{1} \text{ X } a_1 + \mathbf{Browsed}_i \text{ X } a_2$$

$$= a_1 + a_2 \text{ X } \mathbf{Browsed}_i$$

Recall that *Browsed* is a covariate that we defined such that if plot *i* showed signs of browsing the covariate value = 1, and = 0 otherwise. Therefore,

for an unbrowsed plot (where $Browsed_i = 0$),

$$\text{logit}(\psi_i) = \alpha_1 + \alpha_2 \text{ X } 0$$

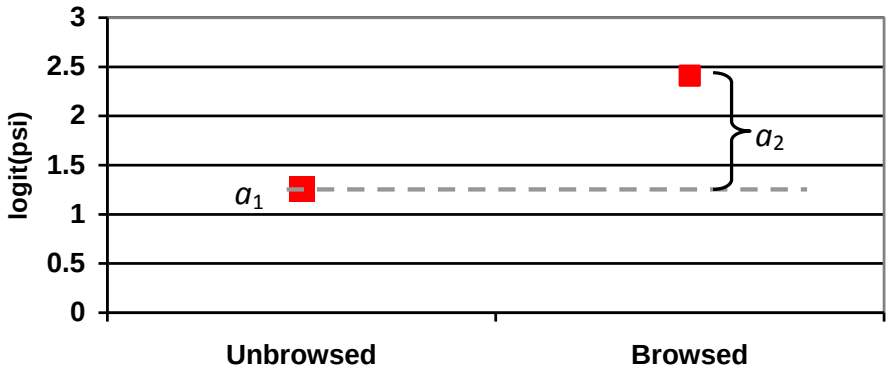
$$= \alpha_1$$

and for a browsed plot (where $Browsed_i = 1$),

$$\text{logit}(\psi_i) = \alpha_1 + \alpha_2 \text{ X } 1$$

$$= \alpha_1 + \alpha_2$$

Note that α_2 is therefore the difference in occupancy between a browsed and unbrowsed plot (on the logit scale), or alternatively what effect browsing has on occupancy compared to plots with no browsing. Graphically, we could represent this as:



Looking at the output from this model, $\hat{\alpha}_2 =$ 1.24 which, as it is >0, indicates that the estimated probability of occupancy is higher at browsed plots. In this fairly simple model we can interpret $\hat{\alpha}_2$ in a fairly straightforward manner, although for more complicated models it can be useful to interpret the effect of a covariate in terms of an odds ratio. Using the logit link, the odds ratio can be calculated as:

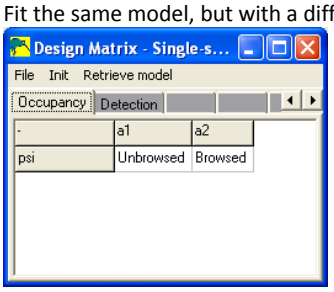
$$\text{OR} = e^{\alpha 2} = e^{1.24} = 3.44$$

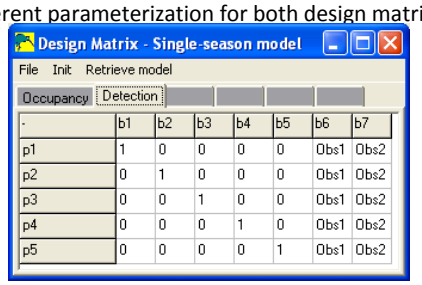
The interpretation of this would be that the odds of occupancy are 3.44 times larger for a browsed plot than an unbrowsed plot. An approximate 2-sided 95% confidence interval for the odds ratio would be:

$$CI = (e^{1.24 - 2 \times 0.787} , e^{1.24 + 2 \times 0.787}) = (e^{-0.27} , e^{2.74}) = (0.76, 15.49)$$

Exercise:

1. Working in small groups, fit the model with the following design matrices. Write out the corresponding equations, draw the graphs and interpret the resulting beta parameter estimates. Fit the same model, but with a different parameterization for both design matrices.





2. Within your groups fit the following set of models to the weta data, and discuss the results.
   i. psi(.)p(.)
   ii. psi(.)p(Day)
   iii. psi(.)p(Obs)
   iv. psi(.)p(Day + Obs)
   v. psi(Browsed)p(.)
   vi. psi(Browsed)p(Day)
   vii. psi(Browsed)p(Obs)
   viii. psi(Browsed)p(Day+Obs).

3. Calculate the summed AIC weights for the *Browsed* covariate for occupancy, and the *Observer* covariate for detection (**hint**: if you right mouse click on the **Results Browser** you can copy the table to the clipboard, and then paste into a spreadsheet).

4. What inferences do you draw about (1) the effect of browsing on occupancy and (2) the detection probabilities of the 3 observers?

5. Calculate the model averaged estimate (and SE) of the probability of occupancy for a browsed plot.