## Project Description and Context

In this module, we'll fit models that involve covariates for occupancy, detection, colonization and extinction probabilities to data that have been collected near Macraes Flat, Otago, NZ by the Department of Conservation. Rock outcrops were surveyed by DOC up to 3 times per year for 5 years. The main interest in this experiment was colonization and extinction probabilities, and whether they differ between areas of tussock and pasture. 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 spreadsheet **Grand_skinks.xls**. This file consists of 1 sheet containing the detection-nondetection data and a covariate indicating whether the site was surrounded by pasture (or not). In this example, not all sites were visited on each day, hence the detection-nondetection data include missing observations that are indicated with a "-". The covariate =1 if the site is surrounded by pasture, and =0 if not.

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

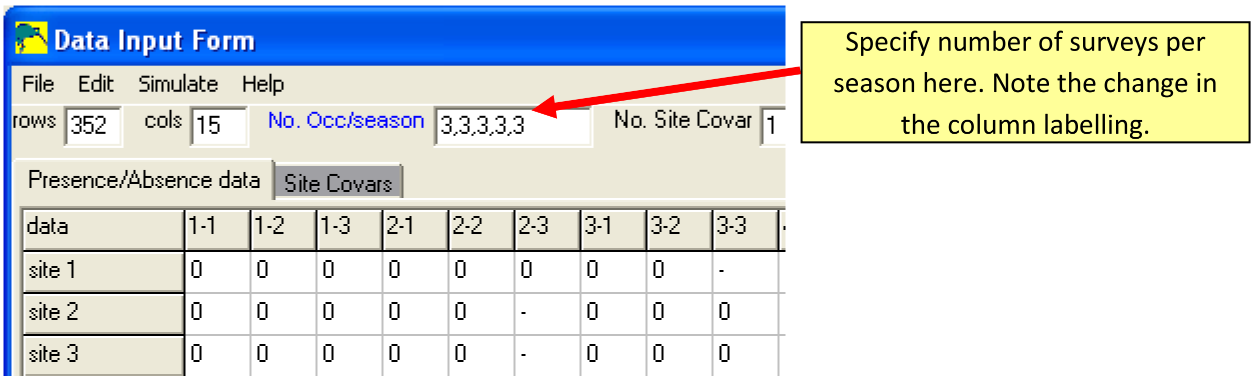
- Learn how to create and run occupancy models where occupancy, colonization, extinction and/or detection is a function of site specific covariates
- Learn to import and analyze temporal data
- Continue to increase comfort level and familiarity with all aspects of analysis in PRESENCE from data exploration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

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

---

## 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 3,3,3,3,3. Next change the number of site-specific covariates to 1.



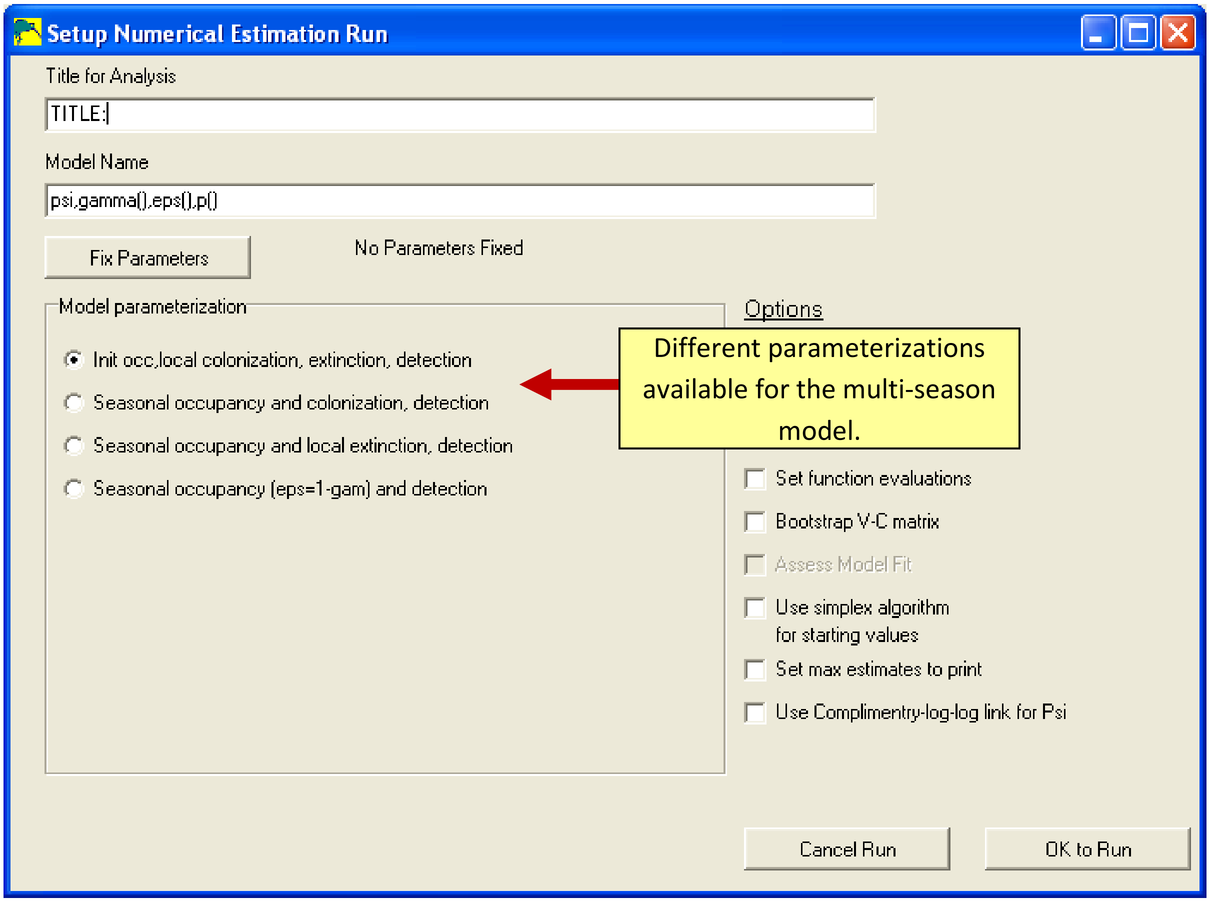

Save the data, then complete setting up the project as for the single-season models.

**Step 2 – Data Exploration:**

Bring back the Data window (**View>Data** ) and examine which cells contain missing values for the Presence/Absence data. Note that there are many cells with missing data ('-'), including some where no data were collected for entire seasons. These cells will not contribute any information towards the estimation of occupancy or detection for those seasons, but occupancy and detection for those sites can be estimated from the model since we assume those sites have the same occupancy and detection probabilities as other sites which do have data. *This can be very handy for generating maps, or inferring occupancy to sites where you didn't survey as you can simply add faux sites to the input file containing nothing but missing data for all occasions.*

**Step 3 – Running a simple model**

Once the **Results Browser** appears, select **Run>Analysis: multi-season>simple multi-season**, which should bring up the following window. Note that it is a similar format to the single-season analysis, but with a different selection of models in the model box. Initially, we will just use the first parameterization and use the other parameterizations in later examples. Note that there are no predefined models, and that all models have to be fit using the design matrix. Looking at the design matrix, previously we only have 2 named tabs in the design matrix window, and note that now we have 4, one for each of our parameter types. There is still only 1 real parameter (row) for occupancy as under this parameterization we are only estimating the probability of occupancy in the first season. Occupancy for the other seasons can be derived from initial occupancy, colonization and extinction rates. There are 4 real parameters for colonization; there are 5 seasons of data so 4 between-season opportunities for the species to colonize a location. Similarly, there are 4 real parameters for extinction probabilities. Finally there are 15 real parameters for detection, 1 for each survey occasion.




The first model we are going to fit is "psi(.),gamma(.),eps(.),p(.)", i.e., each of our parameters are constant across time (survey) and space (sites). This is actually the default model when you open the analysis window so here can just press 'Ok to Run'. Confirm the model, then open the output. Note that the first part of the output is very similar to the single-season output, with the main exception being that there are now 2 additional design matrices, one for colonization (gamma) and one for extinction (eps) probabilities. Focusing on the estimation part of the output, first there are the beta parameters that are reported on the logit-scale, and their variance-covariance matrix. This is followed by the real parameter estimates. From this model we would therefore conclude that;

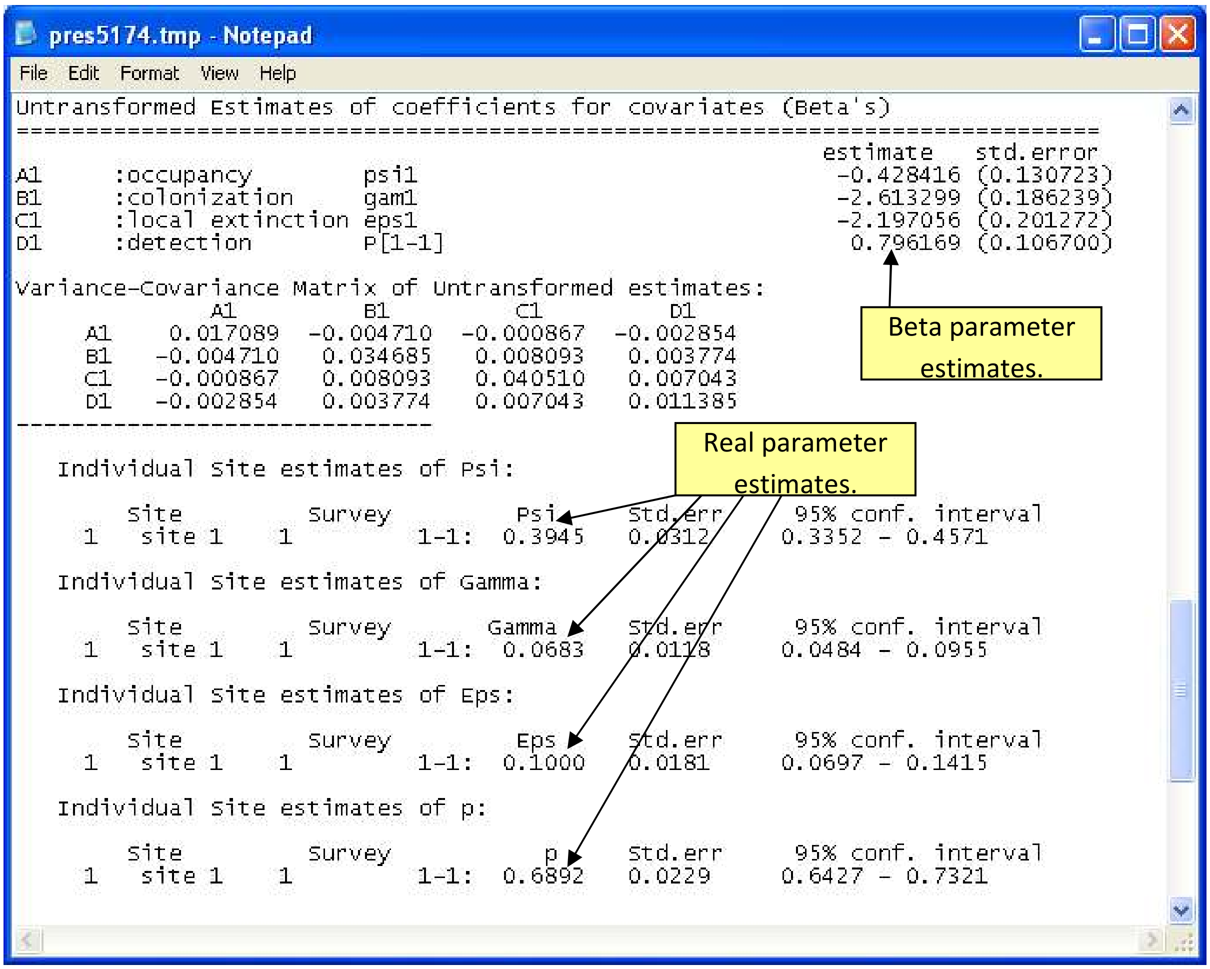
1) the probability of occupancy in the first year was 0.39;

2) between seasons, the probability that skinks colonize a previously unoccupied rocky outcrop is 0.07;

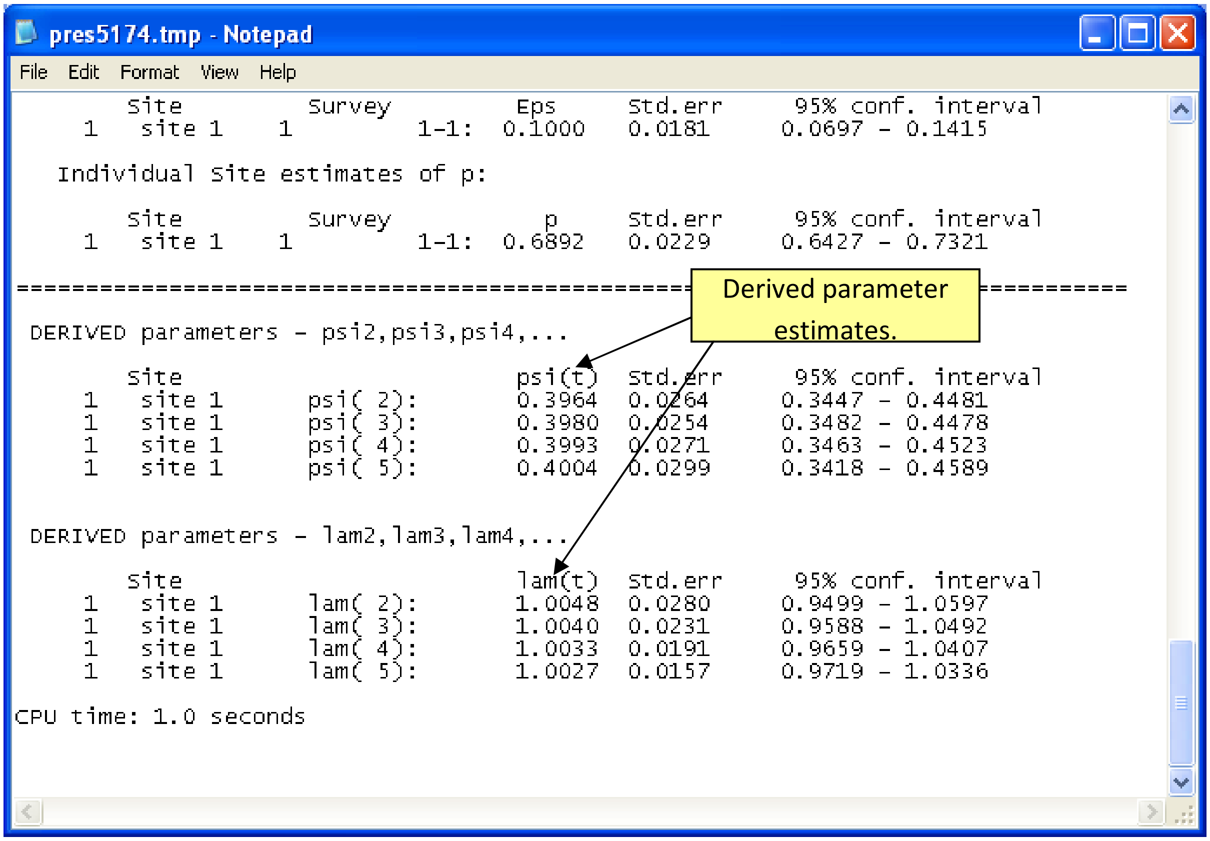
3) between seasons, the probability that skinks go locally extinct from an occupied rocky outcrop is 0.10; and

4) given an outcrop is occupied by skinks, the probability of detecting skinks in a single survey is 0.69.

Below the real parameter estimates, some derived estimates are reported. These are calculated using some equations that shall be covered later, but briefly the derived estimates are the probability of occupancy in years 2-5, and then occupancy-based population growth rates that are calculated as the ratio of successive occupancy estimates. From these estimates one might conclude (based upon this particular model) that the system appears relatively stable, as the derived probabilities of occupancy are very similar and the lamdas are approximately 1.




```
Untransformed Estimates of coefficients for covariates (Beta's)
==============================================================================
                                                          estimate   std.error
A1      :occupancy        psi1                            -0.428416 (0.130723)
B1      :colonization     gam1                            -2.613299 (0.186239)
C1      :local extinction eps1                            -2.197056 (0.201272)
D1      :detection        P[1-1]                           0.796169 (0.106700)

Variance-Covariance Matrix of Untransformed estimates:
              A1         B1         C1         D1
     A1    0.017089  -0.004710  -0.000867  -0.002854
     B1   -0.004710   0.034685   0.008093   0.003774
     C1   -0.000867   0.008093   0.040510   0.007043
     D1   -0.002854   0.003774   0.007043   0.011385
------------------------------

   Individual Site estimates of Psi:

        Site          Survey           Psi     Std.err    95% conf. interval
     1   site 1    1          1-1:  0.3945    0.0312    0.3352 - 0.4571

   Individual Site estimates of Gamma:

        Site          Survey         Gamma     Std.err    95% conf. interval
     1   site 1    1          1-1:  0.0683    0.0118    0.0484 - 0.0955

   Individual Site estimates of Eps:

        Site          Survey           Eps     Std.err    95% conf. interval
     1   site 1    1          1-1:  0.1000    0.0181    0.0697 - 0.1415

   Individual Site estimates of p:

        Site          Survey             p     Std.err    95% conf. interval
     1   site 1    1          1-1:  0.6892    0.0229    0.6427 - 0.7321
```



pres5174.tmp - Notepad

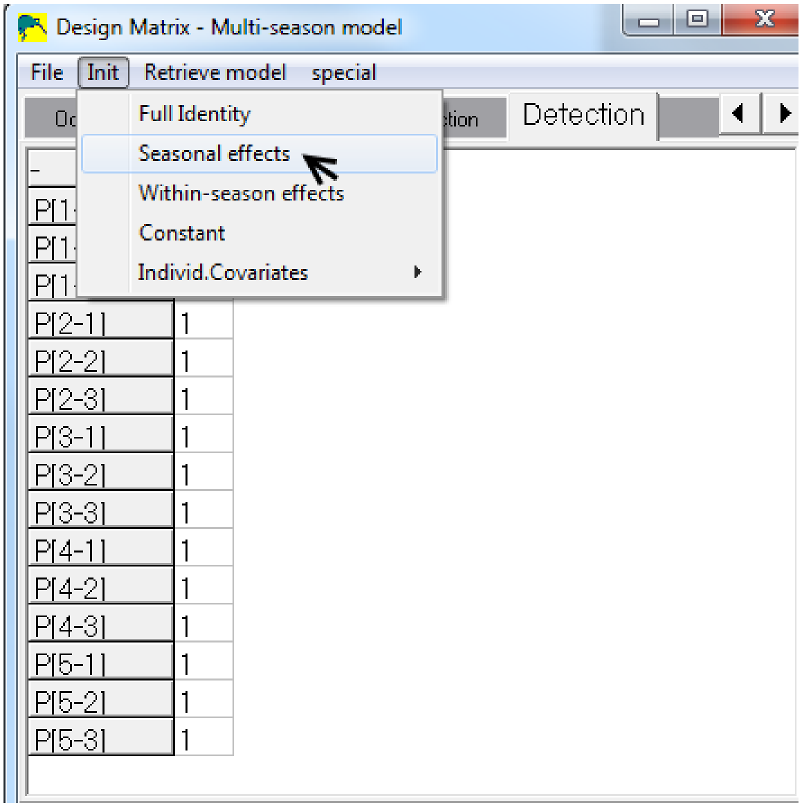
File Edit Format View Help

```
      Site          Survey          Eps     Std.err     95% conf. interval
   1   site 1    1         1-1:  0.1000   0.0181     0.0697 - 0.1415

  Individual Site estimates of p:

      Site          Survey            p     Std.err     95% conf. interval
   1   site 1    1         1-1:  0.6892   0.0229     0.6427 - 0.7321

======================================================================================

 DERIVED parameters - psi2,psi3,psi4,...

      Site                         psi(t)   Std.err     95% conf. interval
   1   site 1       psi( 2):       0.3964   0.0264     0.3447 - 0.4481
   1   site 1       psi( 3):       0.3980   0.0254     0.3482 - 0.4478
   1   site 1       psi( 4):       0.3993   0.0271     0.3463 - 0.4523
   1   site 1       psi( 5):       0.4004   0.0299     0.3418 - 0.4589

 DERIVED parameters - lam2,lam3,lam4,...

      Site                         lam(t)   Std.err     95% conf. interval
   1   site 1       lam( 2):       1.0048   0.0280     0.9499 - 1.0597
   1   site 1       lam( 3):       1.0040   0.0231     0.9588 - 1.0492
   1   site 1       lam( 4):       1.0033   0.0191     0.9659 - 1.0407
   1   site 1       lam( 5):       1.0027   0.0157     0.9719 - 1.0336

CPU time: 1.0 seconds
```

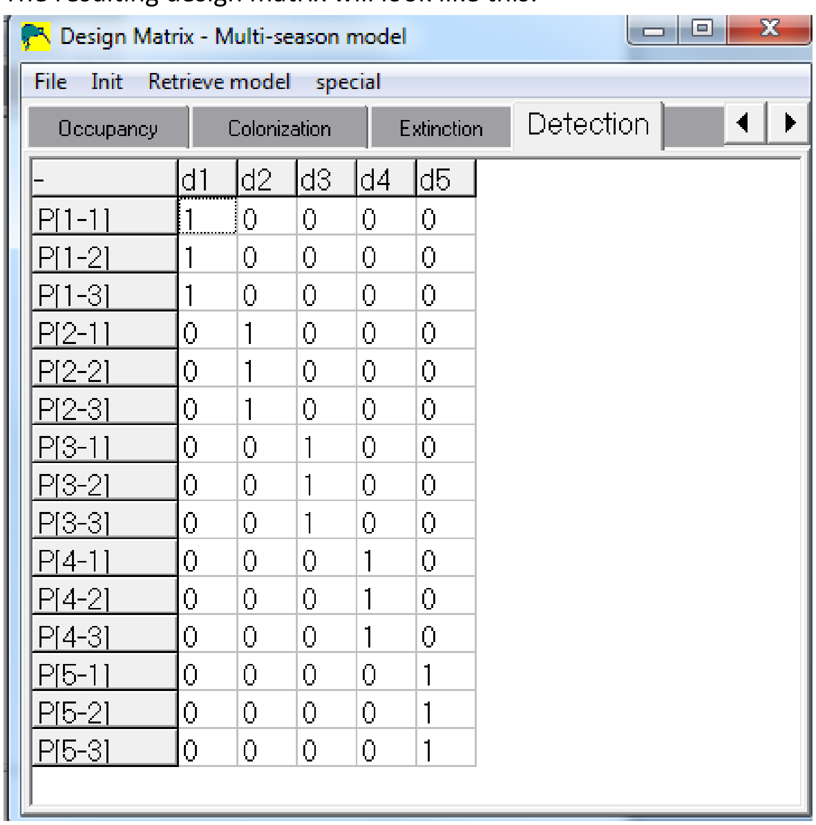


### Step 4 – Running a more general model

Now, fit the model psi(.),gamma(year),eps(year),p(year), that is, we are allowing colonization, extinction and detection probabilities to be different in each year. The design matrices for colonization and extinction can be defined by selecting **Init>Full Identity** in the design matrix window (there are also other parameterisations possible that will give the same model). For detection probability, we are assuming that detection probability varies among years, but within years the probability of detection is constant. A shortcut to defining this type of model for detection is available in the Init menu with the ***Seasonal effects*** option (this option is unavailable for the other design matrices).




The resulting design matrix will look like this:




| - | d1 | d2 | d3 | d4 | d5 |
|---|---|---|---|---|---|
| P[1-1] | 1 | 0 | 0 | 0 | 0 |
| P[1-2] | 1 | 0 | 0 | 0 | 0 |
| P[1-3] | 1 | 0 | 0 | 0 | 0 |
| P[2-1] | 0 | 1 | 0 | 0 | 0 |
| P[2-2] | 0 | 1 | 0 | 0 | 0 |
| P[2-3] | 0 | 1 | 0 | 0 | 0 |
| P[3-1] | 0 | 0 | 1 | 0 | 0 |
| P[3-2] | 0 | 0 | 1 | 0 | 0 |
| P[3-3] | 0 | 0 | 1 | 0 | 0 |
| P[4-1] | 0 | 0 | 0 | 1 | 0 |
| P[4-2] | 0 | 0 | 0 | 1 | 0 |
| P[4-3] | 0 | 0 | 0 | 1 | 0 |
| P[5-1] | 0 | 0 | 0 | 0 | 1 |
| P[5-2] | 0 | 0 | 0 | 0 | 1 |
| P[5-3] | 0 | 0 | 0 | 0 | 1 |

Don't forget to change the model name to "psi(.),gamma(year),eps(year),p(year)", then run the model.

Looking at the results, the estimated probability of;

1) occupancy in the first year is 0.38;

2) colonization is 0.12, 0.01, 0.07 and 0.10 for each of the intervals between years;

3) extinction is 0.07, 0.07, 0.14 and 0.17 for each of the intervals between years

4) detecting the skinks in a single survey is 0.70,0.65, 0.69, 0.84 and 0.66 in each of the 5 years respectively.

**Step 5 – Examining the effects of a covariate**

The next step in this example is to consider how much evidence is there for the importance of the *Pasture* covariate as a predictor for the occupancy-related parameters. We're going to fit a range of models representing different hypotheses about how we believe pasture might influence occupancy and its vital rates. Our model set is generated by considering the different combinations of these hypotheses.

| Parameter | Hypothesis |
|---|---|
| Occupancy | Is the same for all outcrops |
| | Is different for outcrops surrounded by pasture |
| Colonization | Varies in time, does not depend on intervening matrix |
| | Varies in time, intervening matrix has an effect that is consistent each year |
| Extinction | Varies in time, does not depend on intervening matrix |
| | Varies in time, intervening matrix has an effect that is consistent each year |

**Exercise:**

1. Working in small groups, define a candidate set of models based upon these hypotheses. For all models allow detection probability to vary among years, but constant within years, and allow detection to be different dependent upon the surrounding matrix, with the effect being constant for all years, i.e., p(*year + Pasture*) model.
2. Fit these models in PRESENCE.
3. Determine the amount of evidence for each hypothesis.
    a. Copy the PRESENCE AIC table to the clipboard, then paste into Excel. In the spreadsheet, add up the model weights of all models which support each hypothesis (eg., weights of models where initial occupancy is a function of Pasture). Compute an "evidence ratio" for each hypothesis. This is the ratio of the sum of model weights which support the hypothesis versus the sum of models which don't. If this ratio is 1.0, then there is little evidence to support the hypothesis (just as much weight for the hypothesis as against it).