Design and Analysis of Occupancy Studies
Part 2

Carl James Schwarz

Department of Statistics and Actuarial Science
Simon Fraser University
Burnaby, BC, Canada
cschwarz @ stat.sfu.ca

Outline

Single Species; Multi-Season - Sampling Protocol

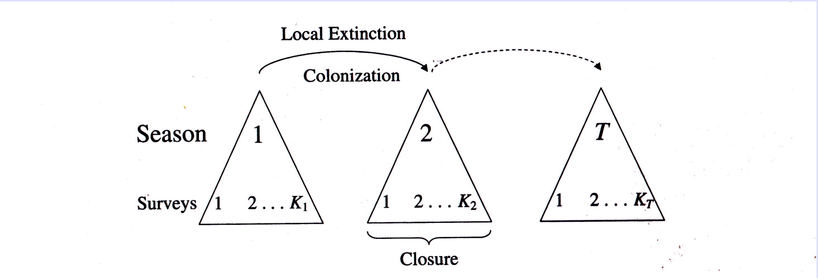
Single-Species Multi-Season Occupancy Studies

Single Species; Multi-Season - Sampling Protocol I

Sampling Protocol:

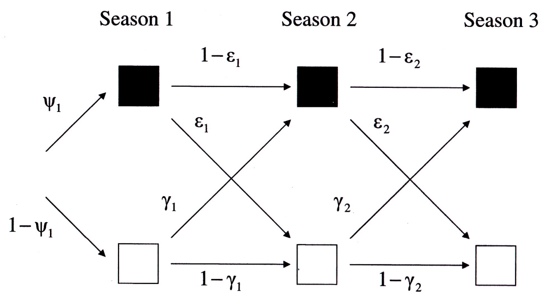
- Landscape divided (artificially or naturally) into S patches or cells or SITES.
- Select $s << S$ sites at random (all sites have equal probability of selection).
- Visit each site K_y times in each of Y (years) seasons.
- Record detection or not detection of species in site i in year y in visit k.
- Create a **Detection/Encounter History** for each visited site e.g. 011 00 0110. [No blanks between season when input into programs.]

Single Species; Multi-Season - Sampling Protocol II



Single Species; Multi-Season - Dynamics

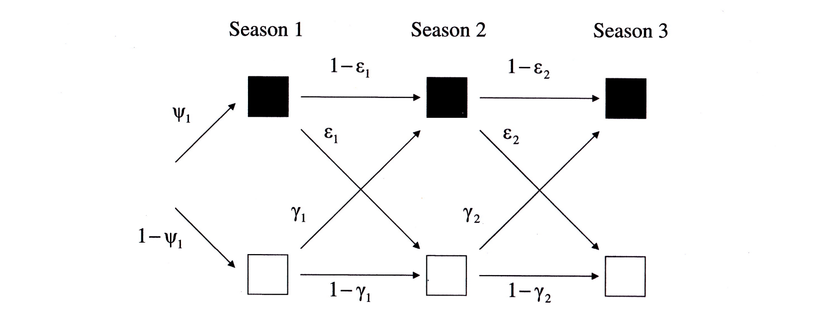
- Occupancy not change WITHIN a season. Initial occupancy ψ_1.
- Occupancy allowed to change ACROSS seasons.
  - Colonization probability γ_y between seasons y and season $y+1$.
  - Local extinction ϵ_y between season y and season $y+1$.
- No false positives; detection $p_{syk} < 1$.



Single Species; Multi-Season - Assumptions

1. Occupancy state of sites is constant during all single-season surveys (closure).
2. **Initial** Probability of occupancy (ψ) is equal across all sites (homogeneity).
3. Probability of detection (p) given occupancy is equal across all sites (homogeneity).
4. Detection of species in each survey of a site is independent of those on other surveys
5. Detection histories at each site are independent
6. No false positives.
7. **First-order Markov process**
    - Extinction/colonization depends on state in season y and not previous seasons (no memory).
    - Sites now occupied, tend to remain occupied in next season and vice-versa.

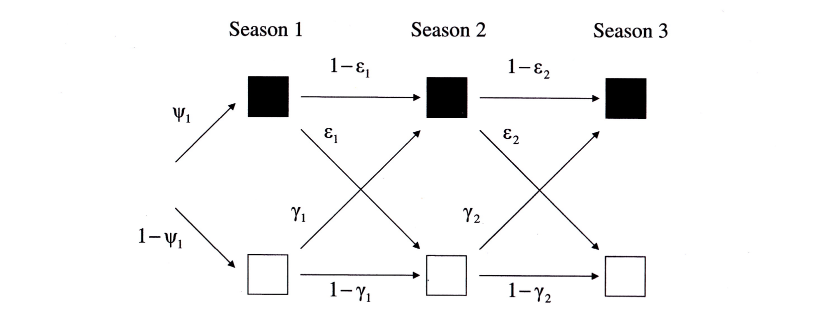
Single Species; Multi-Season - History Probabilities



History 11 01 10

$$\psi_1 p_{11} p_{12}(1-\epsilon_1)(1-p_{21})p_{22}(1-\epsilon_2)p_{31}(1-p_{32})$$

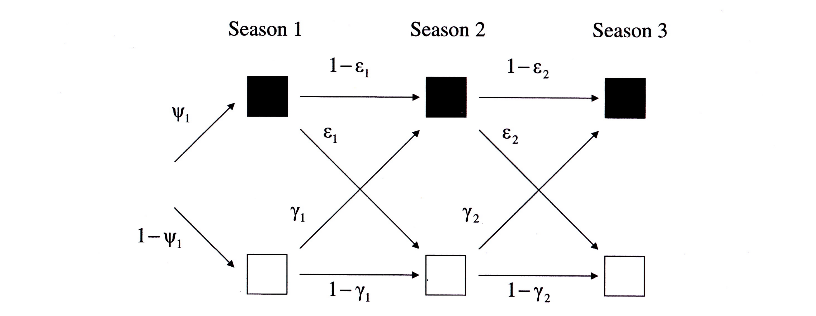
Single Species; Multi-Season - History Probabilities



History 00 10 00

$[\psi_1(1-p_{11})(1-p_{12})(1-\epsilon_1)+(1-\psi_1)\gamma_1] \times p_{21}(1-p_{22}) \times$
$[(1-\epsilon_2)(1-p_{31})(1-p_{32})+\epsilon_2]$

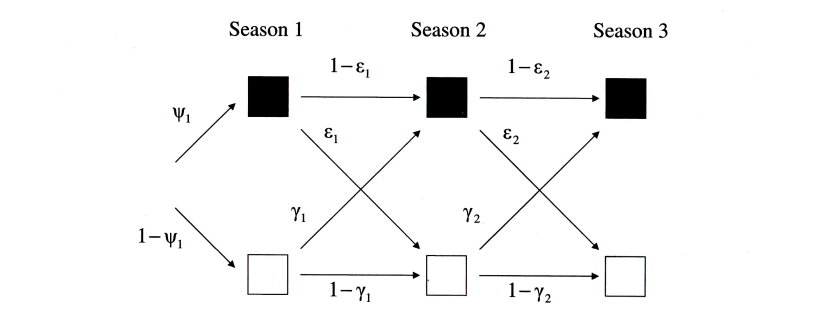
Single Species; Multi-Season - History Probabilities



History 00 10 00

$[\psi_1(1-p_{11})(1-p_{12})(1-\epsilon_1)+(1-\psi_1)\gamma_1]\times$
$p_{21}(1-p_{22})\times$
$[(1-\epsilon_2)(1-p_{31})(1-p_{32})+\epsilon_2]$

Single Species; Multi-Season - History Probabilities



History 00 00 00

You don't want to write this out without using matrices!

Single Species; Multi-Season - Covariate Effects

Covariates can be used to model:

- Initial occupancy probabilities, e.g. habitat effects
- Detection probabilities at global (e.g. weather) or site specific (e.g. habitat) or site*temporal (e.g. observer)
- Extinction/colonization at global (e.g. weather between seasons), site specific (e.g. habitat, patch area

Models fitted and compared using Maximum Likelihood and AIC as before.

Start simple and work to more complex models.

Don't forget model assessment (goodness-of-fit).

Single Species; Multi-Season - Derived Parameters

Seasonal-occupancy probabilities:

$$\psi_{y+1} = \psi_y(1 - \epsilon_y) + (1 - \psi_y)\gamma_y$$

Occupancy change:

$$\lambda_y = \frac{\psi_{y+1}}{\psi_y}$$

Unsatisfactory as occupancy cannot increase indefinitely and so λ tends towards 1.

Odds ratio of Occupancy change:

$$\lambda_y' = \frac{\frac{\psi_{y+1}}{1-\psi_{y+1}}}{\frac{\psi_y}{1-\psi_y}}$$

Odds can increase indefinitely (as occupancy gets closer to 1) and will be linear on $logit(\psi)$ scale.

Single Species; Multi-Season - Example

Single-Species Multi-Season Occupancy Studies

Analysis of Northern Spotted Owl study using *PRESENCE*

Single Species; Multi-Season - NSO - *PRESENCE*

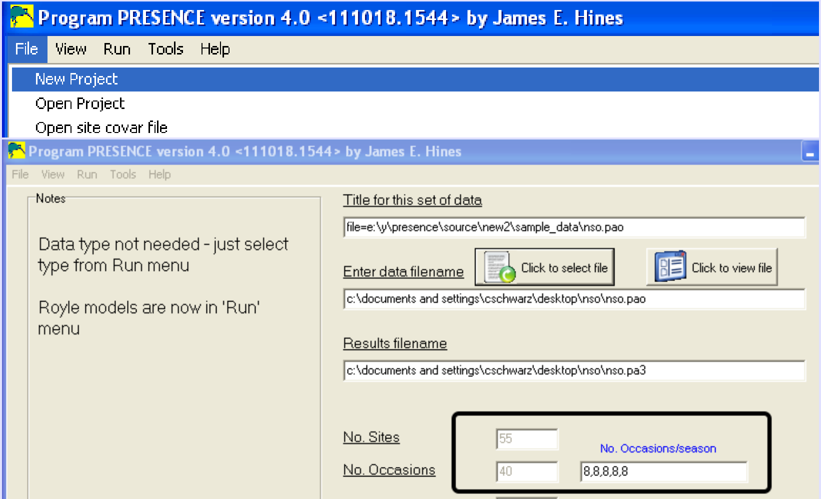
Northern Spotted Owl (*Strix occidentalis caurina*) in California.

$s = 55$ sites visited up to $K = 8$ times per season between 1997 and 2001 ($Y = 5$).

Detection probabilities relatively constant within years, but likely different among years.

Single Species; Multi-Season - NSO - *PRESENCE*

Start a new project and open the *NSO.pao* dataset in the OccupancySampleData folder.



Single Species; Multi-Season - NSO - *PRESENCE*

View the data.

Data Input Form

File Edit Simulate Help

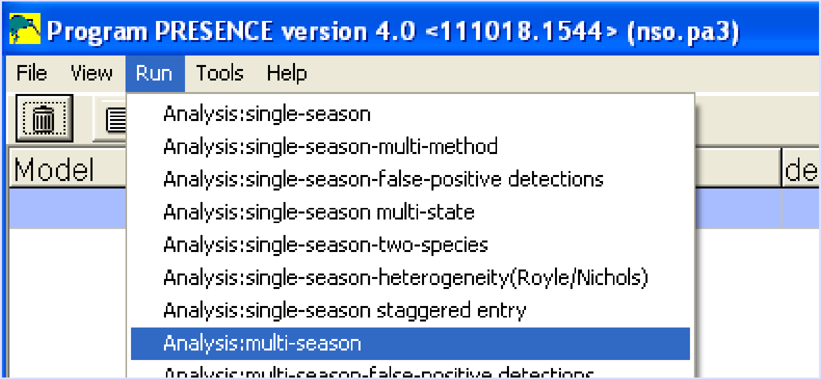
rows 55 cols 40 No. Occ/season 8,8,8,8,8 No. Site Covar 0 No. Sampling Covar 0

Presence/Absence data

data	1-1	1-2	1-3	1-4	1-5	1-6	1-7	1-8	2-1	2-2	2-3	2-4	2-5	2-6	2-7	2-8	3-1
site 1	0	1	1	1	-	-	-	-	0	1	0	-	-	-	-	-	1
site 2	0	0	-	-	-	-	-	-	0	0	0	-	-	-	-	-	0
site 3	1	0	0	0	1	1	1	-	0	1	0	0	-	-	-	-	0
site 4	1	1	1	-	-	-	-	-	1	0	-	-	-	-	-	-	0
site 5	0	0	0	0	0	0	-	-	0	0	0	0	0	0	0	-	0
site 6	1	-	-	-	-	-	-	-	0	0	0	1	1	1	-	-	0
site 7	0	0	0	0	0	0	-	-	0	0	0	0	0	0	0	-	0
site 8	0	0	0	0	0	0	-	-	0	0	1	0	0	0	0	0	0
site 9	0	0	0	0	0	0	-	-	0	0	0	0	0	0	-	-	0
site 10	1	1	1	-	-	-	-	-	0	0	0	0	0	0	0	-	0
site 11	1	1	1	-	-	-	-	-	1	0	0	1	0	1	1	1	1
site 12	0	0	0	0	1	-	-	-	1	1	0	0	0	0	1	-	0
site 13	0	1	1	-	-	-	-	-	0	0	1	1	1	-	-	-	0
site 14	1	1	-	-	-	-	-	-	0	1	0	0	0	-	-	-	0
site 15	1	0	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1

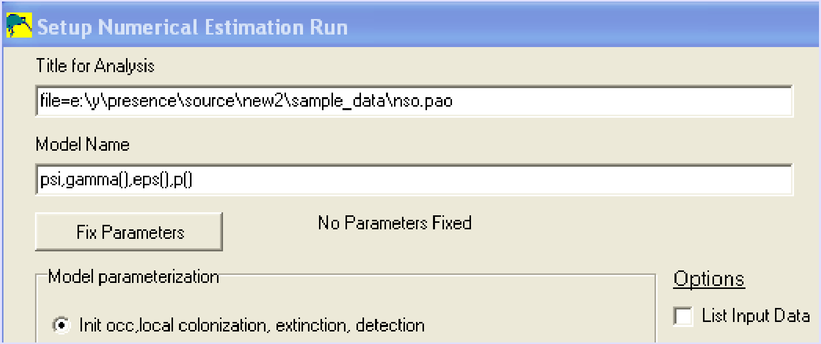
Single Species; Multi-Season - NSO - *PRESENCE*

Select single-species, multi-season models:



Single Species; Multi-Season - NSO - *PRESENCE*

Model with colonization and extinction probabilities:



Single Species; Multi-Season - NSO - *PRESENCE*

Look at DESIGN matrices - What model is being fit?

Occupancy | Colonization | Extinction | Detection

-	a1
psi1	1

Occupancy | Colonization | Extinction | Detection

-	b1
gam1	1
gam2	1
gam3	1
gam4	1

Single Species; Multi-Season - NSO - *PRESENCE*

Look at DESIGN matrices - What model is being fit?

Occupancy | Colonization | **Extinction** | Detection

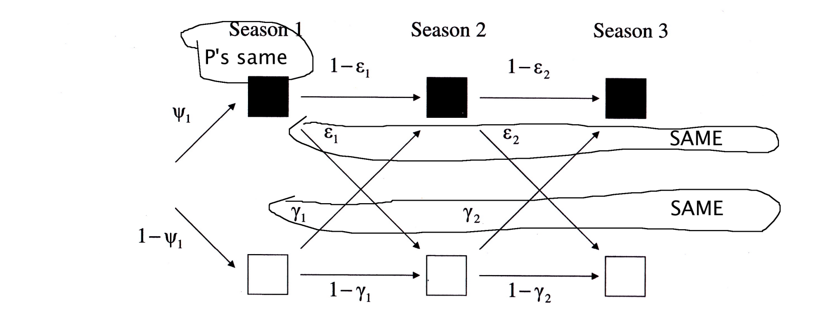
-	c1
eps1	1
eps2	1
eps3	1
eps4	1

Occupancy | Colonization | Extinction | **Detection**

-	d1
P[1-1]	1
P[1-2]	1
P[1-3]	1
P[1-4]	1

Single Species; Multi-Season - NSO - *PRESENCE*

Helpful to draw a diagram of the process model:



Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(*), \epsilon(*), \gamma(*)$ model in the usual way.

File View Run Tools Help

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi,gamma(*),eps(*),p(*)	1363.32	0.00	1.0000	1.0000	4	1355.32

Single Species; Multi-Season - NSO - *PRESENCE*

$\psi(1997), p(*), \epsilon(*), \gamma(*)$ parameter estimates:

```
Real parameters: (computed using covariates from 1st site and 1st survey)

Real parameter :                  estimate    SE(estimate)
     1 psi1                         0.6312       0.0673
     2 gam1                         0.1842       0.0427
     3 gam2                         0.1842       0.0427
     4 gam3                         0.1842       0.0427
     5 gam4                         0.1842       0.0427
     6 eps1                         0.1507       0.0332
     7 eps2                         0.1507       0.0332
     8 eps3                         0.1507       0.0332
     9 eps4                         0.1507       0.0332
    10 P[1-1]                       0.4947       0.0186
    11 P[1-2]                       0.4947       0.0186
    12 P[1-3]                       0.4947       0.0186
    13 P[1-4]                       0.4947       0.0186
    14 P[1-5]                       0.4947       0.0186
```

Single Species; Multi-Season - NSO - *PRESENCE*

$\psi(1997), p(*), \epsilon(*), \gamma(*)$ DERIVED parameter estimates:

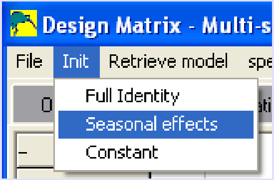
```
==========================================================================

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

       site                          psi(t)  Std.err     95% conf. interval
    1   site 1      psi( 2):         0.6040  0.0510     0.5041 - 0.7039
    1   site 1      psi( 3):         0.5859  0.0511     0.4856 - 0.6861
    1   site 1      psi( 4):         0.5738  0.0566     0.4628 - 0.6848
    1   site 1      psi( 5):         0.5658  0.0621     0.4440 - 0.6876

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

       site                          lam(t)  Std.err     95% conf. interval
    1   site 1      lam( 2):         0.9569  0.0514     0.8562 - 1.0576
    1   site 1      lam( 3):         0.9700  0.0374     0.8968 - 1.0433
    1   site 1      lam( 4):         0.9795  0.0266     0.9274 - 1.0315
    1   site 1      lam( 5):         0.9861  0.0186     0.9497 - 1.0224
```

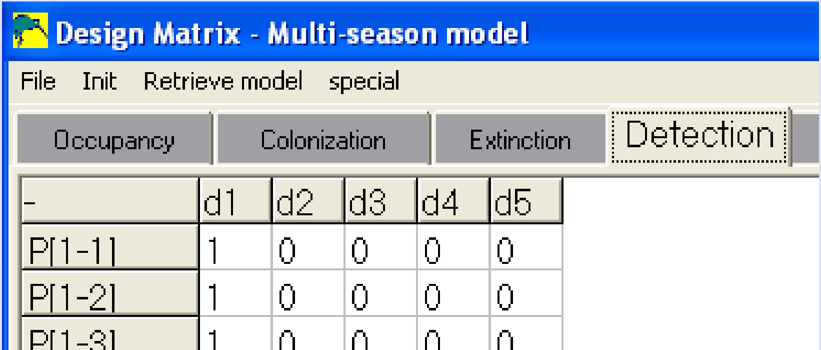
Single Species; Multi-Season - NSO - *PRESENCE*

Modify DESIGN matrix for p to allow for year effects, but equal within each year.

Single Species; Multi-Season - NSO - *PRESENCE*

Modify DESIGN matrix for p to allow for year effects, but equal within each year.





Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ model in the usual way.

File View Run Tools Help

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi,gamma(),eps(),p(year)	1353.52	0.00	0.9926	1.0000	8	1337.52
psi,gamma(*),eps(*),p(*)	1363.32	9.80	0.0074	0.0074	4	1355.32

Single Species; Multi-Season - NSO - *PRESENCE*

$\psi(1997), p(year), \epsilon(*), \gamma(*)$ parameter estimates:

```
Real parameters: (computed using covariates from 1st s

Real parameter :                  estimate    SE(estimate)
     1 psi1                         0.6247       0.0669
     2 gam1                         0.1788       0.0430
     3 gam2                         0.1788       0.0430
     4 gam3                         0.1788       0.0430
     5 gam4                         0.1788       0.0430
     6 eps1                         0.1423       0.0328
     7 eps2                         0.1423       0.0328
     8 eps3                         0.1423       0.0328
     9 eps4                         0.1423       0.0328
    10 P[1-1]                       0.5906       0.0394
    11 P[1-2]                       0.5906       0.0394
    12 P[1-3]                       0.5906       0.0394
    13 P[1-4]                       0.5906       0.0394
    14 P[1-5]                       0.5906       0.0394
    15 P[1-6]                       0.5906       0.0394
    16 P[1-7]                       0.5906       0.0394
    17 P[1-8]                       0.5906       0.0394
    18 P[2-1]                       0.5225       0.0405
```

Single Species; Multi-Season - NSO - *PRESENCE*

$\psi(1997), p(year), \epsilon(*), \gamma(*)$ DERIVED parameter estimates:

```
================================================================================

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

      site                          psi(t)  std.err      95% conf. interval
    1   site 1       psi( 2):       0.6029  0.0514     0.5021 - 0.7038
    1   site 1       psi( 3):       0.5882  0.0515     0.4872 - 0.6891
    1   site 1       psi( 4):       0.5781  0.0570     0.4665 - 0.6897
    1   site 1       psi( 5):       0.5713  0.0626     0.4486 - 0.6940

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

      site                          lam(t)  std.err      95% conf. interval
    1   site 1       lam( 2):       0.9651  0.0507     0.8658 - 1.0645
    1   site 1       lam( 3):       0.9755  0.0370     0.9029 - 1.0480
    1   site 1       lam( 4):       0.9829  0.0265     0.9309 - 1.0349
    1   site 1       lam( 5):       0.9882  0.0188     0.9514 - 1.0250
```

Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ model.

Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ model.

File View Run Tools Help

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi,gamma(),eps(),p(year)	1353.52	0.00	0.7386	1.0000	8	1337.52
psi,gamma(year),eps(year),p(year)	1355.64	2.12	0.2559	0.3465	14	1327.64
psi,gamma(*),eps(*),p(*)	1363.32	9.80	0.0055	0.0074	4	1355.32

Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ estimates

```
Real parameters: (computed using covariates from 1st site and 1st survey)

Real parameter :                  estimate    SE(estimate)
     1 psi1                         0.6295       0.0666
     2 gam1                         0.1076       0.0739
     3 gam2                         0.0692       0.0642
     4 gam3                         0.3862       0.1059
     5 gam4                         0.1163       0.0867
     6 eps1                         0.0886       0.0503
     7 eps2                         0.1340       0.0672
     8 eps3                         0.2391       0.0869
     9 eps4                         0.1188       0.0620
    10 P[1-1]                       0.5893       0.0395
```

Single Species; Multi-Season - NSO - *PRESENCE*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ DERIVED estimates.

```
==============================================================================

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

       site                         psi(t)   Std.err      95% conf. interval
    1   site 1        psi( 2):      0.6136   0.0669      0.4825 - 0.7447
    1   site 1        psi( 3):      0.5581   0.0701      0.4206 - 0.6956
    1   site 1        psi( 4):      0.5953   0.0716      0.4551 - 0.7356
    1   site 1        psi( 5):      0.5716   0.0679      0.4385 - 0.7047

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

       site                         lam(t)   Std.err      95% conf. interval
    1   site 1        lam( 2):      0.9748   0.0692      0.8392 - 1.1103
    1   site 1        lam( 3):      0.9096   0.0794      0.7539 - 1.0652
    1   site 1        lam( 4):      1.0666   0.1511      0.7704 - 1.3628
    1   site 1        lam( 5):      0.9602   0.0904      0.7831 - 1.1374
```

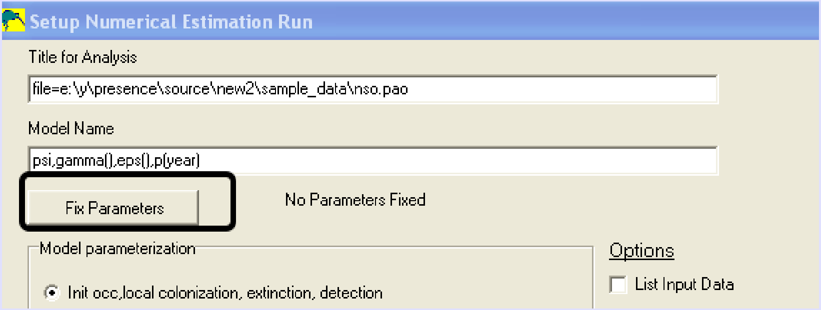
Single Species; Multi-Season - NSO - *PRESENCE*

What model would represent NO CHANGE IN OCCUPANCY in individual sites over the multiple seasons?

Single Species; Multi-Season - NSO - *PRESENCE*

What model would represent NO CHANGE IN OCCUPANCY in individual sites over the multiple seasons, i.e. convert this to an 'equivalent single season model=?

$\gamma_y = 0$ and ϵ_y=0 for all years. Fit the model and FIX some parameters



Single Species; Multi-Season - NSO - *PRESENCE*

What model would represent NO CHANGE IN OCCUPANCY in individual sites over the multiple seasons, i.e. convert this to an 'equivalent single season model=?

$\gamma_y = 0$ and ϵ_y=0 for all years. Fit the model and FIX some

Set Parameters to fix

Parm	Fixed value
psi1	
gam1	0
gam2	0
gam3	0
gam4	0
eps1	0
eps2	0

Single Species; Multi-Season - NSO - *PRESENCE*

$\psi(1997), p(year), \epsilon(NONE), \gamma(NONE)$ has no support.

File View Run Tools Help

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi,gamma(),eps(),p(year)	1353.52	0.00	0.7386	1.0000	8	1337.52
psi,gamma(year),eps(year),p(year)	1355.64	2.12	0.2559	0.3465	14	1327.64
psi,gamma(*),eps(*),p(*)	1363.32	9.80	0.0055	0.0074	4	1355.32
psi,gamma(NONE),eps(NONE),p(year)	1558.56	205.04	0.0000	0.0000	8	1542.56

Single Species; Multi-Season - NSO - *PRESENCE*

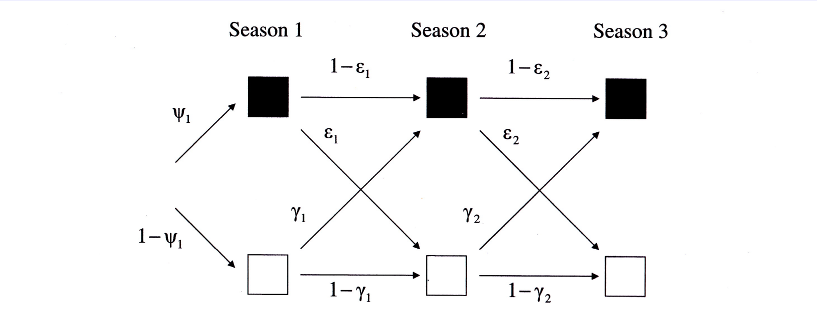
What model would represent RANDOM occupancy over seasons, i.e. no site fidelity between seasons?

A RANDOM occupancy model says that occupancy in each season is a random process with no regard to sites being occupied or not occupied in previous seasons. For example, migrating birds return each year to select nesting sites but the process each year is independent of previous years. Site could be occupied for multiple years in a row, but this the result of random chance rather than any site fidelty.

Single Species; Multi-Season - NSO - *PRESENCE*

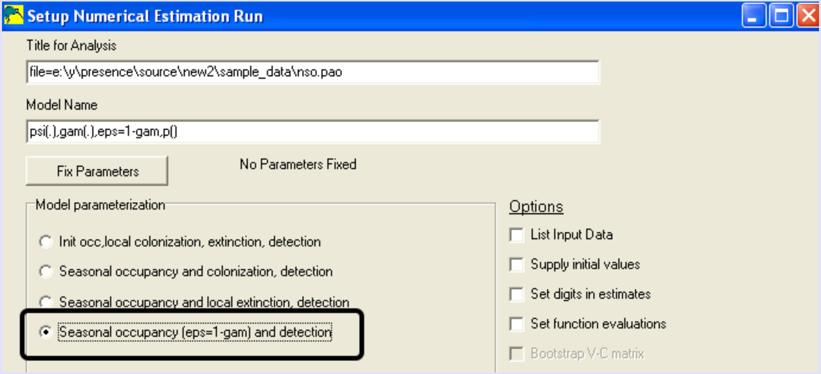
A RANDOM occupancy model implies that an occupied site in season y has the same chance of being occupied in year $y+1$ (i.e. $(1-\epsilon_y)$) as does an unoccupied site in season y being occupied in year $y+1$ (i.e. γ_y).

Or ... RANDOM occupancy $\rightarrow \gamma_y = (1-\epsilon_y)$.



Single Species; Multi-Season - NSO - *PRESENCE*

RANDOM occupancy $\rightarrow \gamma_y = (1 - \epsilon_y)$.



Fit model: $\psi(1997), p(year), \epsilon(year), \gamma = 1 - \epsilon$
Fit model: $\psi(1997), p(year), \epsilon(*), \gamma = 1 - \epsilon$

Single Species; Multi-Season - NSO - *PRESENCE*

Fit model: $\psi(1997), p(year), \epsilon(year), \gamma = 1 - \epsilon$

Fit model: $\psi(1997), p(year), \epsilon(*), \gamma = 1 - \epsilon$

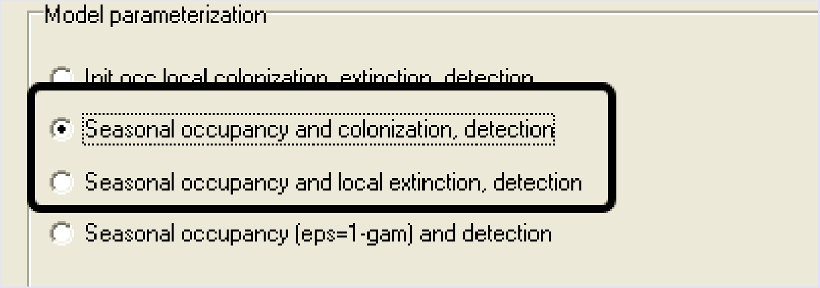
Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi,gamma(),eps(),p(year)	1353.52	0.00	0.7386	1.0000	8	1337.52
psi,gamma(year),eps(year),p(year)	1355.64	2.12	0.2559	0.3465	14	1327.64
psi,gamma(*),eps(*),p(*)	1363.32	9.80	0.0055	0.0074	4	1355.32
psi(.),gam(*),eps=1-gam,p(year)	1443.53	90.01	0.0000	0.0000	7	1429.53
psi(.),gam(year),eps=1-gam,p(year)	1449.32	95.80	0.0000	0.0000	10	1429.32
psi,gamma(NONE),eps(NONE),p(year)	1558.56	205.04	0.0000	0.0000	8	1542.56

Single Species; Multi-Season - NSO - *PRESENCE*

What model would represent a population in equilibrium in occupancy?

Single Species; Multi-Season - NSO - *PRESENCE*

What model would represent a population in equilibrium in occupancy? $\psi_{y+1} = \psi_y \rightarrow \psi_{EQ} = \frac{\gamma}{\gamma+\epsilon}$



Both models are "equivalent".
Specify that ψ is constant via the design matrix.

Single Species; Multi-Season - NSO - *PRESENCE*

Equilibrium:

Fit: $\psi(*), p(year), \gamma(year)$ (implicitly models $\epsilon(year)$)

Fit: $\psi(*), p(year), \epsilon(year)$ (implicitly models $\gamma(year)$)

Fit: $\psi(*), p(year), \epsilon(*)$ (implicitly models $\gamma(*)$)

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi(*),gamma(year),p(year)	1349.34	0.00	0.4101	1.0000	10	1329.34
psi(*),eps(year,p(year)	1349.34	0.00	0.4101	1.0000	10	1329.34
psi(*),gamma(*),p(year)	1351.95	2.61	0.1112	0.2712	7	1337.95
psi,gamma(),eps(),p(year)	1353.52	4.18	0.0507	0.1237	8	1337.52
psi,gamma(year),eps(year),p(year)	1355.64	6.30	0.0176	0.0429	14	1327.64
psi,gamma(*),eps(*),p(*)	1363.32	13.98	0.0004	0.0009	4	1355.32
psi(.),gam(*),eps=1-gam,p(year)	1443.53	94.19	0.0000	0.0000	7	1429.53
psi(.),gam(year),eps=1-gam,p(year)	1449.32	99.98	0.0000	0.0000	10	1429.32
psi,gamma(NONE),eps(NONE),p(year)	1558.56	209.22	0.0000	0.0000	8	1542.56

Delete one of the duplicate models before continuing.

Single Species; Multi-Season - NSO - *PRESENCE*

Equilibrium: $\psi(*), p(year), \gamma(year)$ (implicitly models $\epsilon(year)$)

```
Real parameters: (computed using covariates from 1st site and 1st survey)

Real parameter :                  estimate   SE(estimate)
     1 psi1                         0.5935      0.0496
     2 psi2                         0.5935      0.0496
     3 psi3                         0.5935      0.0496
     4 psi4                         0.5935      0.0496
     5 psi5                         0.5935      0.0496
     6 gam1                         0.1209      0.0540
     7 gam2                         0.1297      0.0626
     8 gam3                         0.3730      0.0893
     9 gam4                         0.1471      0.0630
```

```
DERIVED parameters - eps2,eps3,eps4,...

      site                        eps(t)  std.err     95% conf. interval
   1   site 1       eps( 2):      0.0828  0.0370     0.0103 - 0.1553
   1   site 1       eps( 3):      0.0888  0.0438     0.0030 - 0.1747
   1   site 1       eps( 4):      0.2555  0.0641     0.1298 - 0.3812
   1   site 1       eps( 5):      0.1007  0.0443     0.0139 - 0.1876


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

      site                        lam(t)  std.err     95% conf. interval
   1   site 1       lam( 2):      1.0000  0.0000     1.0000 - 1.0000
   1   site 1       lam( 3):      1.0000  0.0000     1.0000 - 1.0000
   1   site 1       lam( 4):      1.0000  0.0000     1.0000 - 1.0000
   1   site 1       lam( 5):      1.0000  0.0000     1.0000 - 1.0000
```

Single Species; Multi-Season - NSO - *PRESENCE*

Conclusion:

- Changed in occupancy best represented by a Markov process (as expected), i.e. site fidelity at the species level (but not necessarily at the individual level).
- Random occupancy or no changes in occupancy have no support.
- Highest weight given to models in equilibrium (90% of model weights).

Single Species; Multi-Season - Example

Single-Species Multi-Season Occupancy Studies

Analysis of Northern Spotted Owl study using *RPresence*

Single Species; Multi-Season - NSO - *RPresence*

Northern Spotted Owl (*Strix occidentalis caurina*) in California.

$s = 55$ sites visited up to $K = 8$ times per season between 1997 and 2001 ($Y = 5$).

Detection probabilities relatively constant within years, but likely different among years.

Single Species; Multi-Season - NSO - *RPresence*

Read in data and create **.pao* object.

```
input.history <- read.csv("NSO_pg209.csv",
     header=FALSE, skip=2, na.strings="-")
input.history$V1 <- NULL # drop the site number

Nvisits.per.season  <- rep(8,5) # five years with 8 visits

nso.pao <- RPresence::createPao(input.history,
                                nsurveyseason=Nvisits.per.s
                                title='NSO SSMS')
nso.pao
```

Single Species; Multi-Season - NSO - *RPresence*

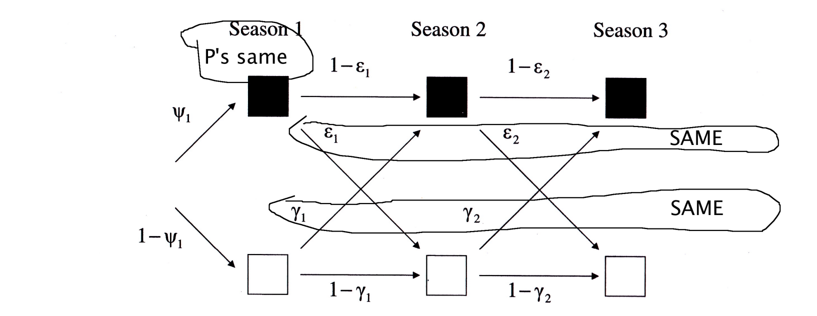
Model with colonization and extinction probabilities parameterization.
Fit the $\psi(1997), \gamma(*), \epsilon(*), p(*)$ model in the usual way.

```
mod.psiDot.gDot.eDot.pDot <-
        RPresence::occMod(
             model=list(psi~1, gamma~1, epsilon~1, p~1),
             type="do.1", data=nso.pao)
```

The *do.1* implies a dynamic occupancy model with the 1^{st} parameterization involving colonization and extinction probabilities.

Single Species; Multi-Season - NSO - *RPresence*

Helpful to draw a diagram of the process model:



Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ Results of model fit.

```
> summary(mod.psiDot.gDot.eDot.pDot)
Model name=psi()gamma()epsilon()p()
AIC=1363.3153
-2*log-likelihood=1355.3153
num. par=4
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ parameter estimates.

```
# Estimates of initial occupancy for unit 1
> mod.psiDot.gDot.eDot.pDot$real$psi[1,]
              est           se lower_0.95 upper_0.95
unit1_1 0.631162 0.06726531  0.4927215  0.7509202

# Estimates of local extinction probability for unit 1
> mod.psiDot.gDot.eDot.pDot$real$epsilon[ seq(1, by=nrow(in
               est           se lower_0.95 upper_0.95
unit1_1 0.1507363 0.03322916 0.09642521  0.2279215
unit1_2 0.1507363 0.03322916 0.09642521  0.2279215
unit1_3 0.1507363 0.03322916 0.09642521  0.2279215
unit1_4 0.1507363 0.03322916 0.09642521  0.2279215
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ parameter estimates.

```
> # Estimate of  local colonization probability for each un
> mod.psiDot.gDot.eDot.pDot$real$gamma[ seq(1, by=nrow(inpu
              est         se lower_0.95 upper_0.95
unit1_1 0.1841758 0.04271845   0.114504   0.282706
unit1_2 0.1841758 0.04271845   0.114504   0.282706
unit1_3 0.1841758 0.04271845   0.114504   0.282706
unit1_4 0.1841758 0.04271845   0.114504   0.282706

# Estimates of detection for unit 1
> mod.psiDot.gDot.eDot.pDot$real$p[ grepl('unit1_', row.nam
                est         se lower_0.95 upper_0.95
unit1_1-1 0.4947399 0.01858221  0.4584116   0.531124
unit1_1-2 0.4947399 0.01858221  0.4584116   0.531124
...
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ derived parameter estimates of occupancy in later years

```
> mod.psiDot.gDot.eDot.pDot$derived$psi[ grepl('unit1_', ro
              est         se lower_0.95 upper_0.95
unit1_2 0.6039540 0.05095837  0.5011031  0.6983648
unit1_3 0.5858583 0.05112748  0.4834709  0.6813277
unit1_4 0.5738231 0.05661527  0.4610093  0.6794431
unit1_5 0.5658186 0.06213786  0.4425225  0.6814739
```

Notice that estimates of ψ are in two different data structures (groan).

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ derived parameter estimates of population growth are NOT given but can be computed. (SE are harder - see me)

```
site_psi <- rbind(
  mod.psiDot.gDot.eDot.pDot$real$psi    [grepl('unit1_', ro
  mod.psiDot.gDot.eDot.pDot$derived$psi[grepl('unit1_', ro
site_psi

logit <- function(x) log(x/(1-x))
expit <- function(x) 1/(1+exp(-x))

lambda <- exp(diff(log(site_psi$est),1))
lambda
prod(lambda) # overall growth in occupancy over entire set

lambda.prime <- exp(diff(logit(site_psi$est),1))
lambda.prime
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ derived parameter estimates of population growth are NOT given but can be computed. (SE are harder - see me)

```
> lambda <- exp(diff(log(site_psi$est),1))
> lambda
[1] 0.9568923 0.9700380 0.9794571 0.9860506
> prod(lambda) # overall growth in occupancy over entire se
[1] 0.8964713
>
> lambda.prime <- exp(diff(logit(site_psi$est),1))
> lambda.prime
[1] 0.8911547 0.9276527 0.9517973 0.9678720
> prod(lambda.prime) # overall growth in occupancy on logit
[1] 0.7615544
```

Single Species; Multi-Season - NSO - *RPresence*

Fit model for p to allow for year effects, but equal within each year.

Single Species; Multi-Season - NSO - *RPresence*

Model with colonization and extinction probabilities parameterization.

Fit the $\psi(1997), \gamma(*), \epsilon(*), p(Year)$ model in the usual way.

```
mod.psiDot.gDot.eDot.pYear <- RPresence::occMod(
      model=list(psi~1, gamma~1, epsilon~1, p~SEASON),
      type="do.1", data=nso.pao)
```

The *do.1* implies a dynamic occupancy model with the 1^{st} parameterization involving colonization and extinction probabilities.

Notice the use of the keyword *SEASON* for seasonal effects.

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ Results of model fit.

```
> summary(mod.psiDot.gDot.eDot.pYear)
Model name=psi()gamma()epsilon()p(SEASON)
AIC=1353.5226
-2*log-likelihood=1337.5226
num. par=8
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ parameter estimates.

```
> # Estimate of initial occupance
> mod.psiDot.gDot.eDot.pYear$real$psi[1,]
              est         se lower_0.95 upper_0.95
unit1_1 0.6247322 0.06689281  0.4876152  0.7443906

> # Estimate of  local colonization probability for each un
> mod.psiDot.gDot.eDot.pYear$real$gamma[ seq(1, by=nrow(inp
              est         se lower_0.95 upper_0.95
unit1_1 0.1788329 0.04301375  0.1092563  0.2788467
unit1_2 0.1788329 0.04301375  0.1092563  0.2788467
unit1_3 0.1788329 0.04301375  0.1092563  0.2788467
unit1_4 0.1788329 0.04301375  0.1092563  0.2788467
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ parameter estimates.

```
> # Estimate of  local colonization probability for each un
> mod.psiDot.gDot.eDot.pYear$real$gamma[ seq(1, by=nrow(inp
              est         se lower_0.95 upper_0.95
unit1_1 0.1788329 0.04301375  0.1092563  0.2788467
unit1_2 0.1788329 0.04301375  0.1092563  0.2788467
unit1_3 0.1788329 0.04301375  0.1092563  0.2788467
unit1_4 0.1788329 0.04301375  0.1092563  0.2788467

> # Estimate of probability of detection at each time point
> mod.psiDot.gDot.eDot.pYear$real$p[ grepl('unit1_', row.na
                est         se lower_0.95 upper_0.95
unit1_1-1 0.5905767 0.03942533  0.5116913  0.6650601
unit1_1-2 0.5905767 0.03942533  0.5116913  0.6650601
...
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ derived parameter estimates of occupancy in later years

```
> # Derived parameters - estimated occupancy for each unit
> mod.psiDot.gDot.eDot.pYear$derived$psi[ grepl('unit1_', r
              est         se lower_0.95 upper_0.95
unit1_2 0.6029468 0.05143318  0.4991699  0.6982204
unit1_3 0.5881573 0.05153023  0.4848551  0.6842353
unit1_4 0.5781171 0.05696025  0.4643893  0.6841224
unit1_5 0.5713011 0.06259530  0.4467618  0.6874203
```

Notice that estimates of ψ are in two different data structures (groan).

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ derived parameter estimates of population growth are NOT given but can be computed. (SE are harder - see me)

```
site_psi <- rbind(
  mod.psiDot.gDot.eDot.pYear$real$psi    [grepl('unit1_', ro
  mod.psiDot.gDot.eDot.pYear$derived$psi[grepl('unit1_', ro
site_psi

logit <- function(x) log(x/(1-x))
expit <- function(x) 1/(1+exp(-x))

lambda <- exp(diff(log(site_psi$est),1))
lambda
prod(lambda) # overall growth in occupancy over entire set

lambda.prime <- exp(diff(logit(site_psi$est),1))
lambda.prime
```

Single Species; Multi-Season - NSO - *RPresence*

Model $\psi(1997), \gamma(*), \epsilon(*), p(YEAR)$ derived parameter estimates of population growth are NOT given but can be computed. (SE are harder - see me)

```
> lambda <- exp(diff(log(site_psi$est),1))
> lambda
[1] 0.9651285 0.9754713 0.9829294 0.9882100
> prod(lambda) # overall growth in occupancy over entire se
[1] 0.9144737
>
> lambda.prime <- exp(diff(logit(site_psi$est),1))
> lambda.prime
[1] 0.9121743 0.9404416 0.9595372 0.9724981
> prod(lambda.prime) # overall growth in occupancy on logit
[1] 0.800498
```

Single Species; Multi-Season - NSO - *RPresence*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ model.

Single Species; Multi-Season - NSO - *RPresence*

Model with colonization and extinction probabilities parameterization.
Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ model.

```
mod.psiDot.gYear.eYear.pYear <- RPresence::occMod(
    model=list(psi~1, gamma~SEASON, epsilon~SEASON, p~SEASO
    type="do.1", data=nso.pao)
```

The *do.1* implies a dynamic occupancy model with the 1^{st} parameterization involving colonization and extinction probabilities. Notice the use of the keyword *SEASON* for seasonal effects.

Single Species; Multi-Season - NSO - *RPresence*

What model would represent NO CHANGE IN OCCUPANCY in individual sites over the multiple seasons?

Single Species; Multi-Season - NSO - *RPresence*

What model would represent NO CHANGE IN OCCUPANCY in individual sites over the multiple seasons, i.e. convert this to an equivalent single season model=?

$\gamma_y = 0$ and ϵ_y=0 for all years. Fit the model and FIX some

parameters
Doesn't seem to work in the current version of RPresence (groan).

Single Species; Multi-Season - NSO - *RPresence*

Here is the current AIC table:

```
                                  Model DAIC    wgt np
1              psi()gamma()epsilon()p(SEASON) 0.00 0.7383
2 psi()gamma(SEASON)epsilon(SEASON)p(SEASON) 2.12 0.2562
3                    psi()gamma()epsilon()p() 9.79 0.0055
>
```

What do you conclude?
Model averaging only works for internal parameters and not derived parameters (groan), but you could write your own model averaging routine (double groan).

Single Species; Multi-Season - NSO - *RPresence*

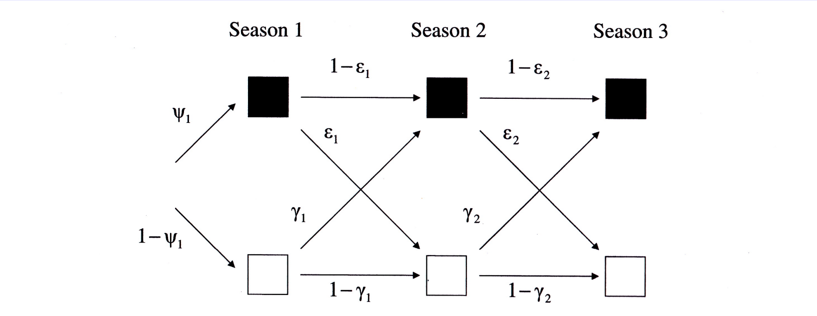
What model would represent RANDOM occupancy over seasons, i.e. no site fidelity between seasons?

A RANDOM occupancy model says that occupancy in each season is a random process with no regard to sites being occupied or not occupied in previous seasons. For example, migrating birds return each year to select nesting sites but the process each year is independent of previous years. Site could be occupied for multiple years in a row, but this the result of random chance rather than any site fidelty.

Single Species; Multi-Season - NSO - *RPresence*

A RANDOM occupancy model implies that an occupied site in season y has the same chance of being occupied in year $y+1$ (i.e. $(1-\epsilon_y)$ as does an unoccupied site in season y being occupied in year $y+1$ (i.e. γ_y).

Or ... RANDOM occupancy $\rightarrow \gamma_y = (1-\epsilon_y)$.



Single Species; Multi-Season - NSO - *RPresence*

A RANDOM occupancy model is fit using *type=do.4*. Now the parameters are ψ (now for each season), and *p* with $\gamma = 1 - \epsilon$ enforced internally depending on estimates of ψ for each year.

Not that this differs from the parameterization adopted by *PRESENCE*.

Single Species; Multi-Season - NSO - *RPresence*

Model: $\psi(1997), p(year), \epsilon(year), \gamma = 1 - \epsilon$
This is equivalent to Model: $\psi(year), p(year)$ where $\gamma = 1 - \epsilon$ is enforced internally).

Single Species; Multi-Season - NSO - *RPresence*

Fit model: $\psi(year), p(year)$ random occupancy model.

```
mod.psiDot.pYear.RO <- RPresence::occMod(
          model=list(psi~SEASON, p~SEASON),
          type="do.4",
          data=nso.pao)
```

Doesn't work properly in this version of *RPresence*. (groan)

Single Species; Multi-Season - NSO - *RPresence*

Fit model: $\psi(1997), p(year), \epsilon(*), \gamma = 1 - \epsilon$

This would represent a steady trend in occupancy over time (why)?
Not possible in *RPresence*.

Single Species; Multi-Season - NSO - *RPresence*

What model would represent a population in equilibrium in occupancy?

Single Species; Multi-Season - NSO - *RPresence*

What model would represent a population in equilibrium in occupancy? $\psi_{y+1} = \psi_y \rightarrow \psi_{EQ} = \frac{\gamma}{\gamma+\epsilon}$

Not possible in *RPresence*.

Single Species; Multi-Season - NSO - *RPresence*

Conclusion:

- Changed in occupancy best represented by a Markov process (as expected), i.e. site fidelity at the species level (but not necessarily at the individual level).
- Random occupancy or no changes in occupancy have no support.
- Highest weight given to models in equilibrium (90% of model weights). (not possible in *RPresence*)

RPresence needs more work!

- Derived estimates don't include population growth.
- Random occupancy models not flexible enough.

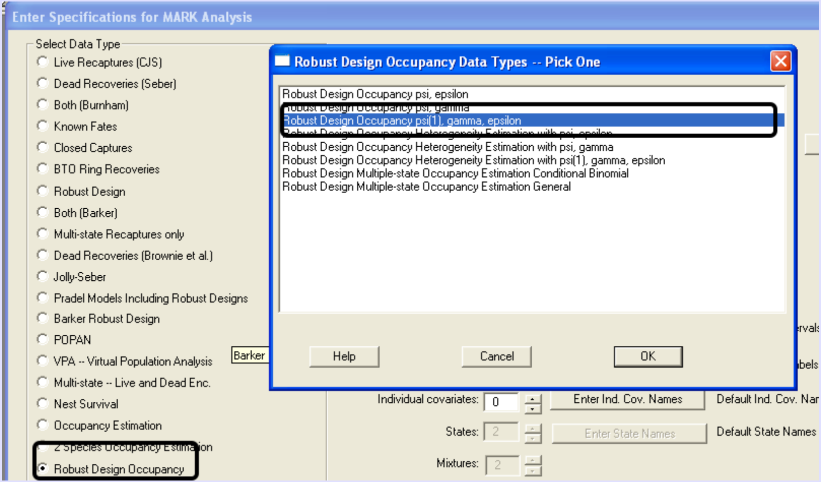
Single Species; Multi-Season - Exampe

Single-Species Multi-Season Occupancy Studies

Analysis of Northern Spotted Owl study using *MARK*.

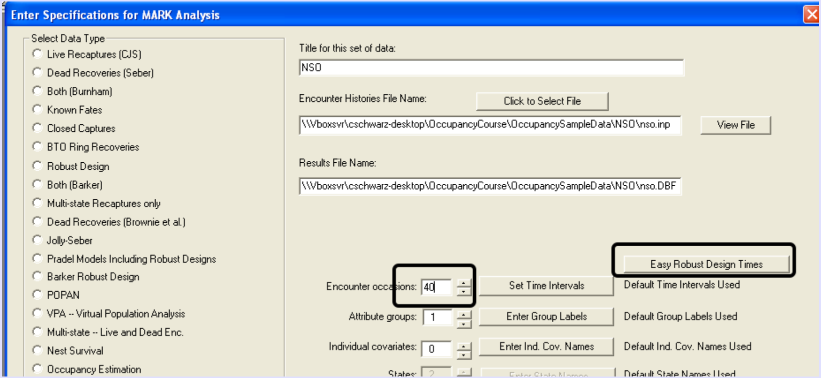
Single Species; Multi-Season - NSO - *MARK*

Start a new project and select the **Robust Design Occupancy** and the parameterization for the multi-season study.



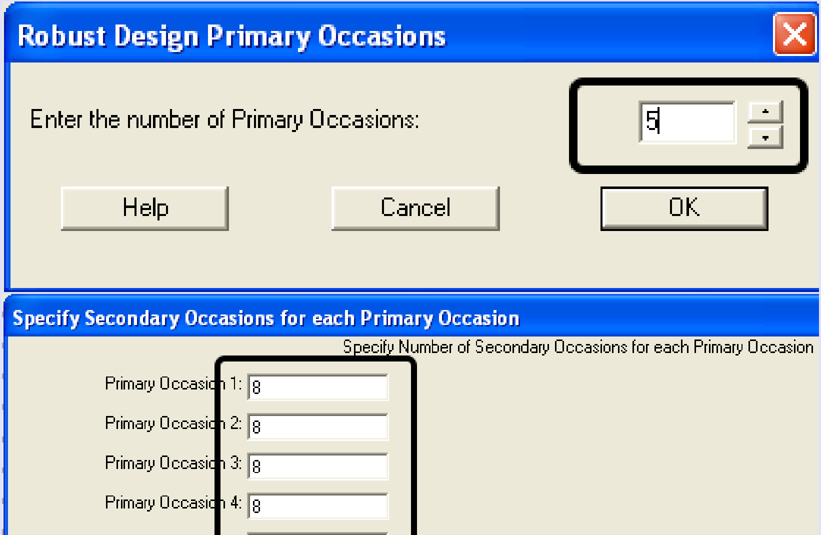
Single Species; Multi-Season - NSO - *MARK*

Open the *NSO.inp* dataset in the OccupancySampleData folder.
View the data.
Enter the TOTAL number of sampling occasions and ...



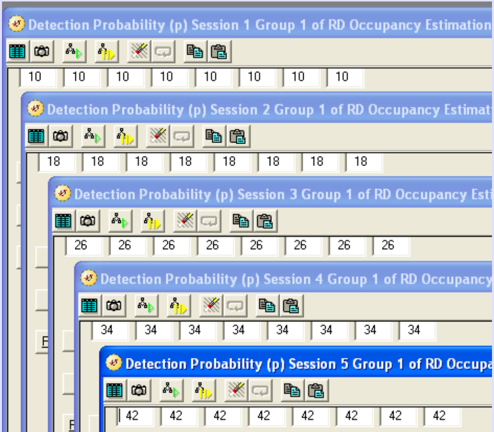
Single Species; Multi-Season - NSO - *MARK*

... split into the primary sessions:



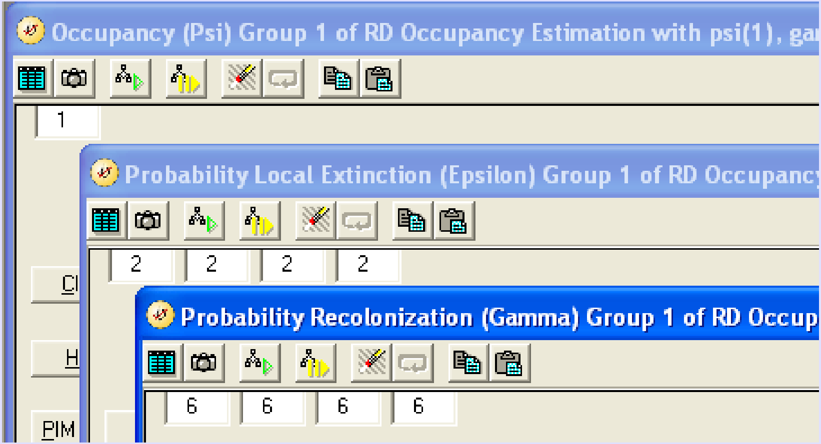
Single Species; Multi-Season - NSO - *MARK*

Fit the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ model in the usual way. LOTS of PIMS! Remember that the actual entries in the PIM are not that important – what is important is the pattern of the entries.



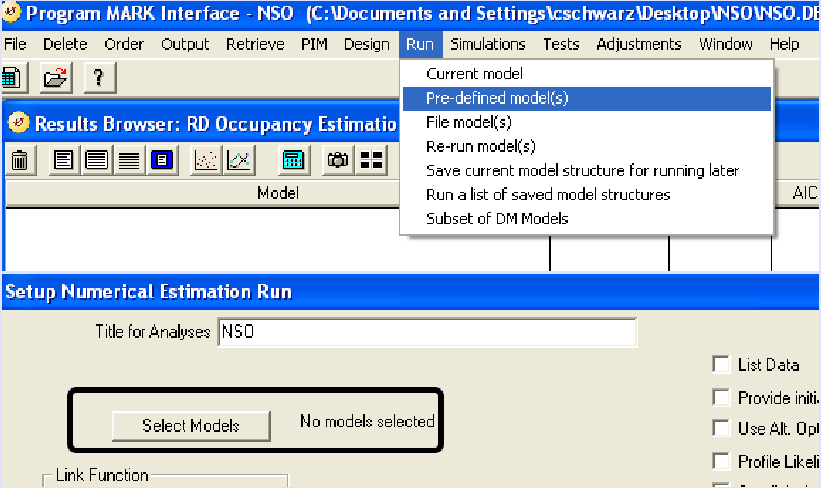
Single Species; Multi-Season - NSO - *MARK*

Fit the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ model in the usual way.
More PIMS!



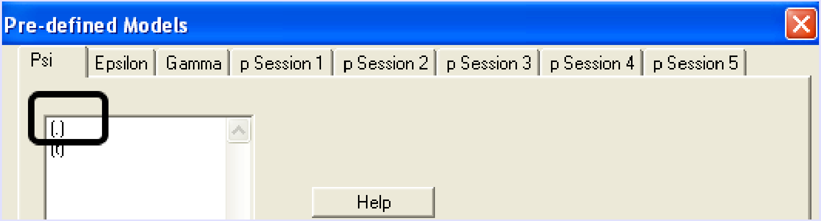
Single Species; Multi-Season - NSO - *MARK*

Specifying the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ model in a different way:



Single Species; Multi-Season - NSO - *MARK*

Specifying the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ model in a different way:



. . . and repeat for other parameter sets.

Single Species; Multi-Season - NSO - *MARK*

Fit the $\psi(1997), p(year), \epsilon(*), \gamma(*)$ and look at estimates:

Results Browser: RD Occupancy Estimation with psi(1), gamma, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1354.0640	0.0000	1.00000	1.0000	8	1337.5226	1337.5226

```
                                   NSO

Real Function Parameters of {Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p S
                                                              95% Confidence Interval
Parameter                   Estimate       Standard Error      Lower          Upper
-------------------------  --------------  --------------  --------------  --------------
   1:Psi                   0.6247321       0.0668928       0.4876127       0.7443925
   2:Epsilon               0.1422939       0.0328023       0.0892233       0.2193294
   3:Gamma                 0.1788328       0.0430139       0.1092550       0.2788491
   4:p Session 1           0.5905766       0.0394282       0.5116838       0.6650665
   5:p Session 2           0.5224659       0.0404720       0.4432412       0.6005762
   6:p Session 3           0.4074917       0.0441805       0.3245375       0.4960771
   7:p Session 4           0.3848502       0.0412514       0.3077762       0.4681714
   8:p Session 5           0.5365859       0.0391612       0.4595708       0.6118943
```

Single Species; Multi-Season - NSO - *MARK*

$\psi(1997), p(year), \epsilon(*), \gamma(*)$ DERIVED parameter estimates:

NSO

Estimates of Derived Parameters

Psi Estimates of {Psi(.) Epsilon(.) Gamma(.) p session 1(.) p session 2(.

Grp.	Occ.	Psi-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.6247321	0.0668928	0.4936223	0.7558420
1	2	0.6029468	0.0514331	0.5021380	0.7037556
1	3	0.5881573	0.0515301	0.4871582	0.6891563
1	4	0.5781171	0.0569602	0.4664751	0.6897591
1	5	0.5713011	0.0625953	0.4486143	0.6939878

Lambda Estimates of {Psi(.) Epsilon(.) Gamma(.) p session 1(.) p session

Grp.	Occ.	Lambda-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.9651285	0.0506787	0.8657983	1.0644587
1	2	0.9754713	0.0370061	0.9029394	1.0480032
1	3	0.9829294	0.0265277	0.9309350	1.0349238
1	4	0.9882100	0.0187782	0.9514047	1.0250152

Lambda' Estimates of {Psi(.) Epsilon(.) Gamma(.) p session 1(.) p sessior

Grp.	Occ.	Lambda'-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	1.0211569	0.0397217	0.9433023	1.0990115
1	2	1.0118058	0.0192922	0.9739931	1.0496185
1	3	1.0068920	0.0099491	0.9873917	1.0263922
1	4	1.0041757	0.0054513	0.9934911	1.0148603

Single Species; Multi-Season - NSO - *MARK*

Fit the $\psi(1997), p(year), \epsilon(year), \gamma(year)$ model.

Results Browser: RD Occupancy Estimation with psi(1), gamma, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1354.0640	0.0000	0.83135	1.0000	8	1337.5226	1337.5226
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1357.2545	3.1905	0.16865	0.2029	14	1327.6391	1327.6391

```
                                   NSO

Real Function Parameters of {Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.)
                                                             95% Confidence Interval
Parameter                  Estimate       Standard Error      Lower           Upper
-------------------------  --------------  --------------  --------------  ----------
   1:Psi                   0.6295098       0.0666177       0.4925683       0.7483738
   2:Epsilon               0.0885534       0.0503042       0.0278409       0.2479005
   3:Epsilon               0.1340050       0.0672305       0.0473582       0.3250840
   4:Epsilon               0.2391323       0.0868726       0.1097536       0.4448196
   5:Epsilon               0.1188086       0.0620459       0.0404908       0.3010767
   6:Gamma                 0.1075657       0.0738889       0.0259703       0.3526955
   7:Gamma                 0.0691959       0.0642430       0.0104145       0.3443152
   8:Gamma                 0.3861798       0.1059457       0.2076074       0.6017136
   9:Gamma                 0.1162734       0.0866811       0.0245636       0.4073849
  10:p Session 1           0.5893491       0.0395458       0.5102508       0.6640827
  11:p Session 2           0.5193206       0.0402749       0.4405559       0.5971365
  12:p Session 3           0.4145854       0.0446560       0.3305567       0.5038963
  13:p Session 4           0.3842570       0.0433030       0.3035953       0.4718298
  14:p Session 5           0.5360559       0.0393535       0.4586769       0.6117368
```

Single Species; Multi-Season - NSO - *MARK*

$\psi(1997), p(year), \epsilon(year), \gamma(year)$ derived estimates

NSO

Estimates of Derived Parameters

Psi Estimates of {Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p

Grp.	Occ.	Psi-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.6295098	0.0666177	0.4989390	0.7600805
1	2	0.6136166	0.0668921	0.4825081	0.7447251
1	3	0.5581251	0.0701563	0.4206187	0.6956315
1	4	0.5953025	0.0715564	0.4550520	0.7355531
1	5	0.5716310	0.0678916	0.4385636	0.7046985

Lambda Estimates of {Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.

Grp.	Occ.	Lambda-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.9747531	0.0691625	0.8391945	1.1103117
1	2	0.9095664	0.0793825	0.7539767	1.0651562
1	3	1.0666113	0.1510669	0.7705203	1.3627024
1	4	0.9602362	0.0904211	0.7830108	1.1374616

Lambda' Estimates of {Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(

Grp.	Occ.	Lambda'-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	1.0165677	0.0472816	0.9238957	1.1092397
1	2	1.0401964	0.0476560	0.9467905	1.1336022
1	3	0.9768712	0.0538263	0.8713716	1.0823709
1	4	1.0164022	0.0416815	0.9347065	1.0980978

Single Species; Multi-Season - NSO - *MARK*

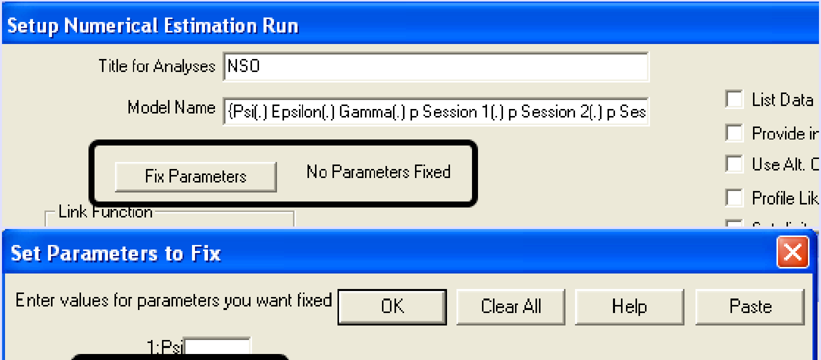
What model would represent NO CHANGE IN OCCUPANCY over the multiple seasons?

Single Species; Multi-Season - NSO - *MARK*

What model would represent NO CHANGE IN OCCUPANCY over the multiple seasons, i.e. convert this to an ‘equivalent single season model=?

$\gamma_y = 0$ and ϵ_y=0 for all years.

Fit the model and FIX some parameters:



Single Species; Multi-Season - NSO - *MARK*

$\psi(1997), p(year), \epsilon(NONE), \gamma(NONE)$ has no support.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1354.0640	0.0000	0.83135	1.0000	8	1337.5226	1337.5226
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1357.2545	3.1905	0.16865	0.2029	14	1327.6391	1327.6391
{Psi(1997) Epsilon(0) Gamma(0) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1554.8776	200.8136	0.00000	0.0000	6	1542.5642	1542.5642

Single Species; Multi-Season - NSO - *MARK*

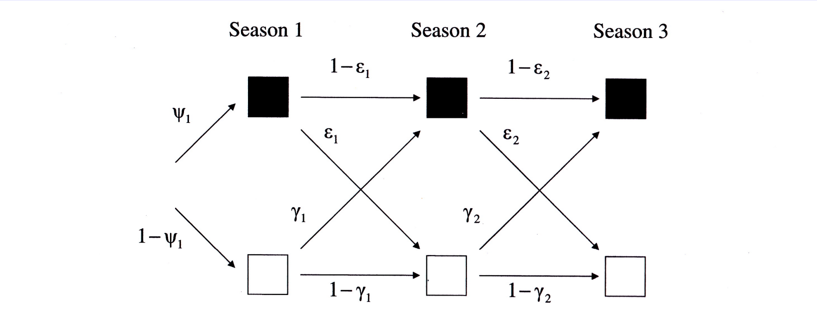
What model would represent RANDOM occupancy over seasons, i.e. no site fidelity between seasons?

A RANDOM occupancy model says that occupancy in each season is a random process with no regard to sites being occupied or not occupied in previous seasons. For example, migrating birds return each year to select nesting sites but the process each year is independent of previous years. Site could be occupied for multiple years in a row, but this the result of random chance rather than any site fidelty.

Single Species; Multi-Season - NSO - *MARK*

A RANDOM occupancy model implies that an occupied site in season y has the same chance of being occupied in year $y+1$ (i.e. $(1-\epsilon_y)$ as does an unoccupied site in season y being occupied in year $y+1$ (i.e. γ_y).

Or, ... RANDOM occupancy $\rightarrow \gamma_y = (1-\epsilon_y)$.



Single Species; Multi-Season - NSO - *MARK*

RANDOM occupancy implies that $\gamma_y = (1 - \epsilon_y)$.
This is tricky!

Recall that

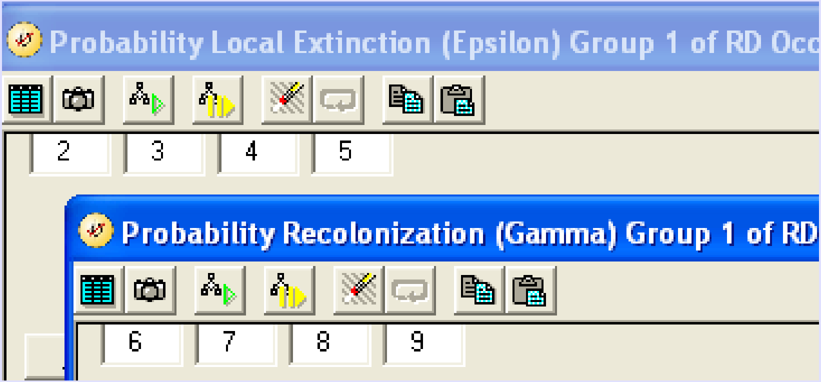
- $logit(\gamma) = \log \frac{\gamma}{1-\gamma}$
- $logit(\epsilon) = \log \frac{\epsilon}{1-\epsilon}$
- $-logit(\epsilon) = -\log \frac{\epsilon}{1-\epsilon} = \log \frac{1-\epsilon}{\epsilon}$

So ... if $logit(\gamma) = -logit(\epsilon)$ then $\log \frac{\gamma}{1-\gamma} = \log \frac{1-\epsilon}{\epsilon}$ and this implies $\epsilon = (1 - \gamma)$ and vice-versa.

Single Species; Multi-Season - NSO - *MARK*

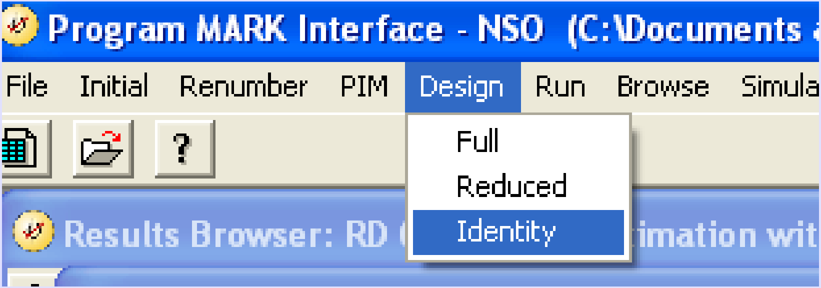
RANDOM occupancy $\rightarrow \gamma_y = (1 - \epsilon_y)$ needs $logit(\gamma) = -logit(\epsilon)$ and this is specified using the DESIGN Matrix of MARK:

Fit model: $\psi(1997), p(year), \epsilon(year), \gamma = 1 - \epsilon$ by starting with the PIMs in the usual way ...:



Single Species; Multi-Season - NSO - *MARK*

RANDOM occupancy $\rightarrow \gamma_y = (1 - \epsilon_y)$ needs $logit(\gamma) = -logit(\epsilon)$ and this is specified using the DESIGN Matrix of MARK:
... and then ask for the design matrix ...:



Single Species; Multi-Season - NSO - *MARK*

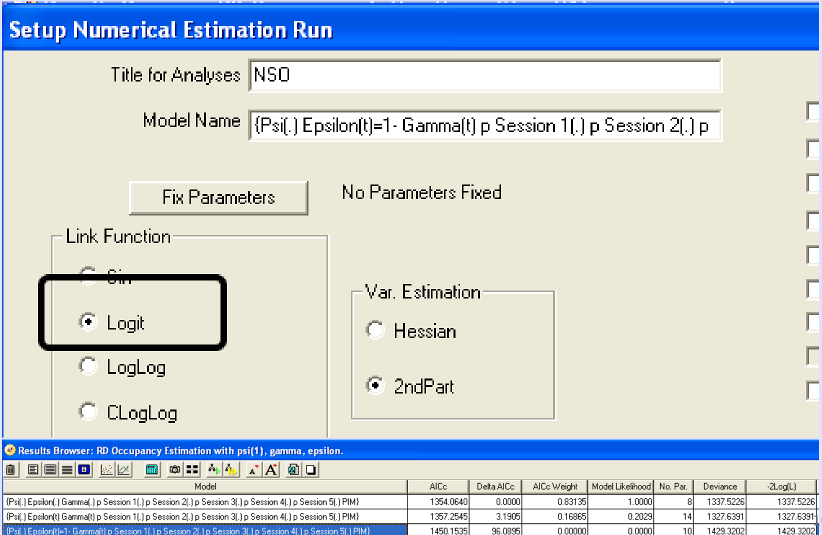
RANDOM occupancy $\rightarrow \gamma_y = (1 - \epsilon_y)$ needs $logit(\gamma) = -logit(\epsilon)$.
... and modify the design matrix to be ... (why?)

Design Matrix Specification: RD Occupancy Estimation with psi(1), gamma, epsi

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

Single Species; Multi-Season - NSO - *MARK*

RUN the model and specify the LOGIT link:



Single Species; Multi-Season - NSO - *MARK*

Fit model: $\psi(1997), p(year), \epsilon(*), \gamma = 1 - \epsilon$

Design Matrix Specification: RD Occupancy Estimation wit

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

Single Species; Multi-Season - NSO - *MARK*

Fit model: $\psi(1997), p(year), \epsilon(*), \gamma = 1 - \epsilon$

Results Browser: RD Occupancy Estimation with psi(1), gamma, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1354.0640	0.0000	0.83135	1.0000	8	1337.5226	1337.5226
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1357.2545	3.1905	0.16865	0.2029	14	1327.6391	1327.6391
{Psi(.) Epsilon(.)=1-Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1443.9512	89.8872	0.00000	0.0000	7	1429.5317	1429.5317
{Psi(.) Epsilon(t)=1- Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1450.1535	96.0895	0.00000	0.0000	10	1429.3202	1429.3202
{Psi(1997) Epsilon(0) Gamma(0) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1554.8776	200.8136	0.00000	0.0000	6	1542.5642	1542.5642

Single Species; Multi-Season - NSO - *MARK*

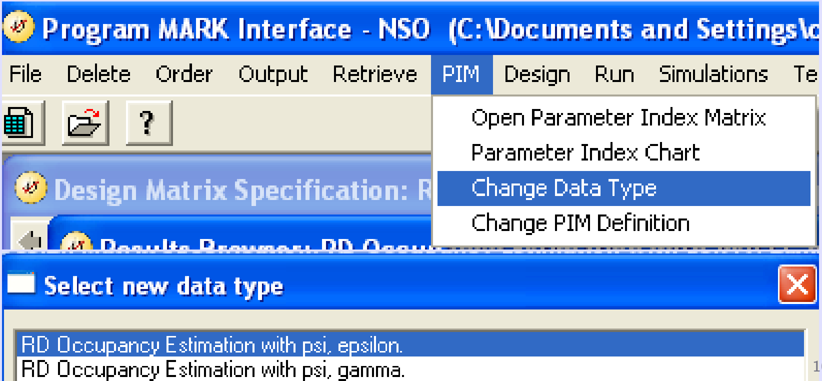
What model would represent a population in equilibrium in occupancy?

Single Species; Multi-Season - NSO - *MARK*

What model would represent a population in equilibrium in occupancy?

$$\psi_{y+1} = \psi_y \rightarrow \psi_{EQ} = \frac{\gamma}{\gamma+\epsilon}$$

Change data type to different parameterization:



Single Species; Multi-Season - NSO - *MARK*

Equilibrium:

Fit: $\psi(*), p(year), \gamma(year)$ (implicitly models $\epsilon(year)$)

Fit: $\psi(*), p(year), \epsilon(year)$ (implicitly models $\gamma(year)$)

Fit: $\psi(*), p(year), \epsilon(*)$ (implicitly models $\gamma(*)$)

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1350.1764	0.0000	0.54350	1.0000	10	1329.3431	1329.3431
{Psi(.) Epsilon(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1352.3704	2.1940	0.18146	0.3339	7	1337.9509	1337.9509
{Psi(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1352.3704	2.1940	0.18146	0.3339	7	1337.9509	1337.9509
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1354.0640	3.8876	0.07781	0.1432	8	1337.5226	1337.5226
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1357.2545	7.0781	0.01578	0.0290	14	1327.6391	1327.6391
{Psi(.) Epsilon(.)=1-Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1443.9512	93.7748	0.00000	0.0000	7	1429.5317	1429.5317
{Psi(.) Epsilon(t)=1- Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1450.1535	99.9771	0.00000	0.0000	10	1429.3202	1429.3202
{Psi(1997) Epsilon(0) Gamma(0) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Session 5(.) PIM}	1554.8776	204.7012	0.00000	0.0000	6	1542.5642	1542.5642

Delete one of the duplicate models before continuing.

Single Species; Multi-Season - NSO - *MARK*

Conclusion:

- Changed in occupancy best represented by a Markov process (as expected), i.e. site fidelity at the species level (but not necessarily at the individual level).
- Random occupancy or no changes in occupancy have no support.
- Highest weight given to models in equilibrium (90% of model weights).

Single Species; Multi-Season - Example

Single-Species Multi-Season Occupancy Studies

Analysis of Northern Spotted Owl study using *RMark*

Single Species; Multi-Season - NSO - *RMark*

Northern Spotted Owl (*Strix occidentalis caurina*) in California.

$s = 55$ sites visited up to $K = 8$ times per season between 1997 and 2001 ($Y = 5$).

Detection probabilities relatively constant within years, but likely different among years.

Single Species; Multi-Season - NSO - *RMark*

Read in data and create the data frame

```
input.data <- read.csv(file.path("..","NSO.csv"), header=F
input.data$V1 <- NULL # drop the site number

input.history <- data.frame(freq=1,
                            ch=apply(input.data,1,paste, co

input.history$ch <- gsub("NA",".", input.history$ch, fixed=
```

```
1     1 0111....010.....11......011.....0101....
2     1 00......000.....000.....000110..0011....
```

Single Species; Multi-Season - NSO - *RMark*

Process the data frame.

```
max.visit.per.year <- 8
n.season <- 5

nso.data <- process.data(data=input.history, model="RDOccup
                         time.intervals=c( rep( c(rep(0,max
                                             rep(0,max.visit.
```

It doesn't matter if you have extra visits that are all NA in a year.

Single Species; Multi-Season - NSO - *RMark*

What parameterization do you want?

```
setup.parameters("RDOccupEG", check=TRUE)
[1] "Psi"     "Epsilon" "Gamma"   "p"

setup.parameters("RDOccupPE", check=TRUE) # psi, epsilon, p
[1] "Psi"     "Epsilon" "p"

setup.parameters("RDOccupPG", check=TRUE) # psi, gamma, p
[1] "Psi"   "Gamma" "p"
```

No random occupancy (but contact me for details)

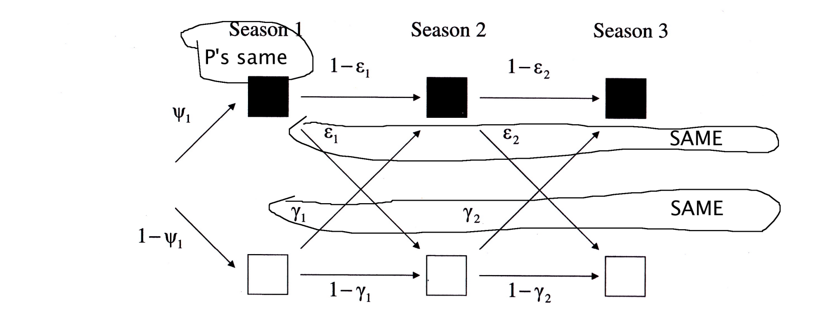
Single Species; Multi-Season - NSO - *RMark*

Model with colonization and extinction probabilities parameterization.
Fit the $\psi(1997), \gamma(*), \epsilon(*), p(*)$ model in the usual way.

```
mod.fit <- RMark::mark(nso.data,
                       model="RDOccupEG",
                       model.parameters=list(
                         Psi    =list(formula=~1), # initial oc
                         p      =list(formula=~1),
                         Epsilon=list(formula=~1),
                         Gamma  =list(formula=~1)))
```

Single Species; Multi-Season - NSO - *RMark*

Helpful to draw a diagram of the process model:



Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ Results of model fit.

```
Output summary for RDOccupEG model
Name : Psi(~1)Epsilon(~1)Gamma(~1)p(~1)

Npar :  4
-2lnL:  1355.315
AICc :  1363.464
```

Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ parameter estimates.

```
get.real(mod.fit, parameter="Psi", se=TRUE)
             all.diff.index par.index estimate        se
Psi g1 a0 t1              1         1 0.631162 0.0672653 0.
             fixed    note group age time Age Time
Psi g1 a0 t1                    1   0    1   0    0
```

Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ parameter estimates.

```
>get.real(mod.fit, parameter="Epsilon", se=TRUE)
                 all.diff.index par.index  estimate
Epsilon g1 a0 t1              2         2 0.1507363 0.03322
Epsilon g1 a1 t2              3         2 0.1507363 0.03322
Epsilon g1 a2 t3              4         2 0.1507363 0.03322
Epsilon g1 a3 t4              5         2 0.1507363 0.03322
                       ucl fixed    note group age time Age
Epsilon g1 a0 t1 0.2279232                  1   0    1   0
Epsilon g1 a1 t2 0.2279232                  1   1    2   1
Epsilon g1 a2 t3 0.2279232                  1   2    3   2
Epsilon g1 a3 t4 0.2279232                  1   3    4   3
```

Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ parameter estimates.

```
>get.real(mod.fit, parameter="Gamma", se=TRUE)

               all.diff.index par.index  estimate         se
Gamma g1 a0 t1              6         3 0.1841758 0.0427184
Gamma g1 a1 t2              7         3 0.1841758 0.0427184
Gamma g1 a2 t3              8         3 0.1841758 0.0427184
Gamma g1 a3 t4              9         3 0.1841758 0.0427184
                     ucl fixed    note group age time Age T
Gamma g1 a0 t1 0.2827079                   1   0    1   0
Gamma g1 a1 t2 0.2827079                   1   1    2   1
Gamma g1 a2 t3 0.2827079                   1   2    3   2
Gamma g1 a3 t4 0.2827079
```

Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ derived parameter estimates of occupancy in later years

```
> mod.fit$results$derived$"psi Probability Occupied"
   estimate         se       lcl       ucl
1 0.6311620 0.06726525 0.4993221 0.7630018
2 0.6039540 0.05095844 0.5040754 0.7038325
3 0.5858583 0.05112760 0.4856482 0.6860684
4 0.5738230 0.05661538 0.4628569 0.6847892
5 0.5658186 0.06213794 0.4440282 0.6876089
```

Single Species; Multi-Season - NSO - *RMark*

Model $\psi(1997), \gamma(*), \epsilon(*), p(*)$ derived parameter estimates of population growth

```
> mod.fit$results$derived$"lambda Rate of Change"
   estimate         se       lcl      ucl
1 0.9568922 0.05137574 0.8561958 1.057589
2 0.9700379 0.03736524 0.8968021 1.043274
3 0.9794571 0.02655675 0.9274059 1.031508
4 0.9860506 0.01858500 0.9496240 1.022477
```

Single Species; Multi-Season - NSO - *RMark*

Model averaging can take place in the usual way. See R code.
You can mix the different model parameterizations. See R code.

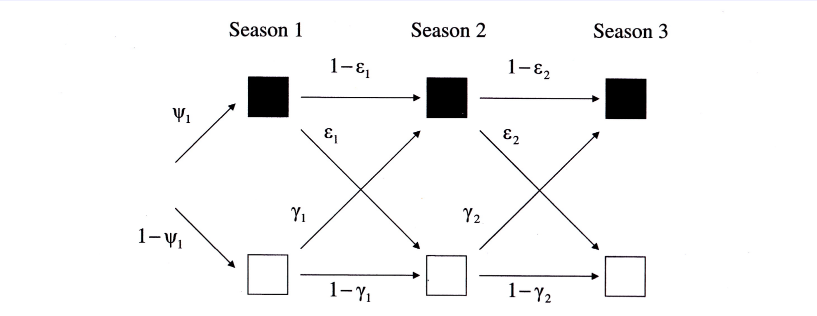
Single Species; Multi-Season - NSO - *RMark*

A RANDOM occupancy model says that occupancy in each season is a random process with no regard to sites being occupied or not occupied in previous seasons. For example, migrating birds return each year to select nesting sites but the process each year is independent of previous years. Site could be occupied for multiple years in a row, but this the result of random chance rather than any site fidelty.

Single Species; Multi-Season - NSO - *RMark*

A RANDOM occupancy model implies that an occupied site in season y has the same chance of being occupied in year $y+1$ (i.e. $(1-\epsilon_y)$ as does an unoccupied site in season y being occupied in year $y+1$ (i.e. γ_y).

Or ... RANDOM occupancy $\rightarrow \gamma_y = (1-\epsilon_y)$.



Single Species; Multi-Season - NSO - *RMark*

This is fit in *RMark* by stacking the data and using a single season model.
See me for details or look at *R* code.

Single Species; Multi-Season - NSO - *RMark*

Conclusion:

- Changed in occupancy best represented by a Markov process (as expected), i.e. site fidelity at the species level (but not necessarily at the individual level).
- Random occupancy is seldom a good model.

Single Species; Multi-season - Skinks - Exercise 01

Grand Skinks.

Data has been collected on 352 tors over a 5 year (the "seasons") period, although not all tors (rock piles) were surveyed each year, with up to 3 surveys of each tor per year.

There is also a site-specific covariate Pasture indicating whether the surrounding matrix is either predominately the modified habitat (farm pasture, Pasture =1) or "native" grassland (tussock, Pasture = 0).

Single Species; Multi-season - Skinks - Exercise 01

Grand Skinks.

Fit $\psi(.), \gamma(.), \epsilon(.), p(.)$ model.
What do you conclude?

Single Species; Multi-season - Skinks - Exercise 01

Grand Skinks.

Fit $\psi(.), \gamma(.), \epsilon(.), p(.)$ model.

- the probability of occupancy in the first year was 0.39;
- between all seasons, the probability that skinks colonize a previously unoccupied rocky outcrop is 0.07;
- between all seasons, the probability that skinks go locally extinct from an occupied rocky outcrop is 0.10;
- given an outcrop is occupied by skinks, the probability of detecting skinks in a single survey is 0.69.

Single Species; Multi-season - Skinks - Exercise 01

Grand Skinks.

Fit $\psi(.), \gamma(year), \epsilon(year), p(year)$

Other models to fit (for later exercises)

- Is the occupancy the same for all outcrops or is it different for outcrops surrounded by pasture?
- Colonization appears to vary over time, but does the intervening matrix have an effect and is it consistent each year?
- Local Extinction appears to vary over time, but does the intervening matrix have an effect and is it consistent each year?
- Is there evidence that occupancy is changing over seasons in a "linear" fashion?

Single Species; Multi-season - Covariates

Covariates can be used for:

- Compare occupancy probabilties among habitat types. For example, is the occupancy in spruce sites different than the occupancy in lodgepole pine units.
- Reduce heterogeneity. For example, detectability may differ between spruce sites and lodgepole sites.
- **Study trends in occupancy over seasons**.
- **Colonization/extinction as function of patch area**.

Single Species; Multi-season - Covariates

Two classes of covariates:

- External covariates that affect all sites simultaneously. For example, rain on a sampling occasion may reduce detectability even though effort is same at all occasions.
- Site covariates specific to a site. For example, habitat would be measured on individual sites. These could vary over time for each site as well, but then need to be measured at each sampling occasion.
- **Season covariates. Apply to all sites in a season**.

Single Species; Multi-season - Covariates

Two types of covariates:

- Continuous, e.g. occupancy is a function of stem density. Use the value directly as the covariate.
- Discrete, e.g. occupancy in spruce or lodgepole pine. Create indicator variables (0,1) for category membership. If there are M categories of the covariate, there are $M-1$ indicator variables. When using *RPresence*, *RMark*, *unmarked*, or *JAGS* you can use categorical variables and *R* will take care of creating indicator variables internally.

Single Species; Multi-season - Covariates

Modeling covariate effects - the *logit* (also known as the *log-odds*) link.
If we model $\psi_{year} = \beta_0 + \beta_1 year$, then as *year* varies, the predicted probability can be < 0 or > 1 which is non-sensical.

Consequently, we model the effects of covariates on the *logit* scale

$$logit(\psi) = \log \frac{\psi}{1 - \psi}$$

where log() is the natural (to the base *e*) logarithm.

The inverse transform is:

$$\psi = \frac{1}{1 + e^{-logit}}$$

Model is then $logit(\psi_{year}) = \beta_0 + \beta_1 year$.

Covariates and Design Matrices

Every covariate creates a design matrix that links the covariate values to the parameter.

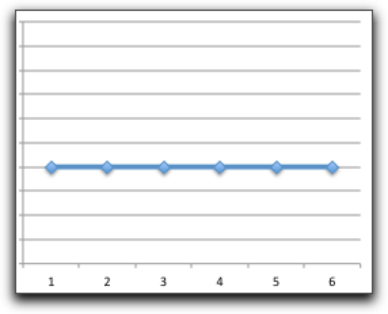
- Intercept (typically the first column) is the baseline.
- Categorical covariates create columns of $1/0$ with $K-1$ columns for K levels.
- Continuous covariates create columns with the covariate value.

Alternate design matrices are possible.

The design matrices are typically hidden from the user when using *RPresence* or *RMark*.

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

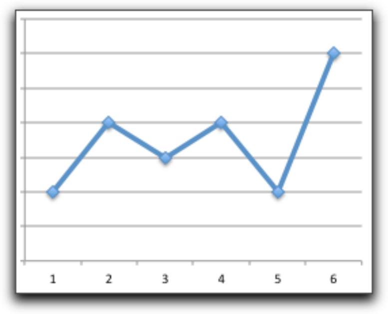


Model $p(*)$.

Index	Design matrix
1	1
2	1
3	1
4	1
5	1
6	1

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

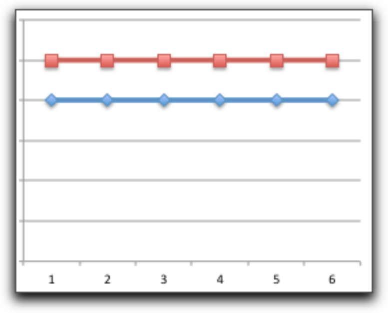


Model $p(t)$.

Index	Design matrix					
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	1	0	0	0
4	0	0	0	1	0	0
5	0	0	0	0	1	0
6	0	0	0	0	0	1

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

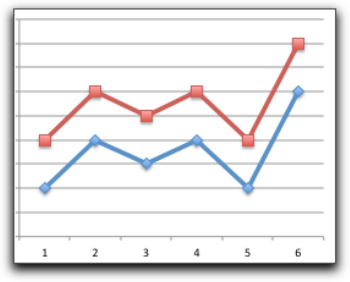


Model $p(H)$.

Index	Design matrix	
1	1	H
2	1	H
3	1	H
4	1	H
5	1	H
6	1	H

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

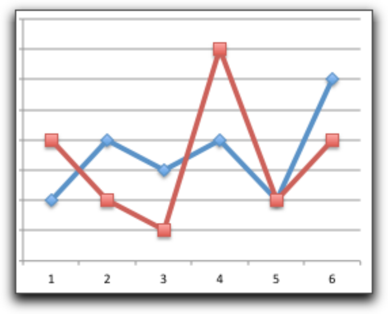


Model $p(t + H)$.

Index	Design matrix					
1	1	0	0	0	0	0	H
2	0	1	0	0	0	0	H
3	0	0	1	0	0	0	H
4	0	0	0	1	0	0	H
5	0	0	0	0	1	0	H
6	0	0	0	0	0	1	H

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

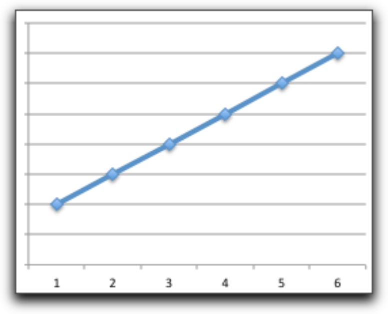


Model $p(t * H)$.

Index	Design matrix							
1	1	0	0	0	0	0	H	0
2	0	1	0	0	0	0	0	H
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	0	1	0	0

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

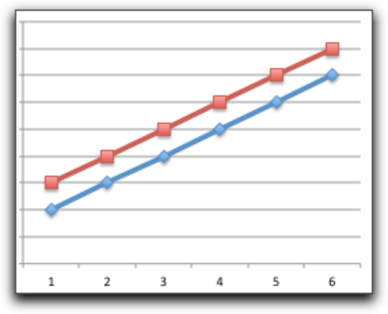


Model $p(Linear)$.

Index	Design matrix
1	1
2	2
3	3
4	4
5	5
6	6

Different Design Matrices

H = indicator of habitat type (2 levels, 0 or 1); modelling detectability over multiple surveys.

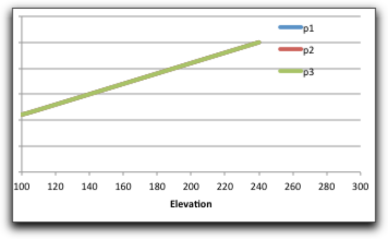


Index	Design matrix	
1	1	H
2	2	H
3	3	H
4	4	H
5	5	H
6	6	H

Model $p(Linear + H)$.

Different Design Matrices

EL =continuous covariate (elevation); modelling detectability as elevation changes

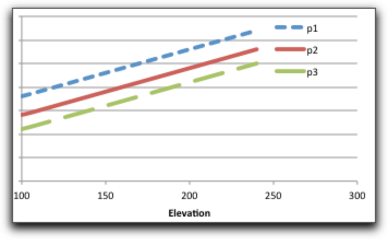


Index	Design matrix	
1	1	EL_1
2	1	EL_2
3	1	EL_3
4	1	EL_4
5	1	EL_5
6	1	EL_6

Model $p(Elev)$ – Lines are co-incident.

Different Design Matrices

EL =continuous covariate (elevation); modelling detectability as elevation changes

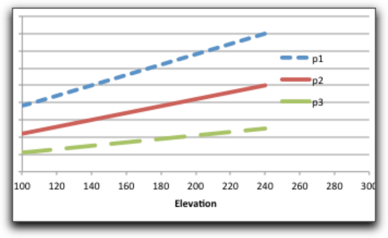


Model $p(t + ELEV)$.

Index	Design matrix					
1	1	0	0	0	0	0	EL_1
2	0	1	0	0	0	0	EL_2
3	0	0	1	0	0	0	EL_3
4	0	0	0	1	0	0	EL_4
5	0	0	0	0	1	0	EL_5
6	0	0	0	0	0	1	EL_6

Different Design Matrices

EL =continuous covariate (elevation); modelling detectability as elevation changes



Index				Design matrix		
1	1	0	0	0	EL_1	0
2	0	1	0	0	0	EL_2
3	0	0	1	0	0	0
4	0	0	0	1	0	0

Model $p(t * Elev)$.

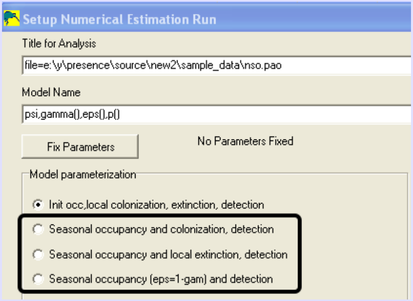
Single Species; Multi-season - Covariates - PRESENCE

Return to the Spotted Owl problem.
Is there evidence that occupancy is changing over time in a "linear" trend? Reopen the Spotted Owl project.

Single Species; Multi-season - Covariates - PRESENCE

Return to the Spotted Owl problem.
Is there evidence that occupancy is changing over time in a "linear" trend?

Need to choose a parameterization where ψ can be modeled as changing over seasons:



Single Species; Multi-season - Covariates - PRESENCE

$logit(\psi_{year}) = \beta_0 + \beta_1 year \rightarrow$ modify the DESIGN matrix for ψ:

File Init Retrieve model special

Occupancy Colonization

-	a1	a2
psi1	1	1
psi2	1	2
psi3	1	3
psi4	1	4
psi5	1	5

Also make proper DESIGN matrix for ϵ (seasonal effects) and p (seasonal effects)

Single Species; Multi-season - Covariates - PRESENCE

$$logit(\psi_{year}) = \beta_0 + \beta_1 year$$

Model	AIC	deltaAIC	AIC wgt	Model Likel	no.Par.	-2*LogLike
psi(*), gamma(year), p(year)	1346.81	0.00	0.6360	1.0000	10	1326.81
psi(LinearTren), gamma(year), p(year)	1347.93	1.12	0.3633	0.5712	11	1325.93
psi(),gamma(),p()	1360.57	13.76	0.0007	0.0010	3	1354.57

No real support for the LINEAR trend model relative to CONSTANT occupancy model.

Single Species; Multi-season - Covariates - PRESENCE

$logit(\psi_{year}) = \beta_0 + \beta_1 year$ estimates of trend and actual ψ_{year} (don't forget that change is on the logit scale).

```
Untransformed Estimates of coefficients for covariates (Beta's)
=======================================================================
                                          estimate     std.error
A1   psi1                             :   0.636681     0.345221
A2   psi1                             :  -0.081870     0.086608
```

```
   Individual Site estimates of <psi1>
                Site               estimate   Std.err    95% conf. interval
psi1            1 site 1         :  0.6353    0.0650      0.5012 - 0.7511
psi2            1 site 1         :  0.6161    0.0546      0.5052 - 0.7161
psi3            1 site 1         :  0.5965    0.0498      0.4964 - 0.6892
psi4            1 site 1         :  0.5767    0.0531      0.4707 - 0.6760
psi5            1 site 1         :  0.5566    0.0638      0.4306 - 0.6757
```

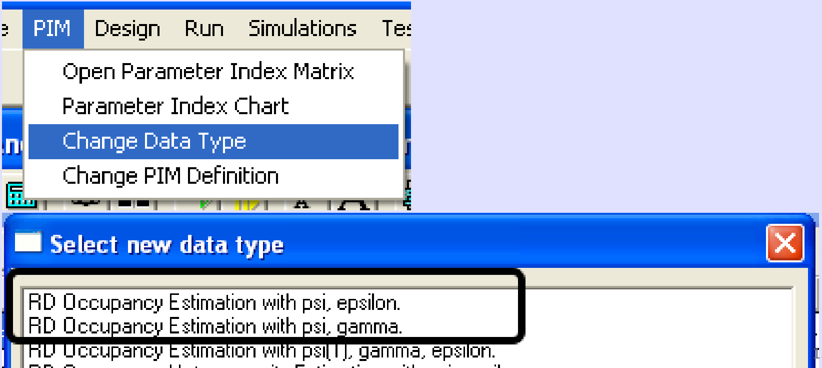
Single Species; Multi-season - Covariates - *RPresence*

The second and third parameterizations of *PRESENCE* are not available in *RPresence*.
Models for *do.1* are fit in a similar way as without covariates.
Models for *do.4* (Random occupancy) has some issues. See me for alternate ways to fit.

Single Species; Multi-season - Covariates - MARK

Return to the Spotted Owl problem.
Is there evidence that occupancy is changing over time in a "linear" trend?

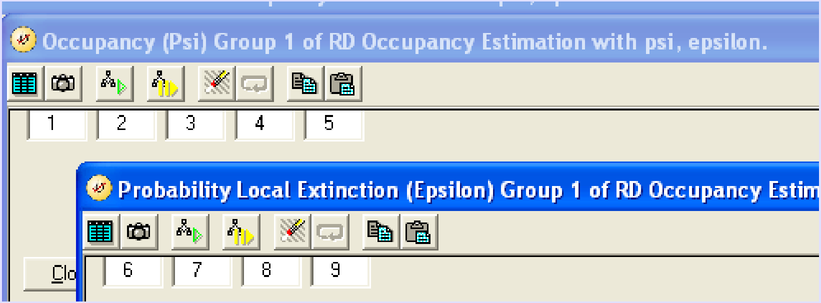
Need to choose a parameterization where ψ can be modeled as changing over seasons:



Single Species; Multi-season - Covariates - MARK

$logit(\psi_{year}) = \beta_0 + \beta_1 year$

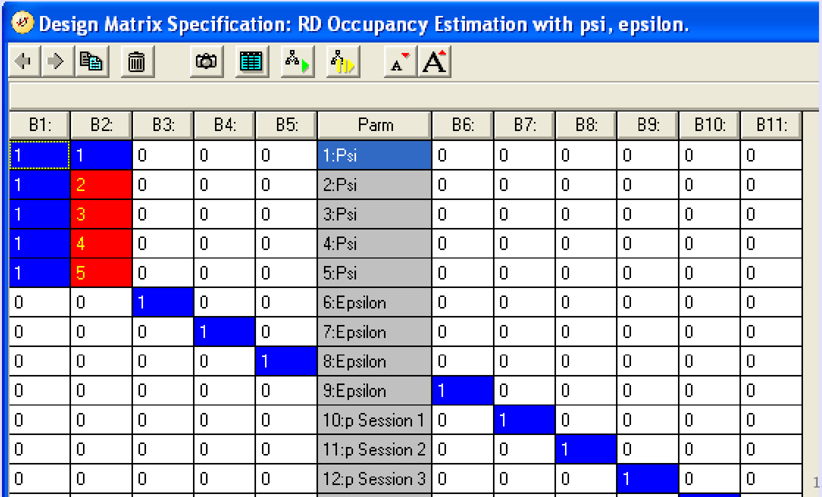
First get the correct PIM structure . . .



Single Species; Multi-season - Covariates - MARK

$$logit(\psi_{year}) = \beta_0 + \beta_1 year$$

ldots and then create the DESIGN matrix as below (why?)



B1:	B2:	B3:	B4:	B5:	Parm	B6:	B7:	B8:	B9:	B10:	B11:
1	1	0	0	0	1:Psi	0	0	0	0	0	0
1	2	0	0	0	2:Psi	0	0	0	0	0	0
1	3	0	0	0	3:Psi	0	0	0	0	0	0
1	4	0	0	0	4:Psi	0	0	0	0	0	0
1	5	0	0	0	5:Psi	0	0	0	0	0	0
0	0	1	0	0	6:Epsilon	0	0	0	0	0	0
0	0	0	1	0	7:Epsilon	0	0	0	0	0	0
0	0	0	0	1	8:Epsilon	0	0	0	0	0	0
0	0	0	0	0	9:Epsilon	1	0	0	0	0	0
0	0	0	0	0	10:p Session 1	0	1	0	0	0	0
0	0	0	0	0	11:p Session 2	0	0	1	0	0	0
0	0	0	0	0	12:p Session 3	0	0	0	1	0	0

Single Species; Multi-season - Covariates - MARK

$logit(\psi_{year}) = \beta_0 + \beta_1 year$ model fit and estimates.

Results Browser: RD Occupancy Estimation with psi, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{Psi(.) Epsilon(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Sessi	1350.1764	0.0000	0.42225	1.0000	10	1329.3431	1329.3431
{Psi(TimeTrend) Epsilon(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4	1351.4526	1.2762	0.22308	0.5283	11	1328.4488	1328.4488
{Psi(.) Epsilon(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Sessi	1352.3704	2.1940	0.14098	0.3339	7	1337.9509	1337.9509
{Psi(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4(.) p Sessi	1352.3704	2.1940	0.14098	0.3339	7	1337.9509	1337.9509
{Psi(.) Epsilon(.) Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4	1354.0640	3.8876	0.06045	0.1432	8	1337.5226	1337.5226
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4	1357.2545	7.0781	0.01226	0.0290	14	1327.6391	1327.6391
{Psi(.) Epsilon(.)=1-Gamma(.) p Session 1(.) p Session 2(.) p Session 3(.) p Session	1443.9512	93.7748	0.00000	0.0000	7	1429.5317	1429.5317
{Psi(.) Epsilon(t)=1- Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Sessio	1450.1535	99.9771	0.00000	0.0000	10	1429.3202	1429.3202
{Psi(1997) Epsilon(0) Gamma(0) p Session 1(.) p Session 2(.) p Session 3(.) p Ses	1554.8776	204.7012	0.00000	0.0000	6	1542.5642	1542.5642

NSO

Real Function Parameters of {Psi(TimeTrend) Epsilon(t) p Session 1(.) p Session 2(.) p S

Parameter	Estimate	Standard Error	95% Confidence Interval Lower	95% Confidence Interval Upper
1:Psi	0.6352511	0.0649330	0.5013865	0.7510231
2:Psi	0.6160796	0.0545485	0.5052291	0.7160526
3:Psi	0.5965404	0.0497796	0.4964216	0.6892158
4:Psi	0.5766901	0.0530951	0.4707600	0.6760099
5:Psi	0.5565891	0.0637712	0.4306688	0.6756351
6:Epsilon	0.0907030	0.0354020	0.0412336	0.1878915
7:Epsilon	0.0963166	0.0425158	0.0393078	0.2173039
8:Epsilon	0.2735640	0.0670179	0.1627922	0.4217408
9:Epsilon	0.1220746	0.0530793	0.0500370	0.2685094

Look at beta parameters for slope and estimate on logit scale.

Single Species; Multi-season - Covariates - *RMark*

Fit in much the same way as single season models. See *R* code.

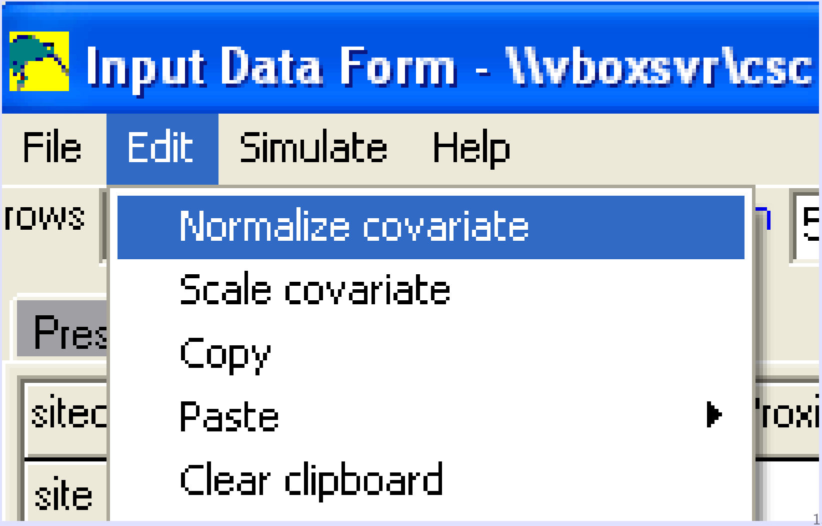
Single Species; Multi-season - Exercise I

Blue-Ridge salamander measured over 2 years.

Investigate effect of elevation and stream proximity on occupancy, local extinction, and colonization.
There are 2 years x 5 visits per year per site. Only one ϵ and only one γ (why?)

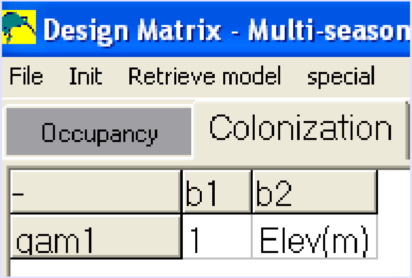
Single Species; Multi-season - Exercise I - PRESENCE

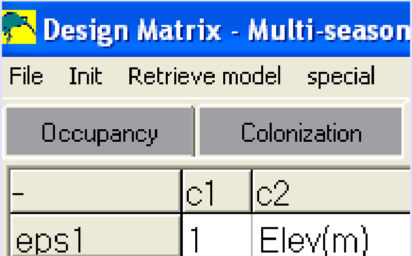
Blue-Ridge salamander. Standardize Elevation covariate (why?)



Single Species; Multi-season - Exercise I - PRESENCE

Blue-Ridge salamander. Specify DESIGN matrix (why)?





Single Species; Multi-season - Exercise I - PRESENCE

Blue-Ridge salamander. Final model results:

Model	AIC	deltaAIC	AIC wgt	Model Likelil	no.Par.	-2*LogLike
psi(StreamProximity),gamma(),eps(),p()	306.74	0.00	0.5181	1.0000	5	296.74
psi(StreamProximity),gamma(),eps(StreamProxim	307.92	1.18	0.2872	0.5543	6	295.92
psi(StreamProximity),gamma(StreamProximity),ep	308.70	1.96	0.1945	0.3753	6	296.70
psi,gamma(),eps(),p()	324.77	18.03	0.0001	0.0001	4	316.77
psi(Elevation),gamma(),eps(),p()	325.68	18.94	0.0000	0.0001	5	315.68
psi,gamma(Elevation),eps(),p()	326.09	19.35	0.0000	0.0001	5	316.09
psi,gamma(),eps(),p(season)	326.19	19.45	0.0000	0.0001	5	316.19
psi,gamma(),eps(Elevation),p()	326.56	19.82	0.0000	0.0000	5	316.56

What do you conclude?

Single Species; Multi-season - Exercise I - MARK

Blue-Ridge salamander measured over 2 years.

Investigate effect of elevation and stream proximity on occupancy, local extinction, and colonization.
There are 2 years x 5 visits per year per site. Only one ϵ or γ (why?)

Single Species; Multi-season - Exercise I - MARK

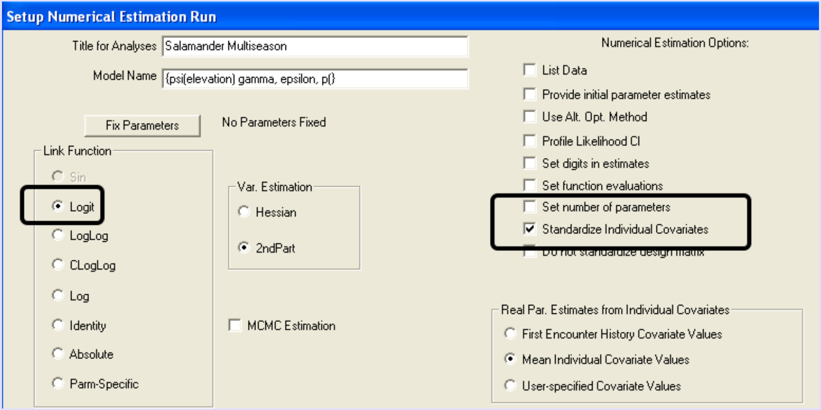
Blue-Ridge salamander. Specify DESIGN matrix (why)?

Design Matrix Specification: RD Occupancy Estimation with psi(

B1:	B5:	B2:	B3:	Parm	B4:
1	Elevatio	0	0	1:Psi	0
0	0	1	0	2:Epsilon	0
0	0	0	1	3:Gamma	0
0	0	0	0	4:p Session 1	1

Single Species; Multi-season - Exercise I - MARK

Blue-Ridge salamander. Standardize the elevation covariate (why)?



Will you standardize the SteamProximity covariate?

Single Species; Multi-season - Exercise I - MARK

Blue-Ridge salamander. Final model results:

Results Browser: RD Occupancy Estimation with psi(1), gamma, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{psi(StreamProx) gamma(StreamProx), epsilon, p()}	307.5325	0.0000	0.41078	1.0000	5	296.6992	296.6992
{psi(StreamProx) gamma, epsilon, p()}	307.5773	0.0448	0.40168	0.9778	5	296.7440	296.7440
{psi(StreamProx) gamma(), epsilon(StreamProx)}	309.1023	1.5698	0.18738	0.4562	6	295.9192	295.9192
{psi, gamma, epsilon, p()}	325.3210	17.7885	0.00006	0.0001	4	316.7730	316.7730
{psi(elevation) gamma, epsilon, p()}	326.5065	18.9740	0.00003	0.0001	5	315.6731	315.6731
{psi() gamma(), epsilon(elevation), p()}	326.9174	19.3849	0.00003	0.0001	5	316.0840	316.0840
{psi, gamma, epsilon, p(year)}	327.0190	19.4865	0.00002	0.0000	5	316.1857	316.1857
{psi() gamma(elevation), epsilon, p()}	327.3919	19.8594	0.00002	0.0000	5	316.5586	316.5586

What do you conclude?

Single Species; Multi-season - Exercise II

Grand Skinks.

Data has been collected on 352 tors over a 5 year (the "seasons") period, although not all tors (rock piles) were surveyed each year, with up to 3 surveys of each tor per year.

There is also a site-specific covariate Pasture indicating whether the surrounding matrix is either predominately the modified habitat (farm pasture, Pasture =1) or "native" grassland (tussock, Pasture = 0).

Single Species; Multi-season - Exercise II

Grand Skinks.

Fit $\psi(.), \gamma(.), \epsilon(.), p(.)$ model.
What do you conclude?

Single Species; Multi-season - Exercise II

Grand Skinks.

Fit $\psi(.), \gamma(.), \epsilon(.), p(.)$ model.

- the probability of occupancy in the first year was 0.39;
- between all seasons, the probability that skinks colonize a previously unoccupied rocky outcrop is 0.07;
- between all seasons, the probability that skinks go locally extinct from an occupied rocky outcrop is 0.10;
- given an outcrop is occupied by skinks, the probability of detecting skinks in a single survey is 0.69.

Single Species; Multi-season - Exercise II

Grand Skinks.

Fit $\psi(.), \gamma(year), \epsilon(year), p(year)$

Other models to fit:

- Is the occupancy the same for all outcrops or is it different for outcrops surrounded by pasture?
- Colonization appears to vary over time, but does the intervening matrix have an effect and is it consistent each year?
- Local Extinction eppears to vary over time, but does the intervening matrix have an effect and is it consistent each year?
- Is there evidence that occupancy is changing over seasons in a "linear" fashion?

Single Species; Multi-season - Exercise III

Redbreast sunfish on Lower Flint River basin in SW Georgia. Fishes were collected seasonally (spring and summer) from 2001-2004 in 26 streams using a quadrat design, i.e., fishes were collected from three 50 - 150 m long segments in each stream via electrofishing. Are $p(t)$ models within a season sensible? (Why?) Investigate effect of streamflow on local extinction and colonization.

Single Species; Multi-season - Exercise III

Redbreast sunfish. Fit the following models:

- $\psi(.), \epsilon(t), \gamma(t), p(season)$
- $\psi(size), \epsilon(lowQ), \gamma(highQ), p(season)$
- $\psi(size), \epsilon(lowQ), \gamma(lowQ), p(season)$
- $\psi(size), \epsilon(highQ), \gamma(lowQ), p(season)$

Single Species; Multi-season - Exercise III

Redbreast sunfish

Results Browser: RD Occupancy Estimation with psi(1), gamma, epsilon.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{psi(size), ext(lowQ), col(high) p(season)}	613.1869	0.0000	1.00000	1.0000	14	583.0108	583.0108
{psi(size), ext(lowQ), col(low) p(season)}	640.0323	26.8454	0.00000	0.0000	14	609.8562	609.8562
{psi(size), ext(highQ), col(low) p(season)}	696.0170	82.8301	0.00000	0.0000	14	665.8409	665.8409
{Psi(.) Epsilon(t) Gamma(t) p Session 1(.) p Session 2(.) p Session 3(.) p Session 4	719.5897	106.4028	0.00000	0.0000	23	667.5897	667.5897

Single Species; Multi-Season - Exercise IV - HARD

In 1942 a small population of house finches (*Carpoldacus mexicanus*) released in Long Island, NY. Population has expanded since then.

NA Breeding Bird Survey (BBS) used to track expansion.

26x 100-km distance bands from LINY established.
Total of 694 sites (multiple in each band) established.
Six 5-year periods (starting in 1976) define the seasons, i.e. season 1 is from 1976-1980, season 2 is from 1981-1985, etc.
50 stops in each season defined as "visit" to the site (not ideal).
Covariates:

- Distance from point of release in units of 1000 kms, i.e. distance band 0.1 = 100-199 km from release point.
- Detection prob as function of distance "crudely" modelled using $f = \{0, 1\}$ if finches detected on more than 10 stops in

Single Species; Multi-Season - Exercise IV - HARD

House Finch. Summary of results from MacKenzie et al. (2006).

Model	ΔAIC	w	$-2l$	$NPar$
$\psi_{76}(d)\gamma(year \times d)\varepsilon(d)p(year \times d + f)$	0.00	78%	44415.64	27
$\psi_{76}(d)\gamma(year \times d)\varepsilon(year + d)p(year \times d + f)$	2.50	22%	44410.14	31
$\psi_{76}(d)\gamma(year \times d)\varepsilon(year)p(year \times d + f)$	13.61	0%	44423.24	30
$\psi_{76}(d)\gamma(year \times d)\varepsilon(year \times d)p(year \times d + f)$	14.70	0%	44422.34	31
$\psi_{76}(d)\gamma(year \times d)\varepsilon(\cdot)p(year \times d + f)$	15.64	0%	44433.27	26
$\psi_{76}(d)\gamma(d)\varepsilon(year \times d)p(year \times d + f)$	18.67	0%	44434.31	27
$\psi_{76}(d)\gamma(year)\varepsilon(year \times d)p(year \times d + f)$	27.19	0%	44436.82	30
$\psi_{76}(d)\gamma(\cdot)\varepsilon(year \times d)p(year \times d + f)$	52.56	0%	44470.2	26

Single Species; Multi-season

Robustness, Planning, and Study Design

Single Species; Multi-season - Robustness

Key Assumptions that lead to bias if violated.

- Homogeneity in ALL parameters (occupancy, colonization, extinction, catchability) across sites.
  - Heterogeneity leads to bias in occupancy (usually negative). Try covariates and similar models as seen in Single-Season case.
- Occupancy status constant in a site within a season.
  - Random movement WITHIN seasons ok, but interpretation of occupancy must be modified to reflect both occupancy and movement.
  - Immigration OR Emigration (not both) leads to biases, but can be handled by pooling (K-1/1, 1/K-1) surveys within a season into two "surveys" and allowing p to vary among these 2 pooled-surveys. Estimate the probability of occupancy at end/beginning. Colonization/extinction bias is small.

Single Species; Multi-season - Planning

Allocation of effort: Number of sites vs. Number of surveys. MacKenzie and Royle (2005). J. Applied Ecology, 42, 1105-1114. doi: 10.1111/j.1365-2664.2005.01098.x

Starting principles:

- Randomization - makes your sample representative of population
- Replication - controls precision
- Stratification - controls for noise, e.g. via covariates

No amount of statistical wizardry can rescue a badly executed survey.

Single Species; Multi-season - Planning

Results similar to those for single season models with similar tradeoffs between number of sites and surveys/site.

Use GENPRES to do simulation validation of experiment!

Single Species; Multi-season - Planning

Defining a "site":

- What is spatial scale were presence/absence is meaningful?
  - E.g. Remnant forest stands and rare species. Is information at stand level sufficient or do you need information on what fraction of stand is occupied?
  - E.g. Large vs. small home ranges.
- "Larger" sites have higher probabilities of occupancy than "smaller" sites (*ceteris paribus*).
  - Rule of thumb: occupancy should be 0.2 $\rightarrow$ 0.8 over your sites.

Single Species; Multi-season - Planning

Site selection

- Randomize, randomize, randomize !
  - ONLY time non-random sampling acceptable is a census.
- Methods of this class assume Simple Random Sample
  - Each site has EQUAL probability of selection
  - Each site selected independently of other sites
- More complex designs possible, buy beyond scope of this course (and current software)
  - Sites are forest stands of different areas and selection is proportion to size of stand (pps).
  - Sites are selected adaptively in waves, i.e. sites near where occupancy found are selected with higher probability.
  - Sites are selected using cluster and multi-state designs, e.g. random select stands, measure all trees in stand.

Single Species; Multi-season - Planning

NEW More on Site Selection

- AVOID selecting sites based on "prior" knowledge of occupancy UNLESS you are only interested in changes in occupancy of these sites. E.g. Selecting previously occupied stands and measuring change in occupancy over time in multi-season models.
- "Regression-to-the-mean" if sites are selected based on historical occupancy rather than a random sample. Extinction probabilities may be ok, but other vital parameterss may be biased.
- Some stratification may be needed, i.e. historical sites = one stratum, new sites = second stratum).

Single Species; Multi-season - Planning

Defining a "Season".

- Critical assumption of closure, i.e. occupancy of site does not change over season.
  - Random movement is analyzable but interpretation of "occupancy" must be modified.
  - Immigration/emigration lead to estimates of occupancy with no direct interpretation.
- Length of season depends on stability of population, i.e. slowly moving animals can be survey over longer "seasons" with closure among sites.
- "Larger" sites can have longer "Seasons" as closure more likely to be satisfied for slowly moving animals, but local deaths may be problematic.

Single Species; Multi-season - Planning

NEW Time between Seasons

- Often this is not under control of experimenter, e.g. single breeding season/year.
- Key consideration is what scale does extinction and colonization take place?
    - Extinction and colonization probabilities are NET, i.e. not possible to separate "temporary" extinction and colonization between seasons (i.e. the "Rescue" hypothesis).
    - Careful not to extrapolate to longer/shorter time intervals because probabilities may not be homogeneous over time.
- Use same time intervals between seasons.
    - Not biologically sensible to calibrate NET colonization /extinction to a per-unit-time basis.

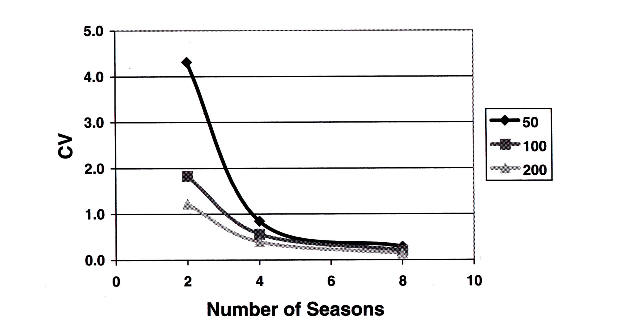
Single Species; Multi-season - Planning

NEW Same or different sites across seasons?

- Can adjust for some missing sites in seasons by missing values.
- Panel designs possible (e.g. 20 sites measured in each of 5 years, vs. 10 sites measured in all years and additional 10 sites measured in year 1, new 10 sites in year 2, etc.
- MacKenzie (2005) found:
  - Panel and Fixed sites both equally efficient in detecting trends.
  - Main consideration is number of SITES.
  - Important than proper sampling be done in panel designs.

Single Species; Multi-season - Planning

NEW More sites vs. more seasons? Is it better to measure 200 sites for 5 years; or 100 sites for 10 years?
CV of trend estimates (0.2/year on logit scale) with 10/100/200 sites in 2, 4, 8 seasons ($K = 3, p = 0.5$).



Mackenzie et al (2005).

- Longer is "better".

Single Species; Multi-season - Planning I

Conducting repeat surveys.

- Many options available:
  - Visit site multiple times with a single survey per visit.
  - Visit site one with multiple INDEPENDENT surveys (e.g. different observers, different transects, different quadrats to look for fecal pellets)
- Key is that repeat surveys need to be INDEPENDENT
  - CAUTION: Detect an animals den on first visit; second visit keys on den location.
    - Use different observer on each visit who does not know location of den.
    - Use "removal" method (see later) where surveys stop after occupancy established.
    - Define "already detected" covariate in modelling

Single Species; Multi-season - Planning II

- CAUTION: Multiple simultaneous surveys with very low density (e.g. one nest per site and several transects are run) are PROBLEMATIC because if one survey detects the nest, the other survey MUST (by definition) not detect the nest.

- How will you align different surveys if models with $p(t)$ are used (e.g. think of the American Toad exercise).

Single Species; Multi-season - Planning

Conducting repeat surveys (continued).

- Avoid confounding observer/ site/ temporal effects.

Design A

	Day								
Site	1			2			3		
1	X	X	X						
2							X	X	X
3				X	X	X			
4				X	X	X			
5							X	X	X
6	X	X	X						
7				X	X	X			
8							X	X	X
9	X	X	X						
p	0.5			0.3			0.8		

Design B

	Day								
Site	1			2			3		
1	X				X				X
2			X	X			X		
3		X				X		X	
4		X		X					X
5			X			X	X		
6	X				X				X
7		X		X				X	
8			X		X		X		
9	X					X		X	
p	0.5			0.3			0.8		

Design B is better. [From MacKenzie et al (2006)]

- Rotate observers among sites and surveys to avoid consistent observer effects.

Single Species; Multi-season - Planning

Optimal number of surveys/site (ignoring costs) for standard design still applicable to multi-season designs.

	ψ								
p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	14	15	16	17	18	20	23	26	34
0.2	7	7	8	8	9	10	11	13	16
0.3	5	5	5	5	6	6	7	8	10
0.4	3	4	4	4	4	5	5	6	7
0.5	3	3	3	3	3	3	4	4	5
0.6	2	2	2	2	3	3	3	3	4
0.7	2	2	2	2	2	2	2	3	3
0.8	2	2	2	2	2	2	2	2	2
0.9	2	2	2	2	2	2	2	2	2

Source: MacKenzie and Royle (2005).

For very low detectability, need to take many surveys in each site!

Single Species; Multi-season - Planning

What drives sample size?

Level of acceptable precision for occupancy estimate

- Preliminary survey: SE of 25% of $\widehat{\psi}$, i.e. if $\widehat{\psi} = 0.80$ then *se* ≈ 0.20.
- Management work: SE of 10% of $\widehat{\psi}$, i.e. if $\widehat{\psi} = 0.80$ then *se* ≈ 0.08.
- Scientific work: SE of 5% of $\widehat{\psi}$, i.e. if $\widehat{\psi} = 0.80$, then *se* ≈ 0.04.

Initial guess of probability of occupancy and detection. (Past surveys; other similar work).

Single Species; Multi-season - Planning

Which is better?

Standard design, 24 sites, 192 surveys

	Season			
# sites	1	2	3	4
24	xx	xx	xx	xx

or

Panel design, 48 sites, 240 surveys

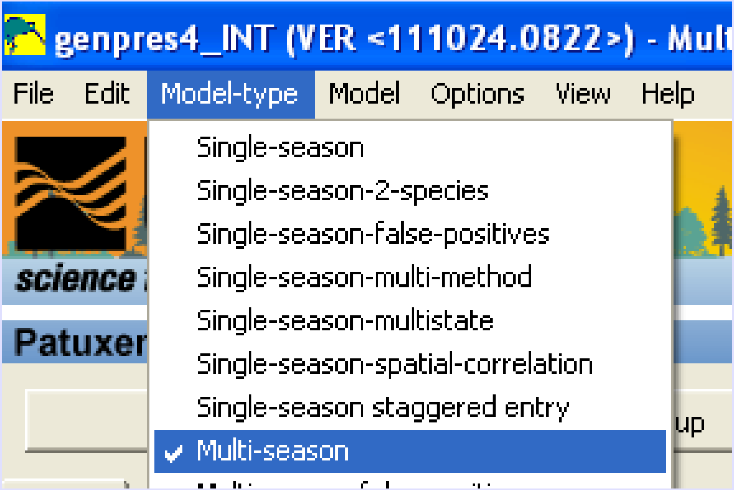
	Season			
# sites	1	2	3	4
12	xx	xx	xx	xx
36	xx	–	-	xx

$\psi_1 = .60, \epsilon = 0.25, \gamma = 0.20, K = 2, p_{s1} = .25, p_{s2} = .75$

ψ declines from 0.60 to 0.40 (how is this computed?).

Single Species; Multi-season - Planning

Launch GENPRES; make sure Single Species; Multi-Season model is selected:

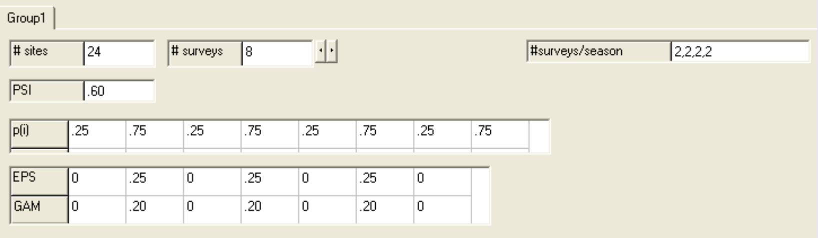


Single Species; Multi-season - Planning

Set up first design:

Standard design, 24 sites, 192 surveys

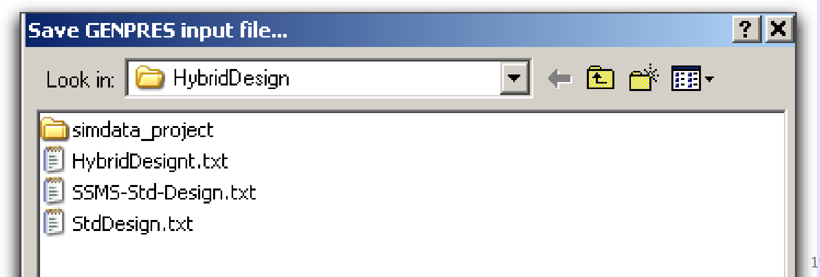
	Season			
# sites	1	2	3	4
24	xx	xx	xx	xx



Why are some entries 0?

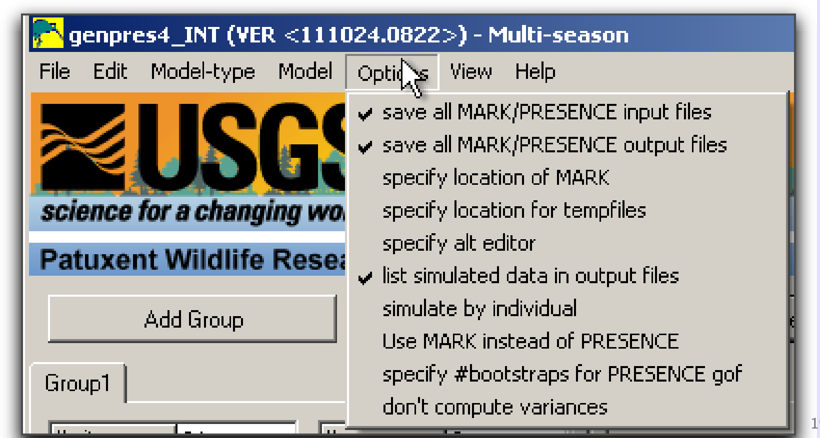
Single Species; Multi-season - Planning

Save the Design:



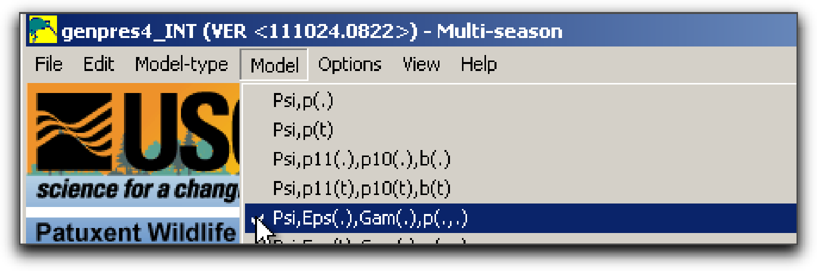
Single Species; Multi-season - Planning

There is no $\psi(TimeTrend)$ model in GENPRES, so we need to save the expected counts and analyze the "data" in PRESENCE ourself. Select what parts of output to see ...



Single Species; Multi-season - Planning

. . . run any of the multi-season model, IGNORE THE RESULTS, . . .



. . . find the generated *.PAO and *.INP files which contain the expected counts . . .

. . . and rename these ‘data’ files

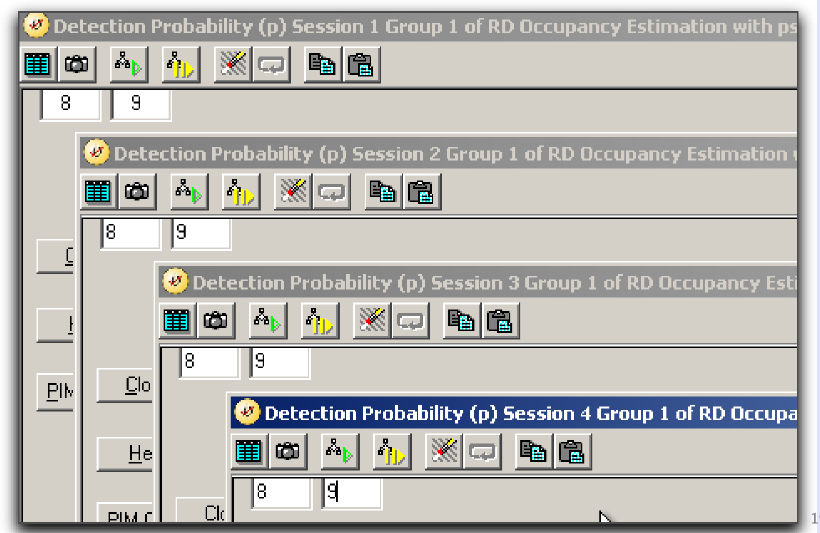
Single Species; Multi-season - Planning

Analyze the expected counts like real data using MARK/PRESENCE. You will want the ψ, ϵ, p or ψ, γ, p parameters in order to model a time trend in ψ.

Start a new project in MARK, select the appropriate models, read in the expected counts, in the usual fashion.

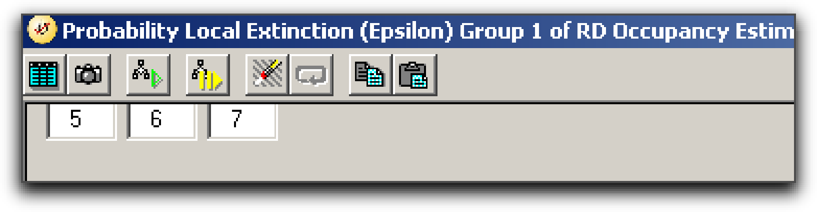
Single Species; Multi-season - Planning

Set up the PIM matrices . . . .



Single Species; Multi-season - Planning

Set up the PIM matrices . . . .



Single Species; Multi-season - Planning

. . . and DESIGN matrices.

B1:	B2:	Parm	B3:	B4:	B5:	B6:	B7:
1	1	1:Psi	0	0	0	0	0
1	2	2:Psi	0	0	0	0	0
1	3	3:Psi	0	0	0	0	0
1	4	4:Psi	0	0	0	0	0
0	0	5:Epsilon	1	0	0	0	0
0	0	6:Epsilon	0	1	0	0	0
0	0	7:Epsilon	0	0	1	0	0
0	0	8:p Session 1	0	0	0	1	0
0	0	9:p Session 1	0	0	0	0	1

Why this format?

Single Species; Multi-season - Planning

Run the model and look at estimates.

Standard Design - test for trend

Real Function Parameters of {psi(TimeTrend), epsilon(t), p(*,survey)}

Parameter	Estimate	Standard Error	95% Confidence Interval Lower	Upper
1:Psi	0.5867630	0.1288577	0.3338087	0.8009445
2:Psi	0.5450202	0.1033862	0.3460149	0.7306145
3:Psi	0.5026349	0.1005072	0.3148678	0.6896617
4:Psi	0.4602117	0.1231734	0.2439168	0.6926086
5:Epsilon	0.2317769	0.1427700	0.0589735	0.5922494
6:Epsilon	0.2439340	0.1440842	0.0652251	0.5986875
7:Epsilon	0.2703375	0.1753951	0.0608993	0.6791525
8:p Session 1	0.2508169	0.0681853	0.1411856	0.4053932
9:p Session 1	0.7524525	0.1084860	0.4925486	0.9049330

Estimates of ψ a bit biased but bias acceptable. Recall that trend is on the logit scale.

Single Species; Multi-season - Planning

Run the model and look at estimates.

Standard Design - test for trend

Estimates of Derived Parameters

Gamma Estimates of {psi(TimeTrend), epsilion(t), p(*,survey)}

Grp.	Occ.	Gamma-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.2280901	0.1747501	-0.1144202	0.5706004
1	2	0.1990499	0.1480365	-0.0911017	0.4892015
1	3	0.1879060	0.1380421	-0.0826565	0.4584685

Lambda Estimates of {psi(TimeTrend), epsilion(t), p(*,survey)}

Grp.	Occ.	Lambda-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.9288591	0.0804008	0.7712736	1.0864446
1	2	0.9222317	0.0965619	0.7329704	1.1114931
1	3	0.9155984	0.1127352	0.6946375	1.1365593

Estimates of γ a bit biased but bias acceptable.

Single Species; Multi-season - Planning

Run the model and look at estimates of TREND (the first 2 beta values) (why?)

Standard Design - test for trend

LOGIT Link Function Parameters of {psi(TimeTrend), epsilion(t), p(*,survey)}

Parameter	Beta	Standard Error	95% Confidence Interval Lower	Upper
1:	0.5206297	0.6981547	-0.8477535	1.8890129
2:	-0.1700300	0.2195502	-0.6003484	0.2602884
3:	[illegible]	[illegible]	[illegible]	[illegible]

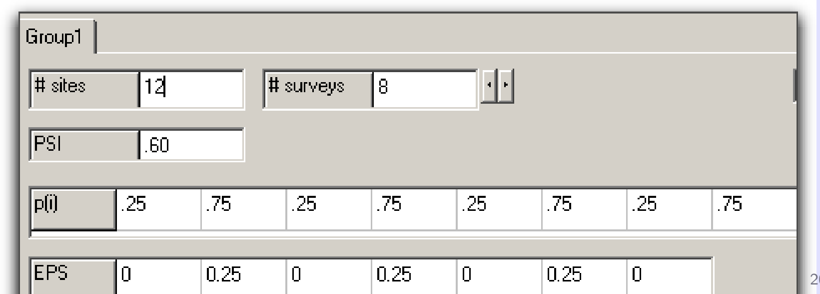
Failed to detect a trend.

Single Species; Multi-season - Planning

Repeat for Hybrid Design.

Panel design, 48 sites, 240 surveys				
	Season			
# sites	1	2	3	4
12	xx	xx	xx	xx
36	xx	- -	- -	xx

$\psi_1 = .60, \epsilon = 0.25, \gamma = 0.20, K = 2, p_{s1} = .25, p_{s2} = .75$

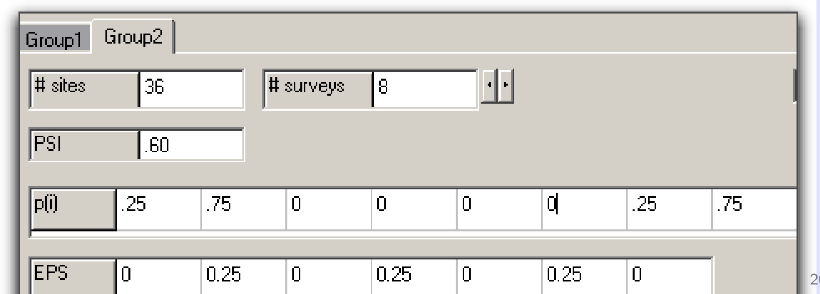


Single Species; Multi-season - Planning

Repeat for Hybrid Design.

Panel design, 48 sites, 240 surveys				
	Season			
# sites	1	2	3	4
12	xx	xx	xx	xx
36	xx	- -	- -	xx

$\psi_1 = .60, \epsilon = 0.25, \gamma = 0.20, K = 2, p_{s1} = .25, p_{s2} = .75$



Single Species; Multi-season - Planning

Set up a new MARK project and refit the same model as the standard design.

Single Species; Multi-season - Planning

Run the model and look at estimates.

SSMS Hybrid Design

Real Function Parameters of {psi(TimeTrend), epsilon(t) p(*,session)}

Parameter	Estimate	Standard Error	95% Confidence Interval Lower	Upper
1:Psi	0.5956028	0.1050281	0.3852104	0.7758847
2:Psi	0.5532950	0.0878843	0.3815587	0.7131892
3:Psi	0.5102010	0.0841627	0.3499405	0.6683916
4:Psi	0.4669547	0.0967619	0.2902201	0.6523913
5:Epsilon	0.2266841	0.1648152	0.0443653	0.6492303
6:Epsilon	0.2398230	0.1684390	0.0490536	0.6586410
7:Epsilon	0.2664281	0.1903424	0.0510933	0.7101301
8:p Session 1	0.2504731	0.0615588	0.1494685	0.3885513
9:p Session 1	0.7514203	0.1021941	0.5084641	0.8983061

Estimates of ψ a bit biased but bias acceptable. Recall that trend is on the logit scale.

Single Species; Multi-season - Planning

Run the model and look at estimates.

SSMS Hybrid Design

Estimates of Derived Parameters

Gamma Estimates of {psi(TimeTrend), epsilon(t) p(*,session)}

Grp.	Occ.	Gamma-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.2292447	0.2242352	-0.2102562	0.6687457
1	2	0.2005772	0.1945093	-0.1806610	0.5818154
1	3	0.1892320	0.1773988	-0.1584697	0.5369337

Lambda Estimates of {psi(TimeTrend), epsilon(t) p(*,session)}

Grp.	Occ.	Lambda-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.9289664	0.0566278	0.8179760	1.0399568
1	2	0.9221138	0.0684288	0.7879933	1.0562343
1	3	0.9152369	0.0803108	0.7578277	1.0726461

Estimates of γ a bit biased but bias acceptable.

Single Species; Multi-season - Planning

Run the model and look at estimates of TREND (the first 2 beta values) (why?)

SSMS Hybrid Design

LOGIT Link Function Parameters of {psi(TimeTrend), epsilon(t) p(*,session)}

Parameter	Beta	Standard Error	95% Confidence Interval Lower	Upper
1:	0.5603593	0.5518647	-0.5212955	1.6420141
2:	-0.1731836	0.1592200	-0.4852548	0.1388876

Failed to detect a trend.

Single Species; Multi-season - Planning

Design effect.

Standard Design - test for trend

Real Function Parameters of {psi(TimeTrend), epsilion(t), p(*,survey)}

Parameter	Estimate	Standard Error	95% Confidence Interval Lower	Upper
1:Psi	0.5867630	0.1288577	0.3338087	0.8009445
2:Psi	0.5450202	0.1033862	0.3460149	0.7306145
3:Psi	0.5026349	0.1005072	0.3148678	0.6896617
4:Psi	0.4602117	0.1231734	0.2439168	0.6926086
5:Epsilon	0.2317769	0.1427700	0.0589735	0.5922494
6:Epsilon	0.2439340	0.1440842	0.0652251	0.5986875
7:Epsilon	0.2703375	0.1753951	0.0608993	0.6791525
8:p Session 1	0.2508169	0.0681853	0.1411856	0.4053932
9:p Session 1	0.7524525	0.1084860	0.4925486	0.9049330

SSMS Hybrid Design

Real Function Parameters of {psi(TimeTrend), epsilon(t) p(*,session)}

Parameter	Estimate	Standard Error	95% Confidence Interval Lower	Upper
1:Psi	0.5956028	0.1050281	0.3852104	0.7758847
2:Psi	0.5532950	0.0878843	0.3815587	0.7131892
3:Psi	0.5102010	0.0841627	0.3499405	0.6683916
4:Psi	0.4669547	0.0967619	0.2902201	0.6523913
5:Epsilon	0.2266841	0.1648152	0.0443653	0.6492303
6:Epsilon	0.2398230	0.1684390	0.0490536	0.6586410
7:Epsilon	0.2664281	0.1903424	0.0510933	0.7101301
8:p Session 1	0.2504731	0.0615588	0.1494685	0.3885513
9:p Session 1	0.7514203	0.1021941	0.5084641	0.8983061

Single Species; Multi-season - Planning

Design effect.

Standard Design - test for trend

Estimates of Derived Parameters

Gamma Estimates of {psi(TimeTrend), epsilion(t), p(*,survey)}

Grp.	Occ.	Gamma-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.2280901	0.1747501	-0.1144202	0.5706004
1	2	0.1990499	0.1480365	-0.0911017	0.4892015
1	3	0.1879060	0.1380421	-0.0826565	0.4584685

Lambda Estimates of {psi(TimeTrend), epsilion(t), p(*,survey)}

Grp.	Occ.	Lambda-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.9288591	0.0804008	0.7712736	1.0864446
1	2	0.9222317	0.0965619	0.7329704	1.1114931
1	3	0.9155984	0.1127352	0.6946375	1.1365593

SSMS Hybrid Design

Estimates of Derived Parameters

Gamma Estimates of {psi(TimeTrend), epsilon(t) p(*,session)}

Grp.	Occ.	Gamma-hat	Standard Error	95% Confidence Interval Lower	Upper
1	1	0.2292447	0.2242352	-0.2102562	0.6687457
1	2	0.2005772	0.1945093	-0.1806610	0.5818154
1	3	0.1892320	0.1773988	-0.1584697	0.5369337

Lambda Estimates of {psi(TimeTrend), epsilon(t) p(*,session)}

95% Confidence Interval

Grp. Occ. Lambda-hat Standard Error Lower Upper

Single Species; Multi-season - Planning

Design effect.

Standard Design - test for trend

LOGIT Link Function Parameters of {psi(TimeTrend), epsilion(t), p(*,survey)}

Parameter	Beta	Standard Error	95% Confidence Interval Lower	Upper
1:	0.5206297	0.6981547	-0.8477535	1.8890129
2:	-0.1700300	0.2195502	-0.6003484	0.2602884

SSMS Hybrid Design

LOGIT Link Function Parameters of {psi(TimeTrend), epsilon(t) p(*,session)}

Parameter	Beta	Standard Error	95% Confidence Interval Lower	Upper
1:	0.5603593	0.5518647	-0.5212955	1.6420141
2:	-0.1731836	0.1592200	-0.4852548	0.1388876

Single Species; Multi-season - Planning - EXERCISE

- What sample size would I need to detect a trend with both designs?
- Is there another design that might be considered?

Hints:

- SE are roughly proportion to $\frac{1}{\sqrt{n}}$, i.e. to halve a SE you need 4x the sample size.
- Data cloning, i.e. replicate an existing dataset multiple times?
- MARK does NOT allow you to "drop" new data into an existing analysis (groan), but a clever use of groups where groups share parameters can "mimic" changing sample sizes.

Single Species; Multi-season - Summary

Single-Species Multi-season Summary

Single Species; Multi-season - Summary

Planning.

- Key parameters ψ, γ, ϵ; nuisance parameter $p_t < 1$.
- Design your study well.
  - What is appropriate spatial scale and temporal scale?
  - Simple Random sample of sites; some relaxation if comparing occupancy between classes.
  - What is a season – assume closure over season.
  - Repeated surveys must be independent.
- Allocate effort between seasons, sites and surveys; usually fewer sites, fewer surveys, and more seasons is better.
- Hybrid designs may be more efficient.

Single Species; Multi-season - Summary

Key assumptions.

1. Occupancy state of sites is constant during each season (closure).
2. Probability of occupancy (ψ) is equal across all sites before first season (homogeneity).
3. Probability of detection (p) given occupancy is equal across all sites (homogeneity).
4. Detection of species in each survey of a site is independent of those on other surveys
5. Detection histories at each location are independent
6. No false positives.
7. Markovian colonization and local extinction.

Single Species; Multi-season - Summary

Software

1. *PRESENCE* and *MARK* have large collection of model types and cleaner model averaging.
2. *RPresence* and *RMark* are useful with scripting.
3. *unmarked* works well for standard model and can get bootstrap SE of any statistic. Help files useless.
4. *JAGS* not for mere mortals.