Statistical methods for incoporating expert elicitation in occupancy modeling frameworks

true true true true

2022-11-09

Introduction

Model framework

Base site-occupancy modeling framework

For now, we layer this framework on a standard occupancy model. Let i index the location, j index the visit to site i, and Z_i denote the partially observed occupancy states of sites $1, \ldots, n$. Then,

$$\begin{aligned} Z_i &\sim \text{Bernoulli}(\psi_i) \\ y_i &\sim \text{Bernoulli}(z_i p_{ij}) \end{aligned}$$

where $\text{logit}(\psi_i) = x_i'\beta$ and $\text{logit}(p_{ij}) = W\alpha$.

Addition of random effects

To accommodate extra sources of variation not captured by X, we include spatially indexed random effects for each site.

$$\begin{aligned} Z_i &\sim \text{Bernoulli}(\psi_i) \\ y_i &\sim \text{Bernoulli}(z_i p_{ij}) \end{aligned}$$

where $\text{logit}(\psi_i) = x_i'\beta + \theta_i$ and $\text{logit}(p_{ij}) = W\alpha$.

Additional structure may be imposed on $\boldsymbol{\theta}$. For example, one could draw $\boldsymbol{\theta}$ from a Gaussian process with spatial covariance structure. Alternatively, $\boldsymbol{\theta}$ may be drawn from a hierarchical normal distribution if expert information is available to inform a grouping structure among sites.

$$\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\tau\Omega\tau}^T)$$

Under this formulation, $\text{logit}(\psi_i) = x_i'\beta + \theta_{c_i}$, where $\boldsymbol{c}$ is a vector of indicators denoting the group associated with ψ_i. Note that the above formulation assume a general correlation structure among $\boldsymbol{\theta}$ (in the form of $\boldsymbol{\Omega}$), but independence may be assumed. Under this framework, expert knowledge may be used to inform $\boldsymbol{\mu}$. For example,

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{s\mathcal{I}s}^T)$$

Expert response structure

The goal is to collect information from experts in a way that can easily inform the prior distributions for the spatial random effects. One option is to solicit feedback directly on the occupancy probabilities, then map those probabilities to the real line.

Daubenmire classes

To aid in describing the uncertainties in this mapping, one could use the Daubenmire coverage class (Daubenmire, 1959).

Table 1: Table 1: Daubenmire coverage classes (Daubenmire, 1959).

Class	Lower	Upper	Midpoint	logitL	logitM	logitU
1	0.00	0.05	0.03	-4.60	-3.66	-3.18
2	0.05	0.25	0.15	-2.75	-1.73	-1.15
3	0.25	0.50	0.38	-1.05	-0.51	-0.04
4	0.50	0.75	0.62	0.04	0.51	1.05
5	0.75	0.95	0.85	1.15	1.73	2.75
6	0.95	1.00	0.98	3.18	3.66	4.60

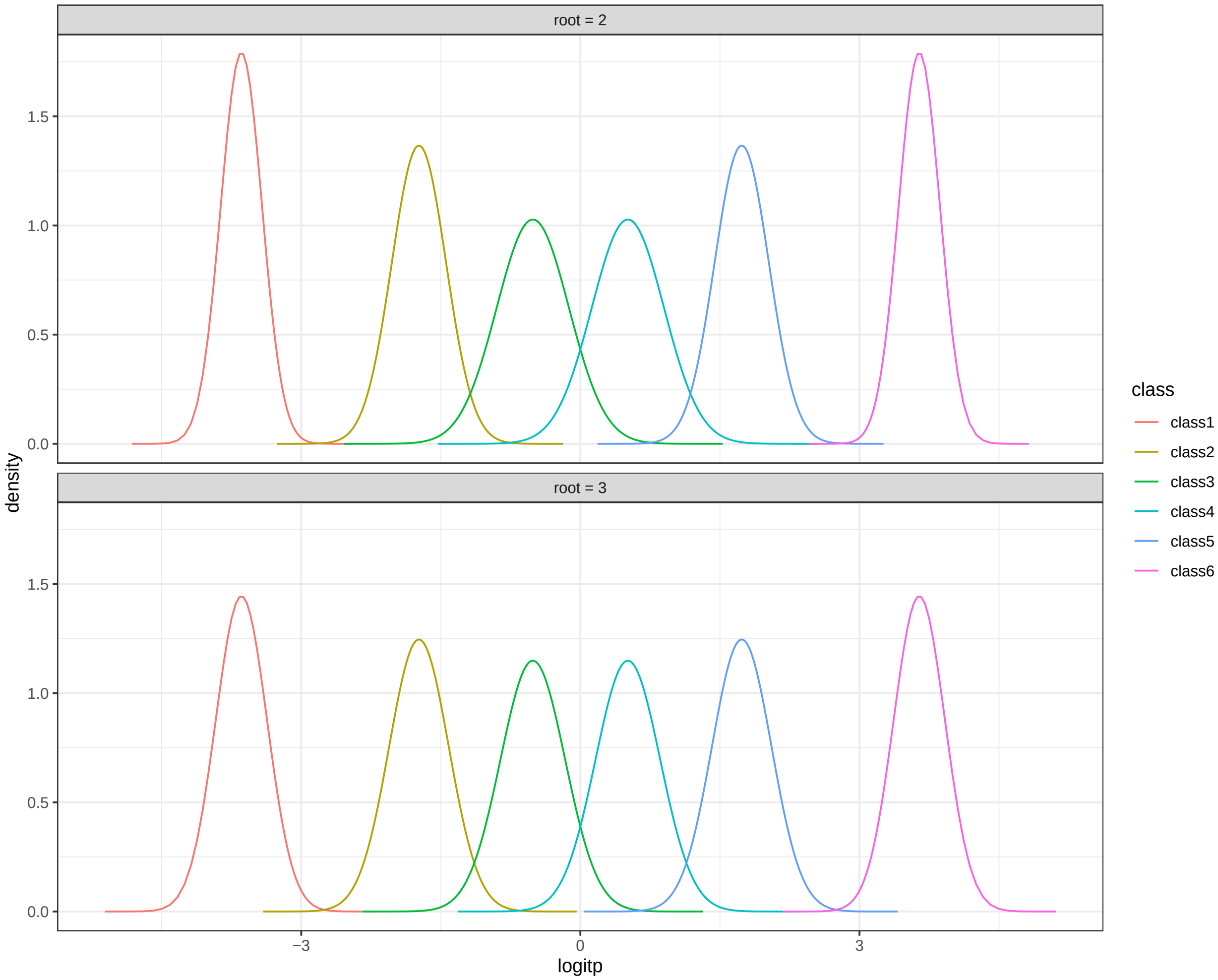
Normal k-root-mean-weighted mapping

There are many potential mappings for expert feedback in the form of Daubenmire classes. Generally, the mapping should imply lesser probability of occurrence for lower classes. Additionally, it may be reasonable to imply greater certainty in large classes (i.e. probabilities close to 0 and 1 are more certain than those close to 0). Finally, it is computationally attractive to restrict mappings to the Normal family of distributions (though this can be relaxed). A mapping that satisfies these requirements is described below.

Let l_i, u_i, and μ_i denote the implied lower, upper, and midpoints of each class on the logit scale. Then the normal mean-weighted mapping proceeds as follows. For each class i and any positive real k, find σ_i such that

$$\int_{l_i}^{u_i} p\left(y \middle| \mu_i, \frac{1}{|\mu_i|^k}\sigma_i\right) dy = .95$$

where $p(y|\mu, \sigma)$ is the density function for a normal random variate with mean μ and standard deviation σ.



Model estimation

To improve computational efficiency, we write the sampler for this model by hand. Leveraging the data augmentation strategy of Polson, Scott and Windle (2013), Gibbs draws are available for all parameters in the model. The full specification of the model is as follows. Let i index the sample location, j index the visit to site i, and $\boldsymbol{c}$ denote a vector group membership informed by expert opinion. Then,

$$
\begin{aligned}
Z_i &\sim \text{Bernoulli}(\psi_i) \\
y_{ij} &\sim \text{Bernoulli}(z_i p_{ij})
\end{aligned}
$$

where $\text{logit}(\psi_i) = \boldsymbol{x}_i^T \boldsymbol{\beta} + \theta_{c_i}$ and

$$\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \tau^2 \boldsymbol{\mathcal{I}})$$

For now, we assume that $p_{ij} = p$ for all $\{i, j\}$ to speed up computation (though this assumption can be easily relaxed).

Polya-gamma data augmentation

For binomial and negative binomial sampling models, we implement the Polya-gamma data augmentation strategy of Polson et al. (2013). This strategy introduces auxiliary Polya-gamma distributed random variables to each likelihood to afford conditionally Gaussian posterior distributions, allowing for Gibbs draws of regression coefficients. This data augmentation is relevant for sampling α, β, and η.

Sampling $\boldsymbol{z}^{(\psi)}$

We let $\boldsymbol{z}^{(\psi)}$ denote the vector of partially observed occupancy states and y_{ij} denote the observed response during visit j to site i. From Dorazio and Erickson (2018), the full conditional distribution of $\boldsymbol{z}^{(\psi)}$ is,

$$Z_i^{(\psi)} \sim \begin{cases} \text{Bernoulli}(1) & \text{if} \sum_{j=1}^{J_i} y_{ij} > 0 \\ \text{Bernoulli}\left(\frac{\psi_i \prod_{j=1}^{J_i}(1-p_{ij})}{1-\psi_i+\psi_i \prod_{j=1}^{J_i}(1-p_{ij})}\right) & \text{if} \sum_{j=1}^{J_i} y_{ij} = 0 \end{cases}$$

Sampling $\boldsymbol{\omega}^{(\beta)}$

Let $\omega_i^{(\beta)} \sim \text{PG}(1, 0)$. Then,

$$z_i = \frac{1}{\omega_i^{(\beta)}}\left(z_i^{(\psi)} - \frac{1}{2}\right) \sim \mathcal{N}\left(\boldsymbol{x}_i'\boldsymbol{\beta} + \theta_{c_i}, \frac{1}{\omega_i}\right).$$

where $z_i^{(\psi)}$ is the latent occupancy state for site i. See Polson et al. (2013) for more detail. Equivalently, let $\boldsymbol{\Omega}_{(\beta)} = \text{diag}(\omega_1^{(\beta)}, \ldots, \omega_n^{(\beta)})$. Then,

$$\boldsymbol{z} \sim \mathcal{N}\left(X\boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{\theta}, \boldsymbol{\Omega}_{(\beta)}^{-1}\right)$$

where $\boldsymbol{C}$ is an indicator matrix denoting group membership. From Polson et al. (2013), the full conditional distribution of $\omega_i^{(\beta)}$ is

$$\omega_i^{(\beta)}|\cdot \sim \text{PG}(1, \boldsymbol{x}_i'\boldsymbol{\beta} + \theta_{c_i})$$

Sampling $\boldsymbol{\beta}$

Let $\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_{0,\beta}, \boldsymbol{\Sigma}_{0,\beta})$.Then,

$$\begin{aligned} &\boldsymbol{\beta}|\cdot \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{V}) \\ &\boldsymbol{V} = \left(\boldsymbol{X}^T\boldsymbol{\Omega}_{(\beta)}\boldsymbol{X} + \boldsymbol{\Sigma}_{0,\beta}^{-1}\right)^{-1} \\ &\boldsymbol{m} = \boldsymbol{V}\left(\boldsymbol{X}^T\boldsymbol{\Omega}_{(\beta)}(\boldsymbol{z} - \boldsymbol{C}\boldsymbol{\theta}) + \boldsymbol{\Sigma}_{0,\beta}^{-1}\boldsymbol{\mu}_{0,\beta}\right) \end{aligned}$$

Proof (omitting boldface text):

Let $z^* = z - C\theta$

$$\begin{aligned} p(\beta|z^*, \omega^{(\beta)}) &\propto p(z^*|\beta, \omega^{(\beta)})p(\beta) \\ &\propto \exp\left\{-\frac{1}{2}(z^* - X\beta)^T\Omega_{(\beta)}(z^* - X\beta)\right\}\exp\left\{-\frac{1}{2}(\beta - \mu_{0,\beta})^T\Sigma_{(0,\beta)}^{-1}(\beta - \mu_{0,\beta})\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\beta^T(X^T\Omega_{(\beta)}X + \Sigma_{0,\beta}^{-1})\beta - 2\beta^T(X^T\Omega_{(\beta)}z^* + \Sigma_{0,\beta}^{-1}\mu_{0,\beta})\right)\right\} \end{aligned}$$

Sampling $\boldsymbol{\theta}$

Let $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \tau^2 \boldsymbol{\mathcal{I}})$. Note that the expert elicitation informs $\boldsymbol{\mu}$. The full conditional posterior distribution of $\boldsymbol{\theta}$ is

$$\begin{aligned}
\boldsymbol{\theta}|\cdot &\sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{V}) \\
\boldsymbol{V} &= \left(\boldsymbol{C}^T \boldsymbol{\Omega}_{(\beta)} \boldsymbol{C} + \frac{1}{\tau^2} \boldsymbol{\mathcal{I}} \right)^{-1} \\
\boldsymbol{m} &= \boldsymbol{V} \left(\boldsymbol{C}^T \boldsymbol{\Omega}_{(\beta)} (\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta}) + \frac{1}{\tau^2} \boldsymbol{\mathcal{I}} \boldsymbol{\mu} \right)
\end{aligned}$$

Proof (omitting boldface text):

Let $z^* = z - X\beta$

$$\begin{aligned}
p(\beta|z^*, \omega^{(\beta)}) &\propto p(z^*|\theta, \omega^{(\beta)}) p(\theta) \\
&\propto \exp\left\{ -\frac{1}{2}(z^* - C\theta)^T \Omega_{(\beta)} (z^* - C\theta) \right\} \exp\left\{ -\frac{1}{2}(\theta - \mu)^T \frac{1}{\tau^2} \mathcal{I} (\theta - \mu) \right\} \\
&\propto \exp\left\{ -\frac{1}{2} \left(\theta^T (C^T \Omega_{(\beta)} C + \frac{1}{\tau^2} \mathcal{I}) \theta - 2\theta^T (C^T \Omega_{(\beta)} z^* + \frac{1}{\tau^2} \mathcal{I} \mu) \right) \right\}
\end{aligned}$$

Sampling $\boldsymbol{\mu}$

Let $\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{s} \boldsymbol{\mathcal{I}} \boldsymbol{s}^T)$, where $\boldsymbol{m}$ and $\boldsymbol{s}$ may be informed by expert opinion. Then

$$\begin{aligned}
\boldsymbol{\mu}|\cdot &\sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{V}) \\
\boldsymbol{V} &= \left(\frac{1}{\tau^2} \boldsymbol{\mathcal{I}} + \boldsymbol{s}^{-1} \boldsymbol{\mathcal{I}} (\boldsymbol{s}^{-1})^T \right)^{-1} \\
\boldsymbol{m}_0 &= \boldsymbol{V} \left(\frac{1}{\tau^2} \boldsymbol{\mathcal{I}} \boldsymbol{\theta} + \boldsymbol{s}^{-1} \boldsymbol{\mathcal{I}} (\boldsymbol{s}^{-1})^T \boldsymbol{m} \right)
\end{aligned}$$

Proof (omitting boldface text):

$$\begin{aligned}
p(\mu|\cdot) &\propto p(\theta|\mu, \tau^2) p(\mu) \\
&\propto \exp\left\{ -\frac{1}{2}(\theta - \mu)^T \frac{1}{\tau^2} \mathcal{I} (\theta - \mu) \right\} \exp\left\{ -\frac{1}{2}(\mu - m)^T s^{-1} \mathcal{I} s^{-1} (\mu - m) \right\} \\
&\propto \exp\left\{ -\frac{1}{2} \left(\mu^T (\frac{1}{\tau^2} \mathcal{I} + s^{-1} \mathcal{I} s^{-1}) \mu - 2\mu^T (\frac{1}{\tau^2} \mathcal{I} \theta + s^{-1} \mathcal{I} s^{-1} m) \right) \right\}
\end{aligned}$$

Sampling τ

Let $\tau^2 \sim$ Inverse-Gamma(a_0, b_0). Then,

$$\begin{aligned}
\tau^2|\cdot &\sim \text{Inverse-Gamma}(a, b) \\
a &= a_0 + \frac{K}{2} \\
b &= b_0 + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu})^T (\boldsymbol{\theta} - \boldsymbol{\mu})
\end{aligned}$$

Proof (omitting boldface):

$$\begin{aligned}
p(\tau^2|\theta, \mu) &\propto p(\theta|\tau^2, \mu) p(\tau^2) \\
&\propto (\tau^2)^{-K/2} \exp\left\{ -\frac{1}{2}(\theta - \mu)^T \frac{1}{\tau^2} \mathcal{I} (\theta - \mu) \right\} (\tau^2)^{-a_0 - 1} \exp\left(-\frac{b_0}{\tau^2} \right)
\end{aligned}$$

Synthetic data scenarios

Idyllic case

In the ideal case, species occupancy arises as a function of some linear relationship with covariates and a regional intercept that is not directly observed. The experts then inform this regional intercept.

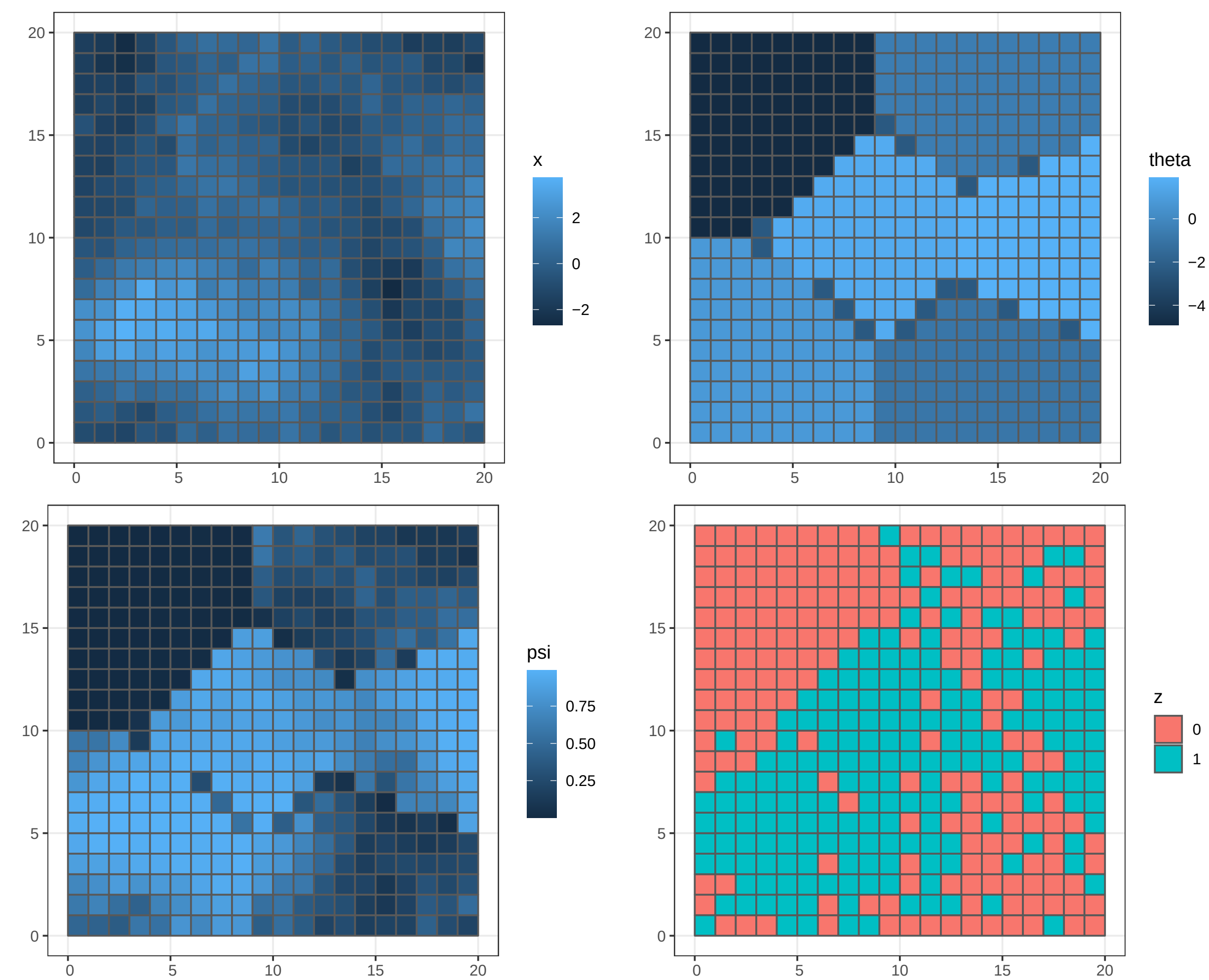
Sampling model

We first imagine the data arising directly from the specified model. That is,

$$
\begin{aligned}
Z_i &\sim \text{Bernoulli}(\psi_i) \\
y_i &\sim \text{Bernoulli}(z_i p_{ij})
\end{aligned}
$$

where $\text{logit}(\psi_i) = x_i'\beta + \theta_{c_i}$ and $\text{logit}(p_{ij}) = W\alpha$. Furthermore, we assume that $\theta_{c_i} \sim N(\mu_{c_i}, \tau^2)$, where c_i is an indicator vector denoting distinct regions assigned by the expert. For now, we assume that p is constant (to speed up simulated studies).

Simulated data



Fit models with Julia

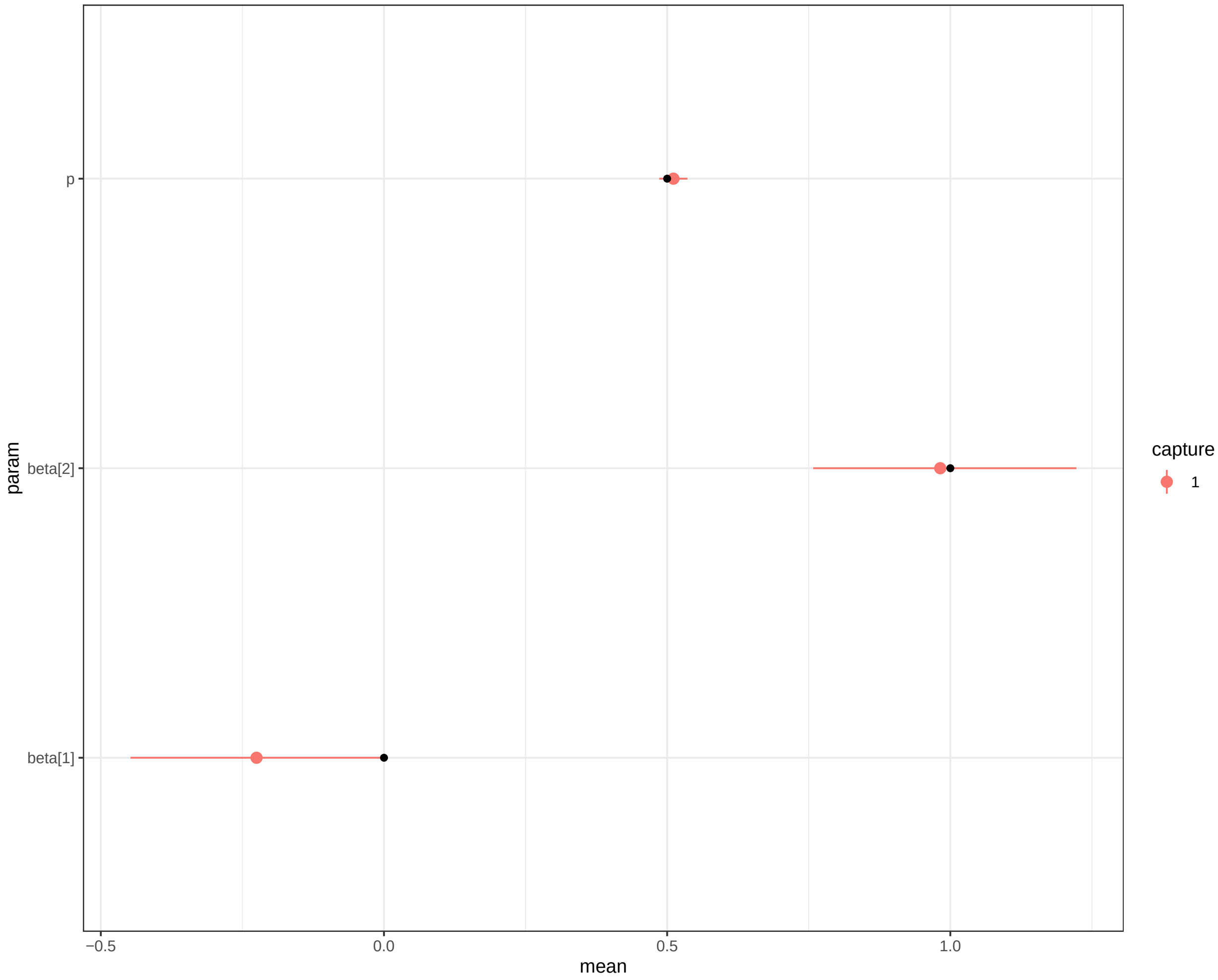
We consider two potential models:

- model0: basic occupancy model without region-specific random effects
- model1: occupancy model with region-specific random effects informed by experts

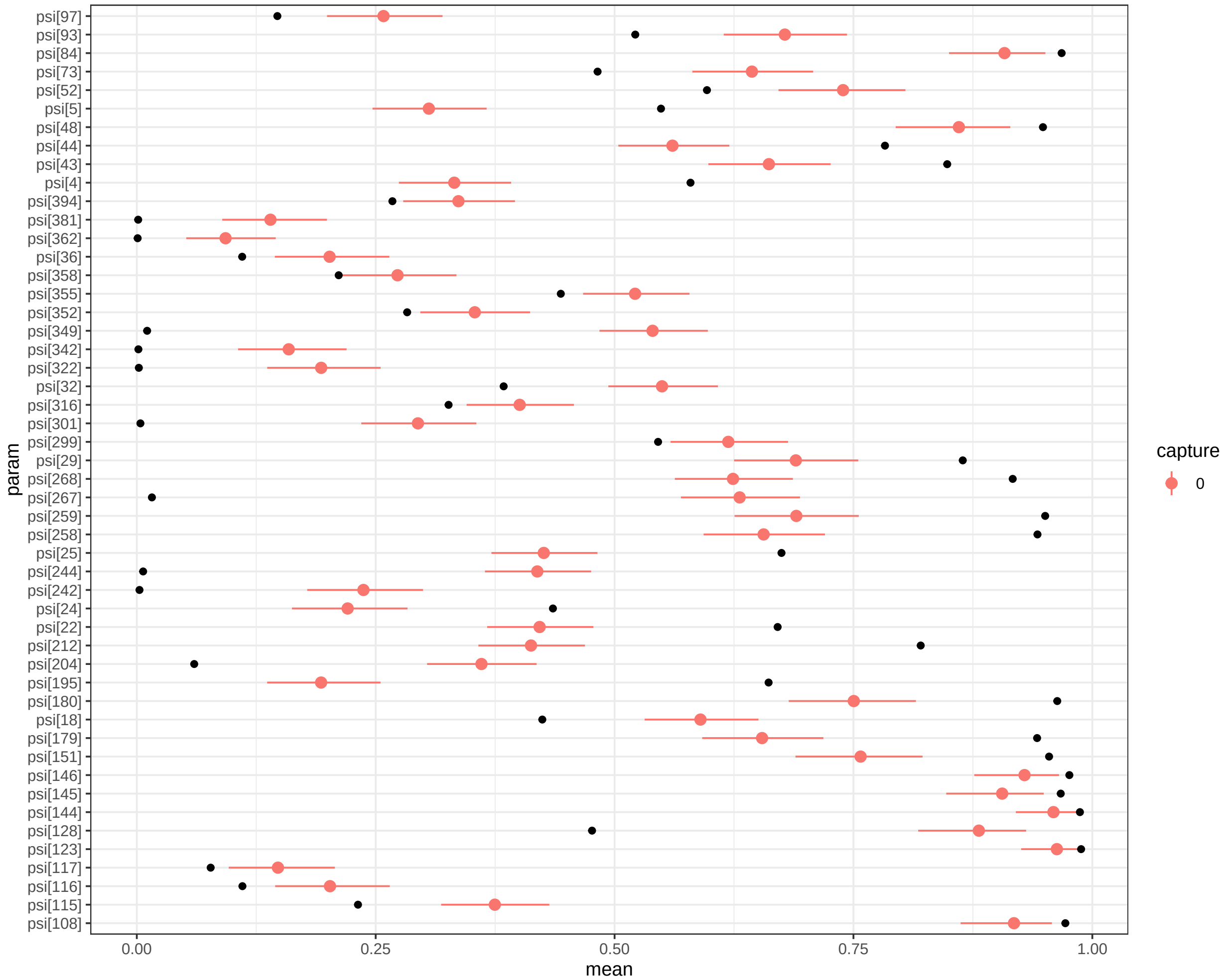
For the time being, we assume the experts are correct.

model0

```
              Mean       2.5%       97.5%      Rhat ess_bulk
p        0.5110552  0.4859887 0.535767859 0.9998710 7525.722
beta[1] -0.2250412 -0.4475441 0.000428128 0.9999461 7340.170
beta[2]  0.9821882  0.7578709 1.222570171 0.9999311 5695.660
psi[1]   0.2394866  0.1806286 0.301809415 0.9999085 6532.857
psi[2]   0.2186129  0.1604826 0.281211272 0.9998611 6441.005
psi[3]   0.1893535  0.1331817 0.251433860 0.9998586 6320.579
psi[4]   0.3322566  0.2742926 0.391719809 0.9998836 7002.069
psi[5]   0.3056476  0.2466830 0.366100585 0.9998817 6877.817
psi[6]   0.5791978  0.5213224 0.639238212 1.0002640 6805.114
psi[7]   0.4558142  0.4017611 0.511929093 0.9999572 7330.942
psi[8]   0.6379023  0.5759368 0.701713565 1.0003803 6482.564
```

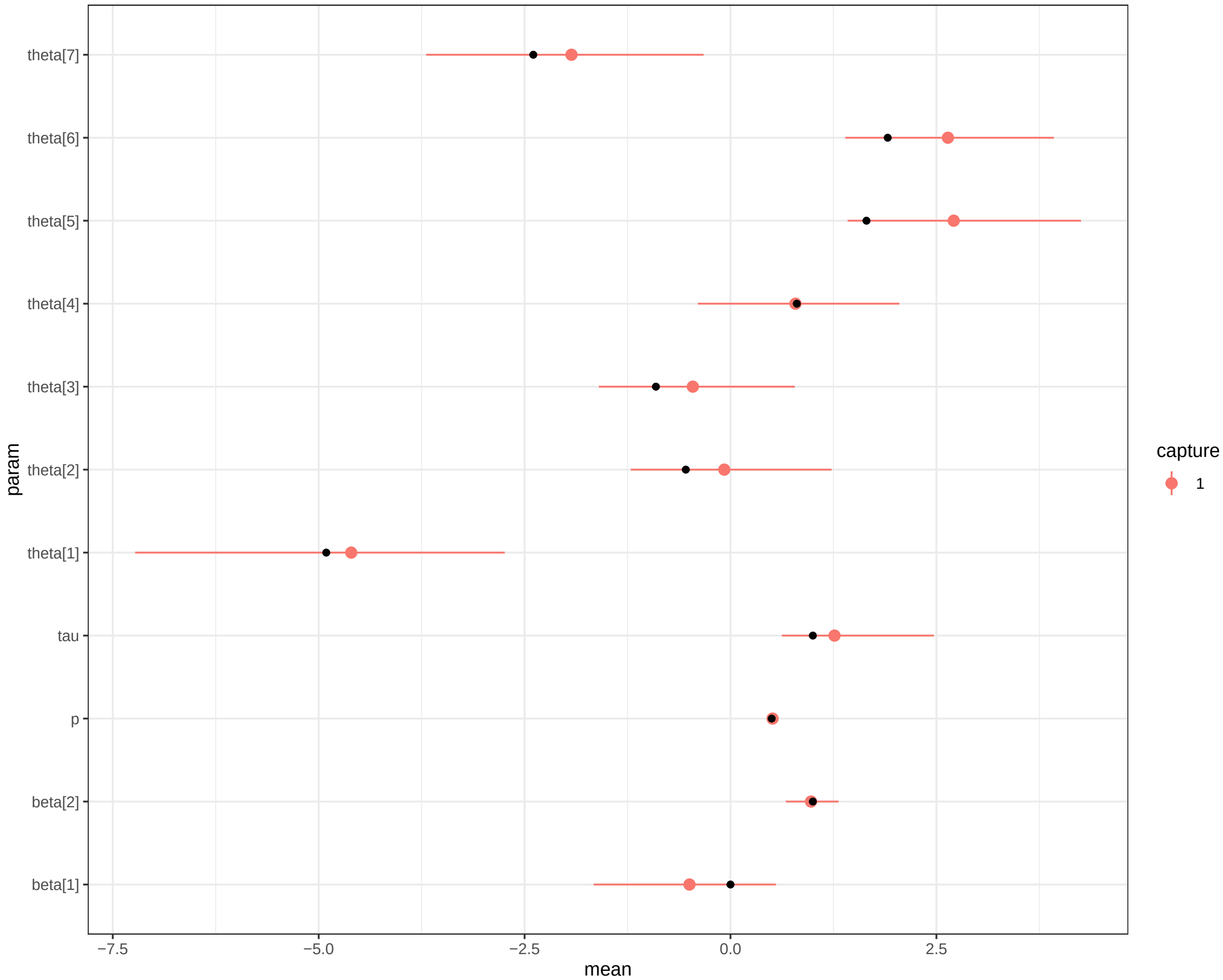


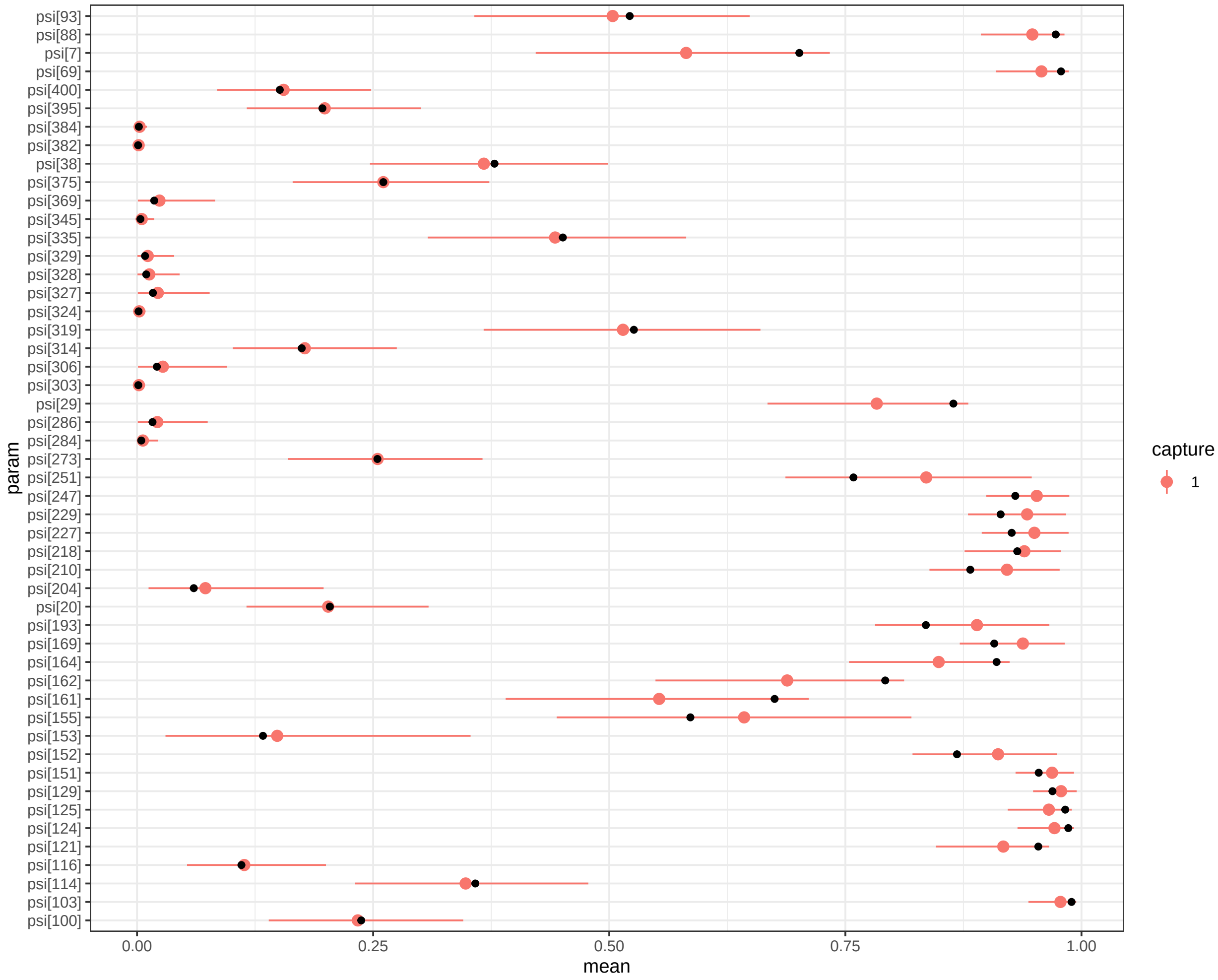
p
beta[2]
beta[1]
param
−0.5
0.0
0.5
1.0
mean
capture
1



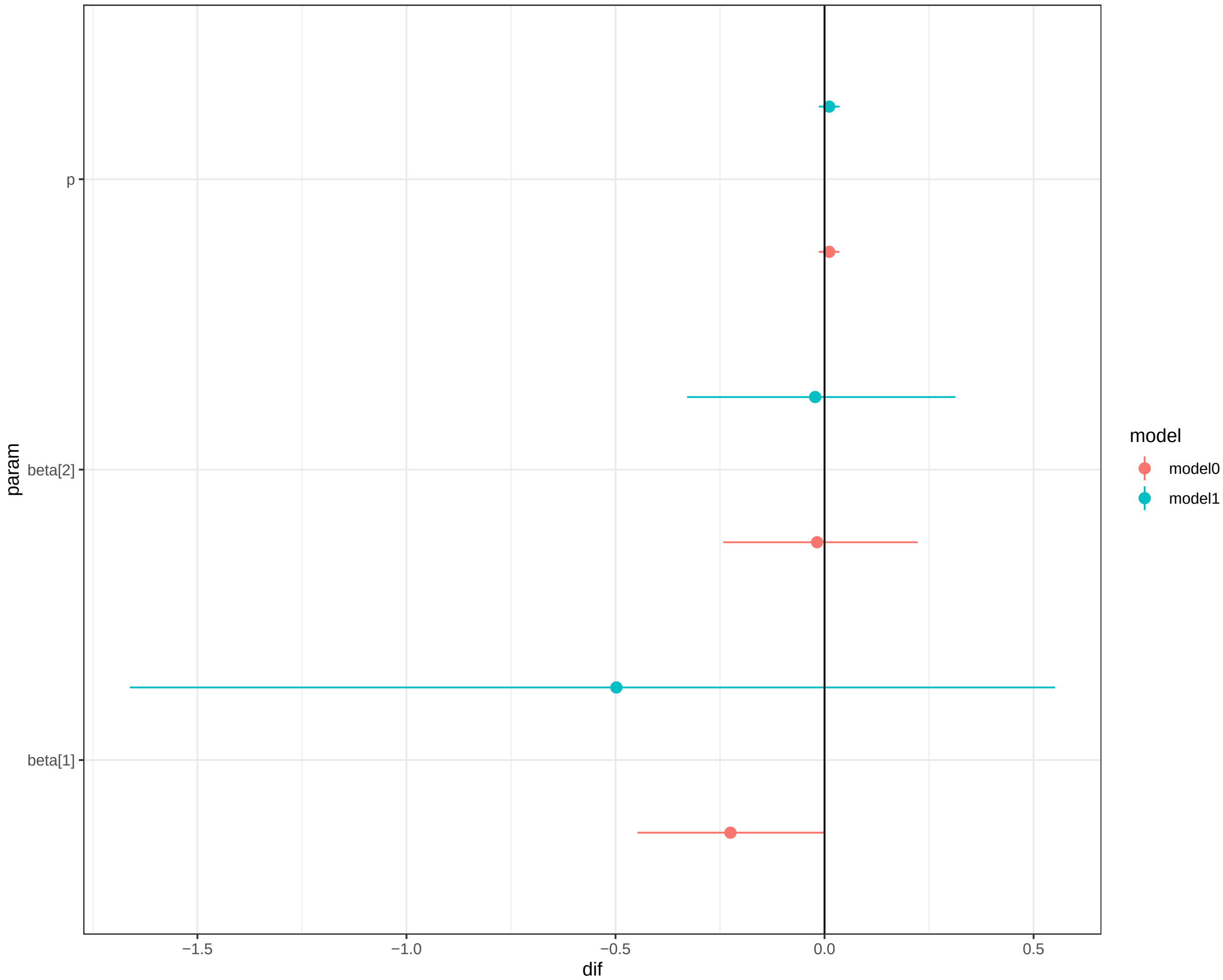
model1

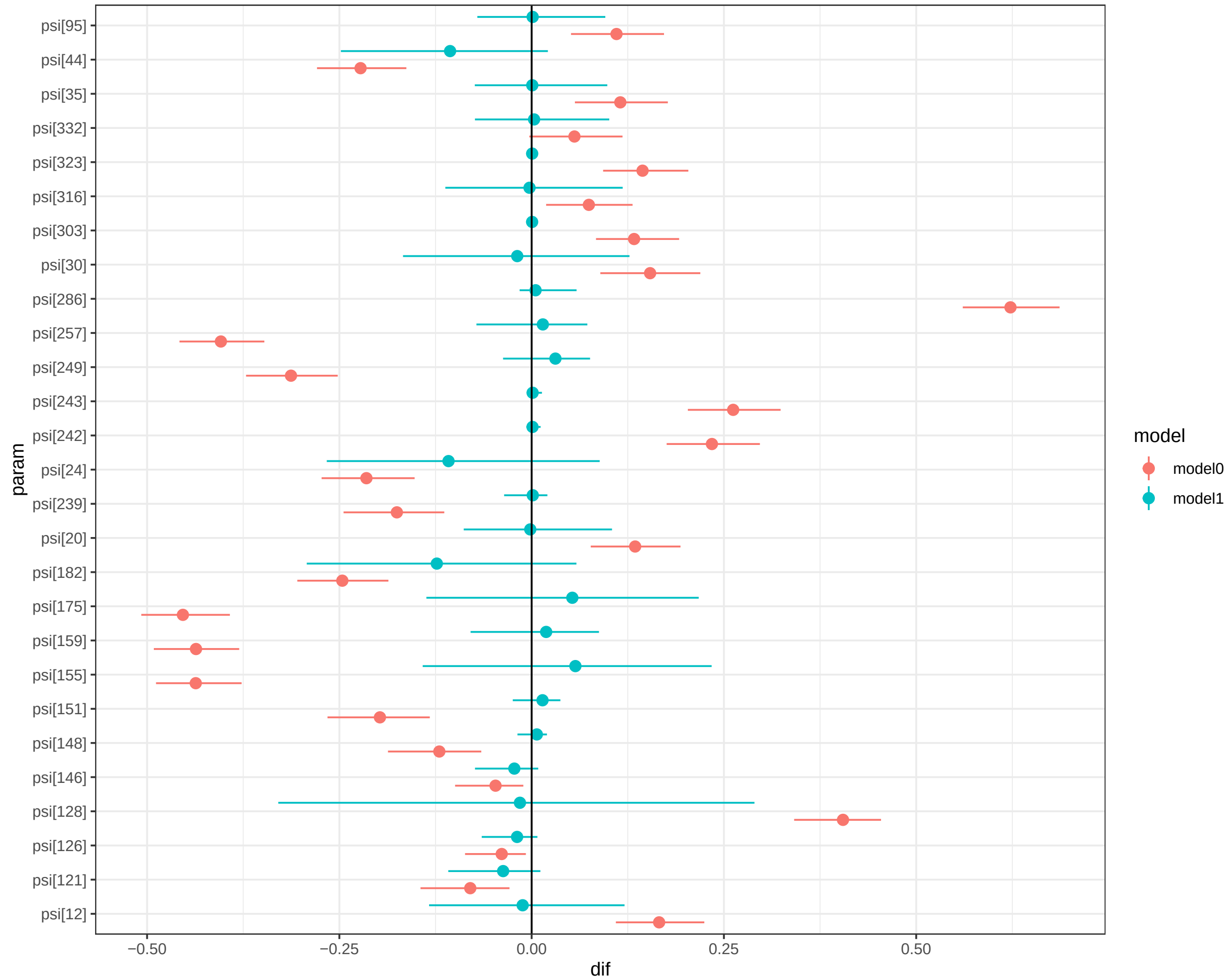
	Mean	2.5%	97.5%	Rhat	ess_bulk
p	0.51108481	0.4862074	0.5360973	1.000194	7372.1701
tau	1.26274914	0.6237233	2.4692426	1.003479	1361.7690
beta[1]	-0.49798292	-1.6616684	0.5515058	1.011493	404.4277
beta[2]	0.97757341	0.6712944	1.3127635	1.000494	3419.8080
theta[1]	-4.60452243	-7.2279204	-2.7408798	1.005345	902.1900
theta[2]	-0.07509478	-1.2112969	1.2282788	1.007554	480.3740
theta[3]	-0.45732525	-1.5981580	0.7808603	1.006037	470.4869
theta[4]	0.78886163	-0.3966938	2.0502357	1.006807	487.3187
theta[5]	2.71026218	1.4220508	4.2555043	1.002956	658.9189
theta[6]	2.63972246	1.3933441	3.9265448	1.003605	618.0339
theta[7]	-1.93068826	-3.6950165	-0.3259983	1.001281	1044.2672


theta[7]
theta[6]
theta[5]
theta[4]
theta[3]
theta[2]
theta[1]
tau
p
beta[2]
beta[1]
param
−7.5
−5.0
−2.5
0.0
2.5
mean
capture
1


psi[93]
psi[88]
psi[7]
psi[69]
psi[400]
psi[395]
psi[384]
psi[382]
psi[38]
psi[375]
psi[369]
psi[345]
psi[335]
psi[329]
psi[328]
psi[327]
psi[324]
psi[319]
psi[314]
psi[306]
psi[303]
psi[29]
psi[286]
psi[284]
psi[273]
psi[251]
psi[247]
psi[229]
psi[227]
psi[218]
psi[210]
psi[204]
psi[20]
psi[193]
psi[169]
psi[164]
psi[162]
psi[161]
psi[155]
psi[153]
psi[152]
psi[151]
psi[129]
psi[125]
psi[124]
psi[121]
psi[116]
psi[114]
psi[103]
psi[100]
param
0.00
0.25
0.50
0.75
1.00
mean
capture
1

Compare


p
beta[2]
beta[1]
param
−1.5
−1.0
−0.5
0.0
0.5
dif
model
model0
model1



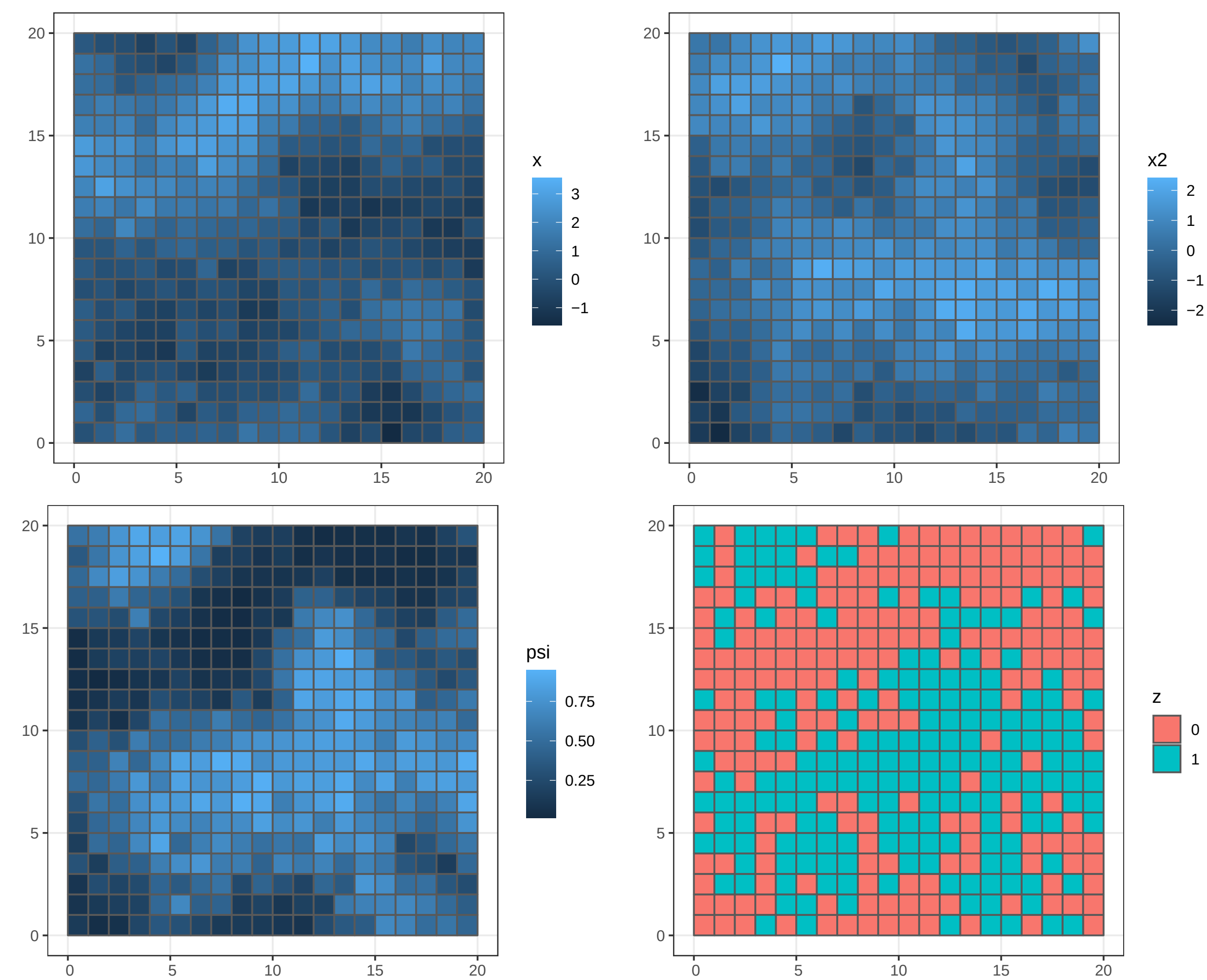
Missing covariate information

Another way to imagine the data generating mechanism is as a consequence of missing covariate information. Suppose that the true data generating model is

$$Z_i \sim \text{Bernoulli}(\psi_i)$$
$$y_i \sim \text{Bernoulli}(z_i p_{ij})$$

where $\text{logit}(\psi_i) = \beta 0 + x'_{1,i}\beta_1 + x'_{2,i}\beta_2$. Further suppose that we only directly observe x_1. In this case, the role of the experts is to adjust our mispecified model by informing hyper-priors on random effects that can be used as a surrogate for the missing covariance information. During data generation, this is achieved by thresholding the true probability surface to match the Daubenmire classes.

Simulated data



Fit models with Julia

We consider two potential models:

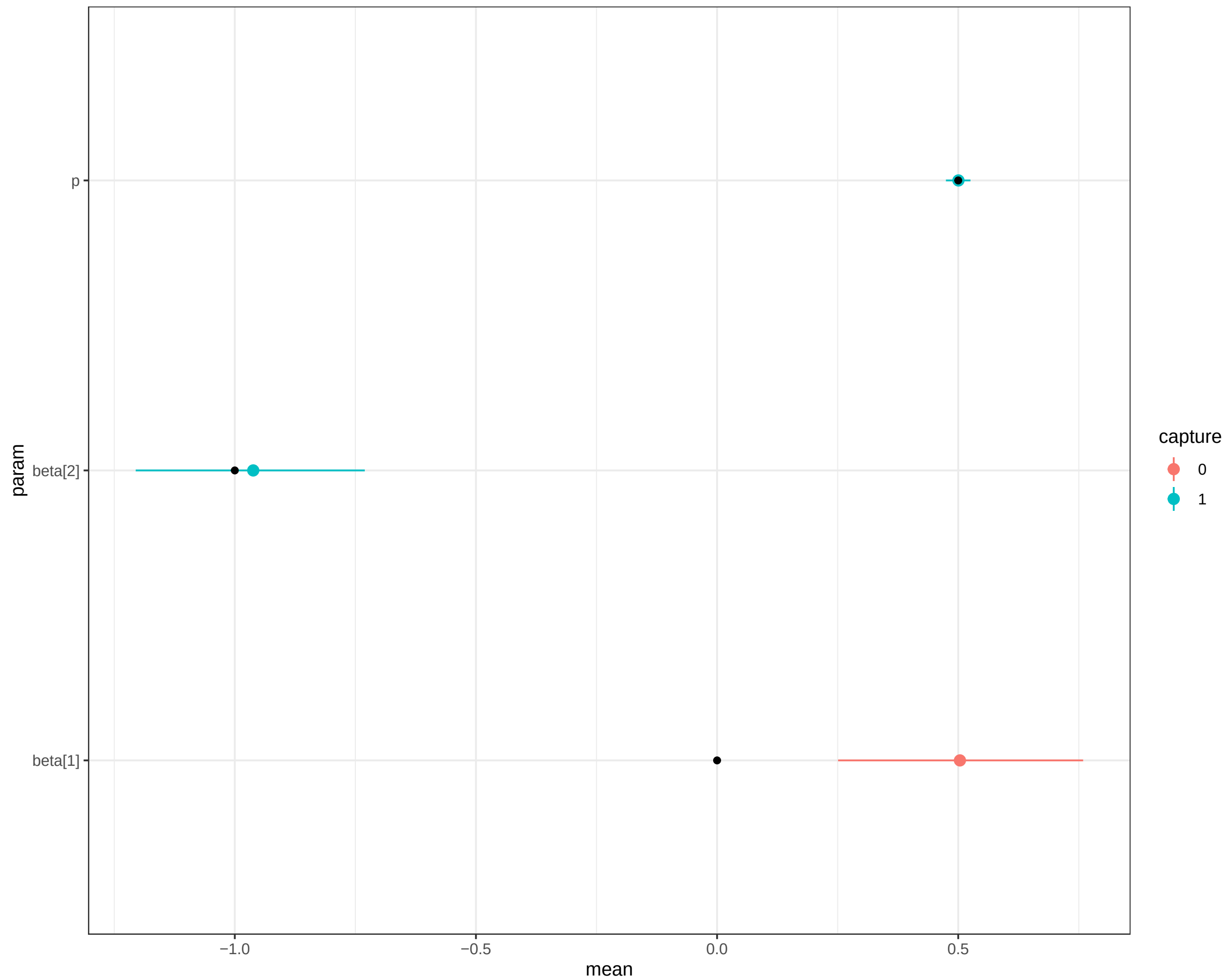
- model0: basic occupancy model without region-specific random effects
- model1: occupancy model with region-specific random effects informed by experts

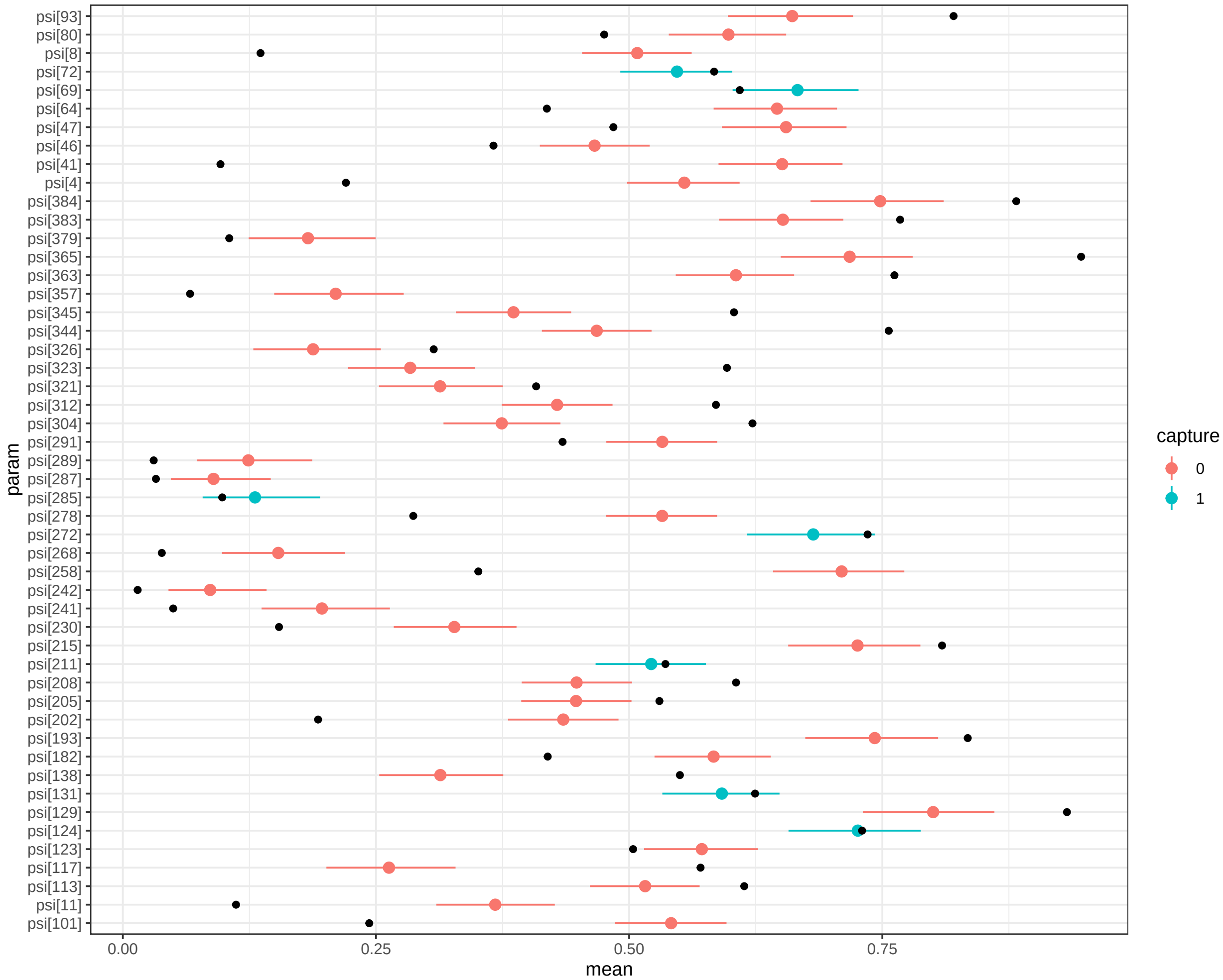
For the time being, we assume the experts are correct.

model0

	Mean	2.5%	97.5%	Rhat	ess_bulk
p	0.5004493	0.4745246	0.5253616	1.0000740	7613.910
beta[1]	0.5035283	0.2508224	0.7590192	0.9998814	6910.613
beta[2]	-0.9617888	-1.2054501	-0.7304798	1.0007261	5893.471
psi[1]	0.6283971	0.5674139	0.6871099	0.9998889	6906.563
psi[2]	0.5027038	0.4483388	0.5564906	0.9998083	6842.481

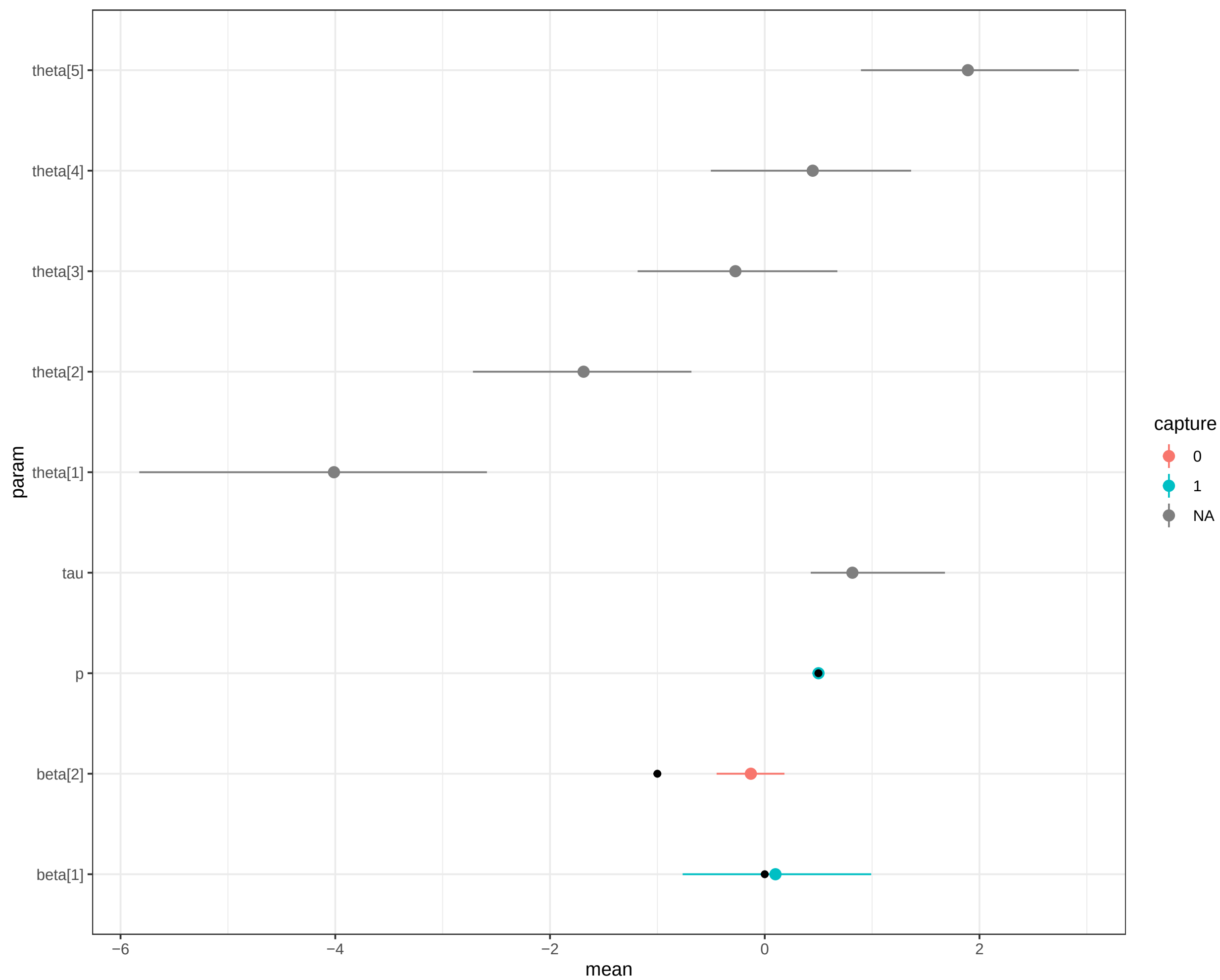
```
psi[3]   0.3641491  0.3056988  0.4229021 1.0008043 6287.028
psi[4]   0.5544383  0.4980193  0.6090761 0.9998260 6945.503
psi[5]   0.4869237  0.4328967  0.5407919 0.9998072 6794.286
psi[6]   0.4912591  0.4371543  0.5450432 0.9998057 6807.877
psi[7]   0.4554299  0.4013616  0.5099638 0.9998159 6521.480
psi[8]   0.5079869  0.4534887  0.5617067 0.9998074 6862.377
```

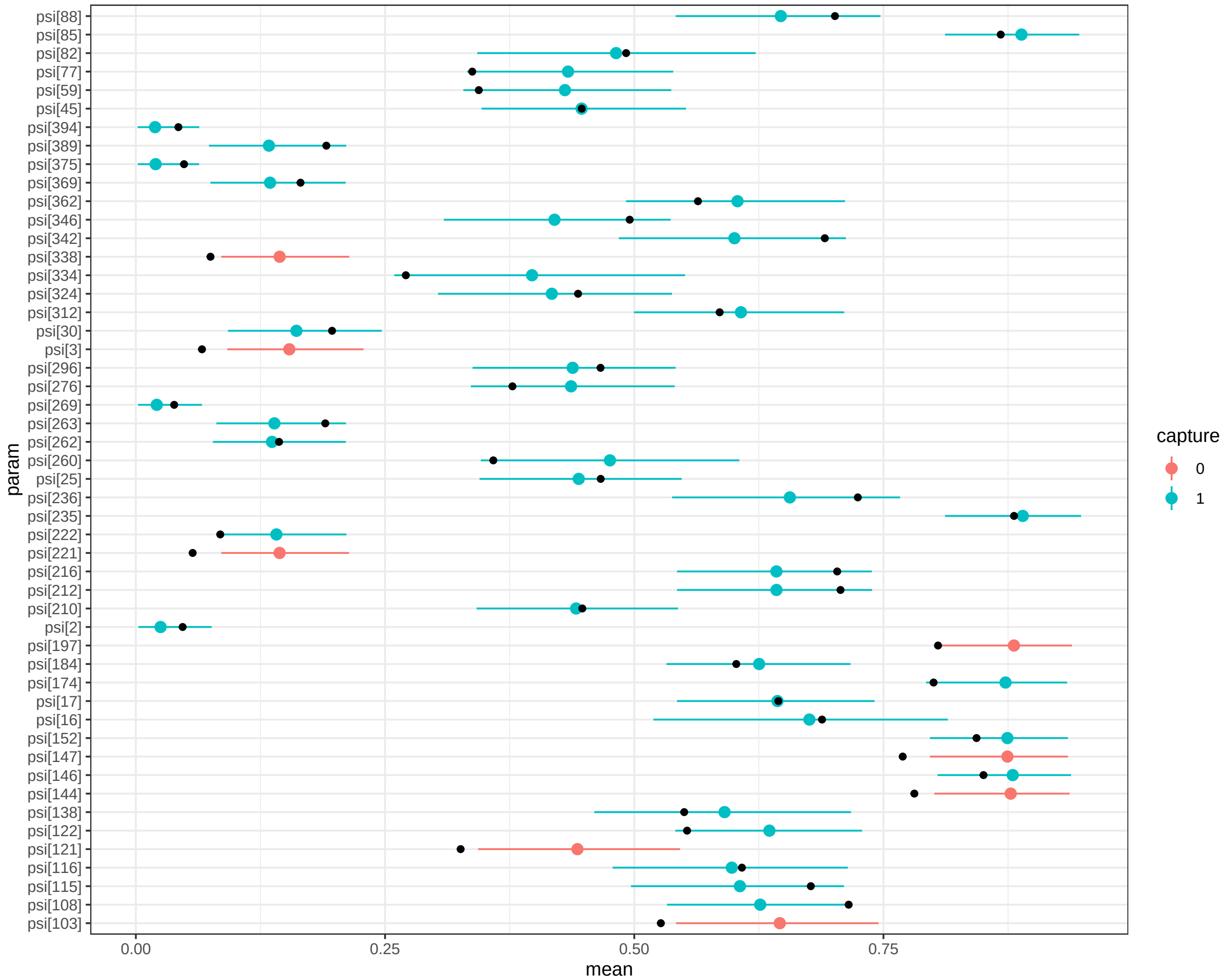




model1

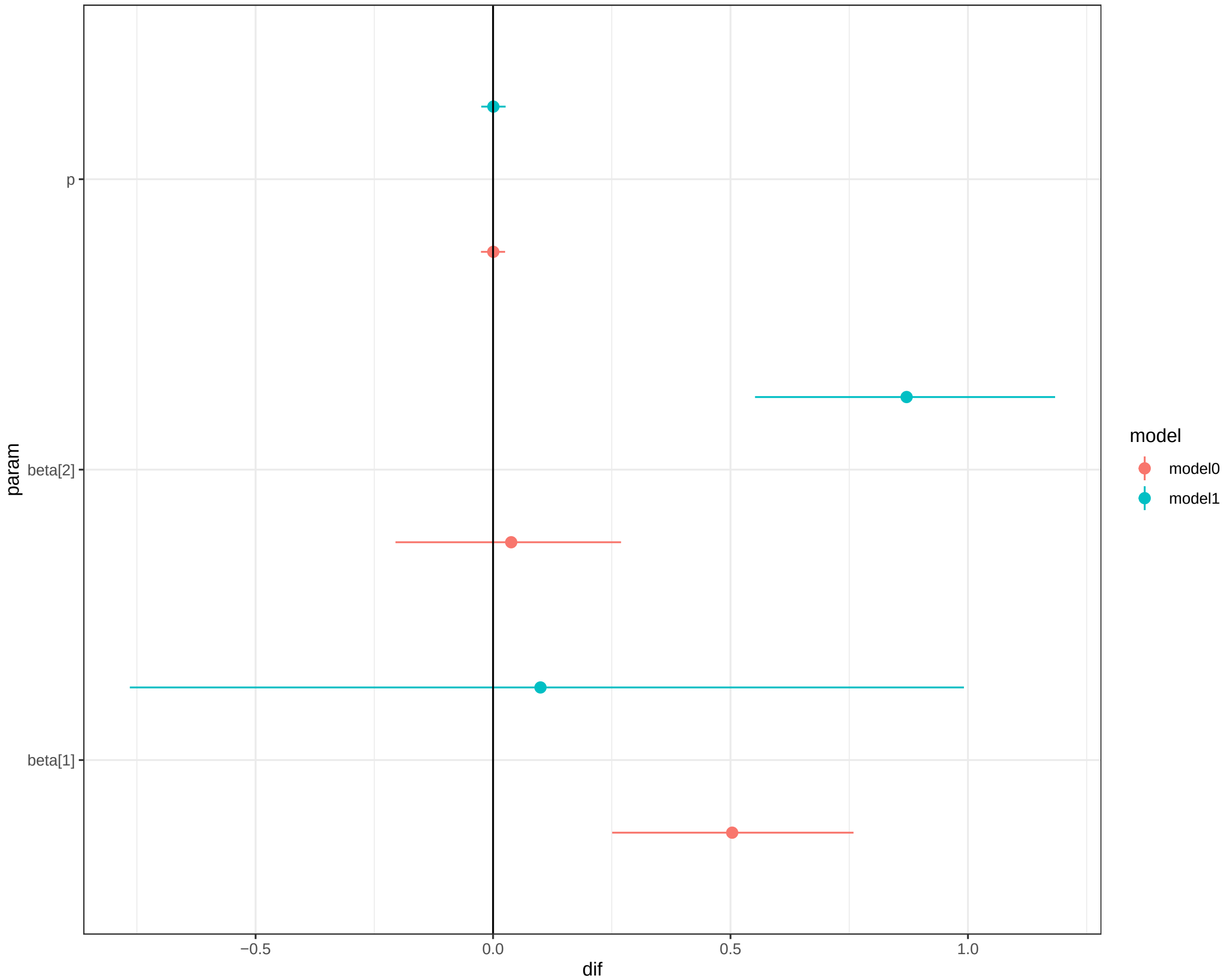
```
               Mean         2.5%      97.5%     Rhat  ess_bulk
p          0.50063780  0.475242693  0.52647355 1.000084 7288.9061
tau        0.81656074  0.428287405  1.67859888 1.001071 2447.6461
beta[1]    0.09986724 -0.764794259  0.99152233 1.005434  487.2515
beta[2]   -0.12918745 -0.448384129  0.18349802 1.002019 3900.7013
theta[1]  -4.01206799 -5.825694904 -2.58734725 1.001361 2361.6896
theta[2]  -1.68721879 -2.717810619 -0.68258461 1.005888  536.3006
theta[3]  -0.27305739 -1.183951766  0.67654796 1.005134  530.0008
theta[4]   0.44696410 -0.502108407  1.36346136 1.004521  538.8751
theta[5]   1.89197815  0.895924091  2.92623146 1.004430  678.3002
psi[1]     0.17629180  0.090133250  0.29198403 1.001163 3973.1363
psi[2]     0.02480570  0.002571172  0.07614703 1.001919 1982.2740
```

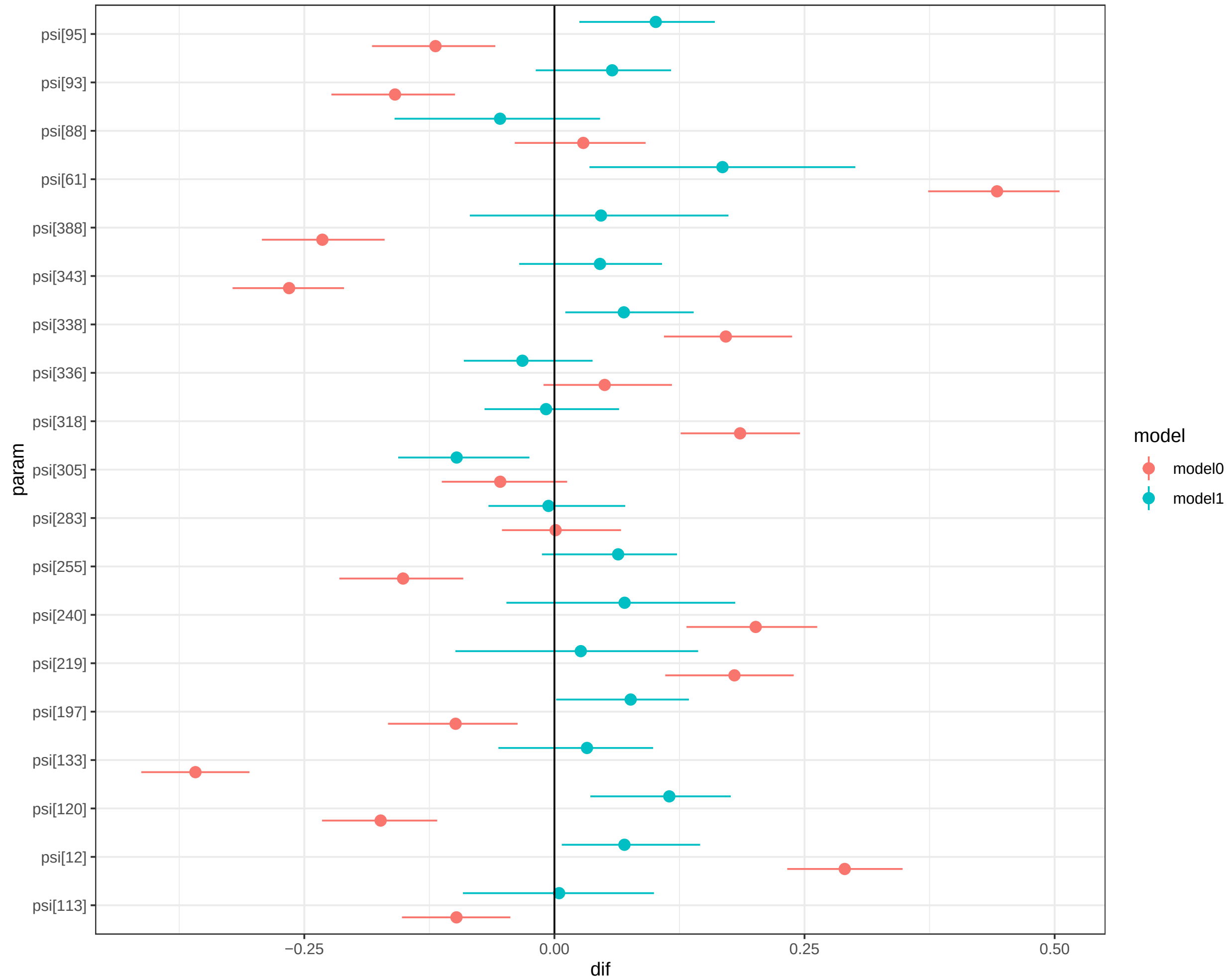

theta[5]
theta[4]
theta[3]
theta[2]
theta[1]
tau
p
beta[2]
beta[1]
param
−6
−4
−2
0
2
mean
capture
0
1
NA

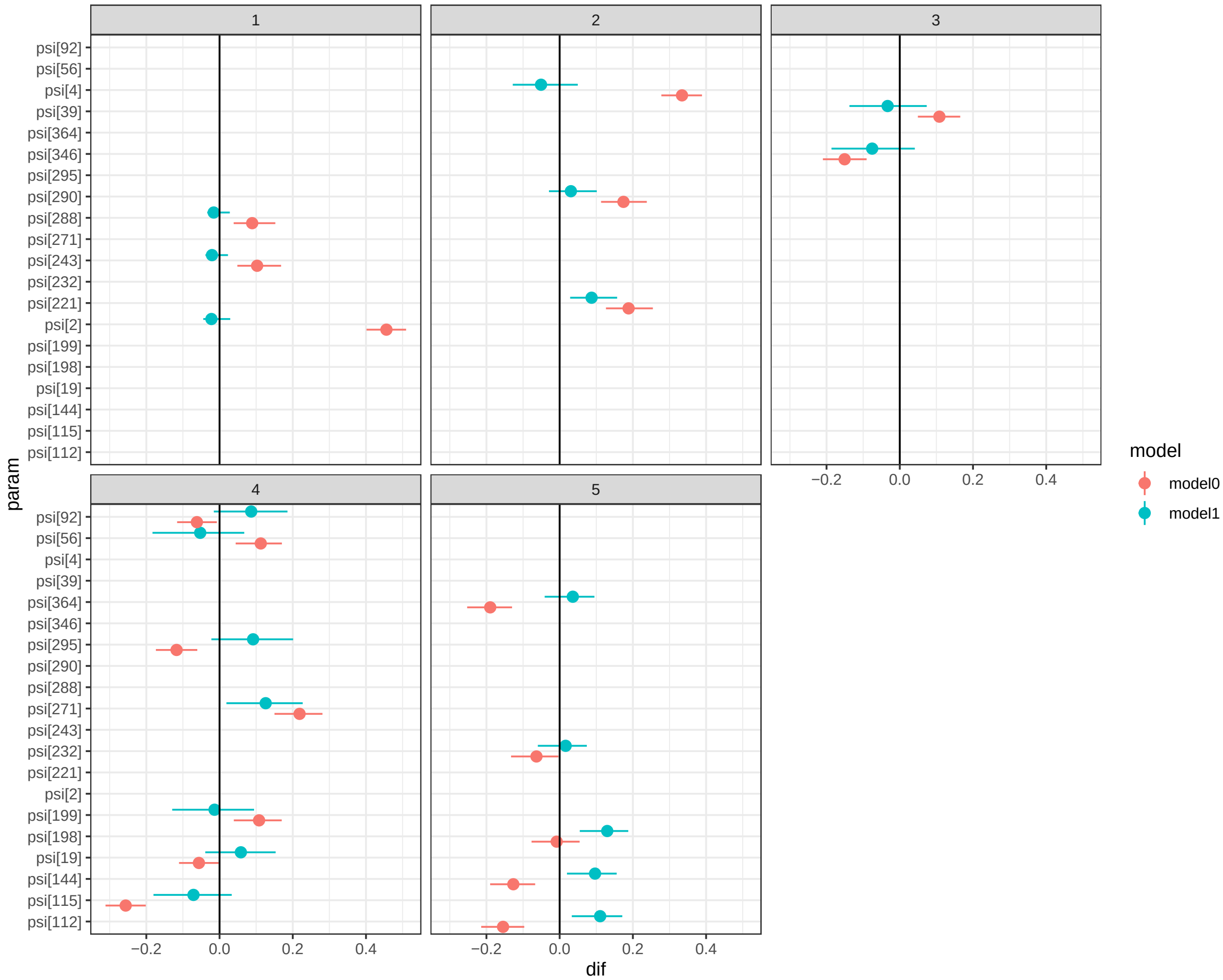


psi[88]
psi[85]
psi[82]
psi[77]
psi[59]
psi[45]
psi[394]
psi[389]
psi[375]
psi[369]
psi[362]
psi[346]
psi[342]
psi[338]
psi[334]
psi[324]
psi[312]
psi[30]
psi[3]
psi[296]
psi[276]
psi[269]
psi[263]
psi[262]
psi[260]
psi[25]
psi[236]
psi[235]
psi[222]
psi[221]
psi[216]
psi[212]
psi[210]
psi[2]
psi[197]
psi[184]
psi[174]
psi[17]
psi[16]
psi[152]
psi[147]
psi[146]
psi[144]
psi[138]
psi[122]
psi[121]
psi[116]
psi[115]
psi[108]
psi[103]
param
0.00
0.25
0.50
0.75
mean
capture
0
1

Compare


p
beta[2]
beta[1]
param
−0.5
0.0
0.5
1.0
dif
model
model0
model1


psi[95]
psi[93]
psi[88]
psi[61]
psi[388]
psi[343]
psi[338]
psi[336]
psi[318]
psi[305]
psi[283]
psi[255]
psi[240]
psi[219]
psi[197]
psi[133]
psi[120]
psi[12]
psi[113]
param
−0.25
0.00
0.25
0.50
dif
model
model0
model1



Appendix: Models in Stan

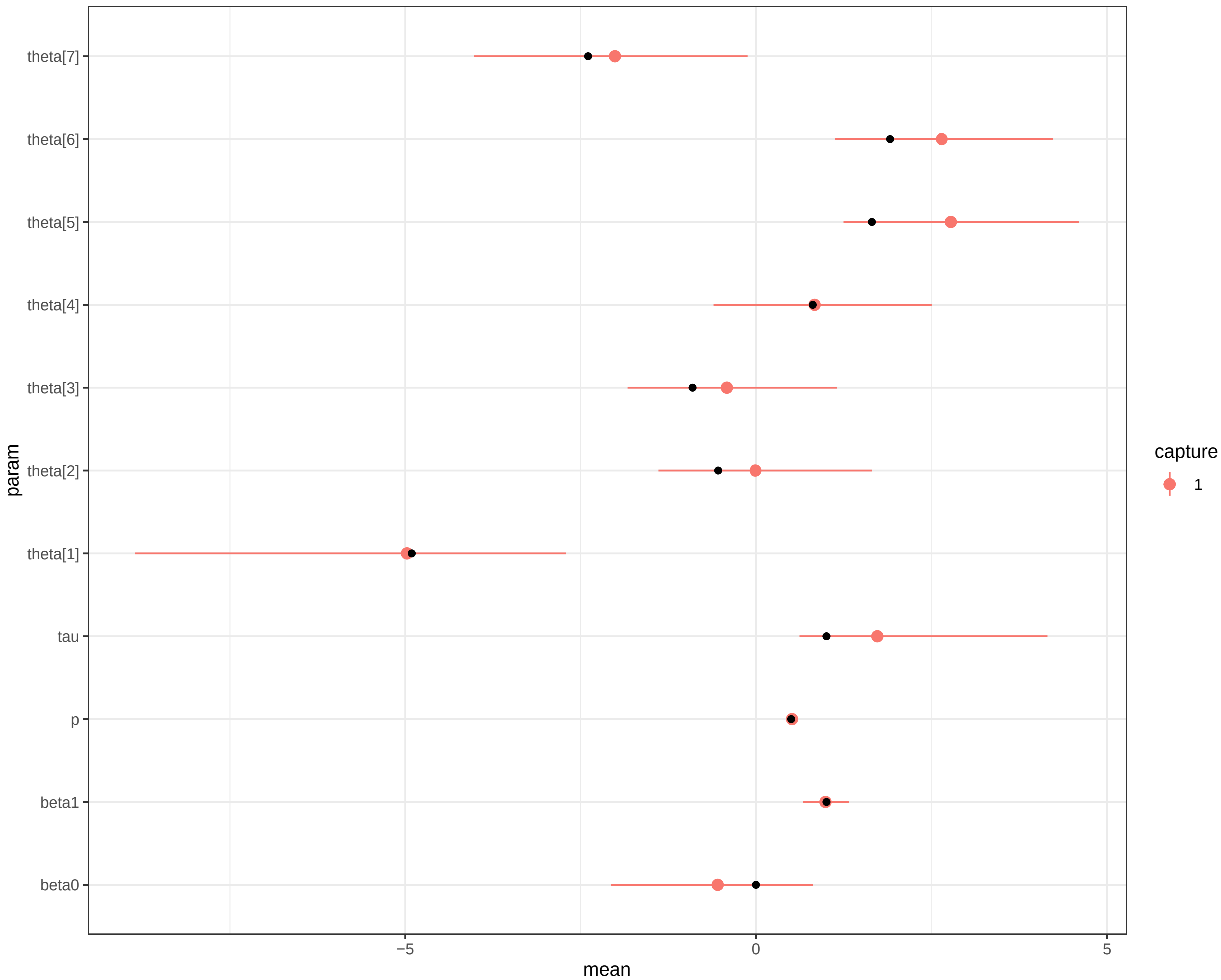
Below we provide **stan** code to fit the above model and showcase it on the data from the idyllic scenario. This model takes approximately 30 seconds to draw 5000 posterior samples (opposed to ~ 3 seconds for the Julia version).

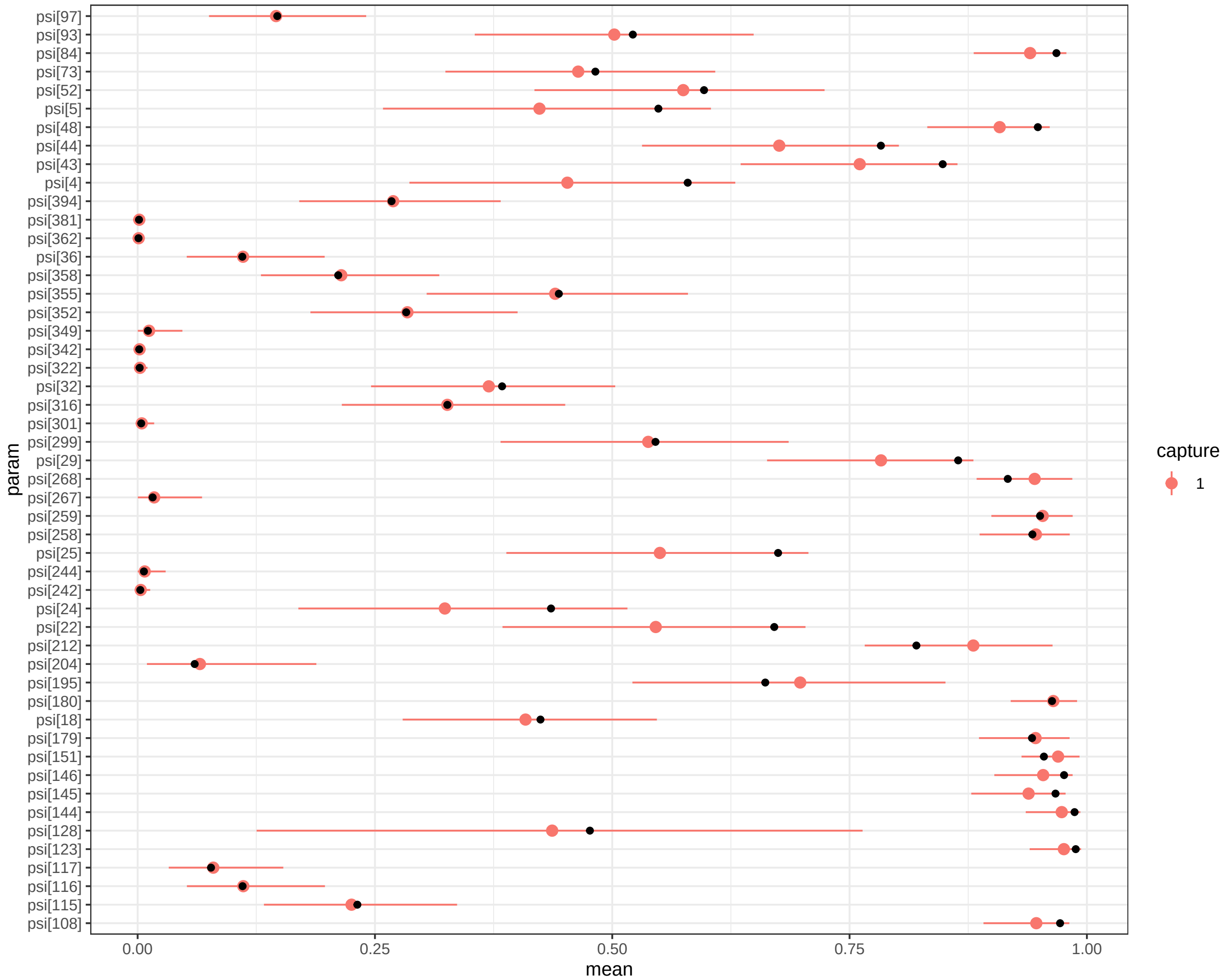
```
S4 class stanmodel 'anon_model' coded as follows:
// consider https://mc-stan.org/docs/stan-users-guide/multivariate-hierarchical-priors.html
// for more complexity and dependence among theta
data {
  int<lower=1> N; // Number of sites
  int<lower=1> J; // Number of visits (fixed for now, doesnt have to be)
  int<lower = 1> K; // Number of categories
  int<lower=0, upper=K> group[N];  // Indicator vector
  int<lower=0> sumy[N]; // Sum of obs for all visits
  real x[N];
  real m[K]; // prior means from mapping the experts
  real<lower = 0> s[K]; // prior sds from mapping the experts
}
parameters {
```

```
  real beta0;
  real beta1;
  real theta[K];
  real mu[K];
  real<lower=0, upper=1> p; // Detection probability
  real<lower=0> tau;
}
model {
  // hyper priors
  for(k in 1:K){
    mu[k] ~ normal(m[k], s[k]);
  }
  // hierarchical prior
  for (k in 1:K){
    theta[k] ~ normal(mu[k], tau);
  }

  // other stuff
  beta0 ~ normal(0, 2);
  beta1 ~ normal(0, 2);

  // Likelihood
  for (i in 1:N) {
    if(sumy[i] > 0) {
      // Occurred and observed
      1 ~ bernoulli(inv_logit(beta0 + beta1 * x[i] + theta[group[i]]));
      sumy[i] ~ binomial(J, p);
    } else {
      target += log_sum_exp(
        // Occurred and not observed
        bernoulli_lpmf(1 | inv_logit(beta0 + beta1 * x[i] + theta[group[i]])) + bernoulli_lpmf(0 | p) *
        // Not occurred
        bernoulli_lpmf(0 | inv_logit(beta0 + beta1 * x[i] + theta[group[i]]))
      );
    }
  }
}
generated quantities{
  real<lower = 0, upper = 1> psi[N];
  for(i in 1:N){
    psi[i] = inv_logit(beta0 + beta1 * x[i] + theta[group[i]]);
  }
}
```

	mean	2.5%	97.5%	n_eff	Rhat
tau	1.72806107	0.6170011	4.1543961	2733.988	0.9999114
p	0.51160411	0.4864588	0.5370601	9404.881	1.0000692
beta0	-0.54986969	-2.0702491	0.8070324	2111.170	1.0002943
beta1	0.98564169	0.6681256	1.3276212	7626.307	1.0002800
theta[1]	-4.97627500	-8.8543627	-2.7048151	3029.120	1.0003897
theta[2]	-0.00939162	-1.3894421	1.6542390	2233.658	1.0004507
theta[3]	-0.41950260	-1.8344061	1.1525231	2357.356	1.0001817
theta[4]	0.83155819	-0.6084308	2.4991407	2254.618	1.0001944
theta[5]	2.77702915	1.2413799	4.6045594	2566.216	1.0001805

```
theta[6]  2.64487948  1.1218330  4.2299176 2505.403 1.0004393
theta[7] -2.01339127 -4.0167960 -0.1252962 3464.901 1.0004062
```





Daubenmire, R.F. (1959) Canopy coverage method of vegetation analysis. *Northwest Science*, **33**, 43–64.

Dorazio, R.M. and Erickson, R.A. (2018) Ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*, **18**, 368–380.

Polson, N.G., Scott, J.G. and Windle, J. (2013) Bayesian inference for logistic models using Pólya–gamma latent variables. *Journal of the American Statistical Association*, **108**, 1339–1349.