Errata:
Hooten, M.B. and N.T. Hobbs. (2015). A guide to Bayesian model selection for ecologists. Ecological Monographs, 85: 3-28.

We have corrected a few programming bugs in the MCMC algorithm for the occupancy model and BMA in the toy example (file: 'occ.aux.mcmc.R'). The updated R code can be found at:

`http://warnercnr.colostate.edu/~hooten/papers/`

The updated code results in the following tables and figures. While the specific values changed, the corresponding model comparison message is very similar to that described in the paper. For extra stability, we also used the integrated likelihood in calculating the scoring functions. We outline the changes below.

For the section on posterior model probabilities (Table 2), the results are similar with M_2 and M_4 weighing more heavily than M_1 and M_3. However, with the new code, M_4 has a higher posterior model probability than before (0.96). Thus, we would expect it to carry most of the weight in a model average (Table 3). As before, the posterior means for β_1 and β_2 are similar, with the intercept (β_0) and detection probability (p) changing due to fixed bug in the MCMC algorithm. However, the model averaged results are as expected, averaged posterior means are most simlilar to the results from M_4 due to the high posterior model probability.

The cross-validation score and CPO agree that the two best predicting models are M_2 and M_4, with model M_3 and M_1 performing less well (Table 4). Similarly, the information criteria support these results with DIC and WAIC concurring that models M_2 and M_4 predict the best, with models M_1 and M_3 performing substantially worse (Table 5). The posterior predictive loss criterion ($D_{\infty,\text{sel}}$) indicates models M_1, M_3, and M_4 are best predicting. However, as stated in the original paper, $D_{\infty,\text{sel}}$ does not use an ideal loss function for binary data.

Finally, Bayesian regularization via an increasingly strong prior for the regression coefficients results in the same pattern as reported in the original paper (Figure 3). That is, as the coefficients shrink toward zero with increasingly stronger priors, the cross-validation score reduces until its minimum occurs at approximately $\sigma^2_\beta = 0.25$. The score then increases as the prior variance converges to zero, indicating a worse predicting null model (as originally reported). Thus, the optimal shrinkage for prediction in this example occurs with a much smaller prior variance than would have been used by default.

Table 2: Willow Tit Occupancy: Prior and posterior model probabilities.

Model	Covariates	$P(M_l)$	$P(M_l \vert \mathbf{y})$
M_1	NULL	0.25	0.00
M_2	ELEV	0.25	0.04
M_3	FOR	0.25	0.00
M_4	ELEV + FOR	0.25	0.96

Table 3: Willow tit occupancy posterior means for p, β_0, and $\boldsymbol{\beta}$ across all models and using BMA.

Parameter	M_1	M_2	M_3	M_4	BMA
p (detection prob.)	0.77	0.74	0.77	0.76	0.76
β_0 (intercept)	-0.44	-0.49	-0.48	-0.63	-0.62
β_1 (elevation)	0.00	1.03	0.00	0.89	0.90
β_2 (forest)	0.00	0.00	0.42	0.43	0.41

Table 4: Willow tit occupancy results for cross-validation and CPO.

Model	Covariates	C-V Score	$-\sum_i \log(\text{CPO}_i)$
M_1	NULL	394.7	196.8
M_2	ELEV	329.3	165.2
M_3	FOR	379.6	188.7
M_4	ELEV + FOR	322.1	160.3

Table 5: Willow tit occupancy results for WAIC, DIC, and $D_{\infty,\text{sel}}$ (posterior predictive loss).

Model	Covariates	WAIC	DIC	$D_{\infty,\text{sel}}$
M_1	NULL	393.6	393.2	80.9
M_2	ELEV	330.4	329.7	95.7
M_3	FOR	377.4	376.8	81.0
M_4	ELEV + FOR	320.7	319.9	85.7

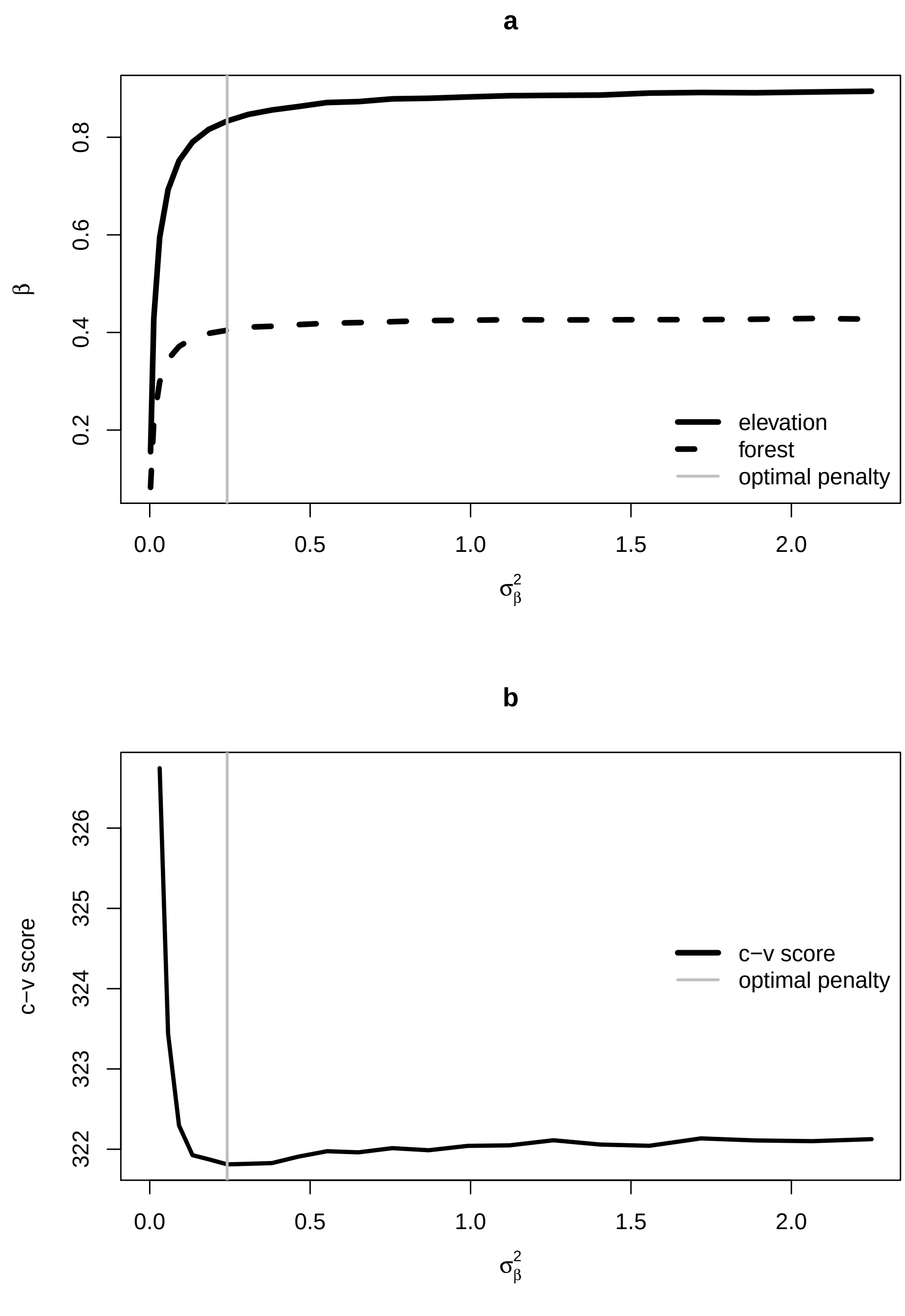
a
β
0.2
0.4
0.6
0.8
0.0
0.5
1.0
1.5
2.0
σ_β^2
elevation
forest
optimal penalty
b
c−v score
322
323
324
325
326
0.0
0.5
1.0
1.5
2.0
σ_β^2
c−v score
optimal penalty

Figure 3: