

Criteria for selecting species distribution models for management decisions

Michael A. McCarthy

School of Botany
The University of Melbourne
Parkville, VIC 3010, Australia
mamcca@unimelb.edu.au

Keywords: occupancy models; spatial ecology; species distribution models; survey design and analysis.

Abstract: Selection of species distribution models is often based on information criteria (e.g., AIC) or discrimination ability (e.g., area under ROC curves). However, when species distribution models are developed for particular management purposes, the model from a set of candidates that best supports management decisions should be selected. Determining this *a priori* is difficult. Species distribution models can assist searches for rare or invasive species (Hauser and McCarthy 2008). In this case, the probability of detecting a species at site i , given it is present, increases with search effort at that site (x_i). For a random encounter model, this probability of detection is $1 - \exp(-bx_i)$, where b is the detection rate when the species is present. The probability the species is present at each site (p_i) is predicted by a species distribution model. When a budget of search effort B is available, the allocation to each of k sites that minimizes the expected number of sites where the species is present but remains undetected is

$$x_i^* = \frac{B}{k} + \frac{\ln p_i}{b} - \frac{\ln \bar{p}}{b},$$

where \bar{p} is the geometric mean of the values of p_i . The expected number of sites with missed detections is

$$L_{\min} = \exp\left(-b \frac{B}{k}\right) \bar{p} k.$$

However, this will be a biased estimate of performance if the predicted probabilities of presence are imprecise. The expected bias in performance can be estimated depending on uncertainty in the values of p_i . Assuming beta distributions for the uncertainty (with parameters a_i and b_i for each site), the proportional bias is expected to be:

$$\text{Bias}_{\text{beta}} = \bar{f} \sum_{i=1}^k a_i (a_i + b_i - 1) / [(a_i - 1)(a_i + b_i)].$$

where \bar{f} is the geometric mean of the estimated probabilities of presence. An approximate estimate of the proportional bias for any assumed distribution is

$$\text{Bias} = \bar{f} \sum_{i=1}^k (1 + c_i^2),$$

where c_i is the coefficient of variation of the prediction at each site i . This helps consider trade-offs between species distribution models with smaller geometric mean predictions (\bar{f}) and those with less uncertainty in those predictions (c_i).

References

Hauser, C.E. and McCarthy, M.A. (2009) Streamlining ‘search and destroy’: cost-effective surveillance for invasive species management. *Ecology Letters*, 12:683-692.