# Occupancy as a surrogate for abundance estimation

# D. I. MacKenzie & J. D. Nichols

**Abstract**

*Occupancy as a surrogate for abundance estimation.*— In many monitoring programmes it may be prohibitively expensive to estimate the actual abundance of a bird species in a defined area, particularly at large spatial scales, or where birds occur at very low densities. Often it may be appropriate to consider the proportion of area occupied by the species as an alternative state variable. However, as with abundance estimation, issues of detectability must be taken into account in order to make accurate inferences: the non–detection of the species does not imply the species is genuinely absent. Here we review some recent modelling developments that permit unbiased estimation of the proportion of area occupied, colonization and local extinction probabilities. These methods allow for unequal sampling effort and enable covariate information on sampling locations to be incorporated. We also describe how these models could be extended to incorporate information from marked individuals, which would enable finer questions of population dynamics (such as turnover rate of nest sites by specific breeding pairs) to be addressed. We believe these models may be applicable to a wide range of bird species and may be useful for investigating various questions of ecological interest. For example, with respect to habitat quality, we might predict that a species is more likely to have higher local extinction probabilities, or higher turnover rates of specific breeding pairs, in poor quality habitats.

Key words: Occupancy, Species distribution, Abundance, Metapopulation, Monitoring.

**Resumen**

*La ocupación como sustituto de la estimación de la abundancia.*— En muchos programas de monitorización puede resultar extremadamente caro estimar la abundancia real de una especie de ave en un área definida, especialmente a grandes escalas espaciales, o donde las aves se dan a densidades muy bajas. A menudo, es posible que resulte conveniente considerar la proporción del área ocupada por la especie como una variable de estado alternativa. Sin embargo, al igual que sucede con la estimación de la abundancia, para poder realizar deducciones exactas es preciso tener en cuenta ciertas cuestiones de detectabilidad: el hecho de que una especie no pueda detectarse no significa que realmente esté ausente. En este estudio analizamos algunos modelos de reciente desarrollo que permiten una estimación no sesgada de la proporción del área ocupada, de la colonización y de las probabilidades de extinción local. Estos métodos permiten un esfuerzo de muestreo desigual, así como la posibilidad de incorporar información sobre covariantes en los emplazamientos de muestreo. También describimos el procedimiento para ampliarlos a fin de incorporar información acerca de individuos marcados, lo que permitiría abordar con mayor detalle cuestiones acerca de la dinámica poblacional (como el índice de rotación de los emplazamientos de los nidos por parte de parejas de reproducción específicas). Consideramos que estos modelos podrían aplicarse a una amplia gama de especies de aves, pudiendo resultar útiles para investigar diversas cuestiones de interés ecológico. Por ejemplo, respecto a la calidad del hábitat, podríamos predecir que una especie presenta más probabilidades de extinción local, o índices de rotación más elevados de determinadas parejas de reproducción, en hábitats de baja calidad.

Palabras clave: Ocupación, Distribución de especies, Abundancia, Metapoblación, Control.

*Darryl I. MacKenzie, Proteus Wildlife Research Consultants, P.O. Box 5193, Dunedin, New Zealand.– James D. Nichols, USGS Patuxent Wildlife Research Center, 11510 American Holly Drive, Laurel, MD 20708–4017, U.S.A.*

Corresponding author: D. I. MacKenzie. E–mail: darryl@proteus.co.nz

## Introduction

One of the primary roles of a wildlife monitoring program should be to track the status of populations so that substantial changes can be identified and appropriate management actions taken. Abundance (the number of individuals in a population) is one measure that can be used to characterise the state of a population for a single species, with changes in abundance reflecting changes in the population's status. However, in order to make accurate conclusions about changes in abundance, it is important that the probability of observing an individual is incorporated into our inferential process (e.g., Yoccoz et al., 2001; MacKenzie & Kendall, 2002; Williams et al., 2002; Schmidt, 2003). This often requires that individuals are identifiable (either by natural patternings such as colour patterns, or by applying unique marks such as rings) to keep accurate records of the number of encounters for each animal. For some bird species, especially those that are difficult to capture, this may require a level of effort that is infeasible to sustain as part of a long–term monitoring program, particularly at a reasonably large spatial scale.

An alternative state variable that could be considered in many situations is the proportion of area occupied by a species (which we refer to henceforth as occupancy). Determining whether a target species is present at a sampling location may be much less costly than collecting the relevant information (if possible at all) for estimating the number of individuals in an area. Such an approach has been considered in the past for a number of bird species including the northern spotted owl (Azuma et al., 1990), marbled murrelet (e.g., see Stauffer et al., 2002) and goshawks (P. Kennedy, pers. comm.). The reasoning behind using occupancy rather than abundance is that at an appropriate scale the two state variables should be positively correlated (i.e., occupancy may increase with increasing abundance), although it should be noted that the two state variables are addressing distinctly different aspects of the population dynamics. While intuitively the questions "What fraction of the landscape does the species occupy?" and "How many individuals of this species are in the landscape?" are similar, it must be recognised that some changes in the size of the population may not be identified using an occupancy approach to monitoring (e.g., changes in animal density) and that some changes in range and occupancy may not be reflected by changes in abundance. However, for certain species the discrepancies between the two state variables may be minimal if the size of sampling unit is chosen appropriately. For example, the number of breeding pairs of a territorial bird species (such as many raptors) may be closely related to occupancy if the sampling unit is chosen to be approximately the same size as a nesting territory. There may also be situations where occupancy is actually the state variable of direct interest, such as when investigating changes in species range and metapopulation incidence functions.

Few species are likely to be so conspicuous that they will always be detected at a sampling unit (site) when present. Dependent upon the survey methods being used, there may be a reasonable chance that the species goes undetected and is declared to be "falsely absent". By not correcting for the fact that the species may go undetected, a naïve count of the number of sites where the species is detected will underestimate the true level of occupancy. Furthermore, inferences about changes in occupancy based upon an observed difference between two (or more) naïve counts should be made with caution, as the difference may be the result of a change in our ability to detect the species rather than a change in occupancy. The arguments against using a naïve count for occupancy are very similar to those given for not using a simple count as an index of abundance (e.g., Yoccoz et al., 2001, MacKenzie & Kendall, 2002; Williams et al., 2002; Schmidt, 2003).

Recently there have been a number of methodological advances for modelling occupancy data while explicitly allowing for the fact that the species may go undetected at a site when present. These can be classified into single season models with homogenous detection probabilities (MacKenzie et al., 2002; Tyre et al., 2003); a single season model with heterogeneous detection probabilities caused by variation in abundance (Royle & Nichols, 2003); and a multiple season model without heterogeneity (Barbraud et al., 2003; MacKenzie et al., 2003). These new likelihood–based approaches provide a statistically robust framework for modelling occupancy data, enabling occupancy to be seriously considered as a surrogate for abundance in monitoring programs. There are strong similarities between these methods and mark–recapture models for individual animals, but there are also some subtle differences in their application.

In this paper we briefly review the multiple season model of MacKenzie et al. (2003). This is very similar to the approach of Barbraud et al. (2003) although the latter approach only models the detection histories following the season in which the species was first detected at the site. The differences between the approaches of MacKenzie et al. (2003) and Barbraud et al. (2003) are in some ways analogous to the differences between the Jolly–Seber and Cormack–Jolly–Seber mark–recapture models (e.g., Seber, 1982; Williams et al., 2002). We also outline how information from marked individuals may be incorporated into the model.

These approaches to the modelling of occupancy dynamics may be very useful for identifying the underlying processes that generate patterns in occupancy (e.g., metapopulations). In particular, we believe that such modelling is likely to be more useful than the common approach of attempting to draw inferences about such processes by observations of occupancy pattern over space at a single point in time (e.g., Hanski, 1992, 1994, 1997). Indeed, there are often many different bio-

logically reasonable processes that can result in the same pattern of occupancy (e.g., Tyre et al., 2001). This should not be surprising. As an analogy, suppose that you are given a randomly selected photograph from a stack of photographs taken throughout a football game. You are then asked to comment on the current state of the game, and how the game has progressed up to that point. It would be possible to tell the current state of play such as which team has the ball and possibly the score; however it would be impossible to make further comment on how the game has progressed. Not until you are able to go through the entire stack of photographs (in order) would you be able to get some idea of how the game progressed. It is the same situation in ecological studies where processes of population dynamics can only be fully understood by observing the population at systematic points in time, noting how the patterns change and modelling these changes in terms of relevant rate parameters.

## Basic sampling scheme

Suppose we wish to estimate the level of occupancy for a target species in some arbitrarily defined "area". The term "area" is used ambiguously here, and may consist of a continuous region such as a forest or national park, or it may be a collection of discrete habitat patches such as ponds or fragmented forest stands. The area can be considered as a collection of subunits that we shall generically refer to as sites. Depending upon the situation and target species, a site may constitute a suitably sized quadrat, potential nesting territory or an individual habitat patch. At $n$ chosen sites, multiple presence/absence (or more correctly detection/nondetection) surveys are conducted for the target species over a relatively short timeframe: a season. During the season all sites are closed to changes in occupancy so that sites are either always occupied or always unoccupied (this may be relaxed as long as the changes are completely random, although it alters the interpretation of the parameters, e.g. proportion of area occupied becomes proportion of area used).

Careful consideration needs to be given to the exact method for selecting the $n$ sites from the area of interest. One of the fundamental rules for statistical inference states that in order to be able to generalize the results from the study sites to the larger area, the sites must be selected from the larger area using a valid probability sampling scheme (e.g., random sampling). This is sometimes overlooked in ecological studies. Failure to select the sites appropriately may lead to estimates that do not correspond to the desired characteristic of the population. We do not give further consideration to the issue of site selection here as the best advice is often situation specific, but we wish to highlight that it is an important issue that is often not given adequate deliberation.

The series of detections and nondetections from the repeated surveys of a generic site $i$ can be recorded as a sequence of 1's and 0's (respectively), which we refer to as a detection history ($H_i$). For example $H_i = \{10\ 00\ 11\}$ would denote that the site has been surveyed for three seasons, with two surveys per season. In this case the species was detected in the first survey of season one; not detected at all in season two; then detected in both surveys during season three. By modelling the underlying stochastic processes that may have caused the observed detection history (just as in much of mark–recapture modelling), we can build a model that will enable us to estimate the quantities of interest.

## A multiple season model

Let $\psi_1$ be the probability a site is occupied by the species in the first season ($t =1$) and $p_{t,j}$ be the probability of detecting the species, given presence, in survey $j$ within season $t$. Further, let $\gamma_t$ denote the probability an unoccupied site becomes occupied by the species between seasons $t$ and $t+1$ (colonization), and let $\varepsilon_t$ denote the probability a site that was occupied by the species in season $t$, is unoccupied in $t+1$ (local extinction). These dynamic parameters enable the modelling of changes in occupancy that may occur between seasons.

For any given detection history, these parameters can be used to describe the process that may have resulted in the observed data. For example, consider the history $H_i = \{01\ 00\}$ indicating that the species was detected in the second survey of the site in season 1, and undetected otherwise. Obviously the site was occupied in the first season with the species being detected, hence the probability of observing the first season's data would be $\psi_1(1 - p_{1,1})p_{1,2}$, but in the second season there are two options. Either the species did not go locally extinct and was not detected in either survey, with probability

$$(1-\varepsilon_1)\prod_{j=1}^{2}(1-p_{2,j})$$

or the species did go locally extinct between seasons so it was not there to be detected (with probability $\varepsilon_1$). The probability of observing the complete history would therefore be:

$$\Pr(H_i = \{01\ 00\}) =$$
$$= \psi_1(1-p_{1,1})p_{1,2}\left\{(1-\varepsilon_1)\prod_{j=1}^{2}(1-p_{2,j})+\varepsilon_1\right\}$$

A slightly more complicated second example would be for the history $H_i = \{00\ 11\}$. Now there are two options for the occupancy state of the site in the first season; therefore we must consider the possible processes that could have resulted in the site being occupied immediately before the start of the surveys for the second season. Either the spe-

cies was present, but undetected, in the first season then did not go locally extinct, or the species was not present at the site in the first season and colonized the site between seasons. The probability of observing this history could therefore be expressed as:

$$\Pr(\mathbf{H}_i = \{00\ 11\}) =$$
$$= \left\{ \psi_1 \prod_{j=1}^{2} \left(1 - p_{1,j}\right)(1 - \varepsilon_1) + (1 - \psi_1)\gamma_1 \right\} p_{2,1}p_{2,2}$$

Generally, however, there could be a large number of possible pathways that would result in the same detection history. It is therefore useful to define a transition probability matrix that details how sites can transition between an occupied and unoccupied state between seasons $t$ and $t+1$ (1). A row vector must also be defined to indicate which occupancy state the site is in the first season (2).

$$\phi_t = \begin{bmatrix} 1 - \varepsilon_t & \varepsilon_t \\ \gamma_t & 1 - \gamma_t \end{bmatrix} \qquad (1)$$

$$\phi_0 = \begin{bmatrix} \psi_1 & 1 - \psi_1 \end{bmatrix} \qquad (2)$$

A detection probability column vector needs to be defined that indicates the probability of observing the portion of the detection history relating to season $t$, $\mathbf{p}_{H_t}$, conditional upon each state. Whenever the species is detected at least once during a season, then the second element must be zero as clearly the site cannot be in the unoccupied state (for example see equation (3). Conversely, when the species is not detected within a season, then there is some probability associated with the occupied state, and if the site is unoccupied then not detecting the species is the only possible history that could be observed that season (4).

$$\mathbf{p}_{\{01\}t} = \begin{bmatrix} (1 - p_{t,1})p_{t,2} \\ 0 \end{bmatrix} \qquad (3)$$

$$\mathbf{p}_{\{00\}t} = \begin{bmatrix} \prod_{j=1}^{2}(1 - p_{t,j}) \\ 1 \end{bmatrix} \qquad (4)$$

The probability of observing any given detection history can now be easily calculated by using the following expression,

$$\Pr(\mathbf{H}_i) = \phi_0 \prod_{t=1}^{T-1} D(\mathbf{p}_{H,t})\phi_t\, \mathbf{p}_{H,T}$$

where $D(\mathbf{p}_{H,t})$ is a diagonal matrix with the elements of $\mathbf{p}_{H,t}$ along the main diagonal (top left to bottom right), zero otherwise, and $T$ is the number of seasons of data collection. The model likelihood is then calculated in the usual manner assuming that the detection histories from the $n$ sites are independent.

$$L\left(\psi_1, \boldsymbol{\varepsilon}, \boldsymbol{\gamma}, \mathbf{p} \mid \mathbf{H}_1, \ldots, \mathbf{H}_n\right) = \prod_{i=1}^{n} \Pr(\mathbf{H}_i)$$

Once the likelihood has been defined then estimation may proceed using either maximum likelihood or Markov chain Monte Carlo.

It is worth noting that based upon the probabilities defined above it is possible to derive two additional biologically relevant quantities, the probability of occupancy in any given year (5) and the rate of change in occupancy between successive years (6).

$$\psi_t = \psi_{t-1}(1 - \varepsilon_{t-1}) + (1 - \psi_{t-1})\,\gamma_{t-1} \quad (5)$$

$$\lambda_t = \frac{\psi_{t+1}}{\psi_t} \qquad (6)$$

The model may even be reparameterized so that these quantities are estimated (or modelled) directly. However, experience to date suggests that it can be difficult to obtain convergence on the estimates for reparameterized models.

## Extensions

### Missing observations

A likely feature of many ecological studies is the existence of missing observations. In some instances it might not be possible to collect the required data: weather conditions may prevent access to some sites; vehicles may breakdown en route; or logistically it may not be possible to sample all sites within a suitably small time frame. MacKenzie et al. (2002) and MacKenzie et al. (2003) show that missing observations can easily be incorporated into the models described above. In effect, the detection probability for the respective survey of a site is set to zero, which fairly reflects the fact that the species could not be detected (even if present) as no survey was conducted at that time. Essentially, this removes the detection probability parameter from the model likelihood (with respect to the site and time in question). The ability of the model to handle missing observations has important ramifications for study designs, as it enables different sites to have different sampling intensities.

### Incorporating covariates

Often researchers may be interested in potential relationships between the model parameters (occupancy, colonization, local extinction and detection probabilities) and characteristics of the sites or generalized weather patterns (e.g., drought years). Further, the surveyor's ability to detect the species during any given survey may also be affected by localized conditions at the sampling site (e.g. weather conditions or intensity of nearby traffic noise). Using the logistic model (7), MacKenzie et al. (2003) detail how such covariate information can be incorporated. The logistic model allows the relationship between the probability of interest for site $i$ ($\theta_i$) and the respective covariate (or covariates; $\mathbf{Y}_i$) to be modelled, where $\boldsymbol{\beta}$ (which may be a vector) is the magnitude or coefficient

for the covariate(s). Analyses of this type could be considered as generalised logistic regression analyses, where allowance has been made for uncertainty in the binary observation of occupancy state (due to imperfect detectability).

$$\theta_i = \frac{\exp(\mathbf{Y}_i \boldsymbol{\beta})}{1 + \exp(\mathbf{Y}_i \boldsymbol{\beta})} \qquad (7)$$

It should be noted that the logistic model is not the only possible method for including covariate information, and that other functional forms may be used if desired.

## Including information from marked individuals

An important question that often arises in various ecological studies is whether sites that are continuously occupied are occupied by the same individuals or whether there is instead turnover of animals at some sites (the "rescue effect" of Brown & Kodric–Brown, 1977). In our opinion, it is not possible to reliably differentiate between the two possibilities from detection/nondetection data, and auxiliary information is required. Such information may be obtained from having uniquely marked individuals in the study population. Below we conceptualize how the above modelling approach could be extended to include this type of information into our inference. We imagine that such an approach would be useful for species where a site is only occupied by a single unit such as a single individual or where small groups effectively exist as a single unit (e.g., breeding pairs).

We could now consider that a site may be in one of three possible mutually exclusive states; i) occupied by the same individual as the previous season (state S); ii) occupied by a different individual from the previous season (D); or iii) not occupied by the species (N). However in the first season, there is no information regarding which sites were occupied by which individuals in the previous year as the sites were not previously being monitored. Therefore, in the first season there are only two states that can be considered; occupied and unoccupied. The transition probability matrices for $t \geq 1$ can then be redefined as (see table 1 for parameter definitions);

$$
\begin{array}{ccc}
\phantom{\phi_1 =} S & D & N
\end{array}
$$
$$
\phi_1 = \begin{bmatrix} \upsilon_1 & \varphi_1 & \varepsilon_1 \\ 0 & \gamma_1 & 1 - \gamma_1 \end{bmatrix}
$$
$$
\phi_t = \begin{bmatrix} \upsilon_t^S & \varphi_t^S & \varepsilon_t^S \\ \upsilon_t^D & \varphi_t^D & \varepsilon_t^D \\ 0 & \gamma_t & 1 - \gamma_t \end{bmatrix} \qquad \text{for } t \geq 2.
$$

Rows of $\boldsymbol{\phi}_t$ denote the occupancy state of sites in season $t$, and columns denote the state in season $t + 1$. Between any two seasons, then all possible transitions are possible except that a site can not go from an unoccupied state to being occupied by the same individual the following season (as no individual was there previously), hence the bottom–left

element of the transition matrix will always be zero. The probabilities in each row of the transition matrix should sum to 1.0, hence not all of the parameters can be independently estimated (i.e., the third probability could be obtained by subtraction), although we have presented the concepts here in terms of a very general model. In practice not all of the parameters may be identifiable; this is a continuing area of research. However, various constraints could be imposed upon the parameters to express (and compare competing) plausible biological hypotheses. For example, is the probability that a new individual occurs at a site different for sites that had an established individual last season (transition: S→D) from those sites that had a new individual last season (D→D). Such a hypothesis could be investigated by comparing sets of models where the constraint $\varphi_t^S = \varphi_t^D$ is imposed against models without such a constraint.

In any given season, however, there are four types of observations that could be made. The same or a different individual may be detected at the site, the species may not be detected at all during the season (which could mean the site is truly in any of the three states because of imperfect detectability), and fourthly, because not all of the individuals may be marked, the species may be observed there but it is unknown whether it is the same or a different individual occupying the site (U). However these can be easily accommodated within redefined capture probability vectors. In the case where the site's state is known with certainty, there is only one non–zero element in the vector, i.e.,

$$
\mathbf{p}_{\{D\}t} = \begin{bmatrix} 0 \\ p'_t \\ 0 \end{bmatrix}
$$

where $\{D\}$ denotes some detection history within season $t$ that indicates the site is occupied by a different individual, and $p'_t$ is the probability of observing the specific sequence of detections within a season. For seasons where the species was not detected, all three elements will be non–zero indicating that the site may have been in any state, i.e.,

$$
\mathbf{p}_{\{0\}t} = \begin{bmatrix} 1 - p'_t \\ 1 - p'_t \\ 1 \end{bmatrix}
$$

For the final situation where the site is known to be occupied, but it unknown whether it is a new or previous occupant, then the first two elements will be non–zero, i.e.,

$$
\mathbf{p}_{\{U\}t} = \begin{bmatrix} p'_t \\ p'_t \\ 0 \end{bmatrix}
$$

Deriving the probability for a given detection history and calculating the model likelihood can then proceed as above.

Table 1. Definition of parameters (P) used to describe the transitions between states for occupancy studies with information from marked individuals.

*Tabla 1. Definición de los parámetros (P) utilizados en los estudios de ocupación para describir transiciones entre estados, con información proporcionada por individuos marcados.*

| P | Definition of parameters |
|---|---|
| $\gamma_t$ | Probability that a site unoccupied in season $t$ is occupied in $t+1$ |
| $\varepsilon_1$ | Probability that an occupied site becomes unoccupied between seasons 1 and 2 |
| $\varepsilon_t^S$ | Probability that a site occupied by an established individual in season $t$, is unoccupied in $t+1$ |
| $\varepsilon_t^D$ | Probability that a site occupied by a new individual in season $t$, is unoccupied in $t+1$ |
| $\upsilon_1$ | Probability that the same individual occupies the site in season 2 as in season 1 |
| $\upsilon_t^S$ | Probability that a site occupied by an established individual in season $t$ is occupied by the same individual in $t+1$ |
| $\upsilon_t^D$ | Probability that a site occupied by a new individual in season $t$ is occupied by the same individual in $t+1$ |
| $\varphi_1$ | Probability that a different individual occupies the site in season 2 than in season 1 |
| $\varphi_t^S$ | Probability that a site occupied by an established individual in season $t$ is occupied by a different individual in $t+1$ |
| $\varphi_t^D$ | Probability that a site occupied by a new individual in season $t$ is occupied by a different individual in $t+1$ |

## Discussion

Abundance or population size has been the traditional state variable used in animal population studies and monitoring programmes. Here we propose the use of an alternative state variable, occupancy or the proportion of area occupied by a species. For some questions involving geographic distribution, range size, and metapopulation dynamics, this is the state variable of primary interest. In other situations, the reduced effort required to estimate occupancy, relative to that required to estimate abundance, may warrant consideration of occupancy as a surrogate for abundance.

Our initial work on occupancy estimation focused on estimation within a single season or other short time period and on a single species (MacKenzie et al., 2002; MacKenzie & Bailey, in press). We have recently extended this single–season work to incorporate multiple species with possible dependencies in both occupancy and detection (MacKenzie et al., 2004). Heterogeneity of detection probabilities among different sites or sampling units beyond that associated with identified covariates is a topic of special interest. Promising approaches developed by Royle (in press) and Royle & Nichols (2002) deal with heterogeneity associated with site–specific variation in abundance and even permit inference about abundances.

Above, we describe our more recent work extending estimation to deal with occupancy dynamics over longer time scales (also see MacKenzie et al., 2003) and outline a proposed approach to building more mechanistic models for the sampling situation in which uniquely marked individuals can sometimes be identified. Our current models permit estimation of rate of change in occupancy, as well as local rates of extinction and colonization, the vital rates of occupancy dynamics. In addition, these rate parameters can be modelled as functions of potentially relevant covariates including site–specific habitat, site isolation or proximity to source locations, etc. We thus believe that this framework permits investigation of a number of interesting ecological hypotheses.

## References

Barbraud, C., Nichols, J. D., Hines, J. E. & Hafner, H., 2003. Estimating rates of extinction and colonization in colonial species and an extension to the metapopulation and community levels. *Oikos*, 101: 113–126.

Bradford, D. F., Neale, A. C., Nash, M. S., Sada, D. W. & Jaeger, J. R., 2003. Habitat patch occupancy by toads (*Bufo punctatus*) in a naturally fragmented desert landscape. *Ecology,*

84: 1012–1023.

Brown, J. H. & Kodric–Brown, A., 1977. Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology,* 58: 445–449.

Hanski, I. , 1992. Inferences from ecological incidence functions. *American Naturalist,* 139: 657–662.

– 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology,* 63: 151–162.

– 1997. Metapopulation dynamics: from concepts and observations to predictive models. In: *Metapopulation biology: ecology, genetics, and evolution*: 69–91 (I. A. Hanski & M. E. Gilpin, Eds.). Academic Press, New York, USA.

MacKenzie, D. I. & Bailey, L. L. (in press). Assessing the fit of site occupancy models. *Journal of Agricultural, Biological and Environmental Statistics.*

MacKenzie, D. I., Bailey, L. L. & Nichols, J. D., 2004. Investigating patterns of species co–occurrence when species are detected imperfectly. *Journal of Animal Ecology,* 73: 546–555.

MacKenzie, D. I., Nichols, J. D. Hines, J. E., Knutson, M. G. & Franklin, A. D., 2003. Estimating site occupancy, colonization and local extinction probabilities when a species is not detected with certainty. *Ecology,* 84: 2200–2207.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A. & Langtimm, C. A., 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology,* 83: 2248–2255.

Royle, J. A. (in press). N–mixture models for estimating population size from spatially replicated counts. *Biometrics.*

Royle, J. A. & Nichols, J. D., 2003. Estimating abundance from repeated presence–absence data or point counts. *Ecology*, 84: 777–790.

Schmidt, B. R., 2003. Count data, detection probabilities, and the demography, dynamics, distribution, and decline of amphibians. *Comptes Rendus Biologies,* 326: S119–S124

Stauffer, H. B., Ralph, C. J. & Miller, S. L., 2002. Incorporating detection uncertainty into presence–absence surveys for marbled murrelet. In: *Predicting species occurances:* 357–365 (J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall & F. B. Samson, Eds.). Island Press, Washington, District of Columbia, U.S.A.

Tyre, A. J., Possingham, H. P. & Lindenmayer, D. B., 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters? *Ecological Applications,* 11: 1722–1737.

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K. & Possingham, H. P., 2003. Improving precision and reducing bias in biological surveys by estimating false negative error rates in presence–absence data. *Ecological Applications*, 13: 1790–1801.

Williams, B. K., Nichols, J. D. & Conroy, M. J., 2002. *Analysis and management of animal populations*. Academic Press, New York.

Yoccoz, N. G., Nichols, J. D. &. Boulinier, T., 2001. Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution,* 16: 446–453.