BioOne COMPLETE

# A Review of Wildlife Abundance Estimation Models: Comparison of Models for Correct Application

Author: Iijima, Hayato

Source: Mammal Study, 45(3) : 177-188

Published By: Mammal Society of Japan

URL: https://doi.org/10.3106/ms2019-0082

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

*Mammal Study* 45: 177–188 (2020) DOI: 10.3106/ms2019-0082
© The Mammal Society of Japan

Award Paper

Review

Series **"Abundance estimation of mammals"**

# A review of wildlife abundance estimation models: comparison of models for correct application

**Hayato Iijima***

*Forestry and Forest Products Research Institute, Matsunosato 1, Tsukuba, Ibaraki 305-8687, Japan*

**Abstract.** In this review, the various models to estimate wildlife abundance are organized for promoting the correct application of them in animal ecology. If individuals of the target wildlife are distinguishable, the capture-recapture (CR) model and the spatially explicit capture-recapture (SECR) model can be applied to the closed population and the Jolly-Seber (JS) model can be applied to the open population. If not, the distance sampling, N-mixture model, random encounter (RE) model, random encounter and staying time (REST) model, and removal sampling can be applied to the closed population, and the harvest-based model can be applied to the open population. Recent advances in the hierarchical model and the integrated population model (IPM) make it possible to model the abundance and demographic rate of the wildlife by considering the ecological process of the target wildlife and observation process of them and to utilize the various but fragmented data. Then, the formalization of the abundance estimation model as a hierarchical model and the construction of the IPM by considering the available data and biological characteristics of the target species are useful for future research.

**Key words:** detection probability, hierarchical model, integrated population model, population closure, robust design.

The abundance of wildlife is one of the key components in animal ecology. However, because humans cannot perfectly control and observe wildlife, the estimation of wildlife abundance is a hot topic in animal ecology. How can we estimate the abundance of wildlife? A direct count of wildlife is the most primitive way to estimate wildlife abundance. The block count (e.g., Jachmann 2002), in which researchers count the wildlife that they see within a fixed area, and the aerial survey (e.g., Pettorelli et al. 2007), in which researchers on an airplane count wildlife, are typical methods of the direct count. However, it is easy to expect that the data of direct count contains large variations. Do individuals of target species exist in the survey area? Can researchers detect individuals that certainly exist in the survey area at the survey timing? The former issue is related with the population closure (or availability), and the latter issue is related with the detection probability. Furthermore, if the survey period is long, the change of abundance by birth, mortality, and migration of the target species should also be considered. Tremendously various models for wildlife abundance estimation have been developed, and the new models are being created. These models differ in the assumption and required data.

The objective of this review is to organize the various models in estimating wildlife abundance and to promote the correct application of these models in animal ecology. First of all, I categorized the models on whether they treat marked or unmarked wildlife. After that, I categorized the models in relation to the population closure, the main target of estimation (abundance or density), and the monitoring methods.

## Abundance estimation for marked wildlife

### *Capture-recapture (CR) model for the closed population*

First, I treat the estimation way of abundance in the closed population by the CR survey because the CR

**E-mail: hayato.iijima@gmail.com*

*Hayato Iijima is the recipient of the promotion award for young scientists of the Mammal Society of Japan in 2017.*

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

model for the closed population is considered to be the most basic of all abundance estimation models. "Closed" means that individuals in the population do not increase or decrease by birth and death, and there is no migration during the study period. If individuals in the population can be identified and detected more than once during several sampling occasions, their abundance can be estimated as below (Williams et al. 2002):

$$\hat{N} = \frac{C}{\hat{p}} \tag{1}$$

$\hat{N}$ is the estimator of abundance, $C$ is the number of counted wildlife, and $\hat{p}$ is the estimator of the detection probability. Although the name of the method is "capture-recapture", researchers do not need to capture wildlife physically if they can distinguish individuals from any patterns or characteristics of their body. Then, the CR method is also known as the "mark-resight" (McClintock et al. 2009) and "mark-recapture" (Greenwood et al. 1985). At the same time, "detection" is also the same meaning as "capture."

Let's imagine the case of two sampling occasions in estimating $\hat{N}$ and $\hat{p}$ under the constant detection probability. $n_1$ is the number of animals caught and marked on the first occasion, $n_2$ is the number of animals caught on the second occasion, $m_2$ is the number of animals caught on both occasions, and $N$ is the true population size. If all animals have equal capture probabilities, the proportion of marked animals in a population after the first sample should approximate the proportion of marked animals in the second sample, i.e.,

$$\frac{n_1}{N} = \frac{m_2}{n_2} \tag{2}$$

Rearrangement of terms in equation (2) leads to the estimator of $N$ ($\hat{N}$) as below, and $\hat{N}$ is known as the Lincoln-Petersen estimator (Petersen 1896; Lincoln 1930):

$$\hat{N} = \frac{n_1 n_2}{m_2} \tag{3}$$

In this case, $\hat{p}$ is derived as $m_2/n_2$.

It is useful to generalize the procedure to obtain the estimators by using probability distribution. The multinomial distribution is as appropriate as the model in estimating abundance and detection probability. In the case of the above example (i.e., two sampling occasions with a constant detection probability), the probability for the data can be expressed below:

$$P(n_1, n_2, m_2 \mid N, p) = \frac{N!}{m_2!(n_1 - m_2)!(n_2 - m_2)!(N - n_1 - n_2 + m_2)!} \times p^{n_1+n_2}(1-p)^{2N-n_1-n_2} \tag{4}$$

The above probability can be extended to $K$ times the sampling occasions as below:

$$P\left(\{x_\omega\} \mid N, p\right) = \frac{N!}{\left[\prod_\omega x_\omega!\right](N - M_{K+1})!} p^{n\cdot}(1-p)^{KN-n\cdot} \tag{5}$$

$$n\cdot = \sum_{j=1}^{K} n_j \tag{6}$$

$x_\omega$ is the number of caught wildlife in each capture history ($\omega$) and $M_{K+1}$ is the total number of marked individuals caught during the study. Specifically, in the case of three sampling occasions, $\omega$ = {(0,0,1), (0,1,0), (0,0,1), …, (1,1,1)}, where the "0" and "1" indicate non-capture and capture in each sampling occasion, respectively.

Recently, the formalization of the CR model for the closed population as a hierarchical model was conducted by Royle et al. (2007). The hierarchical model along with the ecological model describes the ecological and latent process of the targeted wildlife, and the observation model describes the process to observe the latent state of the targeted wildlife (Royle and Dorazio 2008). In the hierarchical CR model for the closed population, the true but latent state about the presence/absence of each individual is estimated. Then, the detection/non-detection of each individual in the population is treated in the hierarchical CR model for the closed population. Implementation of the hierarchical CR model for the closed population is accomplished by the "parameter-expanded data augmentation (PX-DA)" technique (Royle et al. 2007). PX-DA consisted of two things by Royle et al. (2007): (1) adding an arbitrary number of zeros to the data set and (2) analyzing a reparametrized version of the original model. Let's assume that $n$ is the number of marked individuals, $N$ is the latent population abundance, and $M$ is the augmented data size. The relationship among them is $n < N < M$ and $N = \psi M$: $\psi$ is the inclusion probability. Then, the number of zeros that are added to the original data should be larger than $N$. The zeros can be divided into two components, i.e., "real" but unmarked individuals and individuals that do not exist in practice. The sum of $n$ and the "real" but unmarked individuals is $N$. Hierarchical CR

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

model for closed population by using data augmentation is expressed below:

$$\psi \sim \text{Uniform}(0,1) \tag{7}$$

$$z_i \sim \text{Bernoulli}(\psi) \tag{8}$$

$$N = \sum_{i=1}^{M} z_i \tag{9}$$

$$p \sim \text{Uniform}(0,1) \tag{10}$$

$$y_{i,k} \sim \text{Bernoulli}(z_i p) \tag{11}$$

$\psi$ is the inclusion probability as stated above, $p$ is the detection probability, $z_i$ is the latent state of $i$th individual whether $i$th individual truly exists in the target population ("1") or not ("0"), and $y_{i,k}$ is the augmented data of detection (1) or the non-detection (0) of the $i$th individual in the $k$th sampling occasion.

The detection probability can differ during the survey if you obtain the appropriate covariate(s) of the detection probability. Otis et al. (1978) organized various models for detection probability. $M_0$ is the model with the constant detection probability (as above example), $M_t$ is the model with the time-dependent detection probability, $M_b$ is the model with the behavior-dependent detection probability, and $M_h$ is the model with the individual-dependent detection probability. $M_t$ is suitable when the detection probability changes with time. For example, if the detection probability increases in winter because of the food shortage, $M_t$ should be used. $M_b$ is suitable when the detection probability changes by the behavioral response to trapping. For example, if the detection probability differs before and after the first capture of wildlife, $M_b$ should be used. The change of the detection probability by trapping is a well-known phenomenon that is known as "trap happy" and "trap shy" (Tuyttens et al. 1999; Zwolak and Foresman 2008; Augustine et al. 2014). If there is a behavioral response to trapping, the duration time of the behavioral response should be considered.

Because of too many parameters of $M_h$, the parameters of individual effect cannot be identifiable unless each of the parameters is regarded as a random effect that follows any probability distributions (Otis et al. 1978). Then, Huggins (1989) and Alho (1990) reorganized the models as fully-observable covariate model ($M_t$, $M_b$) and individual covariate models. Although the difference between fully-observable covariate models and individual covariate model seems difficult to understand, the former treats the effect of factor (i.e., time or the catch history) as a fixed effect and the latter treats the effect of individual heterogeneity as a random effect. As stated above, we cannot essentially observe "individual heterogeneity". Then, it is reasonable to consider the effect of individual heterogeneity as a random variable. Because all individuals including unobserved individuals are treated in the hierarchical CR model as we see in equations (7) to (11), the effect of individual heterogeneity on detection probability can be easily treated in the hierarchical CR model.

*Spatially explicit capture-recapture (SECR) model*

CR models basically cannot estimate the density of wildlife because the effective sampling area for estimation cannot be defined in the model. To estimate the density of wildlife, the location of individuals in an arbitrary space should be defined. Although there are many studies about SECR models (Hartstack et al. 1971; Efford 2004; Borchers and Efford 2008), I introduce the idea of Royle and Young (2008). The treatment of the activity center of each individual enables the estimation of the effective sampling area. Royle and Young (2008) modeled the detection probability as the relationship between the activity center of each individual and the location of observation.

Here I explain the SECR model to estimate the wildlife abundance and their activity centers for the closed population from the camera-trapping survey data. $J$ camera traps are set in a square space and the photographed wildlife can be identified. The locations of the $J$ camera traps with coordinates are $\boldsymbol{c}_j = (c_{1j}, c_{2j})$; $j = 1, 2, ..., J$. $n$ individuals are totally identified, and the number of photos of the $i$th individual of the $j$th camera in the $k$th sampling occasions ($y_{i,j,k}$) is recorded. However, some individuals may not be detected during the study period. In order to estimate true abundance $N$ in an arbitrary space region, data augmentation is useful in estimating $N$ as explained in the CR model for the closed population. An arbitrary large number of individuals with all zero records are added to the data, and $N$ is modeled as equations (7), (8), and (9).

Next, I introduce another square space $\boldsymbol{S}$ that encloses the region where $J$ camera traps are set. The activity center of each $N$ individual in $\boldsymbol{S}$ such as $\boldsymbol{s}_i = (s_{1i}, s_{2i})$; $i = 1, 2, ..., N$, is modeled as below:

$$s_i \sim \text{Uniform}(\boldsymbol{S}) \tag{12}$$

In this case, $N$ is the estimate of abundance, the area of

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

$\boldsymbol{S}$ is the effective sampling area, and $N/\boldsymbol{S}$ is the estimate of density.

With the location of activity centers of all individuals, the detection probability is modeled as the distance between the activity center of each individual (i.e., $\boldsymbol{s}_i$) and the location of observation (i.e., $\boldsymbol{c}_j$). Although many functions can be used, the half-normal function is a popular function. By using the half-normal function, the detection probability can be modeled as below:

$$p_{i,j} = \exp\left(-\frac{d_{i,j}^2}{2\sigma^2}\right) \tag{13}$$

$p_{i,j}$ is the detection probability of the $i$th individual by the $j$th observation point, $d_{i,j}^2$ is the squared distance between the $i$th individual's activity center and the $j$th observation point (i.e., $(\boldsymbol{s}_i - \boldsymbol{c}_j)^2$), and $\sigma^2$ is the scale parameter of the half-normal function. Therefore, the number of photos of each individual in each sampling occasion can be expressed as below:

$$y_{i,j,k} \sim \text{Poisson}(\lambda_0 p_{i,j} z_i) \tag{14}$$

$\lambda_0$ is the baseline detection intensity.

Here we assume that the population is closed. However, the SECR model can be extended to the open population (Gardner et al. 2010). Furthermore, Chandler and Royle (2013) extended the SECR to be able to be applied to the population without the individual identification or partially identified population (unmark SECR).

*Capture-recapture models for open population*

The estimation of the population abundance by the CR model for the closed population is robust because the relationship between the true abundance and observed counts is obviously expressed by the model. However, the assumption of the population closure is sometimes difficult to accomplish in ecological studies. Furthermore, the temporal change of abundance itself can be the interest of research. Therefore, the CR model for the closed population has been extended to the open population. "Open" means that the individuals in the population can increase or decrease by their natural mortality, birth, and migrations during the study period. To estimate the abundance of open population, the survival and entering of individuals for the target population, in addition to the detection of wildlife, should be treated.

The Jolly-Seber (JS) model (Jolly 1965; Seber 1965) estimates the entry probability in addition to the survival and detection probability. Although there are some variations of the JS model, I explain the super-population approach (Williams et al. 2002) here. For simplicity, I explain the JS model of super-population approach as the hierarchical model. In the super-population approach, we assume that the individuals enter the target population from a "super-population" by each sampling occasion. The entry probability of the individuals from the super-population is defined as below:

$$b_k \sim \text{Dirichlet}(\boldsymbol{\alpha}) \tag{15}$$

$$\eta_1 = b_1,\ \eta_2 = \frac{b_2}{1-b_1}, \cdots,\ \eta_k = \frac{b_k}{1-\sum_{i=1}^{k-1} b_i} \tag{16}$$

$b_k$ is the entry probability of the $k$th survey, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the parameter of Dirichlet distribution, and $\eta_k$ is the conditional entry probability of the $k$th survey. Usually, each component of $\boldsymbol{\alpha}$ is set as 1 (a non-informative prior distribution).

By using the entry probability in addition to the survival probability, we can model the survival and entry of each individual in each sampling occasion below. In this case, we assume the constant survival probability:

$$\varphi \sim \text{Uniform}(0,1) \tag{17}$$

$$z_{i,1} \sim \text{Bernoulli}(\eta_1) \tag{18}$$

$$z_{i,k+1} \mid z_{i,k} \sim \text{Bernoulli}\left(\varphi z_{i,k} + \eta_{k+1} \prod_{k=1}^{k} (1 - z_{i,k})\right) \tag{19}$$

$\varphi$ is the survival probability and $z_{i,k}$ is the latent variable of the presence or absence of the $i$th individual in the $k$th sampling occasion in the target population. Equation (19) indicates the presence or absence of the $i$th individual in the $k$+1th sampling occasion ($z_{i,k+1}$) is determined by the survival of the $i$th individual ($\varphi$) that existed in the target population in the previous sampling occasion ($z_{i,k}$) and the products of absence probability until the $k$th sampling occasion ($\prod_{k=1}^{k}(1 - z_{i,k})$) and the entry probability in the $k$+1th sampling occasion ($\eta_{k+1}$).

Next, I explain the observation model. Since the detection of the individual is generally imperfect in the field survey, the detection probability should be estimated. In this case, I assume that there are no covariates among the detection probability as below:

$$p \sim \text{Uniform}(0,1) \tag{20}$$

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Because the detection is imperfect, the unmarked but existing individuals in the target population should be considered to estimate the population abundance. As explained in the closed population, the data augmentation technique is again useful. The estimation of the latent state of $M$ individuals can be modeled as below:

$$\psi \sim \text{Uniform}(0,1) \tag{21}$$

$$w_i \sim \text{Bernoulli}(\psi) \tag{22}$$

$\psi$ is the inclusion probability as stated above, and $w_i$ is the latent variable among the presence (1) or absence (0) of the $i$th individual in the target population during the whole survey.

By using the above parameters, the observed data including the added zeros can be modeled as below.

$$y_{i,k} \sim \text{Bernoulli}(w_i z_{i,k} p) \tag{23}$$

$y_{i,k}$ is the detection (1) or non-detection (0) of the $i$th individual in the $k$th sampling occasion. Furthermore, abundance of the $k$th sampling occasion ($N_k = (N_1, N_2, ..., N_K)$) and total abundance of the targeted population ($\boldsymbol{N}$) can be derived as below:

$$N_k = \sum_{i=1}^{M} w_i z_{i,k} \tag{24}$$

$$Nind_i = \sum_{k=1}^{K} w_i z_{i,k} \tag{25}$$

$$Nalive_i = \begin{Bmatrix} 1\ (Nind_i > 0) \\ 0\ (Nind_i = 0) \end{Bmatrix} \tag{26}$$

$$\boldsymbol{N} = \sum_{i=1}^{M} Nalive_i \tag{27}$$

$Nind_i$ is the number how many times the $i$th individual is detected, and $Nalive_i$ denotes its state (alive or dead, or included or not-included). These two derived parameters are necessary to calculate $\boldsymbol{N}$.

It is possible that the survival and detection probabilities fluctuate with covariates like time, individual, and capture history, as is case of the detection probability of the CR model for the closed population. However, in this case, there are two parameters. It is necessary to measure covariates in each sampling occasion and to model the survival or detection probability with the measured covariates. However, if there are correlations among the factors of fluctuation of survival or detection probability, parameter identifiability may be lost. Therefore, I recommend researchers to adopt the "robust design" (Pollock et al. 2002). The central philosophy of the robust design is to conduct multiple surveys (i.e., replication) under a constant condition. Under the robust design, the sampling occasions are divided into two hierarchies. As an example of the robust design, let's assume the situation in which you would like to estimate the effects of the climate on the survival and detection probabilities in the count survey of some wildlife species. In the first hierarchy, factors like climate can change in-between the sampling occasions. In the second hierarchy, more than two samplings should be conducted within a short period when the factors can be regarded as stable. Because it is reasonable to assume that there is no mortality among samplings within a second hierarchy, the variation of the number of counts in the second hierarchy reflects the detection probability. Therefore, the usage of the robust design is important in guaranteeing the identifiability of parameters in a model.

## Abundance estimation for unmarked wildlife

*Background*

All models that are explained above coerce the researchers to identify individuals in the population. However, the identification of wildlife is sometimes impossible or practically difficult. The difficulty to identify wildlife has motivated the development of population abundance estimation way without identification of the wildlife. From here, I review such estimation models.

*Distance sampling*

Distance sampling is to count the number of individuals from fixed points or fixed transects. There are many ways to conduct distance sampling, but here I introduce the distance sampling by using the line transect sampling and the abundance estimation method by following Buckland et al. (2001). The basic model for the estimation of the wildlife density by distance sampling is expressed below:

$$\hat{D} = \frac{n}{2wL\hat{P}_a} \tag{28}$$

$\hat{D}$ is the estimator of wildlife density, $n$ is the total number of counted wildlife from a line of $L$ length within $w$ distance from the line, and $\hat{P}_a$ is the probability of detection for an object within an area $a$. $\hat{P}_a$ can be expressed as below:

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

$$P_a = \frac{\int_0^w g(x)dx}{w} \quad (29)$$

$g(x)$ is detection function in relation to $x$ length which is the distance between the line and the object. For simplicity, $\int_0^w g(x)dx$ is sometimes expressed as $\mu$ and equation (28) becomes as below:

$$\hat{D} = \frac{n}{2L\hat{\mu}} \quad (30)$$

$\hat{\mu}$ is the estimator of an effective strip half-width (ESW) that is the distance from the line for which as many objects are detected beyond $\mu$ as are missed within $\mu$ of the line. Various probability functions are available for $g(x)$, and the selection of probability function can be done by the likelihood ratio test or AIC.

The assumptions of distance sampling by the line transect sampling are categorized into experimental design and observation. The assumptions of the experimental design are that 1) objects are spatially distributed in the area to be sampled according to some stochastic process with rate parameter $D$ (i.e., number per unit area) and 2) randomly placed lines or points are surveyed, and a sample of $n$ objects are detected, measured, and recorded. The assumptions of observation are that 1) objects directly on the line or point are always detected (i.e., they are detected with probability 1); 2) objects are detected at their initial location, prior to any movement in response to the observer; and 3) distance (and angles where relevant) are measured accurately (ungrouped data) or objects are correctly counted in the proper distance interval (grouped data).

Recently, the formulation of the distance sampling model as a hierarchical model (hierarchical distance sampling model) was accomplished by Kéry and Royle (2015). Although the details of the hierarchical distance sampling model are not introduced here, the hierarchical distance sampling model is especially useful in modeling the effects of the individual covariates and the spatial heterogeneity of abundance.

*N-mixture model*

The N-mixture model was proposed by Royle (2004) to estimate the local and latent abundance of wildlife and detection probability of the latent abundance by repeated counting in space and time. The N-mixture model is expressed below:

$$N_i \sim \text{Poisson}(\lambda) \quad (31)$$

$$y_{i,k} \sim \text{Binomial}(N_i, p) \quad (32)$$

$\lambda$ is the mean abundance of the target population, $N_i$ is the local abundance in the $i$th location, $p$ is the detection probability, and $y_{i,k}$ is the observed abundance in the $i$th location and the $k$th sampling occasion.

To apply the N-mixture model, researchers need to conduct repeated observations in space and time under closed conditions. The assumptions of the N-mixture model are as follows:

1) In each sampling occasion, the duplicated count of the same individual must be avoided.
2) The detection probability should be constant during the survey.
3) The local abundance ($N_i$) should follow the used probability distribution (in equation (31), Poisson distribution).

Link et al. (2018) examined the biases of the N-mixture model under the three violations of model assumptions and recommended to collect reliable data about detection probability by CR methods because the violation of constant detection probability especially caused the biased estimates.

While, some family models of N-mixture model can estimate the abundance even if the duplicated count of the same individual occurs in a single sampling occasion. Nakashima (2020) showed that the Royle-Nichols model and Poisson/Poisson N-mixture model provided reasonable estimates of the number of animals in such situation. The Royle-Nichols model (Royle and Nichols 2003) treats the duplicated counts by reducing the information of data from abundance to presence/absence. The formulation of the Royle-Nichols model is as below:

$$N_i \sim \text{Poisson}(\lambda) \quad (33)$$

$$y_{i,k} \sim \text{Bernoulli}\left(1-(1-p)^{N_i}\right) \quad (34)$$

In this case, $y_{i,k}$ is the presence/absence of target wildlife in the $i$th location and the $k$th sampling occasion. However, it should be noted that the Royle-Nichols model can estimate realistic abundance only when abundance and detection are low (Kéry and Royle 2015). The Poisson/Poisson N-mixture model (Kéry and Royle 2015) treats the duplicated counts by using Poisson distribution as the alternative of binomial distribution of the observation model. The formulation of the Poisson/Poisson N-mixture model is as below:

$$N_i \sim \text{Poisson}(\lambda) \quad (35)$$

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

$$y_{i,k} \sim \text{Poisson}(N_i \phi_{ik}) \quad (36)$$

In this case, $y_{i,k}$ is the counts of target wildlife in the $i$th location and the $k$th sampling occasion and $\phi_{ik}$ is a rate parameter in the $i$th location and the $k$th sampling occasion.

*Random encounter (RE) model*

Rowcliffe et al. (2008) applied the model about rates of the collision between gas molecules into density estimation of wildlife by using camera-trapping data and named as the RE model. The RE model for wildlife density estimation is expressed as below:

$$D = \frac{y}{t} \times \frac{\pi}{vr(2+\theta)} \quad (37)$$

$D$ is the density of wildlife, $y$ is the observed counts of wildlife, $t$ is the time of sampling, $\pi$ is the ratio of the circumference of a circle to its diameter, $v$ is the animal speed of movement, $r$ is the radius of detection, and $\theta$ is the detective angle from the observer.

Assumptions of the RE model are as follows: (i) animals conform adequately to the model used to describe the detection process; (ii) photographs represent independent contacts between animal and camera; and (iii) the population is closed. Furthermore, as stated above, researchers need to clarify the movement speed of the target wildlife. The movement speed of target wildlife can be obtained by radio-tracking of target wildlife, but it is time-consuming to attach many radio-collars or GPS-collars to wildlife.

*Random encounter and staying time (REST) model*

Nakashima et al. (2018) extended the RE model as researchers can estimate wildlife density by only the data from camera trap. They changed the RE model as below:

$$D = \frac{\text{E}(Y) \times \text{E}(T)}{sH} \quad (38)$$

$D$ is the population density of target animal, E($Y$) is the expected number of encounters, E($T$) is the expected staying time, $s$ is the area of the detection zone, and $H$ is a research period. The important point is to use video mode instead of camera mode to evaluate the animal speed of movement. For evaluation of the animal speed, researchers should set signs in front of the camera trap to clarify when an animal will enter and how long the animal will stay in the detective zone in front of the camera.

There are seven assumptions to apply the REST model. (1) Camera traps must be placed randomly with respect to the spatial distribution of animals. It is the issue of experimental design. (2) Cameras must certainly detect the animals entering the focal area throughout the research period. It is the issue of performance of cameras. (3) Animal density must not vary during the research period; (4) animal movement and behavior are not affected by cameras; (5) observations are independent events. These are the issue of animal behavior. (6) The observed distribution of staying time in the focal area must represent a good fit for the distribution that animal movements actually follow; (7) the observed staying time must follow a given parametric distribution. These are the issue of model structure. Nakashima et al. (2018) examined the violation of some assumptions and showed that the REST model provided unbiased estimates of animal density even when animal movement speeds varied among individuals, and when animals traveled in pairs. However, the REST model is vulnerable to unsynchronized activity patterns among individuals.

*Removal sampling*

Removal sampling is to remove individuals from the target population repeatedly. Although the name is "removal" sampling, the sampled individuals are not necessarily removed eternally from the target population. For example, the removed individuals can be returned to the target population after the survey. Otherwise, an individual can be regarded as "removed" if an observer observes the individual and intentionally never counts the individual again (a mental removal protocol, Kéry and Royle 2015). Under the assumption of closed population and constant detection (catch) probability, the relationship among the number of removed individuals in $K$ times sampling occasions ($\boldsymbol{X}_k = (X_1, X_2, \ldots, X_K)$), abundance ($\boldsymbol{N}_k = (N_1, N_2, \ldots, N_K)$), and catch probability ($p$) can be modeled as below (Dorazio et al. 2005; Bord et al. 2014):

$$N_1 \sim \text{Uniform}\left(\sum_{k=1}^{K} X_k, M\right) \quad (39)$$

$$N_k = N_{k-1} - X_{k-1} \quad (40)$$

$$X_k \sim \text{Binomial}(N_k, p) \quad (41)$$

$N_k$ is the abundance in $k$th sampling occasion, $M$ is the arbitrary large number, and $p$ is the catch probability. In addition to the above assumptions, the removal should reduce the true abundance marginally. In other words, removal sampling cannot be applied when the catch probability is very low or very high. For example, if the

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

removal is not effective to reduce wildlife and then the catch probability is very low, the number of removed individuals may be similar regardless of sampling occasions. The abundance and catch probability can be estimated even in such a situation if researchers conduct many, many samplings and the population closure is certainly satisfied during the survey period, but it is practically difficult to conduct such a sampling especially in the field.

The catch probability per unit effort may decrease with the increase of sampling occasions (Iijima 2017) because individuals with low vigilance (bold individuals) tend to be sampled earlier than those with high vigilance (Honda and Iijima 2016). Mäntyniemi et al. (2005) developed the heterogeneous catchability model, which assumes that the catch probability differs among individuals in the target population and individuals with high catch probability are caught at earlier sampling occasions. The model is expressed as below:

$$X_k \sim \text{Binomial}(N_k, q_k) \tag{42}$$

$$N_k = N_{k-1} - X_{k-1} \tag{43}$$

$$q_k = \mu \frac{\eta}{\eta + k - 1},\ k = 1, \cdots, K \tag{44}$$

$$\mu \sim \text{Beta}(\alpha, \beta) \tag{45}$$

$$\alpha = \mu\eta \tag{46}$$

$$\beta = (1 - \mu)\eta \tag{47}$$

$$X_0 = 0 \tag{48}$$

$X_k$ is the number of caught individuals in the $k$th sampling occasion, $N_k$ is the abundance in $k$th sampling occasion, $q_k$ is the catch probability in $k$th sampling occasion, and $\mu$ is the mean catch probability of individuals in the target population. Because $q_k$ decrease with the sampling occasions from equation (44), the mean catchability declines between consecutive removals but the decrease rate depends on the variation of the catchability in the initial population.

*Harvest-based model*

In recent years, the increase of some wildlife (e.g., sika deer *Cervus nippon*; Iijima et al. 2013) or the expansion of exotic species (e.g., the small Indian mongoose *Herpestes auropunctatus*; Fukasawa et al. 2013) causes the necessity of population management. In the population management, it is necessary to estimate population abundance and the effect of harvest on population abundance. In such situation, the population management rarely finishes within a short period (i.e., closed population). Then, the model to estimate wildlife abundance of open population under harvesting pressure is necessary. Because the sampled individuals are not necessarily removed eternally from the target population and the target population should be closed in the removal sampling model, another model should be defined. Here, I define "the harvest-based model" as the model that estimates wildlife abundance of open population under harvesting pressure and has the ecological process model that explicitly describes the decrease of abundance by the number of harvested wildlife.

Especially in targeting open population, sampling effort tends to differ spatially and temporally. Then, constant catch probability cannot be satisfied. Furthermore, demographic rates should be estimated because the population size can change between any two sampling occasions by birth, natural mortality, or migration. The basic model structure of the harvest-based model is as below:

$$N_{k+1} \sim \text{Poisson}\left((N_k - X_k)r\right) \tag{49}$$

$$X_k \sim \text{Binomial}(N_k, p_k) \tag{50}$$

$r$ is the population growth rate and $p_k$ is the catch probability in $k$th sampling occasion. As shown in equations (49) and (50), we need to estimate $r$ and $p_k$.

The Weibull catch-effort model (Barron et al. 2011; Fukasawa et al. 2013) treats the effect of sampling effort and its change with the increase of sampling occasions on catch probability as below:

$$p_k = 1 - \exp(-aN_k E_k^{\beta}) \tag{51}$$

$p_k$ is the detection probability in $k$th sampling occasion, $a$ is the coefficient of catchability, $N_k$ is the abundance in $k$th sampling occasion, $E_k$ is the effort in $k$th sampling occasion, and $\beta$ is the shape parameter of Weibull distribution. The idea of the Weibull catch-effort model is that the detection probability increases with sampling effort but the degree of detection probability increase with sampling effort differs as the difference of parameter $\beta$.

To estimate demographic rates, it is useful to collect an abundance index in each sampling occasion. The abundance index is any data that is expected to correlate

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

with the true abundance. However, because the harvest-based model explicitly treats the number of harvested individuals to reduce abundance in latent ecological process and the number of harvested individuals is affected by hunting effort and hunting efficiency (Iijima 2017), the number of harvest individuals must not be used as abundance index in the harvest-based model although the number of harvested individuals are known to correlate with abundance (e.g., Ueno et al. 2014).

Let's imagine the case of the three-year survey in estimating the wildlife abundance and demographic rate. The number of harvested wildlife of each year (sampling occasion) is known as $\boldsymbol{X}_k = (X_1, X_2)$, and the wildlife abundance index in each sampling occasion is obtained as $\boldsymbol{I}_k = (I_1, I_2, I_3)$. As the underlying process of wildlife, it is assumed that the wildlife abundance decreases by hunting ($\boldsymbol{X}_k$) and increases by a constant annual population growth rate ($r$). If the true abundance of wildlife is assumed as $\boldsymbol{N}_k = (N_1, N_2, N_3)$ and the initial abundance index is assumed as $I_1 = 1$, population dynamics can be expressed as below:

$$N_2 = (N_1 - X_1)r \tag{52}$$

$$N_3 = (N_2 - X_2)r \tag{53}$$

The abundance index of the second and third sampling occasions can be written as below:

$$I_2 = N_2 / N_1 \tag{54}$$

$$I_3 = N_3 / N_1 \tag{55}$$

By assigning equation (54) to equation (52) and equation (55) to equation (53), $r$ can be obtained as below:

$$r = \frac{X_1 I_3 - X_2 I_2}{2(X_1 I_2 - X_2)} \tag{56}$$

Because we can obtain $X_1$, $X_2$, $I_2$, and $I_3$ as data, we can estimate $r$. With $r$, we can also estimate $\boldsymbol{N}_k$ by assigning equation (56) into equations (52) and (53). The parameter identifiability will greatly improve if the abundance index is the absolute abundance or density. Then, the combination of abundance estimation at certain period by models for the closed population with the harvest-based model is effective to increase parameter identifiability. Such a model can be regarded as an integrated population model (IPM).

The monitoring of covariates that affect the demographic rate is also useful in estimating demographic rates. Sæther et al. (2008) and Iijima et al. (2013) used environmental covariates like rainfall and the type of landscape to model temporal and spatial differences of the annual population growth rate of wildlife. Iijima and Ueno (2016) used the percentages of deciduous forests, evergreen forests, and artificial grasslands to model spatial variation of carrying capacity of sika deer. Use of covariates contributes not only to increase parameter identifiability but also to strengthen the biological meaning of results.

## The implication for future wildlife abundance estimation

As explained above, various models to estimate wildlife abundance have been developed. These models should be selected by considering the type of monitoring data and the assumptions of them (Table 1). Then, researchers need to select the appropriate model to estimate wildlife abundance. Because it is not necessary to assume the distribution of abundance in models for marked individuals, the models should be used if it is applicable. Models for unmarked wildlife can be selected depending on the purpose of the study and the applicable monitoring methods (Fig. 1). If the density of target wildlife is too low, the application of the Royle-Nichols model that can estimate wildlife abundance by the spatially and temporally repeated measure of presence/absence of target species (Royle and Nichols 2003) can be considered (please see the explanation of the N-mixture model).

In recent years, two conceptually important models to estimate wildlife abundance are developed: the hierarchical model and IPM. Regardless of target species or the type of monitoring data, the knowledge of the two models must be useful.

As mentioned above, the hierarchical model is the model with an ecological model that describes the ecological and latent process of target wildlife and an observation model that describes the process to observe the latent state of the target wildlife (Royle and Dorazio 2008). In ecology, observations generally have large variations caused by several factors such as observer, climate condition, and observation equipment. Furthermore, the systematic and ideal observations are sometimes impossible or practically difficult. Therefore, distinguishing the ecological process and observation process is very important to estimate ecological latent parameters like wildlife abundance.

By the distinction of ecological and process models, all

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

**Table 1.** Characteristics of models for wildlife abundance estimation in this paper

| Methods | Individual marking | Population closure | Effective sampling area | Population density | Suitable area for application* | Typical monitoring data |
|---|---|---|---|---|---|---|
| CR[a] model for closed population | Required | Required | Not estimable | Not estimable | Intermediate | Repeated wildlife capture |
| SECR[b] model | Required | Required | Estimable | Estimable | Small | Camera trapping |
| Jolly-Seber model | Required | Not required | Not estimable | Not estimable | Intermediate | Repeated wildlife capture |
| Distance sampling model | Not required | Required | Not estimable | Estimable | Intermediate | Line or point count |
| N-mixture model | Not required | Required | Not estimable | Not estimable | Intermediate | Line or point count |
| RE[c] model | Not required | Required | Not estimable | Estimable | Small | Camera trapping |
| REST[d] model | Not required | Required | Not estimable | Estimable | Small | Camera trapping |
| Removal sampling model | Not required | Required | Not estimable | Not estimable | Large | Repeated wildlife capture |
| Harvest-based model | Not required | Not required | Not estimable | Not estimable | Large | Hunting bag, capture effort and/or abundance index |

a: Capture-recapture.
b: Spatially explicit capture-recapture.
c: Random encounter.
d: Random encounter and staying time.
* It is only relative evaluation among these methods.

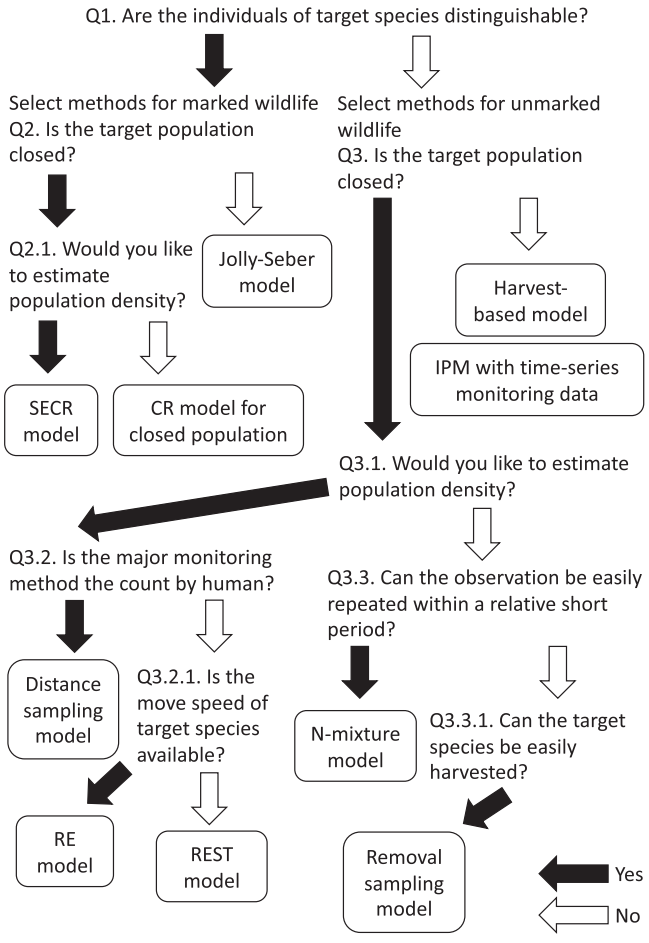



**Fig. 1.** Guide to select models for wildlife abundance estimation.

models for population abundance estimation that are shown above can be described with the same ecological model and different observation models depending on the sampling design and methods (Kéry and Royle 2015). Then, the hierarchical model affords researchers a clear understanding of the relationship between data and the ecological process of target wildlife. For these reasons, the hierarchical model will be the standard framework for population abundance estimation.

IPM is included in the hierarchical model and is the special class model to handle several data about different components of population dynamics (e.g., birth, natural mortality; Kéry and Schaub 2011). Actually, it often occurs that the different parameters about wildlife population dynamics are collected by different researchers or research groups especially in the field of science. IPM can incorporate these various data about population dynamics because IPM can have an ecological model that includes the expected process of population dynamics of target wildlife and observation models about the observation process of each data. Then, the formalization of the abundance estimation model as a hierarchical model and the construction of IPM by considering available data and the biological characteristics of target species are useful for future research.

Acknowledgments: I greatly appreciate Dr. Takuya Shimada for giving me the opportunity to write this

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

review. I especially thank Drs. Ken-ichiro Shimatani, Keiichi Fukaya, and Keita Fukasawa for the discussion with them about statistical modeling. Comments from two anonymous reviewers to the previous manuscript greatly improved the quality of this review. This review is related to my work that is prized as the promotion award for young scientists of the Mammal Society of Japan in 2017.

## References

Alho, J. M. 1990. Logistic regression in capture-recapture models. Biometrics 46: 623–635.

Augustine, B. C., Tredick, C. A. and Bonner, S. J. 2014. Accounting for behavioural response to capture when estimating population size from hair snare studies with missing data. Methods in Ecology and Evolution 5: 1154–1161.

Barron, M. C., Anderson, D. P., Parkes J. P. and Gon III, S. M. 2011. Evaluation of feral pig control in Hawaiian protected areas using Bayesian catch-effort models. New Zealand Journal of Ecology 35: 182–188.

Borchers, D. L. and Efford, M. G. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. Biometrics 64: 377–385.

Bord, S., Druilhet, P., Gasqui, P., Abrial, D. and Vourc'h, G. 2014. Bayesian estimation of abundance based on removal sampling under weak assumption of closed population with catchability depending on environmental conditions. Application to tick abundance. Ecological Modelling 274: 72–79.

Buckland, S. T., Anderson, D., Burnham, K., Laake, J., Borchers, D. L. and Thomas, L. 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, 432 pp.

Chandler, R. B. and Royle, J. A. 2013. Spatially explicit models for inference about density in unmarked or partially marked populations. The Annals of Applied Statistics 7: 936–954.

Dorazio, R. M., Jelks, H. L. and Jordan, F. 2005. Improving removal-based estimates of abundance by sampling a population of spatially distinct subpopulations. Biometrics 61: 1093–1101.

Efford, M. 2004. Density estimation in live-trapping studies. Oikos 106: 598–610.

Fukasawa, K., Hashimoto, T., Tatara, M. and Abe, S. 2013. Reconstruction and prediction of invasive mongoose population dynamics from history of introduction and management: a Bayesian state-space modelling approach. Journal of Applied Ecology 50: 469–478.

Gardner, B., Reppucci, J., Lucherini, M. and Royle, J. A. 2010. Spatially explicit inference for open populations: estimating demographic parameters from camera-trap studies. Ecology 91: 3376–3383.

Greenwood, R. J., Sargeant, A. B. and Johnson, D. H. 1985. Evaluation of mark-recapture for estimating striped skunk abundance. Journal of Wildlife Management 49: 332–340.

Hartstack Jr, A. W., Hollingsworth, J. P., Ridgway, R. L. and Hunt, H. H. 1971. Determination of trap spacings required to control an insect population. Journal of Economic Entomology 64: 1090–1100.

Honda, T. and Iijima, H. 2016. Managing boldness of wildlife: an ethological approach to reducing crop damage. Population Ecology 58: 385–393.

Huggins, R. M. 1989. On the statistical analysis of capture experiments. Biometrika 76: 133–140.

Iijima, H. 2017. The effects of landscape components, wildlife behavior and hunting methods on hunter effort and hunting efficiency of sika deer. Wildlife Biology 2017: wlb.00329. DOI: 10.2981/wlb.00329.

Iijima, H., Nagaike, T. and Honda, T. 2013. Estimation of deer population dynamics using a bayesian state-space model with multiple abundance indices. Journal of Wildlife Management 77: 1038–1047.

Iijima, H. and Ueno, M. 2016. Spatial heterogeneity in the carrying capacity of sika deer in Japan. Journal of Mammalogy 97: 734–743.

Jachmann, H. 2002. Comparison of aerial counts with ground counts for large African herbivores. Journal of Applied Ecology 39: 841–852.

Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52: 225–248.

Kéry, M. and Royle, J. A. 2015. Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS (volume 1 – prelude and static models). Academic Press, San Diego, 808 pp.

Kéry, M. and Schaub, M. 2011. Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective. Academic Press, San Diego, 463 pp.

Lincoln, F. C. 1930. Calculating waterfowl abundance on the basis of banding returns. U.S. Department Agriculture Circular 118: 1–4.

Link, W. A., Schofield, M. R., Barker, R. J. and Sauer, J. R. 2018. On the robustness of N-mixture models. Ecology 99: 1547–1551.

Mäntyniemi, S., Romakkaniemi, A. and Arjas, E. 2005. Bayesian removal estimation of a population size under unequal catchability. Canadian Journal of Fisheries and Aquatic Sciences 62: 291–300.

McClintock, B. T., White, G. C., Antolin, M. F. and Tripp, D. W. 2009. Estimating abundance using mark-resight when sampling is with replacement or the number of marked individuals is unknown. Biometrics 65: 237–246.

Nakashima, Y. 2020. Potentiality and limitations of N-mixture and Royle-Nichols models to estimate animal abundance based on noninstantaneous point surveys. Population Ecology 62: 151–157.

Nakashima, Y., Fukasawa, K. and Samejima, H. 2018. Estimating animal density without individual recognition using information derivable exclusively from camera traps. Journal of Applied Ecology 55: 735–744.

Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62: 3–135.

Petersen, C. G. J. 1896. The yearly immigration of young plaice in the Limfjord from the German sea. Danish Biological Station 6: 1–48.

Pettorelli, N., Côté, S. D., Gingras, A., Potvin, F. and Huot, J. 2007. Aerial surveys vs hunting statistics to monitor deer density: the example of Anticosti Island, Quebec, Canada. Wildlife Biology 13: 321–328. DOI: 10.2981/0909-6396(2007)13[321:ASVHST]2.0.CO;2.

Pollock, K. H., Nichols, J. D., Simons, T. R., Farnsworth, G. L., Bailey, L. L. and Sauer, J. R. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics 13: 105–119.

Rowcliffe, J. M., Field, J., Turvey, S. T. and Carbone, C. 2008. Estimating animal density using camera traps without the need for individual recognition. Journal of Applied Ecology 45: 1228–1236.

Royle, J. A. 2004. N-mixture models for estimating population size

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

from spatially replicated counts. Biometrics 60: 108–115.
Royle, J. A. and Dorazio, R. M. 2008. Hierarchical Modeling and Inference in Ecology. Academic Press, San Diego, 464 pp.
Royle, J. A., Dorazio, R. M. and Link, W. A. 2007. Analysis of multinomial models with unknown index using data augmentation. Journal of Computational and Graphical Statistics 16: 67–85.
Royle, J. A. and Nichols, J. D. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84: 777–790.
Royle, J. A. and Young, K. V. 2008. A hierarchical model for spatial capture-recapture data. Ecology 89: 2281–2289.
Sæther, B.-E., Lillegård, M., Grøtan, V., Drever, M. C., Engen, S., Nudds, T. D. and Podruzny, K. M. 2008. Geographical gradients in the population dynamics of North American prairie ducks. Journal of Animal Ecology 77: 869–882.
Seber, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52: 249–260.
Tuyttens, F. A. M., Macdonald, D. W., Delahay, R., Rogers, L. M., Mallinson, P. J., Donnelly, C. A. and Newman, C. 1999. Differences in trappability of European badgers *Meles meles* in three populations in England. Journal of Applied Ecology 36: 1051–1062.
Ueno, M., Solberg, E. J., Iijima, H., Rolandsen, C. M. and Gangsei, L. E. 2014. Performance of hunting statistics as spatiotemporal density indices of moose (*Alces alces*) in Norway. Ecosphere 5: art13. DOI: 10.1890/ES13-00083.1.
Williams, B. K., Nichols, J. D. and Conroy, M. J. 2002. Analysis and Management of Animal Populations. Academic Press, San Diego, 817 pp.
Zwolak, R. and Foresman, K. R. 2008. Deer mouse demography in burned and unburned forest: no evidence for source–sink dynamics. Canadian Journal of Zoology 86: 83–91.

*Received 18 December 2019. Accepted 31 January 2020.*
*Published online 17 June 2020.*
*Editor was Takuya Shimada.*

Downloaded From: https://bioone.org/journals/Mammal-Study on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use