

ORIGINAL RESEARCH

Covariate and multinomial: Accounting for distance in movement in capture–recapture analyses

Lorelei Guéry¹  | Lauriane Rouan² | Sébastien Descamps³ | Joël Bêty¹ |
Albert Fernández-Chacón⁴ | Grant Gilchrist⁵ | Roger Pradel⁶

¹Département de Biologie, Chimie et Géographie, Université du Québec à Rimouski, Rimouski, Canada – Centre d'Etudes Nordiques, Laval, Québec, Canada

²CIRAD – Biological Systems Department, UMR108 Amélioration Génétique et Adaptation des Plantes méditerranéennes et tropicales, Montpellier, France

³Norwegian Polar Institute, Fram Center, Tromsø, Norway

⁴Research Unit of Biodiversity (UO, CSIC, PA), University of Oviedo, Mieres, Spain

⁵National Wildlife Research Centre, Environment Canada, Ottawa, Canada

⁶CEFE, UMR 5175, CNRS - Université de Montpellier - Université Paul-Valéry Montpellier - EPHE, Montpellier, France

Correspondence

Lorelei Guéry, IRD-UMR MARBEC, CRH, CS 30171, Av. Jean Monnet, 34203 Sète Cedex, France.
Email: lorelei.guery@gmail.com

Abstract

Many biological quantities cannot be measured directly but rather need to be estimated from models. Estimates from models are statistical objects with variance and, when derived simultaneously, covariance. It is well known that their variance–covariance (VC) matrix must be considered in subsequent analyses. Although it is always preferable to carry out the proposed analyses on the raw data themselves, a two-step approach cannot always be avoided. This situation arises when the parameters of a multinomial must be regressed against a covariate. The Delta method is an appropriate and frequently recommended way of deriving variance approximations of transformed and correlated variables. Implementing the Delta method is not trivial, and there is a lack of a detailed information on the procedure in the literature for complex situations such as those involved in constraining the parameters of a multinomial distribution. This paper proposes a how-to guide for calculating the correct VC matrices of dependant estimates involved in multinomial distributions and how to use them for testing the effects of covariates in post hoc analyses when the integration of these analyses directly into a model is not possible. For illustrative purpose, we focus on variables calculated in capture–recapture models, but the same procedure can be applied to all analyses dealing with correlated estimates with multinomial distribution and their variances and covariances.

KEYWORDS

covariate, dependent estimates, link function, multinomial logit, transformations, variance–covariance matrix

1 | INTRODUCTION

Biologists and ecologists routinely study quantities that are not directly measurable (probabilities of survival, capture, etc....), especially in wild populations, and thus need to use models to calculate estimates of these quantities. In addition, ecologists are often interested in assessing the effect of covariates on these quantities, such as the comparison of survival before and after the implementation

of conservation and management measures. Typically these model-produced estimates of say survival probability are calculated from the same model and hence not independent, that is, coming with associated variances and covariances. Thus, using the model-produced estimates in traditional statistical tools, such as for example *t* test (Gossett, 1908), ANOVA (Anderson & Ager, 1978), correlation coefficient, or linear models, highly violates the rule of independence of these tools. The latter therefore becomes inefficient to deal with

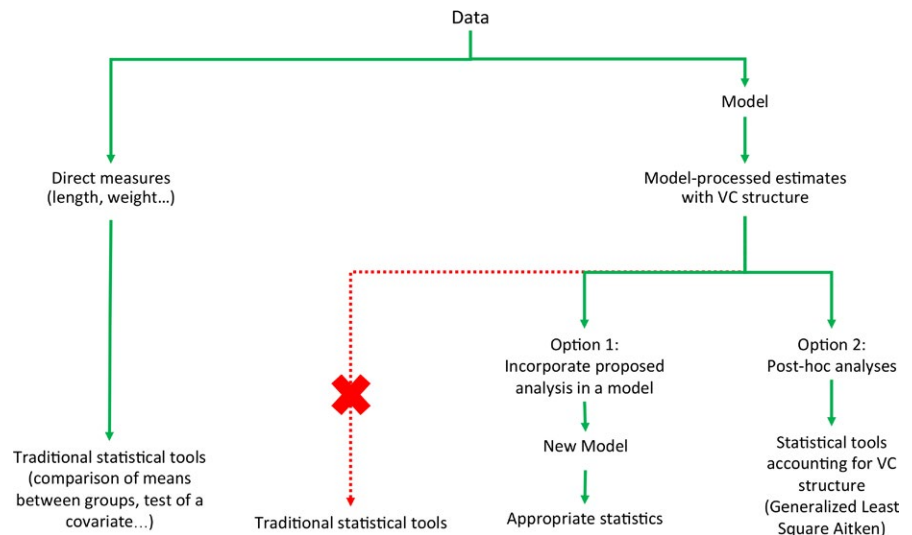


FIGURE 1 Schematic of statistical approach

these dependent estimates because the variance–covariance (VC) structure is ignored.

The ideal option is to build the proposed analysis in the model (option 1, Figure 1). However, it can sometimes become complicated or even not possible to test hypotheses of interest using the currently available models. Biologists are therefore required to run alternative post hoc analyses on the model-produced estimates. Some papers previously discussed this issue of doing statistics on statistics in post hoc analysis either in a frequentist framework (Burnham & White 2002; Grosbois et al., 2008; see Dugger et al., 2015 for an example of the use of these models) or in the Bayesian framework (Brooks & Deroba, 2015; Link 1999; Link & Barker 2004) in which a hierarchical approach was recommended (Cooch, Conn, Ellner, Dobson, & Pollock, 2012; Royle & Dorazio 2008; Sutherland, Brambilla, Pedrini, & Tenan, 2016).

A typical example is the study of movement as a function of the distance. In ecology, dispersal of plant seeds or movement of animals are generally more frequent between neighboring sites. To actually test this idea, one may want to relate the frequency of movements between any two locations to the distance separating them. For animals, frequency of exchanges can be measured from capture–recapture (CR) data gathered on a small number of sites. An appropriate multisite model (Cormack, 1964; Jolly, 1965; Nichols & Kendall, 1995; Pradel, Gimenez, & Lebreton, 2005; Seber, 1965), on which we will focus here for illustrative purpose, produces estimates of the probability that an animal moves or stays from a given site of departure toward a site of arrival. The natural idea is then to regress the probability of movement against the distance between the departure and arrival sites directly in the model (option 1, Figure 1). However, because only the surveyed sites are observable as arrival sites, the multisite models can only estimate conditional probabilities of movement (conditional on the fact that the animals do not move out of the set of surveyed sites). As the different possible transition probabilities (including the probability of not departing) must sum to 1, one of those probabilities chosen arbitrarily, but usually the probability of not departing, is the complement of the others. In this

case, it is not an actual parameter of the model and cannot be forced into the relationship with the distance. The same problem will arise every time that the quantities to be constrained are parameters with a multinomial distribution, that is, variables with more than two modalities such as the breeding status based on the number of offspring produced or the settlement probabilities in different areas.

To circumvent this difficulty, we propose a second option (option 2, Figure 1), that is, applying a generalized least square (GLS) approach to the estimates produced by the unconstrained multisite model. A proper estimation of the VC matrix is therefore needed as a key element for accounting for statistical uncertainty or correlation in parameters estimates that are included in these post hoc analyses. Some scientific articles commonly apply the Delta method (Dorfman, 1938; Ver Hoef, 2012), a method for deriving variance approximations of transformed and correlated variables. The Delta method theory was fully explained and its use recommended to biologists (Cooch & White, 2016; MacKenzie, 2006; Williams, Nichols, & Conroy, 2002). Powell (2007) described the Delta method in a more accessible way to ecologists, but it was less detailed in describing correlated estimates and providing applications with covariates. The use of the Delta method is rarely made the focal point in scientific papers, much less with multinomial variables. This lack of detailed procedures in the literature could lead to incorrect use where dependent estimates are considered as independent ones (Pape Møller & Szép, 2002). Indeed, Fernández-Chacón et al. (2013) was, to our knowledge, the only study that used and described a correct use of the variance–covariance (VC) matrix in post hoc analyses to study impacts of a covariate on a parameter with a multinomial distribution.

The aim of the present work is to propose a how-to guide for calculating the correct VC matrices of dependant estimates with multinomial distribution and how to use them when covariates are tested in post hoc analyses. We first briefly describe the general statistical requirements on link functions and the Delta method and refer to corresponding theoretical papers. Second, we focus on the specificities of CR analyses and how to assess the effect of covariates on parameters involved in the multinomial distribution (e.g.,

TABLE 1 Rules for calculating derivatives of several link functions, and their inverse, used in CMR analyses

Link	Function	Derivative of function	Inverse function (back-transformations)	Derivative of inverse function
Logit	$\beta = \ln\left(\frac{\theta}{1-\theta}\right) \text{ logit}(\beta)$	$\frac{1}{(1-\theta)\theta}$	$\theta = \frac{e^\beta}{1+e^\beta} = \frac{1}{1+e^{-\beta}}$ inverse logit (β) logistic (β)	$\frac{e^{-\beta}}{(1+e^{-\beta})^2} = \left(1 - \frac{1}{1+e^{-\beta}}\right) * \frac{1}{1+e^{-\beta}} = (1-\theta)\theta$
Multinomial logit (Example with β_{A1} or θ_{AB} of the prerequisite section)	$\beta_{A1} = \ln\left(\frac{\theta_{AB}}{1-\theta_{AB}-\theta_{AC}-\theta_{AE}}\right)$	$\frac{\partial \beta_{A1}}{\partial \theta_{AB}} = \frac{1-\theta_{AC}-\theta_{AD}-\theta_{AE}}{(1-\theta_{AB}-\theta_{AC}-\theta_{AD}-\theta_{AE})^2}$ $\frac{\partial \beta_{A1}}{\partial \theta_{AC}} = \frac{\partial \beta_{A1}}{\partial \theta_{AD}} = \frac{\partial \beta_{A1}}{\partial \theta_{AE}} = -\frac{\partial \beta_{A1}}{\partial \theta_{AB}}$	$\theta_{AB} = \frac{e^{\beta_{A1}}}{1+e^{\beta_{A1}}+e^{\beta_{A2}}+e^{\beta_{A3}}+e^{\beta_{A4}}}$	$\frac{\partial \theta_{AB}}{\partial \beta_{A1}} = \left(1 - \theta_{AB}\right) \theta_{AB}$ $\frac{\partial \theta_{AB}}{\partial \beta_{A2}} = -\theta_{AB} \theta_{AC}$ $\frac{\partial \theta_{AB}}{\partial \beta_{A3}} = -\theta_{AB} \theta_{AC}$ $\frac{\partial \theta_{AB}}{\partial \beta_{A4}} = -\theta_{AB} \theta_{AC}$
Arcsin	$\beta = \arcsin(2\theta - 1)$	$\frac{2}{\sqrt{1-(2\theta-1)^2}}$	$\theta = \frac{\sin(\beta)+1}{2}$	$\frac{\cos(\beta)}{2}$
Log: Natural Logarithm	$\beta = \ln(\theta)$	$1/\theta$	$\theta = e^\beta$	e^β
Log-log	$\beta = \ln(-\ln(\theta))$	$\frac{1}{\theta \ln(\theta)}$	$\theta = e^{-e^\beta}$	$-e^\beta - e^\beta$
Complementary log-log	$\beta = \ln(-\ln(1-\theta))$	$\frac{1}{\ln(1-\theta)(\theta-1)}$	$\theta = 1 - e^{-e^\beta}$	$e^\beta - e^\beta$
Identity	$\beta = \theta$	1	$\theta = \beta$	1

Notes. θ refers to any biological parameter (e.g., survival probability) and β represents the transformed θ , that is, a mathematical parameter. Some of these derivatives are needed in the approximation of the variance-covariance matrix using the Delta method. Logit, arcsin, log-log, and complementary log-log link functions constrain parameters to the interval [0, 1], whereas identity and log link functions do not.

multinomial logit link function), a case in which covariates could not be integrated directly in a new model. We provide R code (Team, 2015) in the (Appendix S1) Supporting information to reproduce these post hoc analyses and serve as practical guide for other potential applications.

2 | HOW DOES THE DELTA METHOD WORK IN SIMPLE CASES?

Such as other biological quantities, demographic probabilities (e.g., survival or detection probabilities) can, by definition, vary between 0 and 1. Commonly in statistical analyses, these quantities are transformed with link functions, which provides a mapping of (0,1) space to the real number axis, so between $-\infty$ and $+\infty$, which is useful for most basic statistics (e.g., linear model). In CR analyses, several link functions can be used (Table 1), but the most commonly used is the logit function, such as:

$$\beta = \text{logit}(\theta) = \ln\left(\frac{\theta}{1-\theta}\right),$$

where θ can be estimates of any demographic parameters and β s represent the transformed θ (hereafter named mathematical parameters). To get back the demographic parameters θ , that is, on the [0,1] probability scale, the back-transformation, also called the inverse logit function (Table 1), is:

$$\theta = \text{inverse logit}(\beta) = \frac{e^\beta}{1+e^\beta} = \frac{1}{1+e^{-\beta}} = \text{logistic}(\beta),$$

Then, to assess the impact of a covariate x on demographic parameters, CR models are constructed to constrain estimates of these parameters to be a function of x . In the case of a linear relation, the logit link function, as well as the other link functions of Table 1, can thus be written as:

$$\text{logit}(\theta) = \ln\left(\frac{\theta}{1-\theta}\right) = \beta_i + \beta_s * x,$$

where β_i is the intercept and β_s the slope of this linear relation. Similarly, the back-transformation is:

$$\theta = \text{inverse logit}(\beta_i + \beta_s * x) = \frac{e^{(\beta_i + \beta_s * x)}}{1+e^{(\beta_i + \beta_s * x)}} = \frac{1}{1+e^{-(\beta_i + \beta_s * x)}} = \text{logistic}(\beta_i + \beta_s * x),$$

Several of these biological quantities cannot be measured simply but rather need to be estimated from models. Estimates of these quantities θ derived from models, as well as the mathematical parameters (β), come with variance and covariance. This non-independence is expressed through a VC matrix ($V(\theta)$ and $V(\beta)$ respectively), which needs to be taken into account to perform proper analyses. However, as stated above, the θ are very often transformed to be used in statistical analyses, for example, with link functions in the case of transformations of probability. These scale modifications, when using a link function, influence the shape of the likelihood and

thus the estimate of the variance (Cooch & White, 2016). It is not as simple as back transforming the $V(\beta)$ to obtain $V(\theta)$.

Yet, a convenient and straightforward method for obtaining variance and covariance of one or more transformed demographic parameters is the Delta method that uses the one-dimensional Taylor series approximation (Cooch & White, 2016; Williams et al., 2002). We can approximate $V(\theta)$, the VC matrix of any demographic parameters, by:

$$V(\theta) = D_{\theta} * V(\beta) * D_{\theta}^T,$$

where β represent the mathematical parameters, that is, the transformed θ after applying the link function, $V(\beta)$ is the VC matrix of the β s, D_{θ} is the matrix of the first derivatives of the demographic parameters (θ) and D_{θ}^T is the transpose of D_{θ} .

The logit link function (such as others, see Table 1) is well adapted to variables with binomial distributions, thus convenient in CR analyses where several phenomena are essentially binary (e.g., survived or died, breeder versus non-breeder, captured or not). However, other phenomena, namely transitions in multistate models, are essentially multinomial.

3 | HOW TO ASSESS THE EFFECT OF COVARIATES ON PARAMETERS WITH MULTINOMIAL DISTRIBUTION

In more complex CR models, such as multistate (Nichols & Kendall, 1995) or multi-event (Pradel, 2005) models, the probabilities of interest can have more than two modalities. As an example, we could study the impact of the distance separating any two locations on the frequency of movements between them. In a situation where there are 4 sites, starting from any one site, we have four potential movements including the probability of not moving (Figure 2). Since the associated probabilities must sum to 1, the logit link function is no longer satisfactory and has to be replaced by the multinomial (or generalized) logit (McFadden, 1968) link function. The implementation of this constraint, that the sum of the probabilities of movement from a given departure site is equal to 1, forces in effect a parameterization, where there is no direct correspondence between an actual parameter (β) and a particular movement (θ).

$$\theta_{AB} = \frac{e^{\beta_{A1}}}{1 + e^{\beta_{A1}} + e^{\beta_{A2}} + e^{\beta_{A3}}}; \theta_{AC} = \frac{e^{\beta_{A2}}}{1 + e^{\beta_{A1}} + e^{\beta_{A2}} + e^{\beta_{A3}}};$$

$$\theta_{AD} = \frac{e^{\beta_{A3}}}{1 + e^{\beta_{A1}} + e^{\beta_{A2}} + e^{\beta_{A3}}}; \theta_{AA} = 1 - \theta_{AB} - \theta_{AC} - \theta_{AD},$$

Although there are as many mathematical parameters (3β) as independent demographic parameters (3θ), there is no one-to-one correspondence as can be seen from the equations above (any individual θ is a function of all the components of β through the denominator).

An alternative parameterization separates the probability to disperse from the probabilities of settlement conditional on dispersal

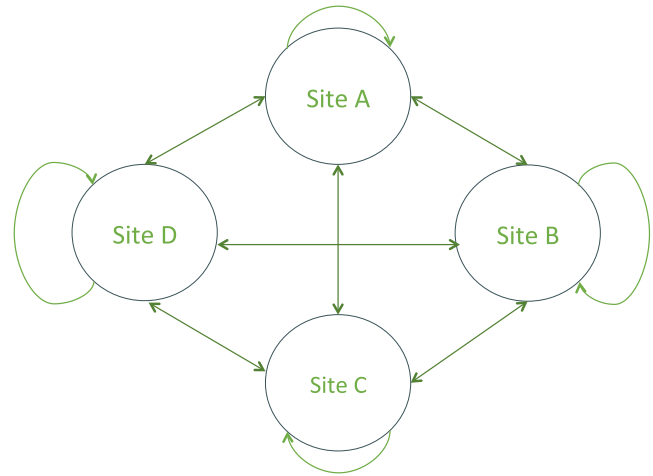


FIGURE 2 Schematic of the possible transitions between different breeding locations

(Grosbois & Tavecchia, 2003). If we take the probability of moving to site D as the complement, then we have:

$$\beta'_{A1} = \ln \left(\frac{\theta'_{AB}}{1 - \theta'_{AB} - \theta'_{AC}} \right) \text{ and } \beta'_{A2} = \ln \left(\frac{\theta'_{AC}}{1 - \theta'_{AB} - \theta'_{AC}} \right),$$

where θ'_{AB} is the transition from A to B, that is, the probability that an individual leaving site A settles on site B the subsequent year. Inverting the multinomial logit function, we obtain the probabilities of settlement as:

$$\theta'_{AB} = \frac{e^{\beta'_{A1}}}{1 + e^{\beta'_{A1}} + e^{\beta'_{A2}}}; \theta'_{AC} = \frac{e^{\beta'_{A2}}}{1 + e^{\beta'_{A1}} + e^{\beta'_{A2}}}; \theta'_{AD} = 1 - \theta'_{AB} - \theta'_{AC};$$

There is still no one-to-one correspondence between the individual β' and the individual θ' .

This absence of a one-to-one correspondence between demographic and mathematical parameters renders obscure, if at all possible, the way to implement a desired constraint on demographic parameters. Therefore, we propose to extract estimates of the demographic parameters from the *unconstrained* model along with their VC matrix, and then run post hoc analyses. As stated above, because the estimates will not be independent, traditional statistical techniques are not appropriate. We propose to use instead the generalized least square (GLS; Aitken 1936) framework. However, for this method to perform properly we need to ensure that all values in the set of real numbers are admissible. This is ensured if we work on the logit scale. The first step is thus to calculate the simple logit of θ' (or θ) above. In this transformation, each element of θ' is mapped one-to-one onto an element of a new vector (thereafter called γ), which corresponds to a transition and only one and which takes its value between $-\infty$ and $+\infty$. Then, we calculate the associated VC matrix ($V(\gamma)$), for example, with the Delta method, which will be used in the GLS analyses. Indeed, the mathematical parameters (β') cannot be used because they do not correspond to any particular transition, and the transitions (θ') themselves have the disadvantage to

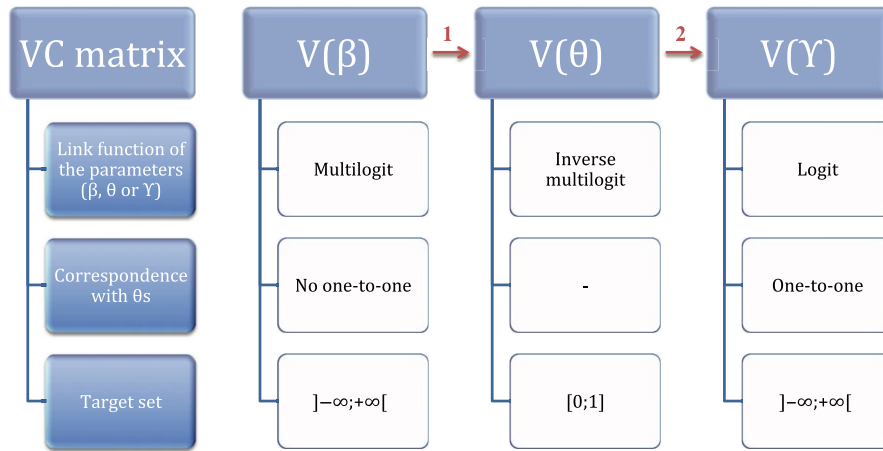


FIGURE 3 General approach for assessing the impact of a covariate on parameters with a multinomial distribution. “VC” means variance–covariance, “ β ” are the mathematical parameters, “ θ ” the estimates of biological quantities and “ γ ” the one-to-one individual transformed estimates. Red arrows indicate the steps (1 and 2) where the Delta method is applied

	β_{A1}	β_{A2}	β_{B1}	β_{B2}	β_{C1}	β_{C2}
θ_{AB}	$(1 - \theta_{AB})\theta_{AB}$	$-\theta_{AB}\theta_{AC}$	0	0	0	0
θ_{AC}	$-\theta_{AC}\theta_{AB}$	$(1 - \theta_{AC})\theta_{AC}$	0	0	0	0
θ_{BA}	0	0	$(1 - \theta_{BA})\theta_{BA}$	$-\theta_{BA}\theta_{BC}$	0	0
θ_{BC}	0	0	$-\theta_{BC}\theta_{BA}$	$(1 - \theta_{BC})\theta_{BC}$	0	0
θ_{CA}	0	0	0	0	$(1 - \theta_{CA})\theta_{CA}$	$-\theta_{CA}\theta_{CB}$
θ_{CB}	0	0	0	0	$-\theta_{CB}\theta_{CA}$	$(1 - \theta_{CB})\theta_{CB}$

For example, θ_{AB} is the transition from A to B, that is, the probability that an individual leaving site A settles on site B the subsequent year.

TABLE 2 Matrix (D_{θ_x}) of the first derivatives of the biological parameters (θ s) (in rows) with respect to the mathematical parameters (β s) (in columns)

vary between 0 and 1, on which regression cannot be made directly (Figure 3).

Program MARK (White & Burnham, 1999) provides the VC matrix for the demographic parameters ($V(\theta')$). We only need to derive the VC matrix $V(\gamma)$ of the logits of θ' , prior to the GLS analysis (Step 2, Figure 3). If instead program E-SURGE (Choquet, Rouan, & Pradel, 2009) is used, only the VC matrix of the mathematical parameters ($V(\beta')$) is directly available, so that there is an additional step in this case (Step 1, Figure 3). We describe the steps required when only the VC matrix of the mathematical parameters is provided as part of step 1 in the analyses.

Step 1: As per the Delta method, we start from $V(\theta') = D_{\theta'}^T * V(\beta') * D_{\theta'}$. Using the same example as above: $V(\beta') = V\left(\begin{matrix} \beta'_{A1} \\ \beta'_{A2} \end{matrix}\right)$ and $V(\theta') = V\left(\begin{matrix} \theta'_{AB} \\ \theta'_{AC} \end{matrix}\right)$. In order to calculate $V(\theta')$, we first need to calculate the matrix of the first derivatives of θ' with respect to β' ($D_{\theta'} = \frac{\partial \theta'}{\partial \beta'}$). Depending on which β' we derived with respect to, the first derivative of θ' will differ. After some algebra based on Table 1, we can write the partial first derivative of θ'_{AB} with respect to each β' such as:

$$\frac{\partial \theta'_{AB}}{\partial \beta'_{A1}} = (1 - \theta'_{AB}) * \theta'_{AB} = -\theta'_{AB} * \theta'_{AC}$$

and any other $\frac{\partial \theta'_{AB}}{\partial \beta'_{ij}} = 0$ with $i \neq A$ and j from 1 to 2. This was applied in all cases to get $D_{\theta'}$ (Table 2), then to take the transpose, $D_{\theta'}^T$, and to consequently calculate $V(\theta')$.

Step 2: In our example, $V(\gamma) = V\left(\begin{matrix} \gamma_{AB} \\ \gamma_{AC} \\ \gamma_{AD} \end{matrix}\right)$. To calculate γ and $V(\gamma)$, an additional transformation on the logit scale of θ' and their corresponding VC matrix had to be calculated. Thus, $\gamma = \text{logit}(\theta') = \ln\left(\frac{\theta'}{1-\theta'}\right)$. Then, using $V(\theta')$ either previously calculated or directly provided by the program MARK, the Delta method gives for example:

$$\text{cov}(\gamma_{AB}, \gamma_{AC}) = \text{cov}\left(\text{logit}(\theta'_{AB}), \text{logit}(\theta'_{AC})\right) = \frac{\text{cov}(\theta'_{AB}, \theta'_{AC})}{\theta'_{AB}(1-\theta'_{AB})\theta'_{AC}(1-\theta'_{AC})}$$

$$\text{and var}(\gamma_{AB}) = \text{var}\left(\text{logit}(\theta'_{AB})\right) = \frac{\text{var}(\theta'_{AB})}{(\theta'_{AB})^2(1-\theta'_{AB})^2}$$

After applying these formula to all cases, we get $V(\gamma)$, presented in Table 3. The one-to-one individual transformed estimates (γ) and their VC matrix ($V(\gamma)$) can now be used to perform a GLS linear regression between the movement probability estimates and a covariate (e.g., distance between sites A, B, C, and D). This procedure

TABLE 3 Variance–covariance matrix $U(\gamma_X)$ of the individual mathematical parameters (γ_X)

	$\gamma_{Y1 \rightarrow Y2}$...	$\gamma_{Y2 \rightarrow Y3}$...	$\gamma_{Y4 \rightarrow Y5}$
$\gamma_{Y1 \rightarrow Y2}$	$\frac{\text{var}(\theta_{Y1 \rightarrow Y2})}{(\theta_{Y1 \rightarrow Y2})^2 (1 - \theta_{Y1 \rightarrow Y2})^2}$...	$\frac{\text{cov}(\hat{\theta}_{Y1 \rightarrow Y2}, \hat{\theta}_{Y2 \rightarrow Y3})}{\hat{\theta}_{Y1 \rightarrow Y2} (1 - \hat{\theta}_{Y1 \rightarrow Y2}) \hat{\theta}_{Y2 \rightarrow Y3} (1 - \hat{\theta}_{Y2 \rightarrow Y3})}$...	$\frac{\text{cov}(\theta_{Y1 \rightarrow Y2}, \theta_{Y4 \rightarrow Y5})}{\theta_{Y1 \rightarrow Y2} (1 - \theta_{Y1 \rightarrow Y2}) \theta_{Y4 \rightarrow Y5} (1 - \theta_{Y4 \rightarrow Y5})}$
...
$\gamma_{Y2 \rightarrow Y3}$	$\frac{\text{cov}(\theta_{Y2 \rightarrow Y3}, \theta_{Y1 \rightarrow Y2})}{\theta_{Y2 \rightarrow Y3} (1 - \theta_{Y2 \rightarrow Y3}) \theta_{Y1 \rightarrow Y2} (1 - \theta_{Y1 \rightarrow Y2})}$...	$\frac{\text{var}(\theta_{Y2 \rightarrow Y3})}{(\theta_{Y2 \rightarrow Y3})^2 (1 - \theta_{Y2 \rightarrow Y3})^2}$...	$\frac{\text{cov}(\theta_{Y2 \rightarrow Y3}, \theta_{Y4 \rightarrow Y5})}{\theta_{Y2 \rightarrow Y3} (1 - \theta_{Y2 \rightarrow Y3}) \theta_{Y4 \rightarrow Y5} (1 - \theta_{Y4 \rightarrow Y5})}$
...
$\gamma_{Y4 \rightarrow Y5}$	$\frac{\text{cov}(\theta_{Y4 \rightarrow Y5}, \theta_{Y1 \rightarrow Y2})}{\theta_{Y4 \rightarrow Y5} (1 - \theta_{Y4 \rightarrow Y5}) \theta_{Y1 \rightarrow Y2} (1 - \theta_{Y1 \rightarrow Y2})}$...	$\frac{\text{cov}(\theta_{Y4 \rightarrow Y5}, \theta_{Y2 \rightarrow Y3})}{\theta_{Y4 \rightarrow Y5} (1 - \theta_{Y4 \rightarrow Y5}) \theta_{Y2 \rightarrow Y3} (1 - \theta_{Y2 \rightarrow Y3})}$...	$\frac{\text{var}(\theta_{Y4 \rightarrow Y5})}{(\theta_{Y4 \rightarrow Y5})^2 (1 - \theta_{Y4 \rightarrow Y5})^2}$

is implemented in MATLAB with the function `lscov` and in R with the function `lm.gls` of the library `MASS`. The R script that allows doing these calculations is detailed in the supporting information (Appendix S1).

By way of illustration, we reproduce here the example of the role of distance on conditional settlement probabilities (Fernández-Chacón et al., 2013), that is, decision on where to go once animals have left their previous site. A long-lived seabird was monitored in four sites. In this study, settlement probabilities were estimated with multi-event mark–recapture models in E-SURGE, taking into account as the complement the settlement on a catchall fifth site (ghost location), which serves for all individuals that go to unmonitored sites or skip breeding altogether. Following the approach presented above with five sites instead of four, we started from the β and $V(\beta)$ to get the θ and $V(\theta)$ (Step 1), then to obtain the γ and $V(\gamma)$ (Step 2), to finally test, with a GLS linear regression, the link between distance separating each of the sites and settlement probabilities. In this numerical example, distance to the destination site did not influence settlement choices (Fernández-Chacón et al., 2013). R routines, numerical steps, variance–covariance matrices, and results are provided in supporting information (Appendix S1).

4 | CONCLUSION

Although integrating the covariate directly into a given model is the preferred method, this is not always possible. The two-step approach presented in this paper works well in the context of CR analyses and has the additional advantage of allowing manipulating the biological quantities themselves rather than the compounded parameters of multinomial logits. Additionally, our method is impervious to the difficulties that may arise in a direct analysis when the same covariate is hypothesized to act on quantities derived from more than one multinomial. For example, when testing the effect of distance on movement probability, the probability to move from site A to site B should be equal to that to move from B to A if distance is the sole determinant. Although the two quantities belong to two different multinomials, this is not an obstacle in our approach.

One may wonder how good is this alternative of post hoc analysis. Preliminary analyses without multinomial distribution (not presented) showed that integrating the covariate into the models or running a posteriori analyses accounting for correlated estimates

provided the same biological conclusions. For example, assessing the effect of the winter NAO on annual survival either comparing models with the integrated covariate (Guéry et al., 2017) or with the GLS technique performing a linear regression between the winter NAO and annual survival led to similar conclusions. However, the interpretation of the details needs to be careful with the extreme case of parameters estimated at the boundaries (0 or 1), where the logit transformation is not linear anymore for these limits ($\text{logit}(0) = -\infty$ and $\text{logit}(1) = +\infty$). Indeed, the Delta method assumes that the transformation of variables is approximately linear over the expected range of the parameter; otherwise, it could fail to correctly approximate the variance (Cooch & White, 2016; Williams et al., 2002).

Our study mostly uses examples from a CR framework but what we propose goes beyond CR models and can be used in any framework as long as one deals with parameter estimates that come from the same model. We hope that this paper will serve as reference and guideline in further needed investigations, not only in population biology but also in all analyses dealing with correlated estimates and their variances and covariances.

ACKNOWLEDGMENTS

The study was supported by the Canadian Wildlife Service–Environment Canada, Nunavut Wildlife Management Board, Greenland Institute of Natural Resources, Polar Continental Shelf Project, Fonds Québécois de la Recherche sur la Nature et les Technologies, Canadian Network of Centres of Excellence ArcticNet, Natural Sciences and Engineering Research Council of Canada, and the Department of Indian Affairs and Northern Canada. We would like to thank Beth Gardner for her work on these EURING proceedings and three anonymous reviewers.

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTION

Lorelei Guéry, Lauriane Rouan and Roger Pradel set up the idea. Lorelei Guéry, Lauriane Rouan, Roger Pradel, Joël Bêty and Sébastien Descamps designed methodology. Lorelei Guéry conducted analyses and wrote the paper. Albert Fernández-Chacón and

Grant Gilchrist provided the data. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Data used are provided in the supporting information (Appendix S1).

ORCID

Lorelei Guéry  <https://orcid.org/0000-0003-3004-8462>

REFERENCES

- Aitken, A. C. (1936). IV.—On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, 55, 42–48. <https://doi.org/10.1017/S0370164600014346>
- Anderson, L. R., & Ager, J. W. (1978). Analysis of variance in small group research. *Personality and Social Psychology Bulletin*, 4, 341–345. <https://doi.org/10.1177/014616727800400237>
- Brooks, E. N., & Derocha, J. J. (2015). When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. *Canadian Journal of Fisheries and Aquatic Sciences*, 72, 634–641.
- Burnham, K. P., & White, G. C. (2002). Evaluation of some random effects methodology applicable to bird ringing data. *Journal of Applied Statistics*, 29, 245–264.
- Choquet, R., Rouan, L., & Pradel, R. (2009). Program E-SURGE: A software application for fitting multievent models. In *Modeling demographic processes in marked populations* (pp. 845–865). Boston, MA: Springer.
- Cooch, E. G., Conn, P. B., Ellner, S. P., Dobson, A. P., & Pollock, K. H. (2012). Disease dynamics in wild populations: Modeling and estimation: A review. *Journal of Ornithology*, 152, 485–509.
- Cooch, E., & White, G. (2016). *Program MARK, 16th edn, A gentle introduction*. Fort Collins, CO: Colorado State University.
- Cormack, R. (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, 51, 429–438. <https://doi.org/10.1093/biomet/51.3-4.429>
- Dorfman, R. (1938). A note on the delta-method for finding variance formulae. *The Biometric Bulletin*, 1, 92.
- Dugger, K. M., Forsman, E. D., Franklin, A. B., Davis, R. J., White, G. C., Schwarz, C. J., ... Doherty Jr, P. F. (2015). The effects of habitat, climate, and Barred Owls on long-term demography of Northern Spotted Owls. *The Condor*, 118, 57–116.
- Fernández-Chacón, A., Genovart, M., Pradel, R., Tavecchia, G., Bertolero, A., Piccardo, J., ... Oro, D. (2013). When to stay, when to disperse and where to go: Survival and dispersal patterns in a spatially structured seabird population. *Ecography*, 36, 1117–1126. <https://doi.org/10.1111/j.1600-0587.2013.00246.x>
- Gossett, W. S. (1908). Student. *Biometrika*, 6, 1.
- Grosbois, V., & Tavecchia, G. (2003). Modeling dispersal with capture–recapture data: Disentangling decisions of leaving and settlement. *Ecology*, 84, 1225–1236. [https://doi.org/10.1890/0012-9658\(2003\)084\[1225:MDWCDD\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[1225:MDWCDD]2.0.CO;2)
- Grosbois, V., Gimenez, O., Gaillard, J. M., Pradel, R., Barbraud, C., Clobert, J., ... Weimerskirch, H. (2008). Assessing the impact of climate variation on survival in vertebrate populations. *Biological Reviews*, 83, 357–399.
- Guéry, L., Descamps, S., Pradel, R., Hanssen, S. A., Erikstad, K. E., Gabrielsen, G. W., ... Bêty, J. (2017). Hidden survival heterogeneity of three Common eider populations in response to climate fluctuations. *Journal of Animal Ecology*, 86, 683–693. <https://doi.org/10.1111/1365-2656.12643>
- Jolly, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration–stochastic model. *Biometrika*, 52, 225–247. <https://doi.org/10.1093/biomet/52.1-2.225>
- Link, W. A. (1999). Modeling pattern in collections of parameters. *The Journal of Wildlife Management*, 63, 1017–1027.
- Link, W. A., & Barker, R. J. (2004). Hierarchical mark–recapture models: A framework for inference about demographic processes. *Animal Biodiversity and Conservation*, 27, 441–449.
- MacKenzie, D. I. (2006). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence*. Cambridge, MA: Academic Press.
- McFadden, D. (1968). *The revealed preferences of a government bureaucracy*. Berkeley, CA: Department of Economics, University of California. unpublished.
- Nichols, J. D., & Kendall, W. L. (1995). The use of multi-state capture–recapture models to address questions in evolutionary ecology. *Journal of Applied Statistics*, 22, 835–846. <https://doi.org/10.1080/02664769524658>
- Pape Møller, A., & Szép, T. (2002). Survival rate of adult barn swallows *Hirundo rustica* in relation to sexual selection and reproduction. *Ecology*, 83, 2220–2228. <https://doi.org/10.2307/3072053>
- Powell, L. A. (2007). Approximating variance of demographic parameters using the delta method: A reference for avian biologists. *The Condor*, 109, 949–954. [https://doi.org/10.1650/0010-5422\(2007\)109\[949:AVODPU\]2.0.CO;2](https://doi.org/10.1650/0010-5422(2007)109[949:AVODPU]2.0.CO;2)
- Pradel, R. (2005). Multievent: An extension of multistate capture–recapture models to uncertain states. *Biometrics*, 61, 442–447.
- Pradel, R., Gimenez, O., & Lebreton, J.-D. (2005). Principles and interest of GOF tests for multistate capture–recapture models. *Animal Biodiversity and Conservation*, 28, 189–204.
- Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*. London: Academic Press.
- Seber, G. A. (1965). A note on the multiple-recapture census. *Biometrika*, 52, 249–259. <https://doi.org/10.1093/biomet/52.1-2.249>
- Sutherland, C., Brambilla, M., Pedrini, P., & Tenan, S. (2016). A multiregion community model for inference about geographic variation in species richness. *Methods in Ecology and Evolution*, 7, 783–791.
- Team, R. C. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician*, 66, 124–127. <https://doi.org/10.1080/00031305.2012.687494>
- White, G. C., & Burnham, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study*, 46, S120–S139. <https://doi.org/10.1080/00063659909477239>
- Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and management of animal populations*. Cambridge, MA: Academic Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Guéry L, et al. Covariate and multinomial: Accounting for distance in movement in capture–recapture analyses. *Ecol Evol*. 2019;9:818–824. <https://doi.org/10.1002/ece3.4827>