

# Improved methods for estimating abundance and related demographic parameters from mark-resight data

Brett T. McClintock<sup>1</sup>  | Gary C. White<sup>2</sup> | Moira A. Pryde<sup>3</sup>

<sup>1</sup>Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA National Marine Fisheries Service, Seattle, Washington

<sup>2</sup>Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, Colorado

<sup>3</sup>New Zealand Department of Conservation, Wellington, New Zealand

## Correspondence

Brett T. McClintock, Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA National Marine Fisheries Service, 7600 Sand Point Way NE, Seattle, WA 98115.  
Email: brett.mcclintock@noaa.gov

## Abstract

Over the past decade, there has been much methodological development for the estimation of abundance and related demographic parameters using mark-resight data. Often viewed as a less-invasive and less-expensive alternative to conventional mark recapture, mark-resight methods jointly model marked individual encounters and counts of unmarked individuals, and recent extensions accommodate common challenges associated with imperfect detection. When these challenges include both individual detection heterogeneity and an unknown marked sample size, we demonstrate several deficiencies associated with the most widely used mark-resight models currently implemented in the popular capture-recapture freeware Program MARK. We propose a composite likelihood solution based on a zero-inflated Poisson log-normal model and find the performance of this new estimator to be superior in terms of bias and confidence interval coverage. Under Pollock's robust design, we also extend the models to accommodate individual-level random effects across sampling occasions as a potentially more realistic alternative to models that assume independence. As a motivating example, we revisit a previous analysis of mark-resight data for the New Zealand Robin (*Petroica australis*) and compare inferences from the proposed estimators. For the all-too-common situation where encounter rates are low, individual detection heterogeneity is non-negligible, and the number of marked individuals is unknown, we recommend practitioners use the zero-inflated Poisson log-normal mark-resight estimator as now implemented in Program MARK.

## KEYWORDS

capture-recapture, capture resight, population size, Program NOREMARK, survival

## 1 | INTRODUCTION

Mark-resight methods are commonly used to estimate abundance and related demographic parameters for animal populations (eg, White and Shenk, 2001; McClintock and White, 2012). They are widely applicable across diverse taxa, including terrestrial mammals (eg, Keech

*et al.*, 2011), marine mammals (eg, Ryan *et al.*, 2011), and birds (eg, Lyons *et al.*, 2016). Mark-resight methods are most often utilized when some segment of the population is marked, and non-invasive sighting data can be simultaneously collected for both marked and unmarked individuals. Non-invasive sighting data can include those arising from ground (eg, McClintock and White, 2007),

aerial (eg, Keech *et al.*, 2011), and camera (eg, Alonso *et al.*, 2015) surveys. Because mark-resight does not require repeated physical capture of individuals, the main advantage of this approach is that costs associated with marking and recapturing can be minimized. In many circumstances, it can therefore be a less-invasive and less-expensive alternative to traditional mark recapture for population monitoring. Relative to capture recapture, mark-resight abundance estimators also tend to be more robust to individual heterogeneity in detection probability because marked individuals are selected independent of the sighting process (ie, “all-zero” detection histories are observed).

Mark-resight estimators have seen much development over the past decade (McClintock and White, 2012; Matechou *et al.*, 2013; Alonso *et al.*, 2015; Lyons *et al.*, 2016), and the models that are currently implemented in the freeware Program MARK (White and Burnham, 1999) are widely used by practitioners. Because it does not require knowledge of the exact number of marked individuals in the population, one of the most popular mark-resight models for estimating abundance and survival parameters is the zero-truncated Poisson log-normal estimator (ztPNE; McClintock *et al.*, 2009; McClintock & White, 2009). Here, we identify several potential deficiencies of the ztPNE currently implemented in Program MARK. Among them are its failure to utilize valuable information about the number of marks ever released into the population and, under Pollock's robust design, its conditioning on first *sighting* instead of first *encounter* (eg., capture). The ztPNE in Program MARK also relies on an ad-hoc adjustment term to account for marks that are sighted but not identified to individual identity (see McClintock *et al.*, 2014). Lastly, while the ztPNE implemented in Program MARK can include individual-level random effects in detection probability, these individual heterogeneity effects are (perhaps unrealistically) assumed to be independent across primary sampling periods.

As a solution to these potential deficiencies of ztPNE, we propose the zero-inflated Poisson log-normal estimator (ziPNE). By allowing for zero inflation in the marked individual sighting data, ziPNE can incorporate all marked individuals ever released into the population and thus does not need to condition on first sighting. We also extend the likelihoods for these models to accommodate unidentified marks and individual-level random effects across primary sampling occasions. After describing ziPNE in the next section, we compare the performance of ziPNE relative to ztPNE under a broad range of simulated sampling scenarios. We then revisit the New Zealand Robin (*Petroica australis*) analysis of McClintock and White

(2009) and compare inferences from the various estimators.

## 2 | METHODS

There are several different study designs for the mark-resight models in Program MARK (McClintock and White, 2012). The mark-resight scenario under consideration consists of  $T$  sampling occasions where sightings of both marked and unmarked individuals are recorded. The observable population size ( $N_t$ ) is defined as the number of individuals that are available to be observed (eg, alive and within the study area) during sighting survey  $t = 1, \dots, T$ . The population is assumed to be geographically and demographically closed within each survey, but not between surveys. To account for temporary emigration off the study area (or any other process that causes a segment of the population to be temporarily unobservable), we define the unobservable population size as the number of individuals that were alive during survey  $t$  but had temporarily emigrated off the study area between times  $t - 1$  and  $t$ .

During each survey, there are  $n_t \in \{0, \dots, M_t\}$  marked individuals and  $U_t = N_t - n_t$  unmarked individuals in the observable population, where  $M_t = \sum_{k=1}^t R_k$  is the total number of marked individuals ever released into the population prior to survey  $t$ , and  $R_t$  is the number of (unmarked) individuals that were newly marked between times  $t - 1$  and  $t$ . The mark-resight data therefore consist of the number of resightings for each marked individual ( $x_{s,t} \in \{0, 1, 2, \dots\}$  for  $s \in \mathbb{M}_t$ ) and the total number of unmarked individual sightings ( $u_t = \sum_{s \notin \mathbb{M}_t} x_{s,t}$ ), where  $\mathbb{M}_t$  is the set of  $M_t$  marked individuals ever released into the population prior to survey  $t$ . In the common circumstance where the identity of an individual detected as marked cannot always be determined (McClintock *et al.*, 2014), then the observed data for the marked segment of the population consist of  $y_{s,t} = x_{s,t} - \epsilon_{s,t}$  for  $s \in \mathbb{M}_t$  and the total number of marked sightings that were not individually identified ( $e_t = \sum_{s \in \mathbb{M}_t} \epsilon_{s,t}$ ). The number of marks alive and within the study area can sometimes be reasonably assumed to be known exactly (eg, if all  $n_t$  marks are introduced immediately prior to sighting survey  $t$ ), but more commonly  $n_t$  is unknown due to attrition of marks over time from mortality and permanent emigration.

When  $n_t$  is unknown for at least one sampling occasion, the zero-truncated Poisson log-normal estimator (ztPNE, McClintock *et al.*, 2009; McClintock and White, 2009) can be used to estimate abundance ( $N_t$ ), apparent survival from time  $t$  to time  $t + 1$  ( $\phi_t$ ), the transition probability from an observable state (eg, in the

study area) at time  $t$  to an unobservable state (eg, temporarily off the study area) at time  $t + 1$  ( $\gamma''$ ), and the transition probability from an unobservable state at time  $t$  to an unobservable state at time  $t + 1$  ( $\gamma'$ ). However, the ztPNE does not capitalize on all of the available information afforded by this mark-resight study design. Notably, the model does not incorporate  $M_t$  into estimation of detection probability parameters, and, under Pollock's robust design (Kendall *et al.*, 1997; McClintock and White, 2009), it conditions on the first *sighting* of each marked individual (instead of first encounter; eg, capture) for estimating the open population parameters ( $\phi$ ,  $\gamma'$ , and  $\gamma''$ ).

To make better use of the available information, we propose the zero-inflated Poisson log-normal estimator (ziPNE). The assumptions of ziPNE are identical to the standard mark-resight assumptions of ztPNE (McClintock *et al.*, 2009; McClintock and White, 2009). Unmodeled individual sighting heterogeneity is well known

$\mathbf{u} = (u_1, \dots, u_T)$ . A further complication for the “across” heterogeneity models is that when  $R_t > 0$  for  $t \in \{2, \dots, T\}$ , then  $y_{s,t:T}$ ,  $e_{t:T}$ , and  $u_{1:t-1}$  may not be independent for  $s \notin \mathbb{M}_{t-1}$ . One possible solution is to formulate a complete data likelihood and use data augmentation (eg, McClintock *et al.*, 2014; Rankin *et al.*, 2016), but such approaches can be time-consuming and difficult to fit (see Web Appendix A). We instead adopt a practical perspective and focus our efforts on a composite likelihood (eg, Varin *et al.*, 2011) that is amenable to maximum likelihood estimation and implementation in Program MARK. Web Table B1 includes definitions for all parameters, latent variables, and random variables in our composite likelihood.

To make the problem tractable, we propose to ignore any covariance between  $\mathbf{Y}$ ,  $\mathbf{e}$ , and  $\mathbf{u}$  and simply integrate over the latent  $\mathbf{Z}$  for each of three components of a composite observed data likelihood as if they were independent:

$$\begin{aligned}
 [\mathbf{Y}, \mathbf{u}, \mathbf{e}|\mathbf{U}, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] &= \prod_{s=1}^{M_T} \int_{\mathbf{z}_s} [y_s|\mathbf{z}_s, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''][\mathbf{z}_s] \\
 &\times \int_{\mathbf{Z}} [\mathbf{e}|\mathbf{Z}, \alpha, \sigma, \mathbf{r}][\mathbf{Z}] \\
 &\times \int_{\mathbf{Z}} [\mathbf{u}|\mathbf{Z}, \mathbf{U}, \alpha, \sigma][\mathbf{Z}],
 \end{aligned}
 \tag{1}$$

to induce bias in abundance estimators, and we develop two ziPNE formulations based on whether or not individual-level effects for sighting rates are assumed independent (“within” heterogeneity *sensu* McClintock *et al.*, 2009) or not (“across” heterogeneity *sensu* Gimenez and Choquet, 2010) between sampling occasions.

Assuming the latent  $x_{s,t}$  sightings follow a Poisson log-normal distribution, we have  $x_{s,t}|\lambda_{s,t} \sim \text{Poisson}(\lambda_{s,t})$ ,  $y_{s,t}|\lambda_{s,t}, r_t \sim \text{Poisson}(\lambda_{s,t}r_t)$ , and  $\epsilon_{s,t}|\lambda_{s,t}, r_t \sim \text{Poisson}(\lambda_{s,t}(1-r_t))$ , where  $\lambda_{s,t} = \exp(\alpha_t + \sigma_t z_{s,t})$  and  $r_t$  is the conditional probability that a mark is individually identified (given detected as marked). The individual sighting rate ( $\lambda_{s,t}$ ) includes a (log-scale) mean ( $\alpha_t$ ), standard deviation ( $\sigma_t$ ), and latent individual-level effect  $z_{s,t} \sim \mathcal{N}(0, 1)$ . When the latent individual-level effects include time dependence, we refer to this as the “within” heterogeneity model. For an “across” heterogeneity model, we simply remove time dependence from the individual effects:  $\lambda_{s,t} = \exp(\alpha_t + \sigma_t z_s)$ , where  $z_s \sim \mathcal{N}(0, 1)$ .

Under this scenario, likelihood-based estimation of  $N_t$  and related demographic parameters is challenging for several reasons. The individual-level effects clearly induce dependencies between  $\mathbf{Y} = (y_{s,t})$  and  $\mathbf{e} = (e_1, \dots, e_T)$ , and, for the “across” heterogeneity models, within  $\mathbf{Y}$ ,  $\mathbf{e}$ , and

where bracket notation  $[data|parameter]$  denotes a probability density (or mass) function. We can account for zero inflation in  $y_s$  by assuming marked individuals are in one of three states at time  $t$  (1 = dead or permanent emigrant; 2 = alive, not a permanent emigrant, off the study area; 3 = alive, not a permanent emigrant, on the study area). Let  $\mathbf{w}_t = [1 - w_t, w_t(1 - g_t), w_t g_t]$  be a row vector of initial state probabilities, where  $w_t$  is the probability that a randomly selected individual from the  $R_t$  new marks introduced between sighting occasion  $t - 1$  and  $t$  was alive and not a permanent emigrant during sighting occasion  $t$ , and  $g_t$  is the conditional probability that this individual was within the study area (given alive and not a permanent emigrant). Let  $\mathbf{P}_{s,t}$  be a  $3 \times 3$  matrix where the diagonal elements denote the probability mass function of  $y_{s,t}$  conditional on the state of individual  $s$  at time  $t$ , and let  $\mathbf{G}_t$  be a  $3 \times 3$  state transition probability matrix with rows corresponding to the state at time  $t$  and the columns corresponding to the state at time  $t + 1$ . We then have

$$[y_s|\mathbf{z}_s, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] = \mathbf{w}_{b_s} \mathbf{P}_{s,b_s} \left[ \prod_{i=b_s}^{T-1} \mathbf{G}_i \mathbf{P}_{s,i+1} \right] \mathbf{1}_3$$

$$\mathbf{P}_{s,t} = \begin{bmatrix} 1 - I(y_{s,t} > 0) & 0 & 0 \\ 0 & 1 - I(y_{s,t} > 0) & 0 \\ 0 & 0 & p_{s,t} \end{bmatrix},$$

$$\mathbf{G}_t = \begin{bmatrix} 1 & 0 & 0 \\ 1 - \phi_t & \phi_t \gamma_t' & \phi_t (1 - \gamma_t') \\ 1 - \phi_t & \phi_t \gamma_t'' & \phi_t (1 - \gamma_t'') \end{bmatrix},$$

where  $b_s \in \{1, \dots, T\}$  is the first sighting occasion for which individual  $s$  was marked (and thus potentially available for resighting),  $I(\cdot)$  is the indicator function,  $\mathbf{1}_3 = [1, 1, 1]^T$ , and  $p_{s,t} = [y_{s,t} | \lambda_{s,t}, r_t]$ . By allowing  $y_{s,t}$  for the  $R_t$  newly marked individuals to be zero inflated (with probability  $1 - w_t g_t$ ), we can include  $M_T$  and condition on first encounter (instead of first sighting) in our composite likelihood.

For both the “within” and “across” models, we wish to integrate over the latent variables  $z_{s,t}$  and  $z_s$ , respectively, in Equation (1). While this is straightforward for  $[y_s | z_s, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma'']$ ,  $[\mathbf{u} | \mathbf{U}, \alpha, \sigma] = \int_{\mathbf{Z}} [\mathbf{u} | \mathbf{Z}, \mathbf{U}, \alpha, \sigma] [\mathbf{Z}]$  and  $[\mathbf{e} | \alpha, \sigma, \mathbf{r}] = \int_{\mathbf{Z}} [\mathbf{e} | \mathbf{Z}, \alpha, \sigma, \mathbf{r}] [\mathbf{Z}]$  pose difficulties because sums of Poisson log-normal random variables are not of standard form (eg, Poisson) and they involve multidimensional integrals that are not amenable to numerical approximation. Similar to ztPNE, we rely on a left-truncated normal distribution for  $[\mathbf{u} | \mathbf{U}, \alpha, \sigma]$  based on the expected value  $\lambda_t = E_z(x_{s,t} | \lambda_{s,t}) = \int E(x_{s,t} | \lambda_{s,t}, z) dz = \int \exp(\alpha_t + \sigma_t z) dz = \exp(\alpha_t + \sigma_t^2/2)$  and variance  $\eta_t = \text{var}_z(x_{s,t} | \lambda_{s,t}) = E_z(\text{var}(x_{s,t} | \lambda_{s,t})) + \text{var}_z(E(x_{s,t} | \lambda_{s,t})) = \lambda_t + \exp(2\alpha_t)(\exp(2\sigma_t^2) - \exp(\sigma_t^2))$  of a Poisson log-normal random variable after integrating over the individual-level effects:

$$[\mathbf{u} | \mathbf{U}, \alpha, \sigma] = \prod_{t=1}^T [u_t^* | U_t, \alpha_t, \sigma_t]$$

where  $u_t^* | U_t, \alpha_t, \sigma_t \sim \mathcal{TN}_{(0, \infty)}(U_t \lambda_t, U_t \eta_t)$ ,  $u_t^* = u_t$  is a non-negative continuous random variable, and  $\mathcal{TN}_{(a,b)}$  represents a normal distribution truncated at  $a$  and  $b$ . Note that for the “across” heterogeneity model, this composite likelihood also ignores the covariance induced in  $\mathbf{u}$  by the latent individual-level effects.

We employ a similar left-truncated normal distribution for  $[\mathbf{e} | \alpha, \sigma, \mathbf{r}]$ :

$$[\mathbf{e} | \alpha, \sigma, \mathbf{r}] = \prod_{t=1}^T [e_t^* | \alpha_t, \sigma_t, r_t] \tag{2}$$

where  $e_t^* | \alpha_t, \sigma_t, r_t \sim \mathcal{TN}_{(0, \infty)}(n_t \lambda_t (1 - r_t), n_t \kappa_t)$ ,  $\kappa_t, \lambda_t (1 - r_t) + \exp(2\alpha_t)(1 - r_t)^2(\exp(2\sigma_t^2) - \exp(\sigma_t^2))$ , and  $e_t^* = e_t$  is a

non-negative continuous random variable. While  $n_t$  can sometimes be reasonably assumed to be known exactly (ie,  $n_t = M_t$  and  $w_t = g_t = 1$ ), for all occasions where the number of marks alive and within the study area is not known exactly, we replace  $n_t$  in Equation (2) with a Horvitz-Thompson-type estimator:  $\hat{n}_t = \sum_{s \in M_t} I(y_{s,t} > 0) p_t^{-1}$ , where  $p_t = 1 - \int \exp(-\exp(\alpha_t + \sigma_t z) r_t) [z] dz$  is the probability of a marked individual being resighted at least once.

We therefore have the following composite likelihood for the “within” heterogeneity ziPNE:

$$[\mathbf{Y}, \mathbf{u}, \mathbf{e} | \mathbf{U}, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] = \prod_{s=1}^{M_T} [y_s | \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] \times \prod_{t=1}^T [e_t^* | \alpha_t, \sigma_t, r_t] [u_t^* | U_t, \alpha_t, \sigma_t], \tag{3}$$

where  $p_{s,t} = \int [y_{s,t} | \lambda_{s,t}, r_t] [z_{s,t}] dz_{s,t}$  and  $\lambda_{s,t} = \exp(\alpha_t + \sigma_t z_{s,t})$ . For the “across” heterogeneity ziPNE:

$$[\mathbf{Y}, \mathbf{u}, \mathbf{e} | \mathbf{U}, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] = \prod_{s=1}^{M_T} \int [y_s | z_s, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] [z_s] dz_s \times \prod_{t=1}^T [e_t^* | \alpha_t, \sigma_t, r_t] [u_t^* | U_t, \alpha_t, \sigma_t], \tag{4}$$

where  $p_{s,t} = [y_{s,t} | \lambda_{s,t}, r_t]$  and  $\lambda_{s,t} = \exp(\alpha_t + \sigma_t z_s)$ . For maximum likelihood analysis, the integrals in Equations (3) and (4) can be easily approximated using Gauss-Hermite quadrature:

$$p_{s,t} = \int [y_{s,t} | \lambda_{s,t}, r_t] [z_{s,t}] dz_{s,t} \approx \frac{1}{\sqrt{\pi}} \sum_{l=1}^L f_l [y_{s,t} | \lambda_{l,t}^*, r_t],$$

and

$$\int [y_s | z_s, \alpha, \sigma, \mathbf{r}, \mathbf{w}, \mathbf{g}, \phi, \gamma', \gamma''] [z_s] dz_s \approx \frac{1}{\sqrt{\pi}} \sum_{l=1}^L f_l \left\{ \mathbf{w}_{b_s} \mathbf{P}_{l,b_s}^* \left[ \prod_{i=b_s}^{T-1} \mathbf{G}_i \mathbf{P}_{l,i+1}^* \right] \mathbf{1}_3 \right\},$$

where

$$\mathbf{P}_{l,t}^* = \begin{bmatrix} 1 - I(y_{s,t} > 0) & 0 & 0 \\ 0 & 1 - I(y_{s,t} > 0) & 0 \\ 0 & 0 & [y_{s,t} | \lambda_{l,t}^*, r_t] \end{bmatrix},$$

$y_{s,t} | \lambda_{l,t}^*, r_t \sim \text{Poisson}(\lambda_{l,t}^* r_t)$ ,  $\lambda_{l,t}^* = \exp(\alpha_t + \sqrt{2} \sigma_t v_l)$ , and  $(f_l, v_l)$  are the corresponding weights and nodes, respectively, for quadrature points  $l = 1, \dots, L$  (Abramowitz and

Stegun, 1964). When  $n_t$  is unknown, we have  $p_t \approx 1 - \frac{1}{\sqrt{\pi}} \sum_{l=1}^L f_l \exp(-\lambda_{l,t}^* r_t)$ .

The size of the observable population (eg, alive and within the study area) during sighting survey  $t$  ( $N_t = U_t + n_t$ ) can be derived from the estimated parameters after fitting Equations (3) or (4) using maximum likelihood methods. As with all derived parameters in Program MARK, standard errors for  $N_t$  are estimated using the delta method and finite-difference approximations of the first derivative. Program MARK also facilitates modeling any of the parameters as a function of covariates using link functions (see Section 4). We note that in a fully time-dependent model for  $w_t$  and  $g_t$  ( $t = 1, \dots, T$ ), only the product  $w_T g_T$  is estimable. When  $T = 1$ , only the product  $w_1 g_1$  is estimable. We also note that because the “within” and “across” heterogeneity models use the same data, they can be compared using model selection criteria such as Akaike’s Information Criterion (AIC; Burnham and Anderson, 2002).

### 3 | SIMULATION STUDY

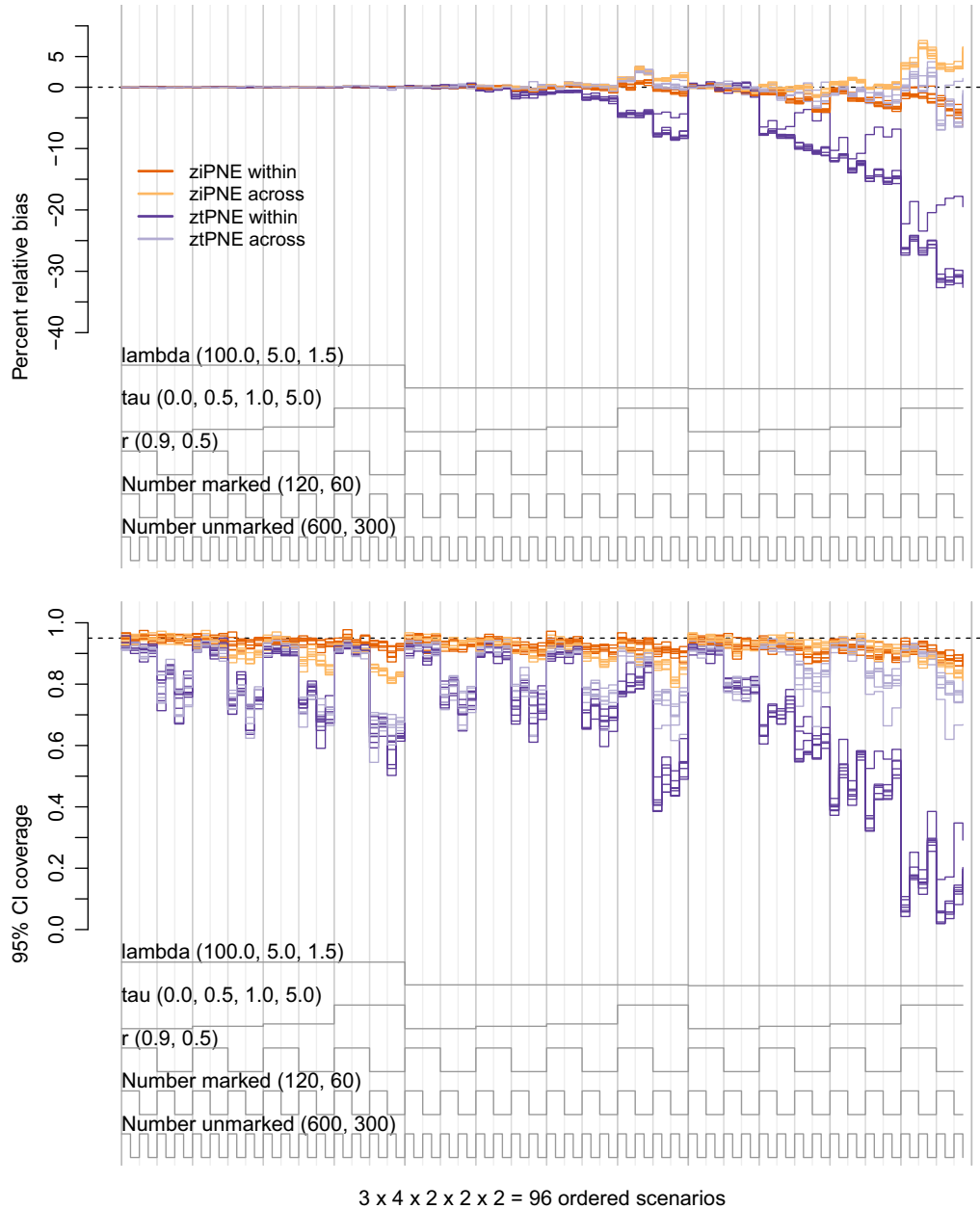
Using data generated from the exact model, we performed a series of simulation experiments to evaluate the performance of our composite ziPNE relative to the ztPNE as implemented in Program MARK (McClintock and White, 2012). Each simulated scenario included  $T = 7$ ,  $\phi_t = 0.95$ ,  $\gamma_t'' = \gamma_t' = 0.1$ ,  $w_t = 0.95$ ,  $g_t = 0.9$ ,  $r_t \in \{0.5, 0.9\}$ ,  $E(R_t) \in \{60, 120\}$ ,  $E(U_t) \in \{300, 600\}$ ,  $\lambda_t \in \{1.5, 5, 100\}$ , and  $\tau \in \{0, 0.5, 1, 5\}$ , where  $\text{var}_z(x_{s,t} | \lambda_{s,t}) = \lambda_t(1 + \tau)$ , such that  $\alpha_t = 2 \log(\lambda_t) - \log(\lambda_t^2 + \lambda_t \tau)/2$  and  $\sigma_t = \sqrt{\log(\lambda_t^2 + \lambda_t \tau) - 2 \log(\lambda_t)}$  for  $t = 1, \dots, T$ . For  $R_t$  ( $t = 2, \dots, T$ ), some of the  $U_{t-1}$  unmarked individuals became newly marked (with probability 0.02) in order to maintain a somewhat consistent marked population size over time. However, no additional unmarked individuals were introduced to the population after the first occasion, so  $N_t$  is expected to steadily decline through time due to mortality and permanent emigration. Scenarios with  $\lambda_t = 100$  represent the “best case” scenario where sighting rates are very high and individuals that are alive and within the study area are virtually guaranteed to be sighted (i.e.  $x_{s,t} > 0$ ). Under this scenario, we would expect ziPNE and ztPNE to perform similarly. Scenarios with both  $\lambda_t = 1.5$  and  $\tau = 5$  represent the pathological case where mean sighting rates are low and individual heterogeneity is high, whereby the vast majority of individuals are never expected to be sighted ( $x_{s,t} = 0$ ) but some individuals will be sighted many times ( $x_{s,t} \gg 0$ ). In terms of  $\alpha$ ,  $\sigma$ , and  $N$  estimation, we would generally expect all estimators to perform less than nominally under this extreme scenario (particularly when  $r_t = 0.5$ ). While we would not expect  $\phi$

estimators to be biased by extreme heterogeneity coupled with low mean sighting rates, we suspect it could also be deleterious to reliable estimation of  $\gamma''$  and  $\gamma'$ .

For each combination, we simulated data from both the “within” and “across” heterogeneity models, thus yielding  $2 \times 2 \times 2 \times 2 \times 3 \times 4 = 192$  simulated scenarios. For each scenario, we fit the two models corresponding to the data-generating heterogeneity model: 1) ziPNE within heterogeneity Equation (3) and ztPNE within heterogeneity; or 2) ziPNE across heterogeneity Equation (4) and ztPNE across heterogeneity. We investigated heterogeneity model mis-specification in an additional series of simulation experiments (see Web Appendix E). We assumed that  $n_1$  was known,  $n_t$  ( $t = 2, \dots, T$ ) was unknown, and  $M_1 = n_1$ . For scenarios with  $\tau = 0$ , we fixed  $\sigma_t = 0$  for each of the fitted models. For simplicity, we also fixed  $w_1 = g_1 = 1$  (because  $n_1$  is known and  $n_1 = M_1$ ) and otherwise assumed no time dependence in parameters (eg,  $\phi_1 = \phi_2 = \dots = \phi_{T-1}$ ). All simulations used  $L = 101$  quadrature points for numerical integration. We evaluated estimator performance based on bias, 95% confidence interval coverage of the true parameter values, and root mean squared error (RMSE). Simulation results are presented using the nested loop plot of Rucker and Schwarzer (2014). Similar to a time series plot, nested loop plots serve to present a large number of simulation results by putting all scenarios into a lexicographical order and arranging them consecutively along the horizontal axis. The quantity of interest (eg, bias, coverage, and RMSE) is then plotted on the vertical axis.

We found the “within” heterogeneity ziPNE to be far superior to ztPNE for the estimation of the closed population parameters ( $N$ ,  $\alpha$ , and  $\sigma$ ), particularly as  $\lambda_t$  decreased,  $\tau$  increased, and  $r$  decreased. In nearly all of the scenarios, ziPNE tended to be less biased with higher 95% confidence interval coverage and lower RMSE (Figure 1, Web Figures C1–C4). Although it performed much better than ztPNE, ziPNE still exhibited an increase in bias and decrease in coverage under the pathological scenarios with low mean resighting rates and extreme heterogeneity. While little difference was found between ziPNE and ztPNE in terms of apparent survival ( $\phi$ ) estimation (Figure 2), ziPNE performed much better than ztPNE in the estimation of both  $\gamma''$  and  $\gamma'$  (Web Figures C5–C8).

We did not find such drastic differences in the performance of the “across” heterogeneity ziPNE relative to ztPNE, but ziPNE did tend to perform better in the estimation of the closed population parameters (Figure 1, Web Figures C2–C4). While the “across” ztPNE tended to exhibit slightly less bias and lower RMSE for  $N$  in the most extreme scenarios, the “across” ziPNE tended to yield higher 95% confidence interval coverage (particularly when  $r_t = 0.5$ ; Figure 1, Web Figure C1). Nevertheless, coverage

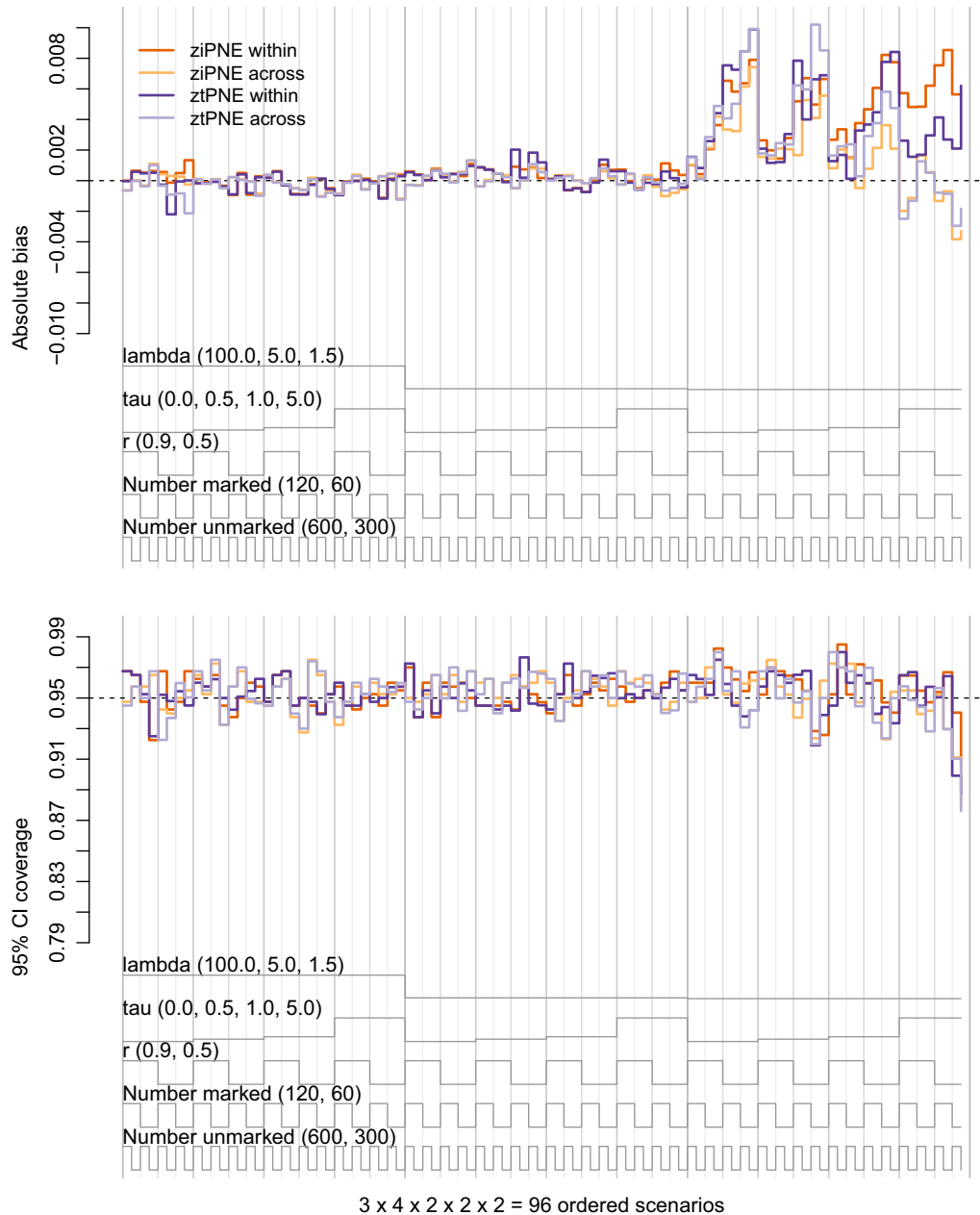


**FIGURE 1** Nested loop plots of percent relative bias (top panel) and 95% confidence interval coverage (bottom panel) of abundance ( $N_t, t = 1, \dots, 7$ ) estimators based on 192 simulated scenarios. Scenarios are ordered from outer to inner loops by  $\lambda_t \in \{100, 5, 1.5\}$  (“lambda”; three levels, decreasing),  $\tau \in \{0, 0.5, 1, 5\}$  (“tau”; four levels, increasing),  $r \in \{0.9, 0.5\}$  (two levels, decreasing),  $E(R_1) \in \{120, 60\}$  (“Number marked”; two levels, decreasing), and  $E(U_1) \in \{600, 300\}$  (“Number unmarked”; two levels, decreasing) such that the left- and right-most scenarios are the best- and worst-case sampling scenarios, respectively. Estimators include the zero-inflated Poisson log-normal estimator (dark orange = ziPNE within, light orange = ziPNE across) and the zero-truncated Poisson log-normal estimator (dark purple = ztPNE within, light purple = ztPNE across). Results are based on 400 simulated data sets for each scenario. This figure appears in color in the electronic version of this article, and color refers to that version

of  $N$  still tended to be less than nominal when  $\tau > 0$  and  $r_t = 0.5$  for all  $\lambda_t$ . For the open parameters ( $\phi, \gamma'', \gamma'$ ), we found that both ziPNE and ztPNE performed equally well in terms of bias and confidence interval coverage (Figure 2), but RMSE for  $\phi$  and  $\gamma'$  in particular tended to be larger for ztPNE (Web Figures C5-C8).

#### 4 | EXAMPLE: NEW ZEALAND ROBIN

As an example application, we return to the New Zealand Robin (*P. australis*) data originally analyzed using ztPNE by McClintock and White (2009). This mark-resight study was conducted by the New Zealand Department of



**FIGURE 2** Nested loop plots of absolute bias (top panel) and 95% confidence interval coverage (bottom panel) of apparent survival ( $\phi$ ) estimators based on 192 simulated scenarios. Scenarios are ordered from outer to inner loops by  $\lambda$ ,  $\in \{100, 5, 1.5\}$  (“lambda”; three levels, decreasing),  $\tau \in \{0, 0.5, 1, 5\}$  (“tau”; four levels, increasing),  $r \in \{0.9, 0.5\}$  (two levels, decreasing),  $E(R_1) \in \{120, 60\}$  (“Number marked”; two levels, decreasing), and  $E(U_i) \in \{600, 300\}$  (“Number unmarked”; two levels, decreasing) such that the left- and right-most scenarios are the best- and worst-case sampling scenarios, respectively. Estimators include the zero-inflated Poisson log-normal estimator (dark orange = ziPNE within, light orange = ziPNE across) and the zero-truncated Poisson log-normal estimator (dark purple = ztPNE within, light purple = ztPNE across). Results are based on 400 simulated data sets for each scenario. This figure appears in color in the electronic version of this article, and color refers to that version

Conservation on three 100 ha study areas (“Knobs Flat”, “Smithy”, and “Walker Creek”) in Fiordland National Park, New Zealand. Between September 2003 and August 2007, capture and banding of both juvenile and adult birds occurred continuously throughout the breeding season (August-March) and intermittently prior to sighting surveys conducted in March (postbreeding)

and August (prebreeding) of each year from 2005 to 2007. Because marked birds could have died or permanently emigrated during the extended capture period, the exact number of marked individuals in each population was unknown for every sighting occasion. For each occasion, the observable population for abundance estimation consisted of all juvenile and adult Robins

within a study area. Sighting surveys did not begin in Smithy until August 2005, so there were  $T = 6$  sighting occasions for both Knobs Flat and Walker Creek, but only  $T = 5$  occasions for Smithy. Notably, ship rat (*Rattus rattus*) control was conducted on the Walker Creek study area from June 2006 to April 2007. Complete details of the study design and data can be found in McClintock and White (2009).

We performed a ziPNE analysis for comparison to the “within” ztPNE analysis of McClintock and White (2009) using Program MARK. While the original ztPNE analysis only included detection histories for  $\sum_{s \in M_T} I(\sum_{t=1}^T y_{s,t} > 0) = 149$  marked birds that were sighted at least once between March 2005 and August 2007, the ziPNE analysis includes detection histories for all  $M_T = 439$  marked birds that were released into the populations prior to the August 2007 sighting occasion. Covariates examined for sighting rate and open population parameters included age class (*age*), sex, age class at time of capture (*agecap*), study area (*area*), time, breeding season (*season*), and rat control on Walker Creek. As in McClintock and White (2009), we also included an interaction between age, sex, and breeding season (*age\*sex\*season*) to investigate potential effects of reproduction on adult survival. For model selection, we used AIC adjusted for small sample sizes (AICc; Burnham and Anderson, 2002). Three covariate models for rat control on Walker Creek were examined: (1) an immediate and constant effect during the entire program (*rat*); (2) an immediate effect with no effect thereafter (*rat1*); and (3) no immediate effect with a delayed effect thereafter (*rat2*). We also investigated models under completely random emigration,  $\gamma'' = \gamma'$  (Kendall *et al.*, 1997), and models ignoring individual heterogeneity ( $\sigma = 0$ ) or temporary emigration ( $\gamma'' = \gamma' = 0$ ).

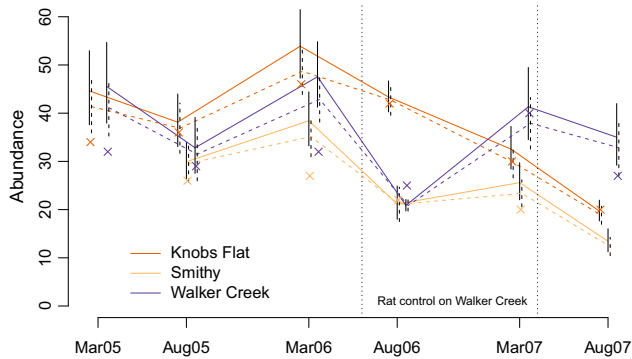
Similar to McClintock and White (2009), we first investigated various models for the sighting rate parameters ( $\alpha$ ,  $\sigma$ , and  $r$ ) under the most general (but identifiable) time- and group-dependent structure for  $U$  and the open parameters ( $\phi$ ,  $\gamma''$ , and  $\gamma'$ ). We proceeded by selecting the model structure for  $\alpha$ ,  $\sigma$ , and  $r$  best supported by AICc and then using only this structure for investigating more parsimonious models for  $\phi$ ,  $\gamma''$ , and  $\gamma'$  (for an evaluation of a similar model selection procedure, see Doherty *et al.*, 2012). We then fit all combinations of the three best supported structures for each parameter based on each stage of this selection process. We assumed the same structure for  $w$  and  $\phi$  and the same structure for  $g$ ,  $\gamma''$ , and  $\gamma'$ . However, because they pertain to the entire capture period beginning in September 2003, we also investigated  $w_1$  and  $g_1$  effects (*cap1*) for the Knobs Flat and Walker Creek study areas, and  $w_2$  and  $g_2$  effects (*cap2*) for the Smithy study area. Unlike the “within” ztPNE analysis

of McClintock and White (2009), we investigated both “within” and “across” heterogeneity models, and we again note that, unlike the ziPNE proposed here, the ztPNE of McClintock and White (2009) does not include the  $r$  parameter and instead relies on an ad-hoc adjustment term for unidentified marks. When fitting the models, we found the Markovian temporary emigration models including additive *age*, *agecap*, or *area* effects to have suspect convergence and thus excluded these from our model set. We based our comparisons on model-averaged parameter estimates and unconditional variances calculated from AICc weights.

AICc overwhelmingly supported “across” heterogeneity models for ziPNE, with the minimum AICc model including: *agecap*, *sex*, *season*, and *area* effects on  $\phi$  and  $w$ ; *cap1* and *cap2* effects on  $w$ ; *agecap* and *season* effects on random emigration (ie,  $\gamma'' = \gamma'$ ) and  $g$ ; *age*, *sex*, and *rat* effects on  $\alpha$ ; *sex* effects on  $\sigma$ ; and *area* by *time* interaction effects on  $r$ . Model selection was somewhat similar for the “within” ztPNE model of McClintock and White (2009), but the ziPNE analysis exhibited considerably more model selection uncertainty (Web Table D1) and weak evidence of rat control effects on  $\phi$ .

For ziPNE, model-averaged mean sighting rates ( $\lambda_t$ ) were the lowest in August 2007 (2.2, 95% CI: 1.9 – 2.5) and the highest on Walker Creek during rat control in March 2007 (3.2, 95% CI: 2.6 – 4). For the “across” ziPNE models, model-averaged  $\sigma = 0.52$  (SE = 0.18) for females, and  $\sigma = 0.28$  (SE = 0.07) for males. Individual mark identification probabilities ( $r$ ) varied by study area and time but were generally high, with model-averaged estimates ranging from 0.81 (SE = 0.07) to 1.00 (SE = 0.00). Point and variance estimates for  $N_t$  tended to be larger for ziPNE, and, except for Walker Creek in August 2006, all estimated 95% confidence intervals for  $N_t$  were greater than the minimum number known alive based on territory mapping that was independent of the mark-resight methodology (Figure 3). As with the original ztPNE analysis, the ziPNE analysis found evidence of higher apparent survival probabilities for individuals that were first captured as adults (Figure 4) when compared to individuals that were first captured as juveniles (Web Figure D1), but ziPNE exhibited less seasonal variation in adult female apparent survival relative to ztPNE. Temporary emigration parameter ( $\gamma''$  and  $\gamma'$ ) point estimates were similar between ziPNE and ztPNE, but ziPNE was generally more precise (Web Figures D2–D4). Estimates for  $w$  after longer capture intervals tended to be the lowest, particularly for individuals that were first captured as juveniles. Estimates for  $g$  were generally high and tended to be somewhat higher for juveniles (Web Figures D5–D6).



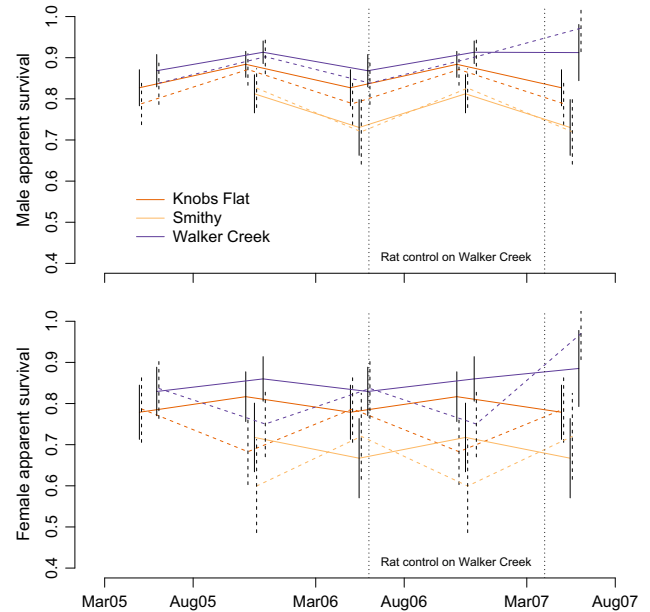


**FIGURE 3** Mark-resight abundance ( $N$ ) estimates and 95% confidence intervals between March 2005 and August 2007 for three New Zealand Robin (*Petroica australis*) populations (dark orange = Knobs Flat, light orange = Smithy, and dark purple = Walker Creek) in Fiordland National Park, New Zealand. Estimates are AICc model averaged based on the zero-inflated model (ziPNE; solid lines) and the zero-truncated model (ztPNE; dashed lines). Crosses (x) indicate the minimum number known alive based on territory mapping that was independent of the mark-resight methodology. Vertical dotted lines indicate a period of rat control on the Walker Creek study area. This figure appears in color in the electronic version of this article, and color refers to that version

## 5 | DISCUSSION

We have proposed a new zero-inflated version of the Poisson log-normal mark-resight estimator in Program MARK. Using simulation, we demonstrated ziPNE has superior performance relative to the zero-truncated model of McClintock and White (2009). By accounting for zero inflation in the marked individual sighting data, ziPNE is able to utilize valuable information about detection rates provided by marked individuals that were released but never subsequently sighted. This not only makes ziPNE more robust to individual heterogeneity for reliable estimation of population size but also allows the likelihood for the open population parameters ( $\phi$ ,  $\gamma''$ ,  $\gamma'$ ) to condition on first encounter (instead of first sighting). ziPNE thus makes better use of available data and, unlike ztPNE, allows  $\gamma'_1$  to be estimated under Pollock's robust design.

In terms of abundance estimation, previous simulation studies also found the “within” heterogeneity zero-truncated model of McClintock and White (2009) to perform increasingly poorly with low sighting rates, high levels of individual heterogeneity, and low-marked individual identification probabilities (McClintock *et al.*, 2009; McClintock and White, 2009; McClintock *et al.*, 2014). Our extensive simulations included a broader range of more extreme scenarios, and the new ziPNE performed near nominally and better than ztPNE in almost every scenario examined. However, the average



**FIGURE 4** Model-averaged four-month apparent survival ( $\phi$ ) estimates ( $\pm$ SE) between March 2005 and August 2007 for three New Zealand Robin populations (dark orange = Knobs Flat, light orange = Smithy, and dark purple = Walker Creek) in Fiordland National Park, New Zealand. Estimates are for males (top panel) and females (bottom panel) that were first captured as adults based on the zero-inflated model (ziPNE; solid lines) and the zero-truncated model (ztPNE; dashed lines). The longer intervals between August and March encompass the breeding season. Vertical dotted lines indicate a period of rat control on the Walker Creek study area. This figure appears in color in the electronic version of this article, and color refers to that version

coverage of  $N$  could still be as low as 81% for the “across” ziPNE when high levels of individual heterogeneity were combined with low mark identification probabilities. While we have not encountered actual mark-resight data as information poor as some of the scenarios included in our simulations, these pathological sampling scenarios provide valuable insights about the limitations of our composite likelihoods. Practitioners should carefully consider study designs that increase sighting rates, decrease individual heterogeneity, and increase marked individual identification whenever possible.

Unless changes in individual behavior (or sighting survey design) effectively eliminates correlations in the marked individual sighting data, the “across” model may often be a more natural choice than the “within” model. Indeed, our New Zealand Robin example demonstrated overwhelming support for the “across” models and thus strong correlations in the individual sighting data across sighting surveys. It deserves note that both the “across” and “within” models tended to perform poorly under heterogeneity model mis-specification (see Web Appendix E). Regardless of the model (ztPNE or ziPNE), it is

therefore important to examine both “within” and “across” heterogeneity models. Fortunately, these models can be compared using standard model selection criteria such as AIC and Bayesian information criterion (BIC). In some cases, it may be desirable to simultaneously model both “within” and “across” heterogeneity; this extension remains an avenue for future research.

Our zipNE is a composite likelihood that ignores some of the dependencies induced by individual heterogeneity among  $\mathbf{Y}$ ,  $\mathbf{e}$ , and  $\mathbf{u}$  in order to facilitate maximum likelihood estimation in Program MARK. While this composite likelihood proved robust under a broad range of simulation scenarios, an alternative (exact) solution can be implemented using a data-augmented complete data likelihood in a Bayesian framework (see Web Appendix A). While relatively easy to code using standard software, poor mixing can cause such implementations to be time consuming and challenging to fit. Thus, while conceptually appealing, they can often be impractical for simulation studies or real-world application (particularly for large  $T$  or  $N$ ). To aid practitioners in determining whether or not the performance of zipNE can be expected to meet their specific objectives under sampling scenarios not covered in our simulations, we provide R code for simulating data from the exact model and fitting zipNE in the Supporting Information.

We recommend practitioners use zipNE for mark-resight analyses in Program MARK whenever the number of marks in the population is unknown due to attrition arising from mortality and permanent emigration. As demonstrated in our New Zealand Robin example, we also recommend previous ztPNE analyses be revisited with zipNE where appropriate. Program MARK includes both individual heterogeneity model formulations (“within” and “across”), and has many additional analysis features, such as covariate modeling via link functions, model selection using AIC or BIC, multimodel inference, and plotting. MARK also provides the ability to simulate data under the mark-resight models, a useful feature for designing studies. For advanced users, these mark-resight models can also be implemented through the R package RMark (Laake, 2013) interface for Program MARK.

## ACKNOWLEDGMENTS

The findings and conclusions in the manuscript are those of the author(s) and do not necessarily represent the views of the National Marine Fisheries Service, NOAA. Any use of trade, product, or firm names does not imply an endorsement by the US Government.

## ORCID

Brett T. McClintock  <http://orcid.org/0000-0001-6154-4376>

## REFERENCES

- Abramowitz, M. and Stegun, I. (1964). Handbook of Mathematical Functions, With Formulas, Graphs and Math Tables (National Bureau of Standards Applied Math Series 55). Washington, DC: U. S. Government Printing Office.
- Alonso, R.S., McClintock, B.T., Lyren, L.M., Boydston, E.E. and Crooks, K.R. (2015). Mark-recapture and mark-resight methods for estimating abundance with remote cameras: a carnivore case study. *PLoS ONE*, 10, e0123032.
- Burnham, K.P. and Anderson, D.R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. New York, NY: Springer-Verlag.
- Doherty, P.F., White, G.C. and Burnham, K.P. (2012). Comparison of model building and selection strategies. *Journal of Ornithology*, 152, 317–323.
- Gimenez, O. and Choquet, R. (2010). Individual heterogeneity in studies on marked animals using numerical integration: capture-recapture mixed models. *Ecology*, 91, 951–957.
- Keech, M.A., Lindberg, M.S., Boertje, R.D., Valkenburg, P., Taras, B.D. and Boudreau, T.A. (2011). Effects of predator treatments, individual traits, and environment on moose survival in Alaska. *Journal of Wildlife Management*, 75, 1361–1380.
- Kendall, W.L., Nichols, J.D. and Hines, J.E. (1997). Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology*, 78, 563–578.
- Laake, J.L. (2013). RMark: An R Interface for Analysis of Capture-Recapture Data with MARK. AFSC Processed Rep 2013-01, 25p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.
- Lyons, J.E., Kendall, W.L., Royle, J.A., Converse, S.J., Andres, B.A. and Buchanan, J.B. (2016). Population size and stop-over duration estimation using mark-resight data and Bayesian analysis of a superpopulation model. *Biometrics*, 72, 262–271.
- Matechou, E., Morgan, B. J. T., Pledger, S., Collazo, J.A. and Lyons, J.E. (2013). Integrated analysis of capture-recapture-resighting data and counts of unmarked birds at stop-over sites. *Journal of Agricultural, Biological, and Environmental Statistics*, 18, 120–135.
- McClintock, B.T., Hill, J.M., Fritz, L., Chumbley, K., Luxa, K. and Diefenbach, D.R. (2014). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology and Evolution*, 5, 1294–1304.
- McClintock, B.T. and White, G.C. (2007). Bighorn sheep abundance following a suspected pneumonia epidemic in Rocky Mountain National Park. *Journal of Wildlife Management*, 71, 183–189.
- McClintock, B.T. and White, G.C. (2009). A less field-intensive robust design for estimating demographic parameters with mark-resight data. *Ecology*, 90, 313–320.
- McClintock, B.T. and White, G.C. (2012). From NOREMARK to MARK: Software for estimating demographic parameters using mark-resight methodology. *Journal of Ornithology*, 152, 641–650.

- 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.
- Rankin, R.W., Nicholson, K.E., Allen, S.J., Krützen, M., Bejder, L. and Pollock, K.H. (2016). A full-capture hierarchical Bayesian model of Pollock's closed robust design and application to dolphins. *Frontiers in Marine Science*, 3, 25.
- Rücker, G. and Schwarzer, G. (2014). Presenting simulation results in a nested loop plot. *BMC Medical Research Methodology*, 14, 129.
- Ryan, G.E., Dove, V., Trujillo, F. and Doherty, P.F. (2011). Irrawaddy dolphin demography in the Mekong river: An application of mark-resight models. *Ecosphere*, 2, 1–15.
- Varin, C., Reid, N. and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, 21, 5–42.
- White, G.C. and Burnham, K.P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study*, 46, S120–S138.
- White, G.C. and Shenk, T.M. (2001). Population estimation with radio-marked animals. In Millspaugh, Radio Tracking and Animal Populations. San Diego, CA: Academic Press. pp. 329–350.

## 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:** McClintock BT, White GC, Pryde MA. Improved methods for estimating abundance and related demographic parameters from mark-resight data. *Biometrics*. 2019;75: 799-809. <https://doi.org/10.1111/biom.13058>