

Calibrating brushtail possum (*Trichosurus vulpecula*) occupancy and abundance index estimates from leg-hold traps, wax tags and chew cards in the Department of Conservation's Biodiversity and Monitoring Reporting System

David M. Forsyth^{1*}, Mike Perry², Paul Moloney³, Meredith McKay⁴, Andrew M. Gormley⁵, Bruce Warburton⁵, Peter Sweetapple⁵ and Robyn Dewhurst²

¹Vertebrate Pest Research Unit, New South Wales Department of Primary Industries, 1447 Forest Road, Orange, NSW 2800, Australia

²Department of Conservation, 28 North Street, Palmerston North 4410, New Zealand

³Arthur Rylah Institute for Environmental Research, Department of Environment, Land, Water and Planning, 123 Brown Street, Heidelberg, Victoria 3084, Australia

⁴Science and Policy Group, Department of Conservation, 161 Cashel Street, Christchurch 8011, New Zealand

⁵Manaaki Whenua Landcare Research, Gerald Street, Lincoln 7640, New Zealand

*Author for correspondence (Email: dave.forsyth@dpi.nsw.gov.au)

Published online: 30 April 2018

Abstract: The Department of Conservation has implemented a Biodiversity and Monitoring Reporting System (BMRS) that estimates occupancy rates and relative abundances of introduced brushtail possums (*Trichosurus vulpecula*) at a representative sample of sites on public conservation land. Leg-hold traps have been used to monitor possums in the BMRS, but wax tags and chew cards have logistical and financial advantages over traps. If possum occupancy rate and/or abundance index estimates differ depending on which of the three methods are used, then correction of the estimates would be required for valid comparisons. We sampled possum occupancy rates and relative abundances using leg-hold traps, wax tags and chew cards in the BMRS sampling design at each of 54 forest and 54 non-forest sites. Possum occupancy rates estimated using each of the three detection methods were similar and hence do not require correction. Median possum abundance index values estimated from traps were lower than those estimated from wax tags and chew cards in forest, but were similar to those from wax tags in non-forest. Calibration is therefore required if possum abundance index values from either chew cards or wax tags are to be validly compared with trap catch abundance index values. We used a zero-inflated negative binomial (ZINB) model to calibrate the chew card and wax tag abundance indices with the trap catch abundance index. The ZINB model allowed us to account for structural zeros (i.e. possums were not present and therefore cannot be detected) and non-structural zeros (i.e. possums were present, but not detected by a particular method). The relationship between possum abundance index estimates from chew cards and leg-hold traps was better calibrated than that between wax tags and leg-hold traps, with the latter particularly poor in non-forest habitat. We therefore conclude that chew cards, rather than wax tags, should replace leg-hold traps for monitoring possum occupancy rates and relative abundances in the BMRS. Our ZINB model can be used to correct chew card abundance index values for valid comparison with previous trap catch abundance index values estimated using the BMRS design.

Keywords: Bayesian; biological invasions; cross-validation; goodness-of-fit; invasive species; monitoring; New Zealand; occupancy–abundance model; pest; predictive accuracy; public conservation land; relative abundance; zero inflation

Introduction

Introduced brushtail possums (*Trichosurus vulpecula*) are widespread in the North, South and Stewart islands of New Zealand (Cowan 2005), where they are a significant threat to some native species (e.g. Innes et al. 2010; Gormley et al. 2012; Byrom et al. 2016) and to agriculture (Nugent 2011). The Department of Conservation (DOC) manages 8.5 million ha of public conservation land (PCL) in New Zealand and together with other agencies such as TBfree New Zealand expends significant resources controlling possums on PCL (DOC 2017; Forsyth et al. (2018).

DOC's mandate to conserve the natural heritage of New Zealand requires it to know where natural heritage outcomes are being achieved and how management

interventions such as possum control can be used to improve poor outcomes. The desired outcome of conserving natural heritage has been defined as maintaining ecological integrity (Lee et al. 2005). Reporting on progress towards achieving this outcome requires a monitoring program to provide unbiased and repeatable ecological-integrity indicator estimates for all PCL. The Biodiversity Monitoring and Reporting System (BMRS) was implemented to enable DOC to report on three natural heritage priority indicators on PCL: indigenous dominance, species occupancy, and ecosystem representation (Allen et al. 2013). Building on the network of carbon-monitoring plots established in New Zealand during the early 2000s (Payton et al. 2004), the BMRS involves monitoring native and non-native taxa at sites located at the vertices of an 8 km grid superimposed over PCL, including North, South and Stewart

steep terrain, there is interest in using an alternative method to monitor possum abundances in the BMRS. Two alternative methods are wax tags (Thomas et al. 2003; Warburton et al. 2004; National Pest Control Agencies 2015) and chew cards (Sweetapple & Nugent 2011a; Ruffell et al. 2015; National Pest Control Agencies 2015), with abundance indices estimated as the percentage of devices bitten by possums (National Pest Control Agencies 2015).

If possum occupancy and abundance index estimates made using wax tags and chew cards are to be validly compared with abundance index estimates made using leg-hold traps, then correction may be required. Previous assessments of relationships between abundance indices have commonly used simple linear regression (e.g. Sweetapple & Nugent 2011a; Ruffell et al. 2015). That approach ignores sampling variation in individual index estimates and assumes that the relationship between indices is linear. A Bayesian approach (Gelman et al. 2004) can simultaneously account for sampling variation and potential non-linear relationships between indices. Because occupancy is an asymptotically unbiased estimator (i.e. the estimate of occupancy converges to the true value as sample size increases; MacKenzie et al. 2006), provided that similar sampling intensities are used and if the false positive and false negative rates are similarly low then possum occupancy rates estimated from each of wax tags, chew cards and leg-hold traps (together with faecal pellet detection/non-detection data) should not require calibration.

This study had two objectives. First, to confirm that possum occupancy rates estimated from each of wax tags, chew cards and leg-hold traps in the BMRS sampling design (including faecal pellet detection/non-detection data) are similar and hence do not require calibration. Second, to determine if possum abundance indices estimated using wax tags and chew cards are similar to those estimated using leg-hold traps; if they are not then we seek to calibrate these relationships such that the wax tag and/or chew card possum abundance estimates could be corrected for valid comparison with trap catch abundance index estimates. We conducted our study in forest and non-forest habitats, and based on our results we recommend which, if either, of wax tags or chew cards should replace leg-hold traps for monitoring possum occupancy and relative abundance in the BMRS.

Materials and methods

Sample sizes and study areas

Based on the results of a precision analysis (see Appendix S1 in Supplementary Material), we sought to calibrate possum relative abundance indices with data from ≥ 50 independent sites in each of forest and non-forest habitats. We sought sites in the North and South Islands in approximate proportion to the percentage of PCL in the two islands (23% and 77%, respectively). We also attempted to sample as wide a range of possum densities in forest as in non-forest habitats. From data collected previously in the BMRS, we sought to sample forest habitat equally within the three TCI classes: $< 5\%$, $5\text{--}15\%$ and $> 15\%$. Since possum densities are substantially lower in non-forest habitat (Gormley et al. 2015), we sought to sample this habitat equally within the following three classes: $< 1\%$, $1\text{--}5\%$ and $> 5\%$. Leg-hold traps were set on the ground, but where kiwi (*Apteryx* spp.) and weka (*Gallirallus australis*) were present we set traps off the ground ('raised sets'; National Pest Control Agencies 2015). Because the probability of a

leg-hold trap set on the ground capturing a possum could differ from that of a leg-hold trap that is set off the ground, in approximately half of the forest and non-forest sites we used raised leg-hold trap sets and in the remainder we used ground leg-hold trap sets.

We attempted to sample a wide range of possum densities in forest and non-forest habitats by subjectively locating our sites based on habitat types and the recent possum-control history. For example, high possum densities were obtained by sampling forest and non-forest habitats that had not been subject to possum control within the last 5 years. Low possum densities were obtained by sampling forest and non-forest habitats in which possums had recently been controlled. Independence of sites was assumed given a minimum distance of 400 m between the closest monitoring devices of two sites (National Pest Control Agencies 2015).

Field methods

Each of our sites consisted of 12×200 m transects and 4×150 m transects (Fig. 1b). Of the 200 m transects, four were sampled using wax tags (PCRWaxTag®, Pest Control Research Ltd, Christchurch), four using chew cards impregnated with 'possum dough' (Trappers Cyanide Ltd, Amberley) to attract possums, and four using No. 1 double-coil-spring leg-hold traps. All three monitoring devices were placed at 20 m intervals on their respective 200 m transects (i.e. 10 per transect; Fig. 1). The 200 m transects were offset by 30° to avoid devices being set at the same site on subsequent nights, which could alter possum behaviour such that detections did not mimic the BMRS sampling design. To reflect the BMRS monitoring options and to minimise the effects of 'night' on our results, two wax tag and two chew card transects were set on Day 1 and checked and removed on Day 2 (i.e. they sampled possums on Night 1), and the remaining two wax tag and two chew card transects were set on Day 2 and checked and removed on Day 3 (i.e. they sampled Night 2). The four trap transects were set on Day 3 and were checked and removed on Day 4 (i.e. they sampled Night 3).

Wax tags and chew cards were set according to the National Pest Control Agencies (2015) protocol, except that they were set for one night rather than three or seven nights (i.e. they were set as would occur if either device was to be used in the BMRS; Appendix S2). Briefly, all wax tags had a luminescent strip to attract possums; in forest habitat they were set with a white flour : icing sugar mixture (5 : 1) lure, and in non-forest habitat they were set on a Corflute backing board without the flour : icing sugar lure. Chew cards were folded in half and in forest habitat were nailed to a tree so that the part of the chew card impregnated with possum dough was 30 ± 5 cm above the ground. In non-forest habitat, the folded chew card was secured on a spike (similarly for wax tags) so that the possum dough was 30 ± 5 cm above the ground. For full details of wax tag and chew card field methods, see Appendix S2.

Leg-hold traps were set at all sites, with raised sets used for all traps at approximately half of the sites and ground sets used for all traps at the other sites. Both types of sets followed the National Pest Control Agencies (2015) protocol, except that the traps were set for one night (as occurs in the BMRS; Gormley et al. 2015; Appendix S2).

Finally, we recorded the detection/non-detection of possum faecal pellets in each of 30 circular plots of 1 m radius located at 5 m intervals along four 150 m faecal pellet transects (Fig. 1; Gormley et al. 2015). These faecal pellet counts followed the BMRS protocol and are used to estimate possum occupancy rates (Gormley et al. 2015).

A site was discarded if a minimum of 75% of any of the four monitoring devices (i.e. wax tags, chew cards, leg-hold traps or faecal pellet plots) could not be sampled in the same habitat type (either forest or non-forest). All possum monitoring devices were set on fine nights, which were defined as < 1 mm of rain in the first 4 h of darkness (commencing 30 min after sunset).

Wax tag and chew card bite mark assessment and estimates of abundance

Bite marks on wax tags and chew cards were recorded for each device as it was collected in the field, following Sweetapple and Nugent (2011a, b). There were six possible outcomes for a wax tag or chew card: bitten at least once by a possum (P); no possum bite marks but identifiable non-target bite mark (NT); no possum bite marks, but unknown bite marks (U); no bite marks (0); lost or otherwise beyond interpretation (L); and not set (NS).

False positives could occur when either bite marks on a wax tag or chew card are attributed to possums when in truth they are not from possums. Conversely, false negatives could occur when bite marks on a wax tag or chew card are not attributed to possums when in truth they are from possums, or when no bite marks are detected but possums are present. We attempted to assess and minimise the misidentification rates by having all wax tags and chew cards scored by field staff as P, NT and U, and a random 10% of wax tags and chew cards without bite marks checked by experts. Furthermore, all wax tags and chew cards scored as zero were double-checked by another field worker. The two experts used hand lenses and stereoscopes in the laboratory to confirm or re-determine the field assessment, with all possum bites assigned a confidence score of either high (no doubt), medium (atypical marks but no plausible alternative) or low (faint or atypical marks that could also have been made by a non-possum agent). If an expert's assessment for a wax tag or chew card differed from that recorded by field staff, then the expert's assessment was used in analyses.

A bite-mark index was subsequently estimated for each transect from the number of devices assessed as having possum bite marks with high or medium confidence divided by the number of interpretable devices checked (i.e. subtracting devices that were L or NS) and multiplied by 100% (National Pest Control Agencies 2015).

Trap catch index of abundance

A trap catch index of possum abundance along each of the four trap transects was estimated following the national protocol (National Pest Control Agencies 2015). Briefly, the total number of possums caught was divided by the number of trap nights minus 0.5 trap nights for each non-target capture and sprung-but-empty trap (Nelson & Clark 1973); that number was multiplied by 100% (National Pest Control Agencies 2015). We assumed that the impact of the loss of a trap is consistent across transects, although it could differ depending on its position relative to the centre of the site.

Statistical analyses

Occupancy models

Occupancy rate was modelled using a Bayesian state-space formulation to account for imperfect detection of possums and overdispersion (Royle & Kéry 2007; Gormley et al. 2015). This model had two components: the true occupancy

and the observation process. Let ψ_h be the probability a site of type h being occupied (where h is the habitat, either forest or non-forest). The true occupancy status Z of site i was modelled by a random variable from a Bernoulli distribution with probability ψ_h :

$$Z_i \sim \text{Bern}(\psi_h) \quad (1)$$

Letting $p_{j,k}$ be the probability that a possum would be detected on transect j by method k , given that the site was occupied, sites were surveyed using four methods (leg-hold traps, chew cards, wax tags, and faecal pellets), each with four transects (i.e. four replicates in space). Thus the observation process ($X_{i,j,k}$) for a site was described as a random variable from a Bernoulli distribution:

$$X_{i,j,k} \sim \text{Bern}(Z_i \times p_{j,k}) \quad (2)$$

The overall occupancy rate (ψ) was estimated as a weighted mean of forest and non-forest occupancy rates. The model allowed for potential differences in detection rates between observation methods, as well as differences in underlying forest, non-forest and overall occupancy rates.

The occupancy models were fitted using Markov chain Monte Carlo (MCMC) sampling in JAGS (Plummer 2003) called from R (R Development Core Team 2015) with R2jags (Su & Yajima 2012). As priors, we used a naïve Beta(1,1) distribution for the detection probabilities $p_{j,k}$. Three chains were used, each with 40 000 iterations, a burn-in of 10 000, and a thinning rate of 3 to ensure convergence (defined as having all Gelman and Rubin's convergence diagnostic potential scale reduction factors [PSRF] < 1.05 [Gelman et al. 2004]). Hence, 30 000 samples were used for inference. Occupancy rates were deemed to be significantly different if the 95% credible interval (CI) for the posterior distribution of their difference did not include 0.

Trap catch, wax tag and chew card abundance index estimates

The trap catch, wax tag and chew card abundance indices were first modelled using a joint occupancy–abundance model (Gormley et al. 2015). The occupancy rate component of the joint model is described above (equations 1 and 2). The abundance (i.e. trap catch, wax tag and chew card abundance indices) component of the model was similar to that used by Gormley et al. (2015), except that we used a negative binomial (NB) model rather than a Poisson model to better account for overdispersion of the non-zero index data. To account for the possibility that possums were detected at a site by at least one of the four methods used to estimate occupancy (i.e. $Z_i = 1$) but were not detected with wax tags, chew cards and/or leg-hold traps, we used a standard negative binomial distribution rather than a truncated negative binomial distribution.

A potential issue with using a negative binomial or Poisson model is that the number of successes drawn from the resulting distribution could exceed the number of detection devices (or conversely the number of failures plus successes could exceed the number of detection devices). Also, the numbers of leg-hold traps, wax tags and chew cards available to detect possums at a site were not always 40. In particular, for each trap that was sprung empty or captured a non-target species, 0.5 was subtracted from the number of available traps, giving a non-integer number (National Pest Control Agencies 2015). If a chew card or wax tag was missing when these devices were checked then we subtracted 1 from the number of those

devices available (National Pest Control Agencies 2015). To address these two issues we used an offset to account for the number of devices available to possums at each site. Following National Pest Control Agencies (2015), all three abundance indices are presented as percentages.

The zero-inflation estimate was derived from the occupancy model (equations 1 and 2). The number of detections (captured possums in traps or wax tags/chew cards with possum bite marks) per site ($Y_{i,k}$), conditional on possums being present (Z_i), was:

$$Y_{i,k} | Z_i=1 \sim NB(v_{i,k}, r_{i,k}) \quad (3)$$

$$v_{i,k} = \frac{r_{i,k}}{r_{i,k} + \mu_{i,k}} \quad (4)$$

$$\log_e(\mu_{i,k}) = \eta_{h,k} + \log_e(\text{Effort}_{i,k}) \quad (5)$$

where $v_{i,k}$ is the scale parameter and $r_{i,k}$ is the shape parameter, $\mu_{i,k}$ is the mean abundance index using method k at site i , $\eta_{h,k}$ is the mean number of expected detections for one device of method k in habitat h (on the \log_e scale), all conditional on possums being present. The offset for effort ($\text{Effort}_{i,k}$) scales the mean number of observations by the number of devices available to possums.

The abundance index models were fitted using MCMC sampling in JAGS (Plummer 2003) called from R (R Development Core Team 2015) with R2jags (Su & Yajima 2012). As priors, we used gamma(0.001, 0.001) distributions for $r_{i,k}$, Beta(1,1) for ψ_h and $p_{j,k}$, and Normal $N(0, 1000)$ distributions for $\eta_{h,k}$. We used three chains with 40 000 iterations, a burn-in of 10 000, and a thinning rate of 3 to ensure convergence (defined as PSRF < 1.05), leaving 30 000 samples for inference. As for occupancy rates, abundance index estimates were deemed significantly different if the 95% CIs for the posterior distributions of their differences did not include 0.

Calibrating the wax tag and chew card abundance indices to the trap catch abundance index

The possum abundance indices were next modelled such that wax tags and chew card abundance estimates at a site could be corrected for valid comparison with trap catch abundance estimates at that site. Zero-inflated negative binomial (ZINB) models were used to correct the wax tag and chew card indices because they explicitly accounted for the large number of zeros in the abundance index. There were two parts to our ZINB model.

The first part of the model estimated the probability of zero-inflation. Zero-inflation relates to structural and non-structural zeros. Structural zeros occur when possums are not present and therefore cannot be detected with any of the detection methods. A site with a structural zero means that all abundance indices for that site should also be zero. Non-structural zeros occur when possums are present but not detected by a detection method, and must be accounted for in the calibration. Let Q_i be a random variable with a Bernoulli distribution that indicates structural zeros (i.e. 0 if structural zero and 1 otherwise) at a site:

$$Q_i \sim \text{Bern}(q_{h,k}), \quad (6)$$

where $q_{h,k}$ is the probability of possums being present dependent on the habitat type h and whether or not they were detected using method k or faecal pellets. If possums were detected using method k or faecal pellets then the site could not be a structural zero and $q_{h,k}$ was set to 1, with Q_i then equivalent to Z_i (i.e. equation 1). If possums were not detected at a site using method k or faecal pellets then the probability of possums being present, given that they were not detected, was estimated using the occupancy rate given non-detection using data from all four detection methods.

The second part of the model estimated the number of times the corrected CCI (CCI_{TCl}) or WTI (WTI_{TCl}) would detect possums, given the number of times the uncorrected method (WTI or CCI) detected possums, given that possums were present ($Q_i=1$). Let R_i be a random variable, conditional on possums being present ($Q_i=1$), for the expected number of possums that would be observed using leg-hold traps at site i . R_i has a negative binomial distribution (conditional on possums being present) and represents a corrected abundance index, and is either:

$$R_i | Q_i=1 \sim NB(\lambda_{i,\text{CCI}_{\text{TCl}}}, t_{i,\text{CCI}_{\text{TCl}}}) \quad (7)$$

$$\lambda_{i,\text{CCI}_{\text{TCl}}} = \frac{t_{i,\text{CCI}_{\text{TCl}}}}{t_{i,\text{CCI}_{\text{TCl}}} + \theta_{i,\text{CCI}_{\text{TCl}}}} \quad (8)$$

or

$$R_i | Q_i=1 \sim NB(\lambda_{i,\text{WTI}_{\text{TCl}}}, t_{i,\text{WTI}_{\text{TCl}}}) \quad (9)$$

$$\lambda_{i,\text{WTI}_{\text{TCl}}} = \frac{t_{i,\text{WTI}_{\text{TCl}}}}{t_{i,\text{WTI}_{\text{TCl}}} + \theta_{i,\text{WTI}_{\text{TCl}}}} \quad (10)$$

where $\lambda_{i,\text{CCI}_{\text{TCl}}}$ and $\lambda_{i,\text{WTI}_{\text{TCl}}}$ are the mean corrected CCI and WTI values, respectively, and $t_{i,\text{CCI}_{\text{TCl}}}$ and $t_{i,\text{WTI}_{\text{TCl}}}$ are the shape parameters for the corrected CCI and WTI, respectively, both conditional on possums being present. Note that in this formulation of the negative binomial the shape parameter can be a non-integer (i.e. Pólya distribution; Hilbe 2011). $\theta_{i,\text{CCI}_{\text{TCl}}}$ and $\theta_{i,\text{WTI}_{\text{TCl}}}$ are given by:

$$\log_e(\theta_{i,\text{CCI}_{\text{TCl}}}) = \beta_{0,\text{CCI}} + \beta_{1,\text{CCI}} \times \text{CCI}_i + \beta_{2,\text{CCI}} \times I_{\text{Non-forest}} + \beta_{3,\text{CCI}} \times \text{CCI}_i \times I_{\text{Non-forest}} + \log_e(\text{Effort}_{i,\text{CCI}}) \quad (11)$$

and

$$\log_e(\theta_{i,\text{WTI}_{\text{TCl}}}) = \beta_{0,\text{WTI}} + \beta_{1,\text{WTI}} \times \text{WTI}_i + \beta_{2,\text{WTI}} \times I_{\text{Non-forest}} + \beta_{3,\text{WTI}} \times \text{WTI}_i \times I_{\text{Non-forest}} + \log_e(\text{Effort}_{i,\text{WTI}}) \quad (12)$$

respectively, where $\beta_{0,\text{CCI}}$ and $\beta_{0,\text{WTI}}$ are the intercepts, $\beta_{1,\text{CCI}}$ and $\beta_{1,\text{WTI}}$ are the coefficients for correcting the WTI and CCI to the TCI, $\beta_{2,\text{CCI}}$ and $\beta_{2,\text{WTI}}$ are the coefficients for the effects of the non-forest habitat, $\beta_{3,\text{CCI}}$ and $\beta_{3,\text{WTI}}$ are the coefficients for the effects of correcting the WTI and CCI to the TCI for non-forest habitat, and $I_{\text{Non-forest}}$ is an indicator variable that was set to 1 if the site was in non-forest habitat and 0 if it was in forest habitat.

The calibration models were fitted using MCMC sampling in JAGS (Plummer 2003) called from R (R Development Core

Team 2015) with R2jags (Su & Yajima 2012). As priors, we used $\text{gamma}(0.001, 0.001)$ distributions for $r_{i,k}$, $\text{Beta}(1, 1)$ for ψ_h and $p_{j,k}$, and $\text{Normal}N(0, 1000)$ distributions for parameters $\beta_{0,CCI}$, $\beta_{1,CCI}$, $\beta_{2,CCI}$, $\beta_{3,CCI}$, $\beta_{0,WTI}$, $\beta_{1,WTI}$, $\beta_{2,WTI}$ and $\beta_{3,WTI}$. We again used three chains with 40 000 iterations, a burn-in of 10 000, and a thinning rate of 3 to ensure convergence (i.e. $\text{PSRF} < 1.05$), leaving 30 000 samples for inference.

The goodness-of-fit of the two abundance index calibration models to the data were assessed by comparing the discrepancy of the posterior predictive distribution with the observed data. The posterior predictive distribution consisted of 30 000 replicated data sets drawn from the posterior distribution conditioned on the model parameters. The proportion of times the test statistic for the replicated data was greater than or equal to the value for the observed data is the Bayesian P-value, with values close to 0 or 1 indicating lack of fit (Gelman et al. 2004).

We also assessed the predictive accuracy of the two calibration models using k-fold cross-validation (see also Falcy et al. 2016). Cross-validation involves randomly selecting a proportion of the data to fit the model (the training data) and then testing predictive performance using the data excluded

from the fitting process (the ‘out-of-bag’ validation data). We constructed training and validation datasets using a fold size of $k = 6$ (i.e. $n = 18$ sites for the validation data). The models were fitted to the remaining training data ($n = 90$ sites) and the parameters from those models were used to predict the expected values for the validation datasets. In a Bayesian setting, prediction of the validation dataset was achieved by setting the observed abundance index values for the validation data to NA (missing) and then allowing the models to predict the values for the missing observations, given the covariate and other values, using full Bayesian inference. We repeated this for all possible folds of size $n = 90$ sites and compared the predicted expected values to the observed data by calculating the root mean square error (RMSE) as a measure of predictive accuracy.

Data and code availability

The data and code used in our analyses are available from the corresponding author upon request.

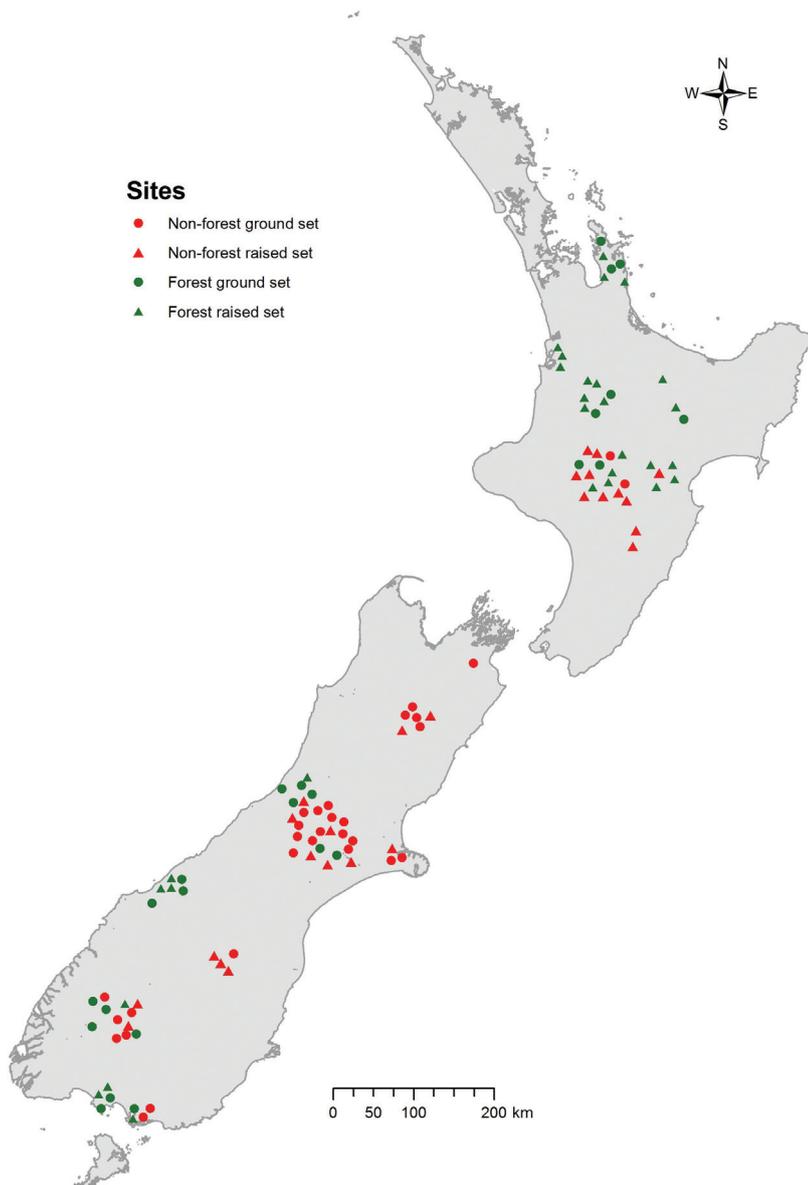


Figure 2. The 108 sites ($n_{\text{forest}} = 54$, $n_{\text{non-forest}} = 54$) at which brushtail possum monitoring was conducted using wax tag, chew card and trap catch index methods during February–May 2015. For clarity, sites are offset so that they do not overlap.

Results

Data overview

Data were obtained during February–May 2015 from a total of 108 independent sites: 54 sites in forest and 54 sites in non-forest (Fig. 2). Of a total of 12 960 devices that could have been set at the 108 sites, only 34 devices were not set (nearly all of these due to the presence of impassable or dangerous terrain) and two devices were lost. Hence, our data consisted of outcomes for a total of 12 924 devices (4309 traps, 4309 wax tags and 4306 chew cards). Out of a possible 12 960 faecal pellet plots, 30 plots (one complete transect) were not sampled.

The trap catch index of possum abundance at sites differed from what we sought for non-forest habitat (Appendix S3, S4). In particular, there were fewer high-TCI sites and more medium-TCI sites in non-forest habitat.

Misidentification rates for wax tags and chew cards

Experts checked 1190 wax tags and 1085 chew cards (Appendix S5). Of these, 30.25% of wax tags and 36.96% of chew cards were assessed as having possum bite-marks. The false negative rates for wax tags and chew cards (i.e. devices assessed in the field as not being bitten but subsequently assessed as possum) were 1.11% and 3.20%, respectively. The percentage identified as having possum marks in the field, but determined as not having possum marks by experts (i.e. false positives), was higher for chew cards (6.85%) than for wax tags (1.71%). A higher percentage of wax tags (20.67%) was assessed in the field as having non-target bite marks compared with chew cards (14.19%), but of these, chew cards had a slightly higher percentage of false negatives than wax tags (3.25% and 1.22%, respectively). The percentages of wax tags and chew cards assessed in the field as ‘unknown’ were similar (4.54% and 6.08%, respectively), but a much higher percentage of chew cards (60.61%) than wax tags (12.96%) were subsequently determined by experts as being from possums.

Detecting possums

Possums were detected at 86 of the 108 sites (44 of 54 forest sites and 43 of 54 non-forest sites; Appendix S6). Possums were not solely detected by traps at any forest or non-forest site (Table S6a). Wax tags and chew cards each detected possums at one forest site in which possums were not detected by traps and/or faecal pellets. Overall, chew cards detected possums at more sites (70) than wax tags (61) or traps (59) (Appendix S6). Faecal pellets were not detected at all sites where possums were detected, and vice versa. Occupancy estimates would be biased high if faecal pellets remained from what we sought for possums. In forest habitat, however, no faecal pellets were detected without at least one other method also detecting possums, indicating that all forest sites were truly occupied.

Occupancy rates in forest and non-forest habitats

The possum occupancy rates estimated using traps, wax tags and chew cards (each together with faecal pellets) were similar (i.e. 95% credible intervals for each method easily included the median estimate from each of the other methods) in forest habitat, non-forest habitat and overall (Fig. 3). The possum occupancy rate estimated from leg-hold traps in non-forest habitat was lower than that estimated from chew cards and from wax tags, but still within the 95% credible intervals. When all four detection methods were combined, the occupancy estimates were marginally higher and consistent across forest habitat, non-forest habitat and overall. However, although the combined estimates were higher, they were still well within the 95% credible intervals for the estimates derived from each of wax tags, chew cards, leg-hold traps and faecal pellets.

Abundance index estimates in forest and non-forest habitats

Median possum abundance index values from leg-hold traps, chew cards and wax tags were always higher in forest than non-forest habitat (i.e. 95% CIs for their difference excluded zero; Table 1; Fig. 4). There were also some large differences

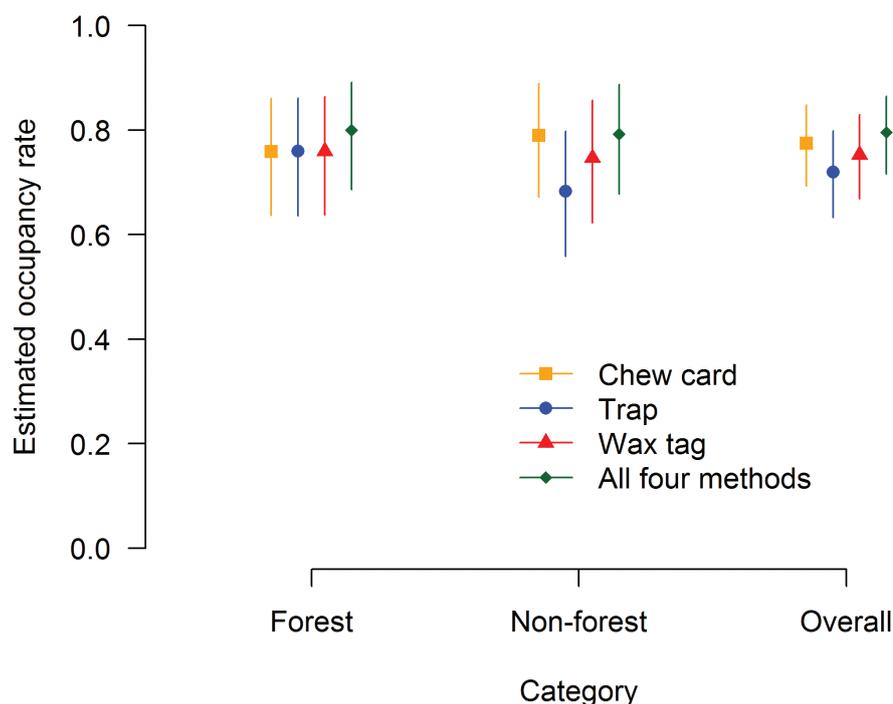


Figure 3. Possum occupancy rates estimated from each of chew cards, leg-hold traps and wax tags used separately with faecal pellets (traps and pellets; chew cards and pellets; wax tags and pellets) and when all three methods and faecal pellets were used together. Median values and their 95% credible intervals (from posterior distributions) are shown for sites in forest and non-forest habitats, and overall.

Table 1. Median and maximum values for the trap catch, chew card and wax tag indices of possum abundance in forest and non-forest habitats. All three indices are expressed as percentages, with the maximum possible value being 100.0%.

Index	Habitat			
	Forest		Non-forest	
	Median	Maximum	Median	Maximum
Trap catch	5.2	35.1	0.0	32.0
Chew card	11.2	42.5	2.5	40.0
Wax tag	10.0	62.5	0.0	22.5

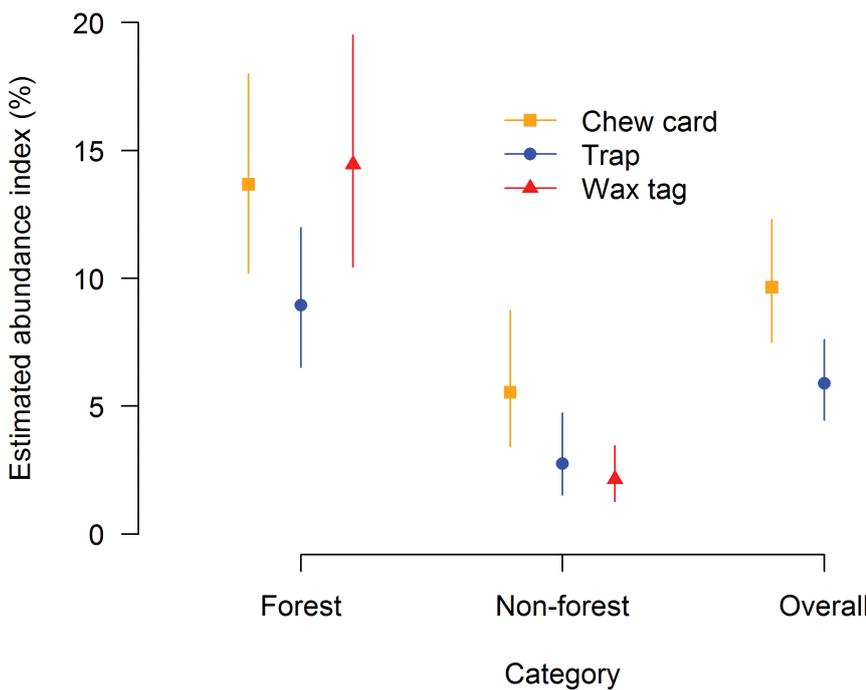


Figure 4. Possum abundance index values (%) in forest and non-forest habitats, and overall. Medians and 95% credible intervals obtained from posterior distributions are shown.

in abundance index estimates within habitats. Median chew card and wax tag abundance index values were similar in forest habitat, but different in non-forest habitat. The median trap catch abundance index value was much lower than the median chew card and wax tag abundance index values in forest habitat, but was similar to the median chew card and wax tag abundance index values in non-forest habitat. Median overall (i.e. forest and non-forest sites pooled) wax tag and chew card abundance index values were similar, but the overall trap catch abundance index value was significantly lower (Fig. 4). These results indicate that calibration is required for valid comparison of possum abundance index estimates made using either wax tags or chew cards with those made with leg-hold traps.

Calibrating the chew card and wax tag abundance indices to the trap-catch abundance index

Inspection of the scatter plots revealed that much data were around zero, and an ordinary linear regression line highlighted the slope and the variability of the data, with most of the points some distance from the regression line (Fig. S4). The Bayesian P-values of 0.544 and 0.543 for the models calibrating the wax tag abundance index and the chew card abundance index to the trap catch abundance index, respectively, indicate that these two models provided reasonable approximations of the observed data.

Both abundance index calibration models performed better in forest than non-forest habitat (Fig. 5). The calibration models indicated weakly non-linear relationships between each of the wax tag and chew card abundance indices and the trap catch abundance index in forest habitat, with uncertainty increasing at higher abundance index values. The relationships between each of the wax tag and chew card abundance indices and the trap catch abundance index in non-forest habitat were more strongly non-linear. The relationship between the wax tag abundance index and the trap catch abundance index was particularly poor for sites in non-forest habitat, with the strong non-linearity leading to saturation (i.e. a maximum abundance index value of 100%) at relatively low trap catch abundance index values.

The chew card abundance index calibration model estimated that at sites without possum detections from either chew cards or pellets, zero-inflation was 74% (95% CIs: 41–94%) at sites in forest habitat and 69% (26–94%) at sites in non-forest habitat (Table 2). In other words, possums were estimated to be present but not detected by chew cards or pellets at 26% and 31% of forest and non-forest sites, respectively. The wax tag abundance index calibration model estimated that at sites without possum detections from either wax tags or pellets, zero-inflation was 93% (76–100%) at forest sites and 96% (87–100%) at non-forest sites (Table 2).

The predictive accuracy of the two calibration models

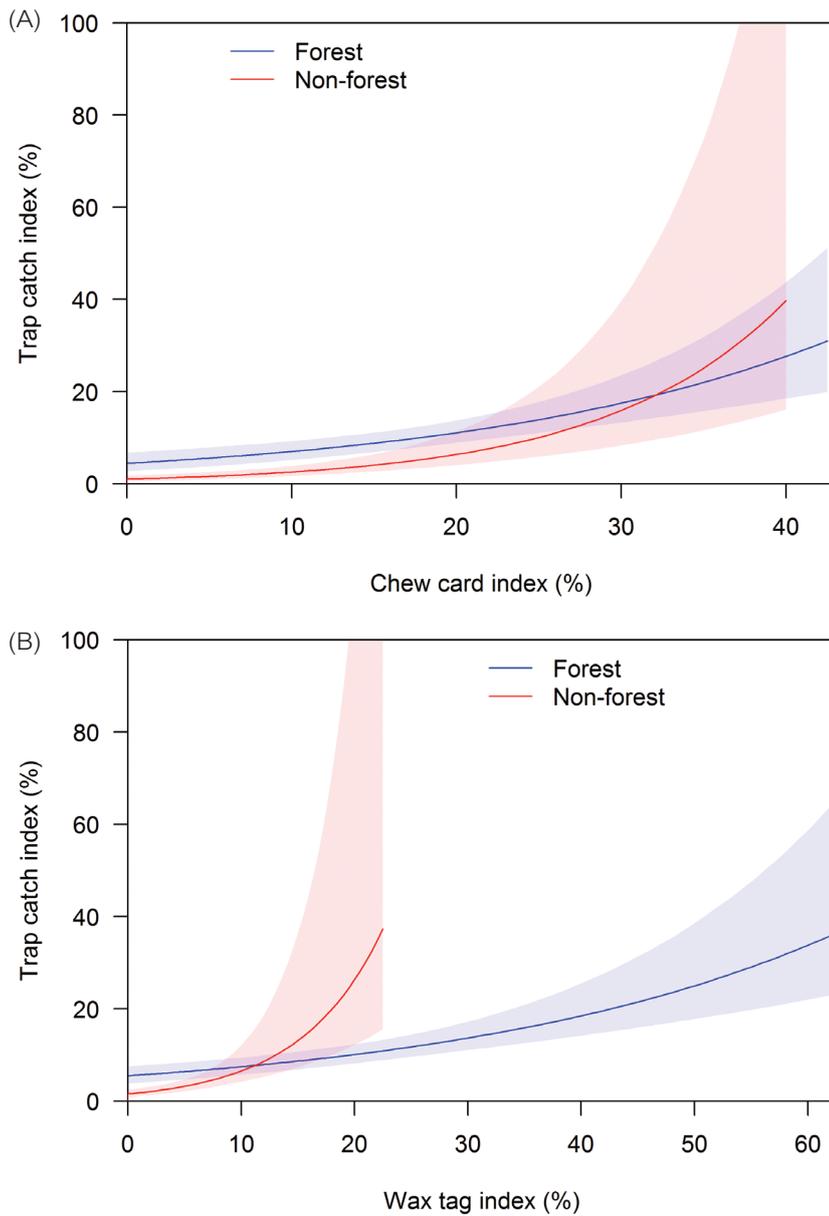


Figure 5. Calibrated relationships between each of the chew card abundance index (A) and the wax tag abundance index (B) and the trap catch abundance index in forest and non-forest habitats. Shaded regions represent 95% credible intervals. The highest observed chew card abundance index values were 42.5% (forest) and 40.0% (non-forest). The highest observed wax tag index abundance values were 62.5% (forest) and 22.5% (non-forest).

Table 2. Posterior density summaries (mean, SD, 2.5th, 50th [median] and 97.5th percentiles) for parameters in the models calibrating the wax tag abundance index (WTI) and chew card abundance index (CCI) to the trap catch abundance index (TCI). $q_{h,ND,k}$ is the probability of possums being present at a site in habitat h but not detected using method k and faecal pellets. Beta values are on the ln scale.

Parameter	Mean	SD	2.5%	Median	97.5%
WTI to WTI_{TCI}					
$q_{F,ND,WTI}$	0.069	0.066	0.002	0.049	0.243
$q_{NF,ND,WTI}$	0.036	0.036	0.001	0.025	0.131
$\beta_{0,WTI}$	-2.711	0.146	-3.010	-2.707	-2.436
$\beta_{1,WTI}$	2.671	0.476	1.793	2.654	3.676
$\beta_{2,WTI}$	-0.659	0.264	-1.19	-0.659	-0.141
$\beta_{3,WTI}$	7.732	2.014	3.853	7.695	11.813
CCI to CCI_{TCI}					
$q_{F,ND,CCI}$	0.262	0.135	0.060	0.241	0.586
$q_{NF,ND,CCI}$	0.314	0.179	0.064	0.281	0.745
$\beta_{0,CCI}$	-3.087	0.226	-3.554	-3.076	-2.661
$\beta_{1,CCI}$	4.458	0.956	2.662	4.412	6.496
$\beta_{2,CCI}$	-1.303	0.395	-2.106	-1.298	-0.528
$\beta_{3,CCI}$	3.736	2.057	-0.121	3.662	8.007

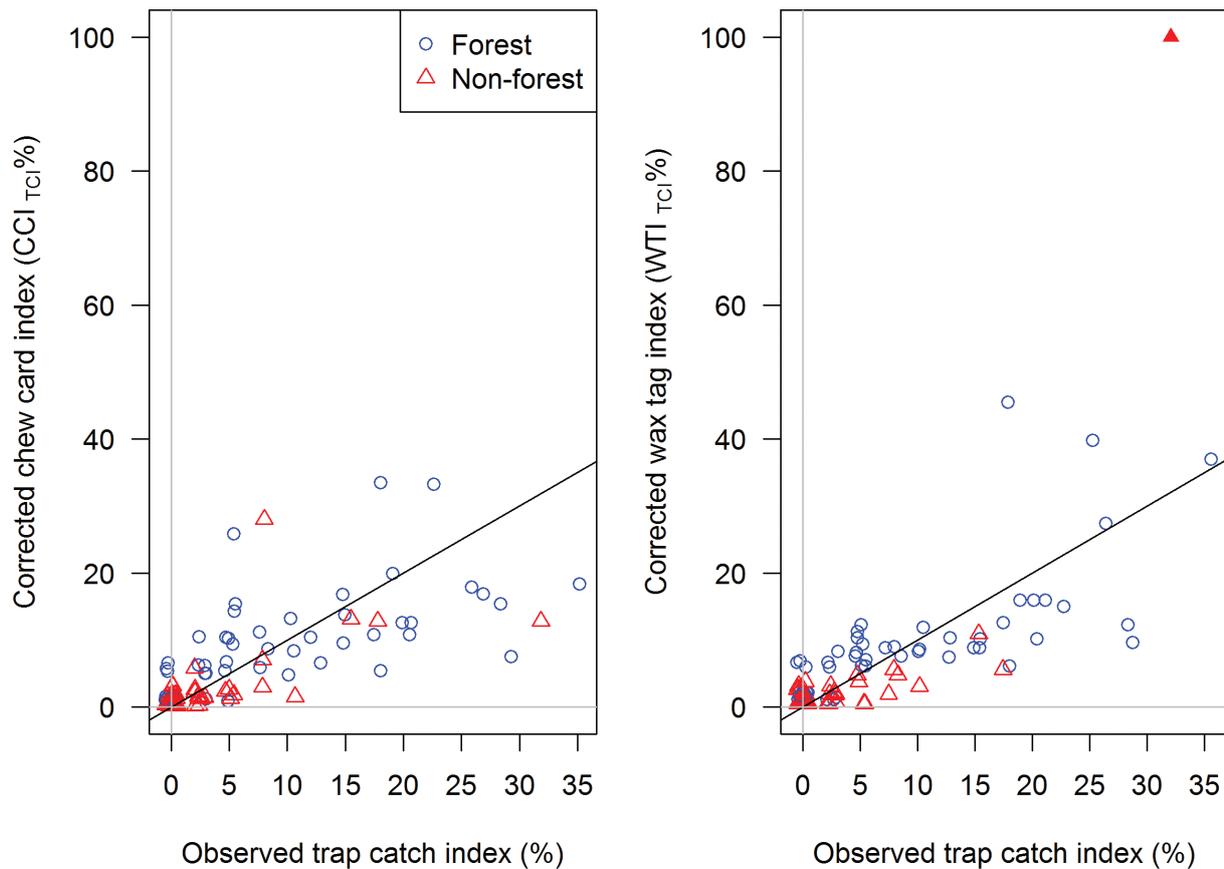


Figure 6. Corrected chew card (left) and wax tag (right) abundance index values predicted by the calibration models fitted to the test (out-of-bag) data from each fold, following 6-fold cross-validation. The diagonal line represents the 1:1 relationship. The filled triangle in the right figure indicates a predicted corrected wax tag value of the maximum 100%.

varied greatly under the 6-fold cross-validation (Fig. 6). The cross-validation estimated the RMSE for the chew card abundance index calibration (i.e. CCI to CCI_{TCI}) to be 6.38, much smaller than the estimated RMSE for the wax tag abundance index calibration (i.e. WTI to WTI_{TCI}) of 13.67. The accuracy of predictions differed between sites in forest and non-forest habitats, particularly for the wax tag abundance index calibration. The estimated RMSE for the chew card abundance index calibration in forest habitat was 7.42 (c.f. a mean TCI value in this habitat of 9.21%) and in non-forest habitat was 5.19 (c.f. a mean TCI value in this habitat of 2.57%). The estimated RMSE for the wax tag abundance index calibration in forest habitat was 6.85 and in non-forest habitat was 17.94. Highlighting the particularly poor predictive ability of the wax tag abundance index calibration model in non-forest habitat was one site having a predicted value of the maximum 100%, a large distance from the observed trap catch abundance index value of 32% (Fig. 6). These results indicate that the calibration of the chew card abundance index to the trap catch abundance index is more accurate than that from the wax tag abundance index to the trap catch abundance index.

Discussion

Using data collected at 54 sites in each of forest and non-forest habitat, our analyses confirmed that there is no need to correct estimates of possum occupancy rate made using any one of

leg-hold traps, chew cards or wax tags in the Department of Conservation’s Biodiversity and Monitoring Reporting System. This is because occupancy is an asymptotically unbiased estimator (MacKenzie et al. 2006), and sampling effort was identical for leg-hold traps, wax tags and chew cards. In contrast, correction is necessary if possum abundance index values from either chew cards or wax tags are to be validly compared with trap catch abundance index values from sites that have been sampled using the latter method.

Correcting wax tag or chew card abundance index values for valid comparison with trap catch abundance index values is more complicated than multiplying by a constant correction factor (e.g. Sweetapple & Nugent 2011; Appendix S1) for three main reasons. First, possums were sometimes detected at a site by one or two but not all three abundance monitoring methods. In these cases one or two of the abundance index values were 0% and this would not change if it was multiplied by a constant correction factor. We accounted for both structural zeros (i.e. no possums were detected by any detection method and therefore all abundance indices should be zero) and non-structural zeros (i.e. possums were not detected using chew cards and/or wax tags, but were detected by another detection method and therefore should have non-zero abundance index values) by estimating the corrected abundance index values when the chew card or wax tag abundance index value was zero using a zero-inflated negative binomial model. Second, the relationships between abundance indices sometimes saturated at values much less than the maximum 100%. This

meant that corrected abundance index values could be 100% at relatively low trap catch abundance index values (e.g. the corrected wax tag abundance index in non-forest habitat). Third, uncertainty in the relationships between abundance indices increased with increasing possum abundance, particularly in non-forest habitat. We accounted for this uncertainty in the calibration and prediction models by using a Bayesian framework. A consequence of these three issues, and the inherent variability of the possum abundance index data, is that the precision of the corrected index values will always be less than the uncorrected index value (i.e. larger 95% CIs relative to the mean).

The modelled relationships between each of the wax tag and chew card abundance indices and the trap catch abundance index were much poorer in non-forest than forest habitat. The predictive accuracy of both calibration models was therefore also poorer in non-forest than forest habitat. In particular, the relationship saturated too rapidly for sensible conversion of wax tag abundance index values to trap catch abundance index values in non-forest habitat when wax tag abundance index values were $\geq 12\%$. Although correcting chew card abundance index values $\geq 30\%$ in non-forest habitat is similarly problematic, such high values would be rarely observed in the BMRS (Gormley et al. 2015). There are several potential reasons why the abundance index relationships were poorer in non-forest than forest habitat. First, although we attempted to sample non-forest sites with a wide range of possum abundances we obtained fewer sites than desired in the medium- and high-density abundance classes, and more sites in the low-density class. Hence, only a few sites determined the form of the relationship at high possum abundances in non-forest habitat. Second, possums may interact with wax tags and chew cards differently in non-forest than forest habitat, perhaps because home range sizes are larger and food resources more patchily distributed in the former compared to the latter (Glen et al. 2012; Rouco et al. 2013). Third, melting of wax tags in the high daytime temperatures that sometimes occurred in non-forest sites may have obscured some possum bites, leading to the particularly poor relationship between wax tag abundance index values and trap catch abundance index values in this habitat.

The parameter values in Table 2 can be used to correct wax tag and chew card abundance index values for valid comparison with trap catch abundance index values in the BMRS. Another approach is to use our data and code (available from the corresponding author) in JAGS to predict the corrected abundance index value from the observed abundance index value, with the corrected abundance index value specified as a missing value: this is the method we used in our cross-validation procedure and is favoured because it fully preserves the uncertainty in the modelled relationships. We emphasise that only interpolation (i.e. estimates from within the range of observed values) should be used for any correction of wax tag or chew card abundance index values to trap catch abundance index values, with forest and non-forest habitats corrected separately. The maximum chew card abundance index values that can be corrected for valid comparison with trap catch abundance index values were 42.5% in forest and 40.0% in non-forest, and the maximum wax tag abundance index values were 62.5% in forest and 22.5% in non-forest (but see the caveats above for non-forest habitat).

There is a much greater potential for false negative and false positive results for wax tags and chew cards than for traps. Hence, all wax tags and chew cards assessed in the field as being bitten or potentially bitten were checked by

experts; their assessment was considered ‘truth’ and used in our analyses. Misidentification rates were low for wax tags and chew cards identified in the field as either bitten by possums or bitten by non-targets. However, for devices classified as ‘unknown’ in the field, a much greater percentage of chew cards than wax tags was subsequently determined by experts to be possum. Also, the percentage of randomly collected chew cards that were assessed as being not bitten in the field and later classified as bitten by possums (i.e. false negatives) was three times greater for chew cards (3.20%) than for wax tags (1.11%). These differences may be due to possums sometimes leaving a light single impression that is more readily detected using a lens on chew cards than on wax tags; the latter were more often compromised by random scrapes, smears and melt issues, which can obscure possum marks. Also, since wax tags are not flat, there is a much greater range of angle of approach and resulting shape of marks, making light bite interpretation by field staff and experts more difficult than for chew cards. Chew cards had a higher probability of detecting possums on a transect, and hence at sites, than did wax tags or leg-hold traps. The higher detection probability of chew cards compared to wax tags may be due to the former having a larger surface area, increasing their visibility and making them easier to bite. The possum dough impregnated into the chew cards may also be a better olfactory lure than the wax tag. For a possum to be detected by a leg-hold trap, the trap must be available to be sprung (i.e. set correctly and not sprung) and the possum must step on the trap plate (Warburton 2000). Possums can eat the flour : icing sugar lure placed behind the trap without being caught (BW, unpubl. data).

Our occupancy and abundance models assumed that false positives did not occur. This assumption was valid because the devices assessed in the field were subsequently assessed by experts. Our expert assessments may sometimes have been incorrect, and this could be assessed in blind trials (e.g. by using devices monitored with remote cameras to confirm the source of bite marks, with those devices being subsequently assessed by experts). If false positives prove to be important, then this source of uncertainty could be accounted for in occupancy and abundance models (see Royle & Link 2006).

Conclusion

Chew cards and wax tags would have major logistical and financial advantages over leg-hold traps in the Department of Conservation’s Biodiversity and Monitoring Reporting System. Using the BMRS field sampling design, possum occupancy estimates obtained from chew cards and wax tags do not require correction for valid comparison with occupancy estimates from leg-hold traps. In contrast, both chew card and wax tag abundance index values do require correction for valid comparison with trap catch abundance index values. We conclude that chew cards rather than wax tags should replace leg-hold traps in the BMRS due to the better predictive accuracy of the chew card calibration model in non-forest habitat. The maximum chew card abundance index values that can be corrected for valid comparison with trap catch abundance index values were 42.5% in forest and 40.0% in non-forest, although correcting chew card abundance index values $\geq 30\%$ in non-forest habitat is problematic due to rapid saturation.

Acknowledgements

This study was funded by the Department of Conservation. We thank the following DOC staff and contractors for organising and/or conducting field work: Adrian Braaksma, Chris Brausch, Sam Cave, Patrick Crowe, Mike Detlaff, John Donald, Peter Doonan, Keven Drew, Mike Elliott, Aemon Farrell, Sarah Forder, Joel Henton, Billie Herries, Gerard Hill, Charlie Hogg, Paul Jacques, Peter Kirkman, Nick Moss, Tim O'Neill, Danielle Pearson, Moana Robb, David Soper, Jane Tansell, Pete Thomas, Paul Van Klink, Karipa Wall and Troy Watson. Benno Kappers and Sam Rowland helped manage data, and David Burlace, Richard Ewans and Kat Manno (all Department of Conservation) provided technical support. Morgan Coleman (Landcare Research) assisted with the expert assessment of chew cards and wax tags, and Jeanette Birtles (Organic Editing) edited the manuscript. We also thank David Ramsey, Michael Scroggie (both Arthur Rylah Institute for Environmental Research), Elaine Wright, Ian Westbrooke (both DOC), Dean Anderson (Landcare Research) and three anonymous reviewers for constructive comments on manuscript drafts.

References

- Allen RB, Wright EF, MacLeod CJ, Bellingham PJ, Forsyth DM, Mason NWH, Gormley AM, Marburg AE, MacKenzie DI, McKay M 2013. Designing an inventory and monitoring programme for the Department of Conservation's Natural Heritage Management System. Landcare Research Contract Report LC1730. Lincoln, Landcare Research. 244 p.
- Byrom AE, Innes J, Binny RN 2016. A review of biodiversity outcomes from possum-focused pest control in New Zealand. *Wildlife Research* 43: 228–253.
- Cowan PE 2005. Brushtail possum. In: King CM ed. *Handbook of New Zealand mammals*, 2nd edn. South Melbourne, Oxford University Press. Pp. 56–80.
- Department of Conservation (DOC) 2017. Annual report for the year ended 30 June 2017. Wellington, Department of Conservation. 113 p.
- Falcy MR, McCormick JL, Miller SA 2016. Proxies in practice: calibration and validation of multiple indices of animal abundance. *Journal of Fish and Wildlife Management* 7(1): 117–128; e1944-687X.
- Forsyth DM, Link WA, Webster R, Nugent G, Warburton B 2005. Nonlinearity and seasonal bias in an index of Brushtail Possum abundance. *Journal of Wildlife Management* 69: 976–984.
- Forsyth DM, Ramsey DSL, Perry M, McKay M, Wright EF (2018). Control history, longitude and multiple abiotic and biotic variables predict the abundances of invasive brushtail possums in New Zealand forests. *Biological Invasions* 20: 2209–2225.
- Gelman AB, Carlin JS, Stern HS, Rubin DB 2004. *Bayesian data analysis*, 2nd edn. Boca Raton, Chapman and Hall/CRC. 667 p.
- Glen AS, Byrom AE, Pech RP, Cruz J, Schwab A, Sweetapple PJ, Yockney I, Nugent G, Coleman M, Whitford J 2012. Ecology of brushtail possums in a New Zealand dryland ecosystem. *New Zealand Journal of Ecology* 36: 29–37.
- Gormley AM, Holland EP, Pech RP, Thomson C, Reddiex B 2012. Impacts of an invasive herbivore on indigenous forests. *Journal of Applied Ecology* 49: 1296–1305.
- Gormley AM, Forsyth DM, Wright EF, Lyall J, Elliott M, Martini M, Kappers B, Perry M, McKay M 2015. Cost-effective large-scale occupancy–abundance monitoring of invasive Brushtail Possums (*Trichosurus vulpecula*) on New Zealand's public conservation land. *PLoS ONE* 10(6): e0127693.
- Hilbe JM 2011. *Negative binomial regression*, 2nd edn. Cambridge, Cambridge University Press. 553 p.
- Innes J, Kelly D, Overton J McC, Gillies C 2010. Predation and other factors currently limiting New Zealand forest birds. *New Zealand Journal of Ecology* 34: 86–114.
- Lee W, McGlone M, Wright E 2005. Biodiversity inventory and monitoring: a review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. Landcare Research Contract Report: LC0405/122. Lincoln, Landcare Research. 213 p.
- MacKenzie DI, Lachman GB, Droegge S, Royle JA, Langtimm CA 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83: 2248–2255.
- MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LA, Hines JE 2006. *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Burlington, Elsevier Academic Press. 344 p.
- National Pest Control Agencies (NPCA) 2015. A1: Possum population monitoring using the trap-catch, waxtag and chewcard methods. Wellington, National Pest Control Agencies. 44 p.
- Nelson L, Clark FW 1973. Correction for sprung traps in catch/effort calculations of trapping results. *Journal of Mammalogy* 54: 295–298.
- Nugent G 2011. Maintenance, spillover and spillback transmission of bovine tuberculosis in multi-host wildlife complexes: a New Zealand case study. *Veterinary Microbiology* 151: 34–42.
- Payton IA, Newell CL, Beets P 2004. *New Zealand carbon monitoring system: indigenous forest and shrubland data collection manual*. Lincoln, Landcare Research/Rotorua, Forest Research Institute. 68 p.
- Pekelharing CJ, Parkes JP, Barker RJ 1998. Possum (*Trichosurus vulpecula*) densities and impacts on fuchsia (*Fuchsia excorticata*) in South Westland, New Zealand. *New Zealand Journal of Ecology* 22: 197–203.
- Plummer M 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Hornik, K., Leisch, F. and Zeileis, A. eds *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Vienna, Austria, 20–22 March 2003. Vienna, Austrian Association for Statistical Computing (AASC) and the R Foundation for Statistical Computing. Pp. 1–20.
- RDevelopmentCoreTeam 2015. *R: a language and environment for statistical computing*. Vienna, R Foundation for Statistical Computing. <http://www.R-project.org/>
- Rouco C, Norbury GL, Smith J, Byrom AE, Pech RP 2013. Population density estimates of brushtail possums (*Trichosurus vulpecula*) in dry grassland in New Zealand. *New Zealand Journal of Ecology* 37: 12–17.
- Royle JA, Kéry MA 2007. Bayesian state–space formulation of dynamic occupancy models. *Ecology* 88: 1813–1823.
- Royle JA, Link WA 2006. Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87: 835–841.
- Ruffell J, Innes J, Didham RK 2015. Efficacy of chew-track-

- card indices of rat and possum abundance across widely varying pest densities. *New Zealand Journal of Ecology* 39: 87–92.
- Su Y-S, Yajima M 2012. R2jags: using R to run ‘JAGS’. R package, version 0.03–08. <http://CRAN.R-project.org/package=R2jags> (accessed 13 July 2015).
- Sweetapple P, Nugent G 2011a. Chew-track-cards: a multiple-species small mammal detection device. *New Zealand Journal of Ecology* 35: 153–162.
- Sweetapple P, Nugent G 2011b. Chew-track-cards: a guide to the interpretation of animal tooth impressions. http://www.landcareresearch.co.nz/publications/researchpubs/chew-card_a_guide_to_the_interpretation_of_animal_tooth_impressions.pdf (accessed 5 April 2016).
- Thomas MD, Brown JA, Maddigan FW, Sessions LA 2003. Comparison of trap-catch and bait interference methods for estimating possum densities. *New Zealand Plant Protection* 56: 81–85.
- Warburton, B. 2000. Monitoring possum populations. In: Montague TL ed. *The brushtail possum: biology, impact and management of an introduced marsupial*. Lincoln, Manaaki Whenua Press. Pp. 56–80.
- Warburton B, Barker R, Coleman M 2004. Evaluation of two relative-abundance indices to monitor brushtail possums in New Zealand. *Wildlife Research* 31: 397–401.
- Williams BK, Nichols JD, Conroy MJ. 2002. *Analysis and management of animal populations*. San Diego, Academic Press. 817 p.

Received 11 May 2016; accepted 26 March 2018

Editorial board member: Dean Anderson

Supplementary Material

Additional supporting information may be found in the online version of this article:

Appendix S1. Sample size required for estimating the relationship between each of the new abundance index methods (chew cards and wax tags) and the trap catch abundance index.

Appendix S2. Protocol for setting wax tags, chew cards and leg-hold traps in our study.

Appendix S3. Summary of sites at which sampling was conducted during February–May 2015.

Appendix S4. Scatter plots of observed chew card, wax tag and trap catch abundance index values at sites in forest and non-forest habitats.

Appendix S5. Outcomes of independent expert assessment of wax tag and chew card field assessments.

Appendix S6. Detection of possums at sites using chew cards, wax tags, leg-hold traps and faecal pellets.

The *New Zealand Journal of Ecology* provides online supporting information supplied by the authors where this may assist readers. Such materials are peer-reviewed and copy-edited but any issues relating to this information (other than missing files) should be addressed to the authors.