R-10590
Statistical Limits of RTCI Monitoring

Dave Ramsey and Steve Ball

R-10590
Statistical Limits of RTCI Monitoring

Dave Ramsey
Landcare Research
PO Box 11052
Palmerston North
New Zealand

and

Steve Ball
Landcare Research
PO Box 69
Lincoln 8152
New Zealand

Landcare Research Contract Report: LC0304/068

PREPARED FOR:
Animal Health Board
P O Box 3412
Wellington
New Zealand

DATE: January 2004

Reviewed by:	Approved for release by:
Bruce Warburton Scientist Landcare Research	Phil Cowan Science Manager Biosecurity and Pest Management

DOI: https://doi.org/ 10.7931/x3ct-hr91

Contents

Summary

Project and Client

Landcare Research used simulation modelling to evaluate the statistical limits of the residual trap-catch monitoring index (RTCI) for measuring the abundance of low-density possum populations. This work was carried out between September 2002 and December 2003 for the Animal Health Board (Contract No. R-10590).

Objectives

- To estimate the probability that existing monitoring protocols correctly identify instances when the true mean RTCI or degree of clustering is above particular target thresholds.
- To explore the effectiveness of modifications to current monitoring methods, in terms of detecting instances of high mean RTCI or excessive clustering.
- To characterise patterns of possum clustering within a managed vector risk area.

Results

As expected, for a given population density there was a wide range in the simulated values of RTCI, especially as density increased. However, the relationship between median RTCI and density was linear over the range of densities examined (0.1 to 2.0 possums/ha) with a one-unit increase in population density (i.e. 1 possum/ha) equivalent to a 5.4-unit increase in RTCI.

Analysis of RTCI monitoring data from the Wellington and Southland regions indicated that the majority of contract areas had a distribution of line RTCIs with higher variation than expected from a strictly random distribution of possums. We conclude from this that there is a tendency for possums to be non-randomly distributed over the landscape following possum control. Simulation analysis revealed that the most plausible range for the scale of clustering required to generate the level of variability in line RTCIs observed in the Wellington/Southland data was cluster sizes between 12.6 and 50 ha when the mean RTCI was 1% or 2% and between 50 and 100 ha when the mean RTCI was 5%. In addition, the possum densities within these clusters were predicted to be about 4–8 possums/ha.

A clustered distribution of possums in an area following possum control means that the sensitivity of monitoring, using a mean RTCI threshold, is reduced. When the residual population is clustered, monitoring is more likely to produce a false negative result (incorrectly concluding that the mean RTCI is below the threshold) than when the residual population is randomly distributed. This effect is also accentuated as the mean RTCI threshold reduces. Thus, if the target for monitoring is just a mean RTCI, then more monitors (RTCI surveys) will incorrectly pass if the residual possum population is clustered. The most effective modification for improving the sensitivity of monitoring to detect levels of RTCI higher than the threshold was to increase spatial coverage (specifically, by halving the length of trap lines and doubling the number of lines).

Our analysis of the sensitivity of various line threshold targets in combination with a mean RTCI revealed that the standard monitoring protocol, as well as various modifications, was

unlikely to correctly identify cluster sizes of 12.6 ha or smaller (i.e. smaller than the cluster size typical in the Wellington/Southland data at a mean RTCI of 1%), no matter what line threshold was used. Line thresholds had moderate sensitivity to correctly identify cluster sizes of 50–100 ha (i.e. the cluster size typical at a mean RTCI of 5%). Monitoring over 6 nights instead of 3 appeared to correctly identify clusters more consistently than the other monitoring designs explored.

The ability to successfully monitor possum populations at very low densities is limited by the inherent variability of trap catch data. In particular, we considered the example of zero trap catch as a monitoring result. Because this result can occur at a range of possum densities, zero trap catch from a single survey using the existing monitoring protocol provides a very poor indication of the probability of local eradication. However, a result of zero trap catch can provide the basis for statements about the likelihood of population density being below 1% RTCI.

Conclusions

- Vector managers need to be aware that any method for auditing the outcome of possum control operations represents a compromise between two competing interests: sensitivity (probability of correctly identifying instances that exceed some threshold for population abundance or clustering), and specificity (probability of correctly identifying instances below some population threshold).

- Data on the RTCI monitoring operations obtained from regional councils suggest that possum populations following control are more likely to be clustered than not. Monitoring programmes must, therefore, take clustering into account, in terms of trying to detect clustering within individual operations, and in understanding how clustering affects estimates of mean abundance.

- Monitoring using the standard protocol is more likely to result in a false negative result (incorrectly concluding that the mean RTCI is below the threshold) when the residual population is clustered. This effect is accentuated as the mean RTCI threshold reduces. Increasing the spatial coverage (increasing the number of lines) improved the robustness of monitoring to the effects of clustering.

- The standard monitoring protocol is not well equipped to detect clustering of possum populations. Cluster sizes of 12.6 ha or less are almost impossible to detect. Line thresholds could detect moderately large, high-density clusters (50-ha cluster within a 2500-ha area) with moderate sensitivity (50–80%) and high specificity (>95%) if trapping occurred over 6 nights.

- There is an inherent conflict in the statistical requirements of estimating mean possum abundance and detecting clusters of possums. Whereas estimates of the mean are improved by increasing spatial coverage (i.e. lots of trap lines, but few traps per line), the detection of clusters is improved by increasing the sampling intensity within individual lines (e.g. doubling the number of nights of sampling). If managers are seeking to increase both the precision of mean RTCI estimates and the probability of detecting clusters, their needs to be an increase in the total sampling effort. However, if there is little opportunity to increase total sampling effort, managers will need to decide between these two aims.

- Obtaining zero trap catch as a result of monitoring provides a very poor indication of the probability of local eradication. However, if a zero trap catch is obtained, it provides very strong evidence ($P = 0.999$) that the actual residual population is less than 1% RTCI and strong evidence ($P = 0.95$) that the residual population is less than 0.5% RTCI.

Recommendations

- The reliability of the threshold mean RTCI criterion as used with current monitoring protocol/intensity should be improved by increasing spatial coverage with more lines, especially when aiming for a low threshold mean RTCI (e.g. 1%). Decreasing the number of traps per line would achieve this most cost effectively.

- The sensitivity of the standard monitoring protocol to identify true clustering in possum populations following control would be increased by modifying the line threshold criterion to allow trapping to be extended over 6 nights. The following 6-night thresholds (for a standard 10-trap line) provide the greatest sensitivity (probability of correctly identifying clustering) over other line thresholds considered, while having a specificity (probability of correctly identifying lack of clustering) of at least 95%:
    - Mean RTCI of 1%: no line over 5%
    - Mean RTCI of 2%: no line over 6.7%
    - Mean RTCI of 5%: no line over 13.3%.

  Alternatively, if a 5-trap line is adopted (and double the usual number of lines is used), the following 6 night thresholds provide the greatest sensitivity and specificity:
    - Mean RTCI of 1%: no line over 6.7%
    - Mean RTCI of 2%: no line over 10%
    - Mean RTCI of 5%: no line over 16.7%

  If 6-night sampling is adopted for cluster detection, it is important that mean RTCI is still based on only the first 3 nights of sampling. Otherwise (i.e. for a mean based on all 6-nights) RTCI will be biased low (and hence contracts more likely to be passed) because removal trapping decreases the per-night catch over time.

- If the objective is to have a high sensitivity of detecting clustering of possum populations following control, then the AHB should consider developing an alternative monitoring system designed specifically for this purpose. Within the constraints of line-based monitoring, we found that the biggest improvement in cluster detection was through extending the trapping period to 6-night samples. Nonetheless, spatial coverage is likely to be an important aspect of cluster detection. Given that increasing the number of monitoring lines is a costly way of improving spatial coverage, we recommend that the AHB explore cluster-detection methods that do not rely on line-based monitoring.

1. Introduction

Landcare Research used simulation modelling to evaluate the statistical limits of the residual trap-catch monitoring index (RTCI) for measuring the abundance of low-density possum populations. This work was carried out between September 2002 and December 2003 for the Animal Health Board (Contract No. R-10590).

2. Background

The statistical power of possum monitoring programmes using the nationally recognised residual trap-catch index (RTCI) (NPCA 2000) is ultimately limited by sampling variation. The effect of sampling variation is especially apparent when agencies attempt to monitor possum populations at the very low densities achieved under current vector control strategies (typically RTCIs of 2–4%). For example, the use of confidence limits around the RTCI appears to have fallen out of favour, due partly to the difficulties associated with setting defendable targets for contract purposes, but also due to the often skewed nature of the data collected at very low possum densities, making calculation of conventional confidence intervals problematic. Presently, targets for vector control operations usually consist of two thresholds, either (a) reduction to below an overall mean RTCI or (b) reduction to below an overall RTCI and no individual line RTCI to exceed some threshold value. Part (b) may also include a modification that states some percentage of monitoring lines may not exceed a target threshold. The use of these individual line thresholds is an attempt to obtain uniformity of control over the area. However, the use of individual line thresholds may also be affected by sampling variation irrespective of the uniformity of control. For example, the more monitoring lines used, the more likely it is that one line will exceed the target threshold by chance alone even when possums are randomly distributed across the landscape.

Despite this drawback, uniformity of control is believed to be essential in ensuring that Tb cannot persist in controlled possum populations. The use of individual line RTCI thresholds has some advantages as it can provide spatial information about the uniformity of control and is relatively simple to use and interpret. However, it is unknown how sensitive (i.e. statistically powerful) the mean RTCI/individual line RTCI criteria are for detecting where control has resulted in an excessively clustered or “patchy” distribution of remaining possums. Neither is it clear how effective these criteria are in determining the scale of such clustering, nor whether these criteria can reliably identify real changes in possum density when the pre-control density is already low, as is nowadays often the case. To investigate the sensitivity of the current criteria used to evaluate vector control operations, we used a computer-based approach to simulate the monitoring process, and assess the statistical power of RTCI to discriminate between various levels for known (i.e simulated) true population density. In addition, we also assessed the sensitivity of monitoring to detect different scales of clustering in residual possum populations.

3. Objectives

- To estimate the probability that existing monitoring protocols correctly identify instances when the true mean RTCI or degree of clustering is above particular target thresholds.
- To explore the effectiveness of modifications to current monitoring methods, in terms of detecting instances of high mean RTCI or excessive clustering.
- To characterise patterns of possum clustering within a managed vector risk area.

4. Methods

4.1 Model description

We used a spatially explicit, Monte Carlo simulation model of the trapping process to simulate RTCI data for possums (Efford et al. 2004; Ramsey et al. 2004). A brief description of the model follows. Possums are assumed to occupy home ranges that are fixed for the duration of trapping. The probability, *P,* of an individual animal, *i*, being caught in a particular trap, *j*, declines with the distance, *d*, between its home range centre and the trap. For simplicity the relationship, a detection function in the sense of (Buckland et al. 1993)), is assumed to be half-normal:

$$P_{ij} = g(0)e^{\left(-d_{ij}^2/2\sigma^2\right)}, \tag{1}$$

where *g*(0) is the probability of capture when the trap is located exactly at the centre of the home range, and σ is a measure of home range size. Other detection functions such as uniform or negative exponential can also be used (e.g. Efford et al. 2004). A uniform detection function takes the form:

$$\begin{aligned} P_{ij} &= g(0) \quad \{0 < d_{ij} \le \sigma\} \\ P_{ij} &= 0 \quad \{d_{ij} > \sigma\} \end{aligned} \tag{2}$$

When there are several traps, the total probability of capture for a single individual (the probability it is caught during the time interval) is

$$P_i = 1 - \prod_{j=1}^{N} 1 - P_{ij} \ , \tag{3}$$

where the product is over *N* traps (*j* = 1 … *N*).

In reality, the realised capture probability will be less than the total individual capture probability as capture of a particular animal in a trap excludes this trap from further capture. No simple expression is available for the P_{ij} over a time interval when animals compete for traps, and traps simultaneously compete for animals. Therefore, to handle these competing events, we simulated the sequence of captures in continuous time by treating each combination of X animals and *N* empty traps as a competing Poisson process. Thus, the finite capture probability has an exponential distribution,

$$p_{ij} = 1 - e^{-\lambda_{ij} t} \quad , \tag{4}$$

where λ_{ij} is the instantaneous capture (or hazard) rate for animal *i* and trap *j* over one time interval *t*. Since we are usually interested in the capture probability over a single night, setting $t = 1$ gives

$$\lambda_{ij} = -\ln(1 - p_{ij}) \, . \tag{5}$$

The algorithm is then:

1. Calculate λ for each animal+trap combination from eqns (1) or (2), and (5)
2. Simulate the time to first capture for each combination by drawing a pseudorandom number from an exponential distribution with rate λ
3. Find the next capture (remaining animal+trap with minimum time to first capture)
4. If time exceeds 1 then ignore this capture and exit
5. Record capture and remove combinations involving this animal or this trap
6. If at least one animal and one trap remain then go to 3 else exit.

Exponential pseudorandom numbers are obtained as $-\log(U)/\lambda$ where $0 \le U \le 1$ has a uniform distribution. Step 3 is expedited by sorting the list of combinations by ascending values of the simulated time to first capture.

Realistic simulation of the capture process relies on the estimation of the parameters of the half-normal detection function. These parameters are related to the encounter and interaction rate of possums with traps (*g*(0)) and possum home range size (σ) and were estimated from both live-trapping studies using leghold traps (Steve Ball unpublished data) and cage-trapping studies (Murray Efford unpublished data). We have estimates of these detection function parameters for brushtail possums from seven populations sampled both extensively and intensively over a wide range of environmental conditions. For example, we have estimates of *g*(0) and σ from 64 closed capture sessions spanning 21 years of mark-recapture data from the Orongorongo Valley. These data would adequately sample any likely seasonal and/or annual time trends evident in these parameters. This large body of mark-recapture estimates of *g*(0) and σ for brushtail possums captured in traps allow us to confidently simulate the biologically feasible range of heterogeneity in capture probability likely to be experienced during possum trapping. Based on these field estimates, values of the detection parameters were generated randomly according to a normal mixture (σ) and a gamma distribution (*g*(0)) for each realisation (i.e. replicate for each combination of parameters – see Appendix 1 for a description of the stochastic algorithm).

4.2 Simulation of RTCI for possums

Using field estimates of the detection parameters, we attempted to estimate the relationship between known possum density and the likely frequency distribution of RTCI estimates. We simulated RTCI for known (i.e. simulated) population densities of 0.1, 0.2, 0.4, 0.8, 1.0 & 2.0 possums/ha. All simulations occurred on a 2500-ha square arena with the predetermined number of possums distributed randomly within this area. We then simulated random placement of 20 10-trap monitoring lines in the area according to the NPCA guidelines for possum monitoring (NPCA 2000) (Fig. 1). All traplines were set for 3 nights and all captured possums were removed (i.e. an individual could only be caught once during a particular monitor). For each known population density, 1000 iterations of the simulation model were run, and the mean RTCI for each simulation was determined. We used the median of these

1000 mean RTCI values as the estimate of the "true" or expected value of RTCI for a given population density. To extrapolate to other population densities, we fitted a linear regression between true population density and median RTCI. Median RTCI was the dependent variable in this analysis and we predicted a true population density given a median RTCI using inverse prediction, so that a "true" density could be predicted for any particular median RTCI:

$$D = \frac{(RTCI - C)}{\beta}, \qquad (6)$$

where D is true population density, $RTCI$ is the median RTCI and C and β are the intercept and slope parameters respectively estimated from the linear regression equation.

4.3 Characterisation of possum clustering using existing monitoring data

To characterise the likely scale of clustering that might be encountered during monitoring operations we analysed the extensive GIS databases of possum RTCI monitoring data from two vector risk areas held by Environment Southland and Wellington Regional Council. The raw data were values of total trap catch per trap line (number of possums plus the number of "escapes" – i.e. when a trap was found sprung with fur). These data were analysed separately for each 'contract' or "subcontract" area (i.e. an area monitored at a single point in time as an independent unit for contracting purposes). For simplicity we herein use the term "contract" to refer collectively to data sets originally labelled as contracts or subcontracts. Of the total 141 contract areas for which we had data, 23 (16%) were from Wellington, and 118 (84%) from Southland.

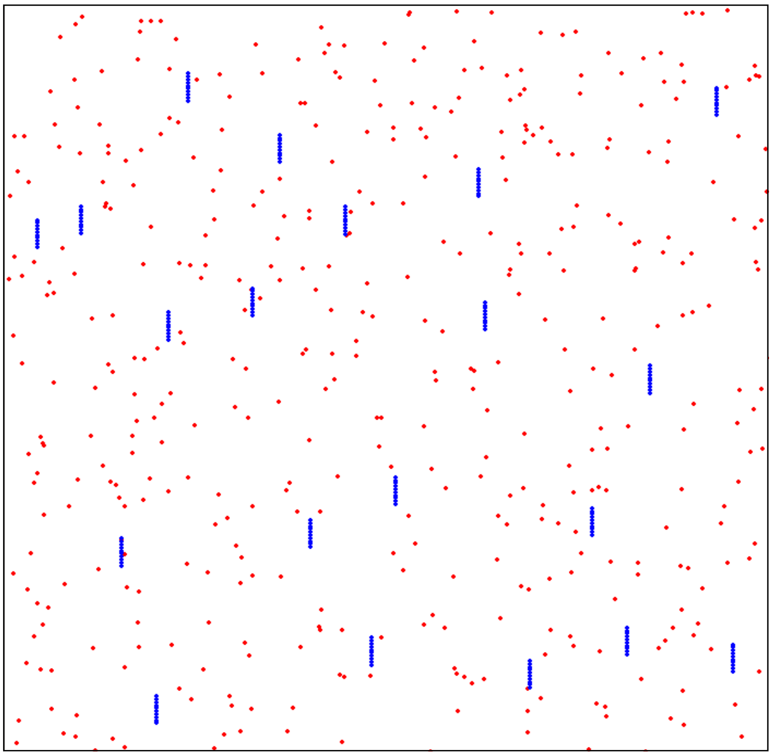

Fig. 1 Example scenario of simulation of RTCI monitoring of possums. Possum density is 0.18 possums/ha corresponding to an expected value for RTCI of 1%. Size of the simulated area is 2500 ha (5 km x 5km).

The data from each contract were collected over a relatively short period, generally within several weeks. Therefore, any evidence of spatial patchiness is unlikely to represent artefacts of seasonal variation in trappability. Beyond this temporal stratification we do not go further towards accounting for between-line variation in trap catch. So, for example, we ignore the possible effect of habitat stratification, and take the view that even if patchiness was caused by habitat variation we were simply interested in characterising the magnitude of the resultant variation in trap catch rather than defining its causes.

The financial years 2000/01, 2001/02 and 2002/03 had by far the greatest amounts of monitoring data (compared with previous years), so were targeted for analysis. Furthermore, analyses were restricted to post-control data sets, as too few pre-control data sets were available for comprehensive analysis.

The data were initially screened to select the first monitoring operation following annual control. This was indicated in the Wellington data as a separate field, and in the Southland data as the earliest sampling date in a given financial year. All contract areas analysed contained at least 10 trap lines in a given financial year, with most between 10 and 30 but some containing as many as 70. Data were available for all three years for many of the Southland operation areas, with fewer repeats available in the Wellington data. To standardise the sampling unit of individual trap lines, data were only used if a line was based on 30 trap nights (i.e. 10 traps over 3 nights). However, we allowed for the loss of up to five effective trap nights due to sprung-and-empty traps, and non-target captures.

The monitoring data were used to estimate between-line coefficient of variation (CV%) in %RTCI for each contract area. The CV(%) was estimated separately for three levels of mean RTCI: 1%, 2% and 5%, grouping the data according to mean RTCI as follows:

Nominal	Actual	Number of contracts
1% RTCI	0.75 – 1.25	83
2% RTCI	1.75 – 2.25	30
5% RTCI	3.00 – 7.00	28

The CV(%) of %RTCI expected from randomly generated data can be calculated based on the binomial distribution (Sokal & Rohlf 1995). This "null expectation" assumes that each trap night has an equal independent probability of capturing a possum. This probability equals 0.01 for a mean RTCI of 1%, 0.02 for 2% RTCI and 0.05 for 5% RTCI. The expected variance of RTCI% can then be calculated using the delta method (Seber 1982):

$$\mathrm{Var(RTCI\%)} = (npq).(100/n)^2 , \qquad (7)$$

where n = 30 (the number of trap nights), p = %RTCI/100, and q = 1 – %RTCI/100.

The factor of $(100/n)^2$ is applied to account for scaling the data from counts (i.e. numbers of captures) to %RTCI.

According to this equation we have an expected CV(%) of 181.7% for a mean RTCI of 1%, 127.8% for a mean of 2%, and 79.6% for a mean of 5%. Thus we expect higher **relative** variation in line RTCI% as mean RTCI% decreases. (Note: these predictions are based on the assumption of equal trapping probability for every trap night.) However, because possums are removed upon capture, trapping probability will on average be slightly higher on the first night of trapping and lowest on the third night. To confirm that this removal effect has a negligible impact on the expected CV(%), we estimated the CV through simulation (in which captured possums were removed), for scenarios of mean RTCI equal to 1%, 2%, or 5%. From the 1000 simulations for each scenario, a subset was selected where the mean RTCI was within the bounds set for empirical data (0.75–1.25% for 1% nominal RTCI, 1.75–2.25% for 2%, and 3–7% for 5%). The CV(%) for these simulations was very close to the CV predicted from the binomial distribution, with an expected CV of 181.7% for a mean RTCI of 1%, 129.1% for a mean of 2%, and 80.3% for a mean of 5% (see above for comparison). We therefore concluded that the effect of removal trapping on the expected coefficient of variation is negligible for the low densities considered here.

Once the observed relative variation in line RTCI% was determined for the Wellington/Southland data set, we attempted to determine the most likely scale of clustering in possum distribution that matched the observed distribution of empirical (field) RTCIs. This was undertaken by simulating clustered possum populations at "true" overall mean RTCIs of 1%, 2%, and 5% and identifying the most plausible combination of cluster size (ha) and cluster density (i.e., possum density within the cluster/ha) that could generate the observed CV(%). For each value of "true" RTCI%, a single circular cluster varying between 3.1 ha (100-m radius) and 201 ha (800-m radius) was simulated with densities of possums within the cluster varying between 2 and 10 possums/ha (Fig. 2). All possums not residing within the cluster were distributed at random in the surrounding area so that the **overall** density of possums did not vary for any particular value of "true" RTCI. As above, simulations were conducted in a 2500-ha area and monitored using 20 lines of 10 traps over 3 nights. One hundred replicate surveys were carried out for each combination of overall mean RTCI, cluster size and cluster density. The mean CV(%) of the line RTCI over the 100 monitors was used as the expected value of CV% for that particular combination of overall mean RTCI, cluster size and cluster density.

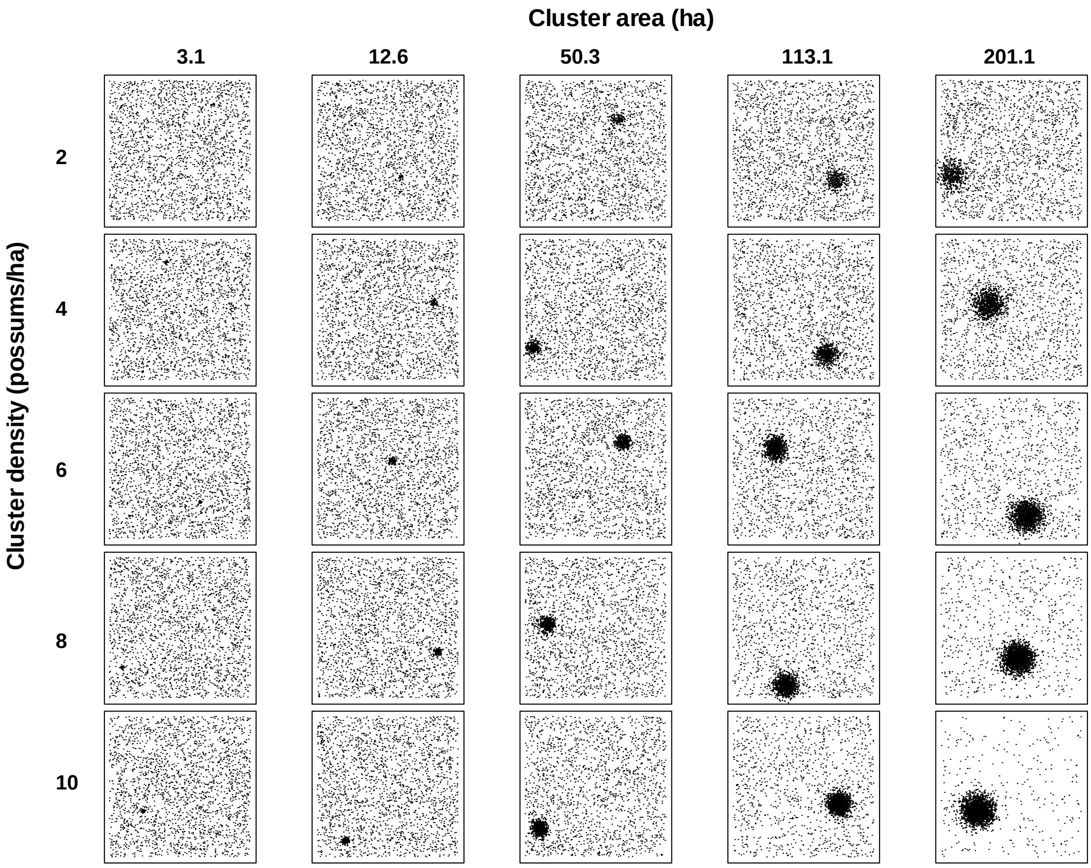

Fig. 2 Matrix of combinations of cluster density (possums/ha) and cluster area (ha) used to define the plausible range for the spatial scale of clustering present in the Wellington and Southland RTCI data. Overall possum density in this example is 0.92 possums/ha giving an

expected value of "true" RTCI of 5%.

4.4 Statistical limits of RTCI monitoring

To determine the ability of the standard RTCI monitoring protocol to identify differences between the "true" RTCI as well as different levels of clustering, we simulated scenarios where the "true" RTCI (= median RTCI) and level of possum clustering were known. We explored the performance of two criteria commonly used for assessing the outcome of control operations using RTCI monitoring. These criteria were (a) threshold mean RTCI and (b) threshold mean RTCI and threshold individual line RTCI.

(a) Evaluation of mean RTCI threshold

For the criterion of mean RTCI, we determined what range of "true" RTCI would be likely to be detected as exceeding a particular threshold mean RTCI. For each of three thresholds or "target" RTCI of either1%, 2% or 5%, we conducted simulated monitors where the "true" RTCI was changed systematically between 0.5% and 10% RTCI. For each combination of threshold mean RTCI and "true" RTCI we determined the probability that a "true" RTCI would be judged as not exceeding the threshold, by calculating the proportion of 1000 simulated monitors where the mean RTCI was equal to or less than the designated threshold mean. Simulations were undertaken both for possums distributed randomly and for clustered possum populations. For the clustering scenario we chose values for cluster area and cluster density that best represented the observed patterns of clustering in the Wellington/Southland data.

We also examined the effect of modifications to the standard monitoring protocol on the probability of detecting an excessively high "true" RTCI by altering the number of lines, the number of traps per line, and the number of days of monitoring in the following combinations:

1. 20 lines of 10 traps for 3 nights (Standard)
2. 40 lines of 5 traps for 3 nights
3. 10 lines of 20 traps for 3 nights
4. 20 lines of 10 traps for 6 nights.

Thus, the total effort (number of trap nights) was equivalent for options 1,2 and 3 while option 4 doubled the number of trap nights but kept a similar spatial coverage to the standard protocol.

(b) Evaluation of individual line RTCI threshold

We explored the performance of the individual line RTCI threshold at detecting scenarios where the spatial distribution of possums was clustered. We initially constructed scenarios with random possum distributions with "true" RTCIs of 1%, 2%, and 5%. In order to evaluate just the individual line threshold we set the target mean RTCI threshold equal to the "true" RTCI. Only simulated monitors that "just passed" on the criteria of mean RTCI were then selected to evaluate the line threshold RTCI criteria. In practice "just passing" on mean RTCI threshold meant that the mean RTCI for a particular monitor was equal to or up to 0.5% less that the target mean RTCI (i.e. for a target of 2% RTCI, a particular monitor was included if the mean was within the range 1.5–2%). For the individual line threshold criteria, we examined various line threshold targets ranging from 2 to 7 possums on any line for mean RTCIs of 1% or 2%, and 5 to10 possums per line for mean RTCI of 5%. One thousand simulated monitors were undertaken for each of these "true" RTCIs and the probability of the monitor reaching the threshold calculated as the proportion of the 1000 simulated monitors with at least one line equalling or exceeding the threshold number of possums.

We then simulated various levels of possum clustering for each value of "true" RTCI using the same criteria as used in the characterisation of possum clustering (Section 4.3) (Fig. 2). For each combination of "true" RTCI, cluster size, and cluster density we simulated RTCI monitoring for evaluation against the specified target thresholds. The results obtained for each combination of cluster size and cluster density were compared with those obtained for a strictly random distribution of possums. As the true level of clustering was known, we could determine the prediction errors by classifying the result of any particular monitoring in the following manner.

		Clustering actually present	
		+	-
Monitor indicates clustering present	+	a	b
	-	c	d

Thus, the true positive rate (= sensitivity) of correctly identifying the presence of clustering (based on a particular line threshold) is equal to a/(a+c) while the true negative rate (= specificity) of correctly identifying the absence of clustering is equal to d/(d+b). The optimal line threshold for a particular mean RTCI was selected as the one that had the maximum sensitivity while having a specificity of no less than 95%.

In addition, we compared results with modifications of the standard monitoring protocol as detailed in evaluation of mean RTCI threshold (Section 4.4 (a)). Following discussions with the AHB, we also explored a further modification to the protocol by examining the sensitivity and specificity of line thresholds for 40 lines of 5 traps over 6 nights.

4.5 Interpretation of zero trap catch

With the reduction of possum densities to increasingly lower levels, there is a temptation to reduce management targets (mean RTCI) accordingly. However, because of the inherent variability in trap catch data, it is not clear that such targets can be monitored effectively among the statistical "noise" of sampling variation. To quantify the uncertainty associated with monitoring possum populations at very low abundance, we considered the scenario of a data set with zero trap catch, and asked: "What does zero trap catch tell us about population abundance?"

We estimated the probability of obtaining zero catch for a range of possum densities, from 0.0004 possums/ha to 0.2 possums/ha (Note: 0.2 possums/ha is equivalent to 1.125% RTCI in the simulation model). The probability, P, of zero trap catch with one possum (equivalent to a "true" RTCI of 0.05%) was estimated from the proportion of simulations without capture from 100 000 simulations with a single possum randomly placed within 2500 ha, sampled using 20 lines (10 traps per line over 3 nights). This probability was then used to estimate the probability of zero trap catch with two possums (P^2), and of obtaining zero trap catch with N possums (P^N), based on a random distribution of possums across the landscape.

5. Results

5.1 Simulation of RTCI for possums

The relationship between RTCI and true population density was linear over the range of densities examined, but was highly heteroscedastic (i.e. the variance in RTCI increased as population density increased) (Fig. 3). A linear regression between the median RTCI and true population density showed an almost perfect linear fit, with $R^2 = 0.999$ (Fig. 3). A one-unit increase in population density (i.e. 1 possum/ha) resulted in an increase in median RTCI by 5.36 (Table 1). Using inverse prediction an RTCI of 5%, say, was equivalent to a population density of 0.92 possums/ha.

Table 1 Parameter estimates from a simple linear model fitted to the median RTCI estimates from 1000 simulated RTCI monitors at different values for true population density.

	Parameter estimates		
Model	β (SE)	*t*	*P*
Intercept	0.053 (0.0733)	0.72	0.51
RTCI	5.36 (0.0743)	72.1	0.00

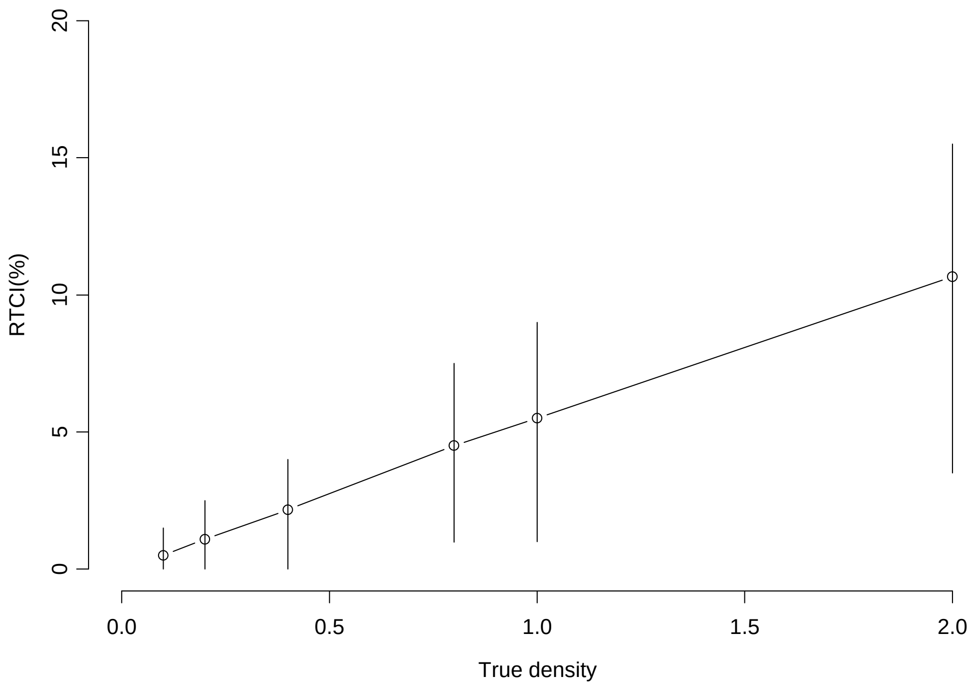

Fig. 3 Relationship between population density of brushtail possums and the median value of 1000 simulated RTCI estimates made using the trapping simulation algorithm. True

population densities were simulated in a 2500-ha arena and simulated RTCI surveys consisted of 20 lines of 10 traps at 20-m spacing over 3 nights. The parameters of the half-normal spatial detection function were drawn randomly for each survey, using the algorithm in Appendix 1. Error bars represent the 2.5% and 97.5% quantiles of the distribution of RTCI values for each value of true population density.

5.2 Characterisation of possum clustering using existing monitoring data

In general, the relative variance in observed RTCI data from Wellington/Southland exceeded the level expected by chance (Fig. 4; Table 2), with 66% of 1% mean RTCI contracts, 87% of 2% mean RTCI contracts, and 93% of 5% mean RTCI contracts having a higher CV(%) than expected by chance. Although the observed CV was not significantly higher than expected for any of the 141 contract areas (i.e. all confidence intervals overlapped the null expectation), the high proportion of values above expected lends strong support to clustering in the observed data. Overall, the CV of 76% of contracts exceeded the null expectation (83% of Wellington contracts and 75% of Southland contracts). There was no significant difference between Wellington and Southland data in the proportion of CV estimates that exceeded the null expectation ($\chi^2 = 0.68$, d.f. = 1, $P = 0.41$).

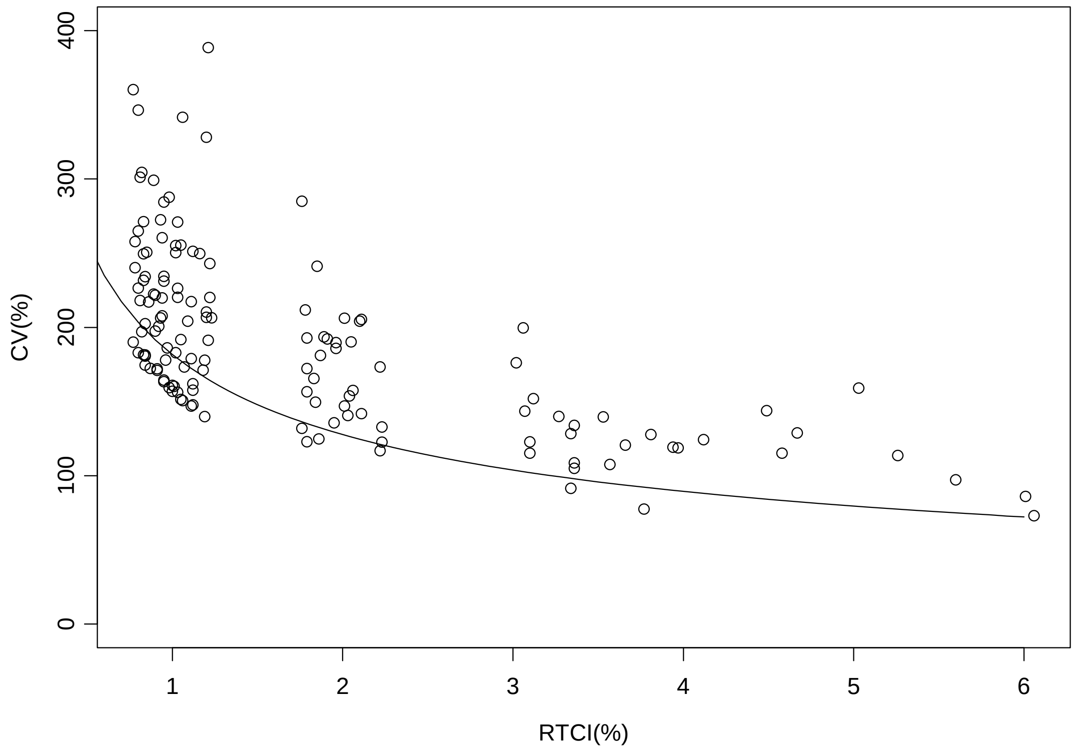

Fig. 4 Coefficient of variation (CV(%)) among line RTCI from 141 contract areas from the Wellington/Southland regions for nominal mean RTCIs of either 1%, 2%, or 5%. The line represents the CV(%) expected from a strictly random distribution of possums (i.e. based on the binomial distribution whereby all trap nights have equal probability of capturing a possum).

Simulated combinations of cluster size and cluster density for values of "true" RTCI of 1%, 2%, or 5% were synthesised to provide contours of expected CV(%) for particular combinations of cluster size and cluster density (Fig. 5). Combinations that could plausibly generate the observed CV(%) in the Wellington/Southland data suggested that for mean RTCI of 1%, cluster sizes of approximately 12–50 ha with densities of 4–6 possums/ha could have produced the observed variation among line RTCI. For a mean RTCI of 2%, cluster sizes of approximately 50 ha with densities of 4–8 possums/ha or a cluster size of approximately 100 ha with a density of 2–4 possums/ha could have produced the observed variation among line RTCI. For a mean RTCI of 5%, cluster sizes of approximately 100 ha with densities of 4–8 possum/ha could have produced the observed variation among line RTCI (Fig. 5).

Table 2 Frequency of coefficient of variation (CV%) values among line RTCIs from 141 contract areas from the Wellington/Southland regions for nominal mean RTCIs of 1% ,2%, or 5%.

	Nominal mean RTCI		
CV(%) category	1%	2%	5%
0–100	0	0	5
101–150	3	11	19
151–200	33	13	4
201–250	26	5	0
251–300	14	1	0
301+	7	0	0
Total	83	30	28

5.3 Evaluation of mean RTCI threshold

The probability of correctly identifying instances when the “true” mean RTCI was above 1%, 2% and 5% thresholds depended greatly on the distribution of the possum population, with less sensitivity when the population was clustered compared with a random distribution (Figs 6 and 7). For example, in randomly distributed populations, a “true” mean RTCI of 3% was failed reliably (sensitivity > 0.95, false negative rate < 0.05) when judged against a 1% threshold (Fig. 6). However, if the population was clustered, then the “true” mean RTCI needed to be at least 5% in order to consistently distinguish it from the 1% threshold with similar reliability (Fig. 7). The effect of clustering declined with increasing values for the threshold RTCI, such that a 5% threshold had only a minor increase in the false negative rate when the population was clustered.

Overall, the effect of modifications to the standard monitoring protocol had minor effects on the false negative rate of different threshold RTCIs. However, doubling the number of lines and reducing their length generally reduced the false negative rate for a given threshold RTCI when the population was clustered. For example, using a 1% threshold and monitoring with the standard 20 lines of 10 traps over 3 nights resulted in a false negative rate of 0.04 for a “true” RTCI of 3% when the population was randomly distributed, and false negative rate of 0.3 when the population was clustered. By comparison, using the same threshold and “true” RTCI and monitoring with 40 lines of five traps over 3 nights resulted in a false negative rate of 0.03 when the population was randomly distributed and only 0.17 when the population was clustered.

Increasing the number of nights to 6 generally resulted in an increase in the false negative rate for a given “true” RTCI, because there is a different relationship between a 6-night RTCI and population density than the one based on 3 nights due to the decline in the catch rate as the number of nights increases.

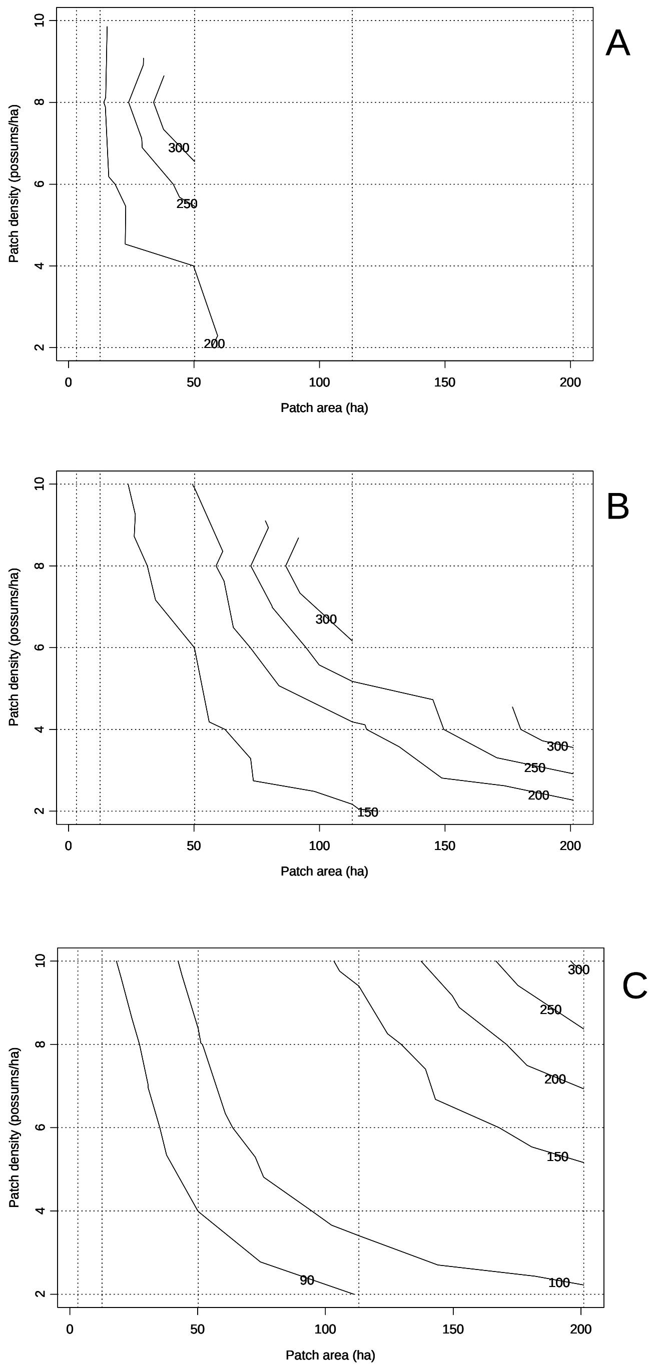

Fig. 5 Contours of coefficients of variation (CV%) among line RTCI for simulated RTCI monitoring of populations with different levels of clustering (cluster size (ha) and cluster population density (possums/ha)) for overall mean "true" RTCIs of (A) 1%, (B) 2% or (C) 5%.

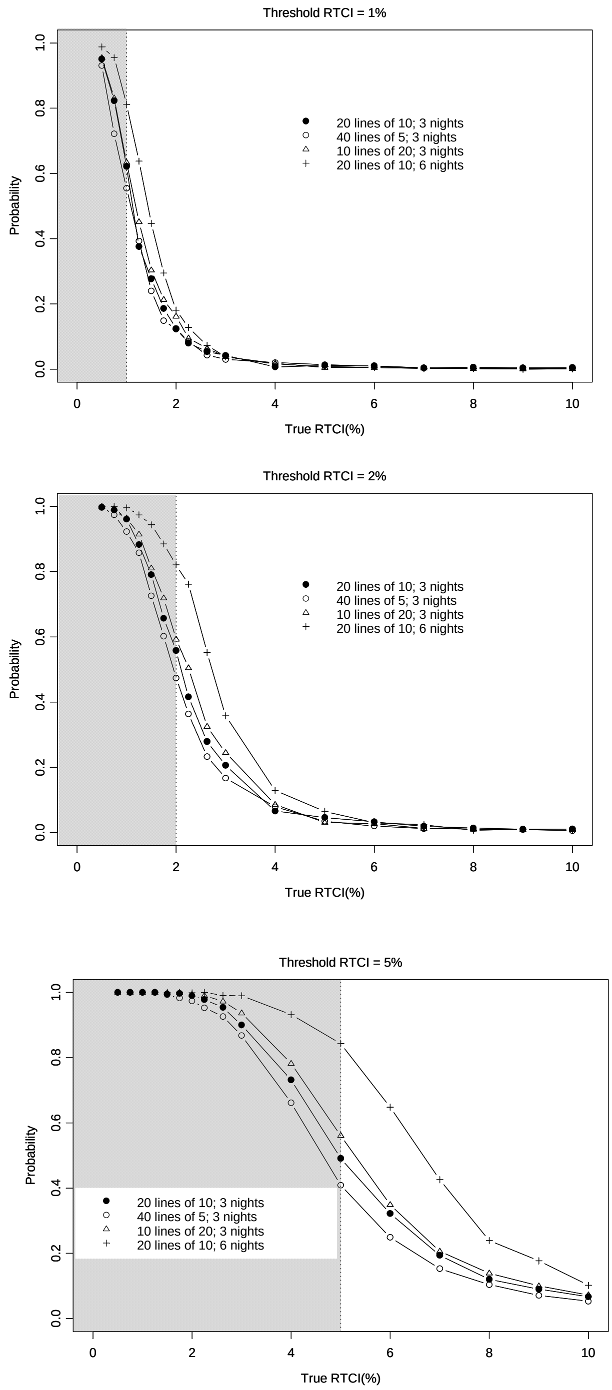

Fig. 6 Probability of a value for "true" RTCI being judged to be equal to or less than a threshold RTCI for threshold RTCIs of 1%, 2%, or 5%. **Populations are distributed randomly.** The shaded portion of each graph is the probability of a true negative result (correctly concluding that the "true" RTCI is equal to or below the threshold) and the unshaded portion of the graph is the false negative rate (incorrectly concluding that the "true" RTCI is equal to or below the threshold). Values for different lines are results for modifications to the standard monitoring protocol.

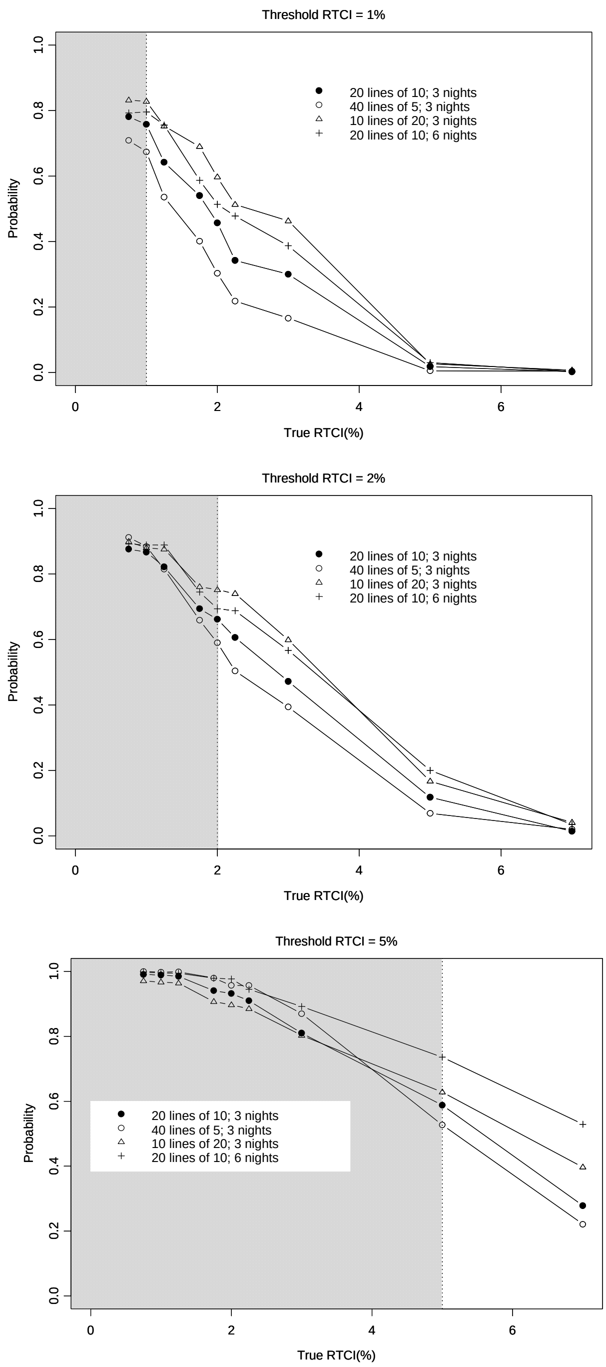

Fig. 7 Probability of a value for "true" RTCI being judged to be equal to or less than a threshold RTCI for threshold RTCIs of 1%, 2%, or 5%. **Populations are clustered.** The shaded portion of each graph is the probability of a true negative result (correctly concluding that the "true" RTCI is equal to or below the threshold) and the unshaded portion of the graph is the false negative rate (incorrectly concluding that the "true" RTCI is equal to or below the threshold). Values for different lines are results for modifications to the standard monitoring protocol.

5.4 Evaluation of individual line RTCI threshold

Figures 8-10 show the sensitivity and specificity of different values for the line threshold for each mean RTCI. Each graph also gives results for the different modifications to the standard monitoring protocol. For possum populations at a "true" mean RTCI of 1%, none of the monitoring scenarios considered correctly identified cluster areas of ≤12.6 ha even if the density within the cluster was high (10/ha). Clusters of 50 ha could be detected with moderate sensitivity (40–60%) if their density was at least 4–8 possums/ha. The line threshold that had the highest sensitivity while still having a specificity of 95% or more was four possums on a standard 10-trap line over 6 nights. This would be equivalent to assigning a monitor failure if any line exceeded 5% over 6 nights for a 10-trap line (Fig. 8).

For possum populations at a "true" mean RTCI of 2%, cluster sizes of ≤12.6 ha were also unlikely to be correctly identified. The optimal line threshold was achieved by trapping a standard 10-trap line over 6 nights with five possums being the optimal threshold. This is equivalent to assigning monitor failure if any line exceeded 6.7% over 6 nights (Fig. 9).

For possum populations at a "true" mean RTCI of 5%, the overall optimal line threshold was nine possums caught on a standard 10-trap line over 6 nights. This was equivalent to assigning monitor failure if any line exceeded 13.3% over 6 nights (Fig. 10).

In summary, the standard monitoring protocol, or the modifications to the standard protocol considered here are unlikely to detect clustering in possum populations where the size of clusters is about 12 ha or less. Even comparatively larger clusters of ≈ 50 ha will only have a moderate level of detectability within a 2500-ha area provided the density within the cluster is relatively high (≥ 6 possums/ha).

The modification to the standard monitoring protocol of 40 lines of 5 traps over 6 nights was compared against 20 lines of 10 traps over 6 nights to determine whether there was any benefit to doubling the number of lines and halving their length when trapping over 6 nights in terms of the sensitivity and specificity of cluster detection. Generally, 40 lines of 5 traps had a higher sensitivity, for a given specificity, than 20 lines of 10 traps. For a true RTCI of 1%, the line threshold that had the highest sensitivity and a 95% specificity when trapping over 6 nights was 3 or more possums on a 5-trap line (Fig. 11). For a true RTCI of 2%, the optimal line threshold when trapping over 6 nights was 4 or more possums on a 5-trap line (Fig. 12). For a true RTCI of 5% the optimal line threshold when trapping over 6 nights was 9 or more possums on a 10-trap line (Fig. 13). However, if a specificity of 92% could be tolerated, then a line threshold of 6 or more possums on a 5-trap line gave equivalent sensitivity.

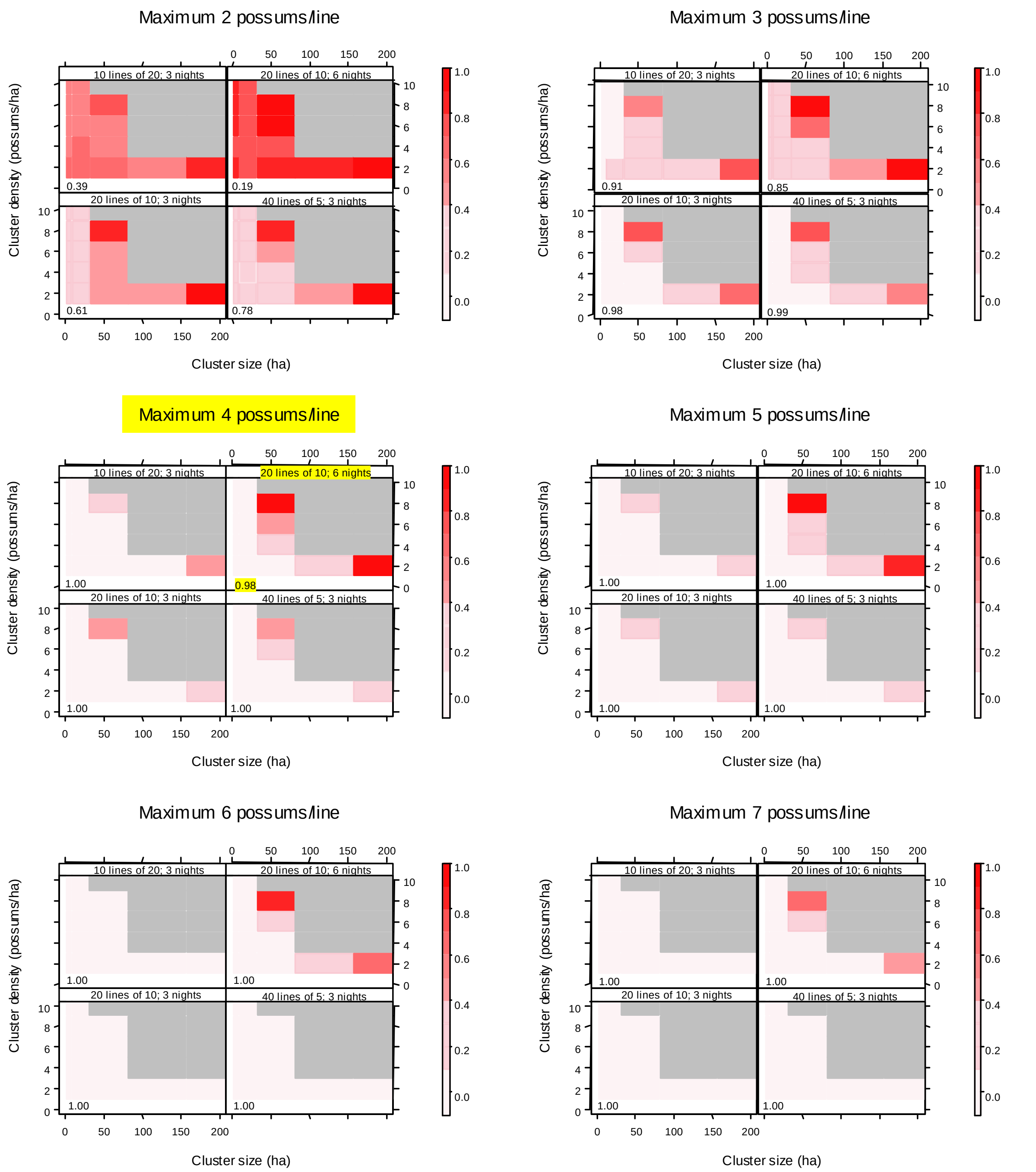

Figure 8. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) for simulated surveys at a **true mean RTCI of 1%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Grey shading indicates that the particular combination of clustering is not compatible with a true mean of 1% RTCI. Yellow highlighting represents the optimal line threshold and monitoring protocol.

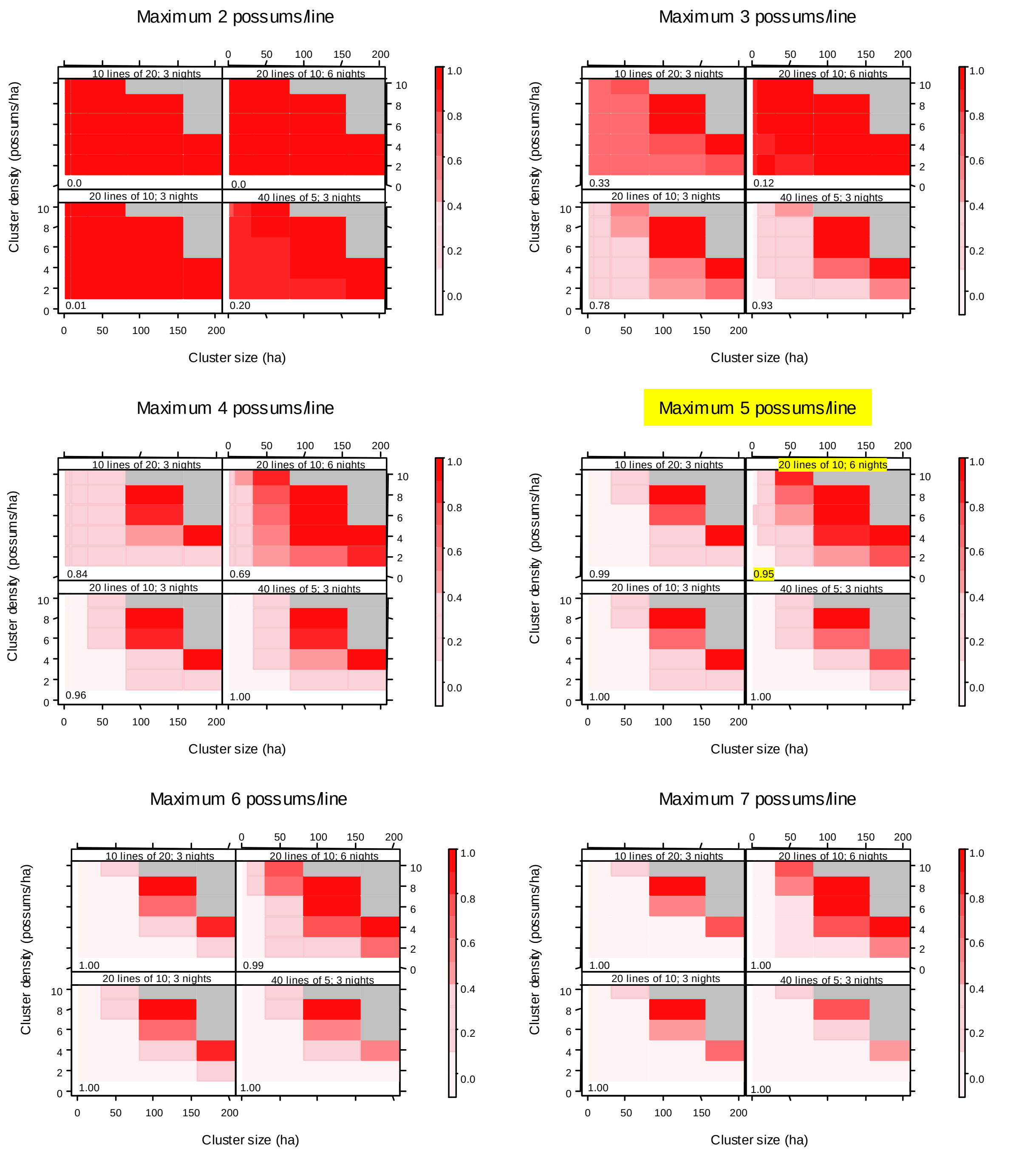

Figure 9. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) for simulated surveys at a **true mean RTCI of 2%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Grey shading indicates that the particular combination of clustering is not compatible with a true mean of 2% RTCI. Yellow highlighting represents the optimal line threshold and monitoring protocol.

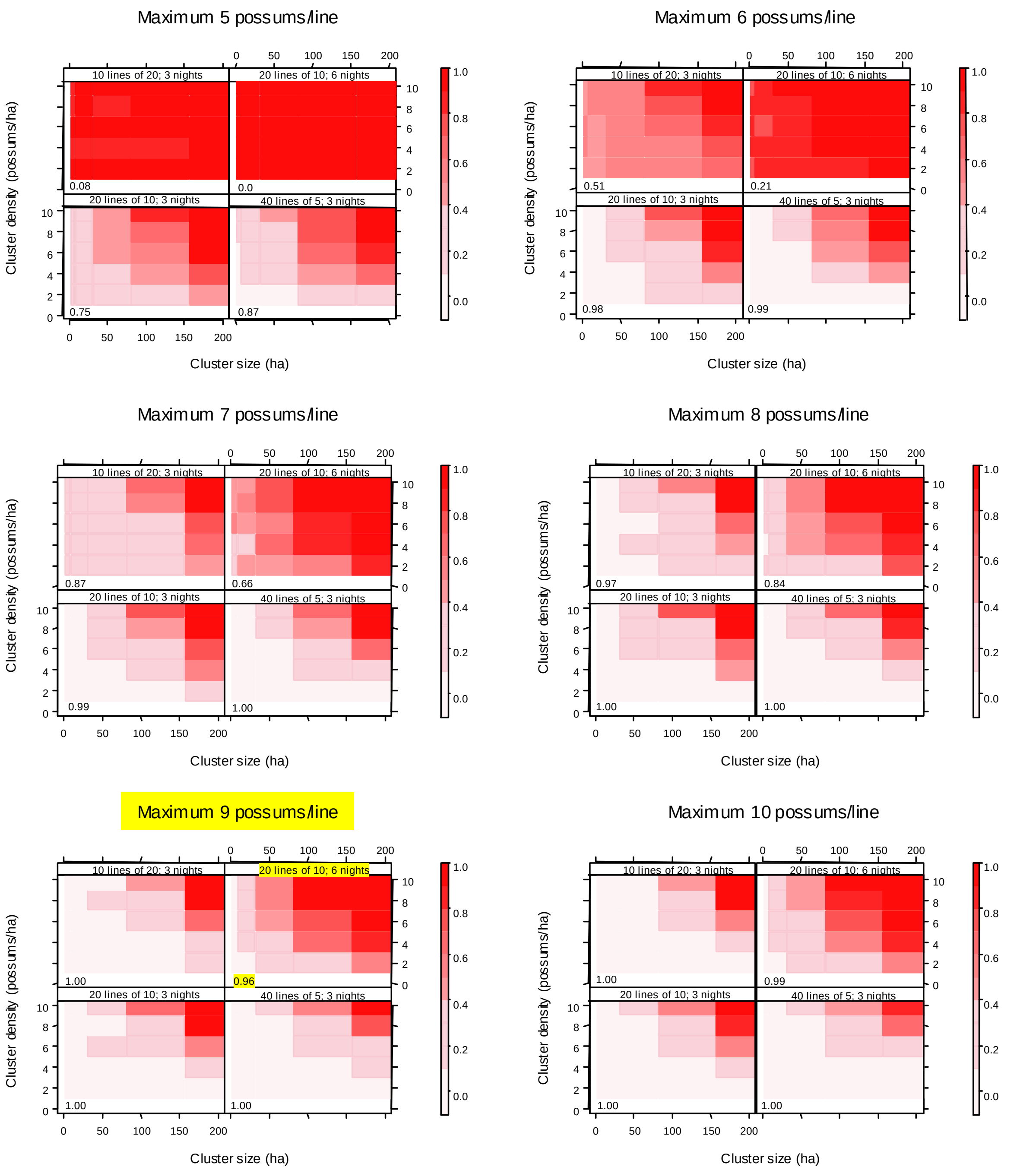

Figure 10. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) for simulated surveys at a **true mean RTCI of 5%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Yellow highlighting represents the optimal line threshold and monitoring protocol.

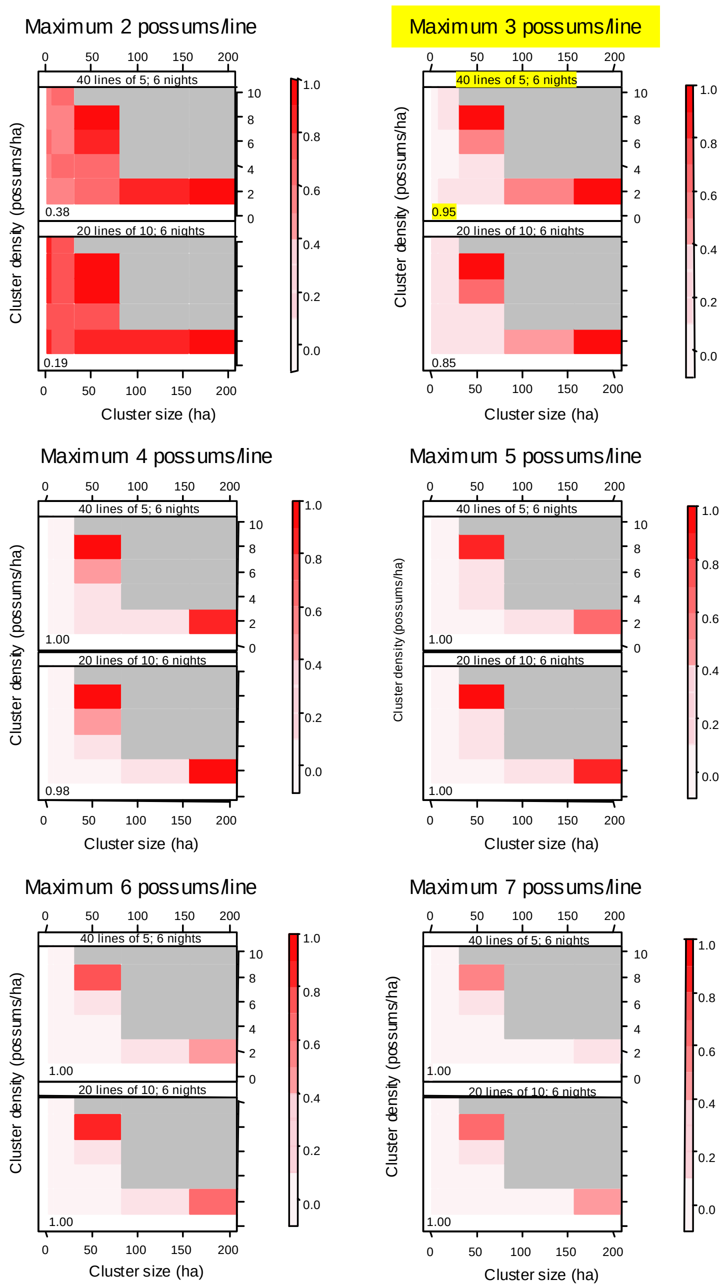

Figure 11. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) when trapping over **6 nights** for simulated surveys at a **true mean RTCI of 1%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Yellow highlighting represents the optimal line threshold and monitoring protocol.

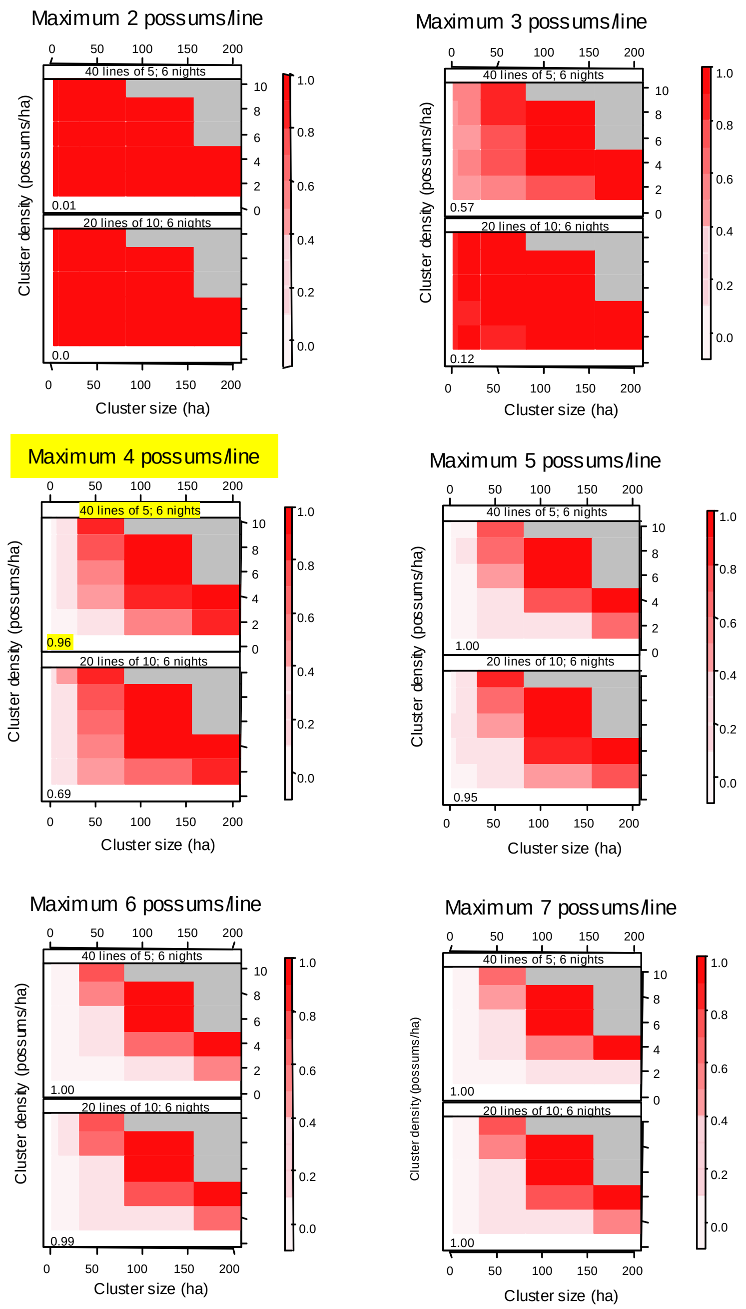

Figure 12. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) when trapping over **6 nights** for simulated surveys at a **true mean RTCI of 2%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Yellow highlighting represents the optimal line threshold and monitoring protocol.

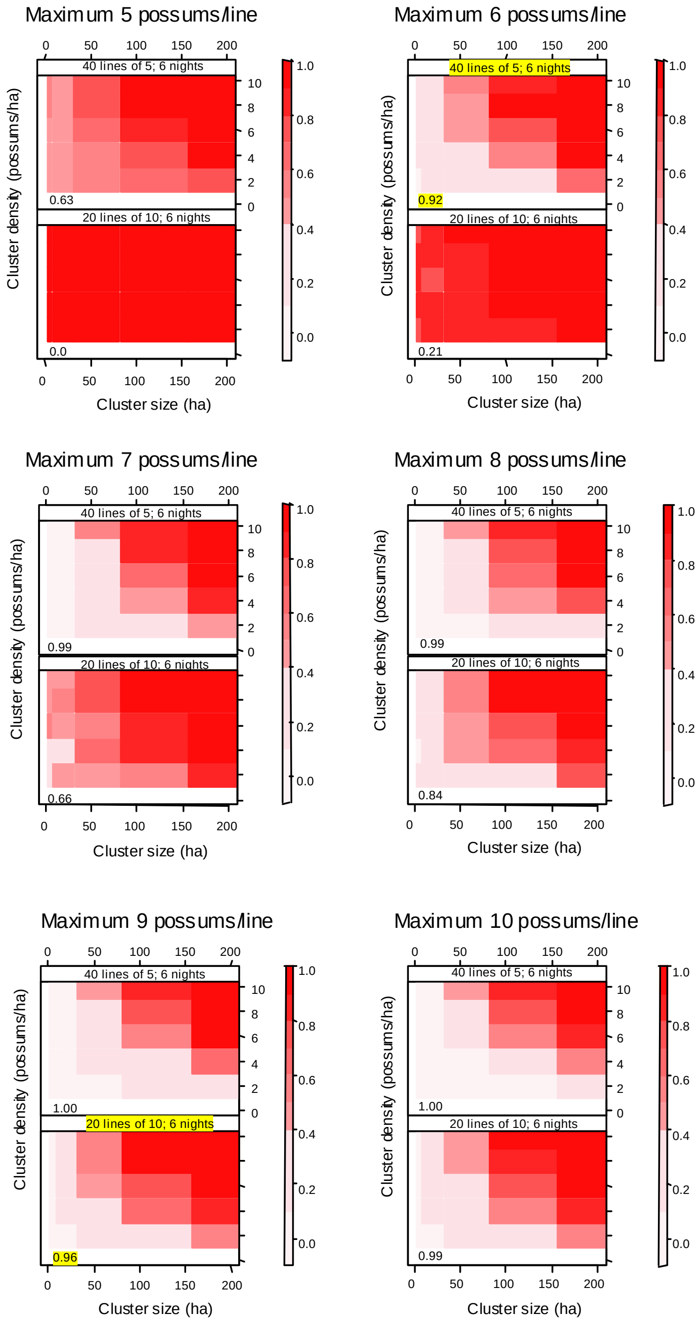

Figure 13. The sensitivity of different values of individual line thresholds (probability of at least one line equaling or exceeding the stated number of possums – red shading) to detect different levels of clustering (cluster size (ha) and cluster density (possums/ha)) when trapping over **6 nights** for simulated surveys at a **true mean RTCI of 5%**. The value in the lower left corner of each quadrant indicates the specificity (i.e. true negative rate – probability of not equaling or exceeding the stated line threshold when there is no clustering). Yellow highlighting represents the optimal line threshold(s) and monitoring protocol.

5.5 Interpretation of zero trap catch

We found there was a relatively high probability of obtaining a result of zero trap catch over a range of densities (Fig. 14). For example, the probability of recording zero trap catch from 20 trap lines at density of 0.0004 possums/ha (1 possum in 2500 ha) was estimated at 0.98582 (95% confidence interval of +/− 0.000738, calculated by normal approximation of the binomial distribution). Based on this value we estimated that the probability of zero trap catch from two possums (density = 0.0008) was 0.972 (i.e. 0.98582^2), and that even with 100 possums (density = 0.04) there is greater than a 0.20 probability of zero trap catch.

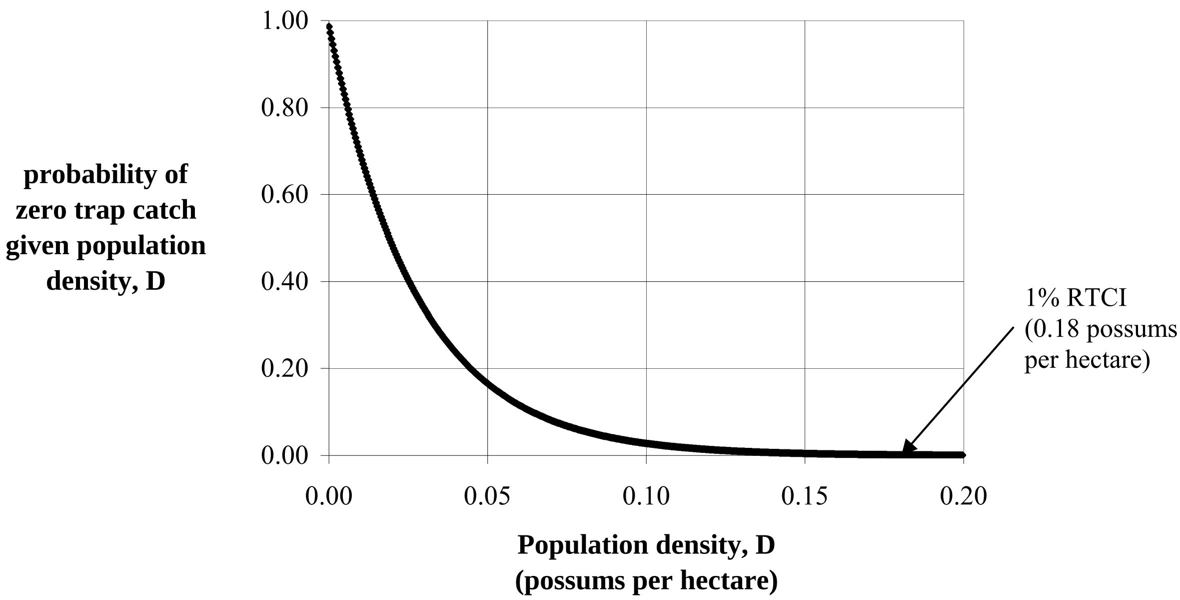

Fig. 14 Probability of obtaining a result of zero trap catch from a population density (possums/ha) of D, using 20 trap lines with 10 traps per line over 3 days. (Note: a density of 0.18 possums/ha is equivalent to 1% RTCI in our simulations. A density of 0.084 is equivalent to 0.5% RTCI).

Thus, a result of zero trap catch is likely over a range of densities. Accordingly, a single zero trap catch should not be taken as a strong indication of eradication (at least for the equivalent sampling effort of 20 lines in 2500 ha). This statement can be extended using Bayesian analysis (Hilborn & Mangel 1997). The mathematics of Bayes Theorem allows us to convert a statement about the "probability of zero trap catch given possum density, D" to a statement of the "probability of possum density being less than D given a result of zero trap catch." To use Bayes Theorem, we need to initially specify a prior distribution of our "beliefs" in what the density is. This essentially places a weighting on each of the alternate possibilities for what the density equals (i.e. 1 possum; 2 possums; N possums). We assumed a uniform prior that places equal weight in any density, but bounded this at 500 possums/2500 ha (0.2 possums/ha). This is equivalent to saying that we are initially unsure of the underlying possum density, and therefore assume that all possibilities (up to N = 500) are equally likely. On this basis, a Bayesian interpretation of zero trap catch would estimate only a 0.014 probability (i.e. a 1 in 70 chance) of there truly being zero possums. Figure 15 shows the probability that the population density is less than D. This shows that we are only 75% sure (probability = 0.75) that the population density is less than 0.04, and 95% sure that the population density is less than 0.084.

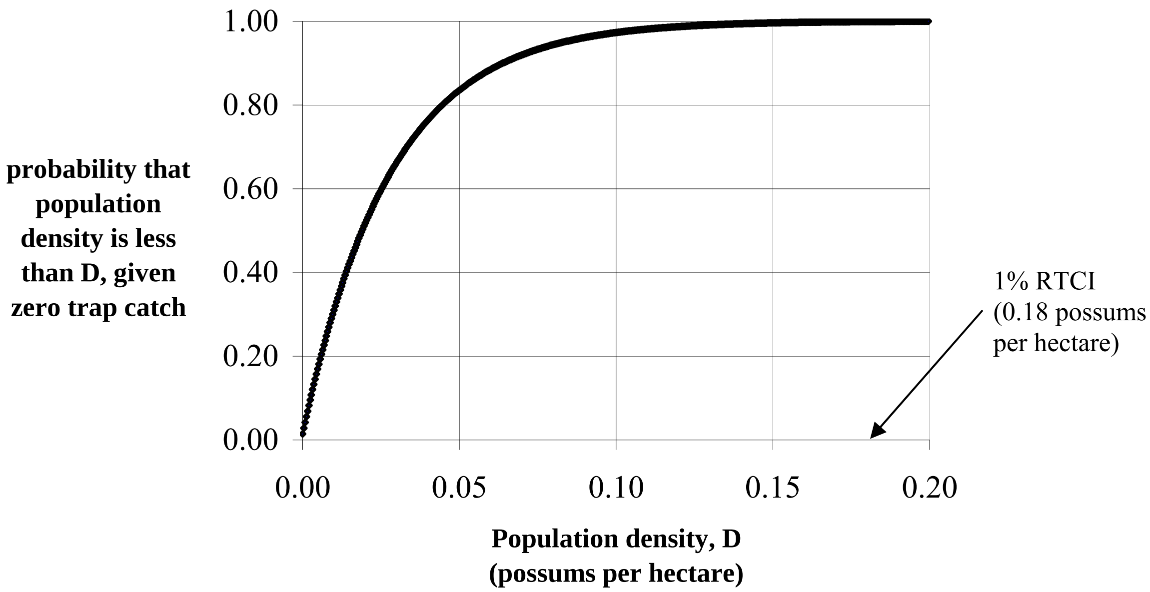

Fig. 15 Bayesian interpretation of a result of zero trap catch (using 20 trap lines with 10 traps per line over 3 days in 2500 ha). This shows the probability that the observed density of possums is below D, given zero trap catch. For example, there is a 0.75 probability that population density is less than 0.04 (i.e. 1 in 4 times we would mistakenly say the density is below 0.04 when it is actually above it). (Note: a density of 0.18 possums/ha is equivalent to 1% RTCI in our simulations. A density of 0.084 is equivalent to 0.5% RTCI.)

The interpretation of zero trap catch can also be phrased in terms of %RTCI, where 1% RTCI is equivalent to a density of 0.18 possums/ha and 0.5% is equivalent to 0.084 possums/ha. A monitoring result of zero trap catch is extremely unlikely at 1% RTCI (probability = 0.002), and moderately unlikely at 0.5% RTCI (probability = 0.051). Applying a Bayesian interpretation of zero trap catch, we would estimate a 0.950 probability that RTCI is 0.0–0.5%, a 0.049 probability that it is 0.5–1.0%, and 0.001 probability that RTCI is larger than 1.0%.

6. Discussion

We introduced a new conceptual model of animal trapping studies that incorporates a detection function to model the interaction of animals with traps. The use of a spatial detection function means that both “trappability” and animal movement are effectively integrated into the calculation of overall capture probability for a particular animal in a particular trap. This formulation has theoretical advantages in allowing for individual heterogeneity in capture probability to be handled naturally while being robust to trap-density and trapping layout.

As expected, for a given population density there was a wide range in the simulated values of RTCI, especially as density increased. However, the relationship between median RTCI and population density was linear over one order of magnitude of density with a one-unit increase in population density (i.e. 1 possum/ha) equivalent to a 5.4-unit increase in RTCI. Further simulation work has shown that the predicted simulated relationship of RTCI with population density is also linear over two orders of magnitude of true population density (0.1–10 possums/ha) with the slope of the relationship slightly lower than presented here at 4.9-unit increase in RTCI per unit increase in density (Ramsey et al. 2004) .

Analysis of RTCI monitoring data from the Wellington/Southland regions indicated that the majority of contract areas had a distribution of line RTCIs with higher variation than expected from a strictly random distribution of possums. We conclude from this that there is a tendency for possums to be non-randomly distributed over the landscape following possum control (however, taken individually some of these results may not be statistically distinguishable from randomness). As possum habitat is never likely to be completely uniform with respect to either possum carrying capacity or the ease with which possums can be controlled, it seems unlikely that the pre-control distribution of possums will be random. Furthermore, the likelihood of clustering increases as the mean RTCI increases (clustering is more likely at high RTCIs). Simulation analysis revealed that the most plausible range for the scale of clustering required to generate the level of variability in line RTCIs observed in the Wellington/Southland data was cluster sizes between 12.6 and 50 ha when the mean RTCI was 1% or 2% and between 50 and 100 ha when the mean RTCI was 5%. In addition, the likely densities of possums within these clusters were predicted to be around 4–8 possums/ha. Well over half of the contract data from the Wellington/Southland regions exhibited variation in line RTCIs consistent with these scales of clustering. We conclude from this analysis that possum control either induces the residual population to form aggregations; that is the effect of possum control itself is likely to result in an aggregated distribution of possums post-control; or that the control is not effective in eliminating the clustering that is present prior to control.

In randomly distributed populations, the mean RTCI criterion is “fair” in the sense that there is a similar probability for the two statistical “mistakes” to be made: (1) false negatives where monitoring incorrectly passes operations in which RTCI is truly higher, or (2) false positives where monitoring incorrectly fails an operation in which the true RTCI was in reality lower than the threshold. This is indicated by the symmetrical shape of the curves around the threshold values in Fig. 6. For example, with standard monitoring (20 lines of 10-trap lines over 3 nights) and a 5% target mean threshold, the probability of incorrectly passing an operation that is 1% in excess of the threshold is equal to 0.322, very similar to 0.268, the probability of incorrectly failing an operation that is 1% below the threshold. This means that vector managers can be as confident of correctly failing true failures as possum controllers can be of having operations pass when they have actually reduced the possum population below the threshold.

Most of the modifications in monitoring protocol that we examined did not markedly change the probabilities of failing operations on the threshold mean RTCI criterion for randomly distributed populations. An ideal monitoring programme would be one that rapidly switches from a high chance of passing operations immediately below the target threshold, to a low chance of passing operations immediately above the threshold (i.e. equivalent to a very steep negative gradient in Fig. 6). However, none of the modifications examined resulted in notable improvements in this gradient, and we would therefore not recommend changes in the protocol when monitoring for mean abundance in randomly distributed possum populations. Furthermore, the modification of using 6 nights sampling instead of 3 resulted in elevated probabilities of passing operations. This difference is attributable to removal trapping decreasing the local abundance of possums around trap lines in the latter part of the trapping session, and therefore reducing the average per-night trap catch. This essentially changes the nature of the relationship between RTCI and the actual density of possums (i.e. the slope of the regression between RTCI and possum density), and any proposal to change the monitoring protocol by increasing the number of sampling nights should consider that the resultant RTCI is not comparable with an RTCI based on fewer nights.

The implication of a clustered distribution of possums in an area following possum control is that the sensitivity of monitoring (using a mean RTCI threshold), is reduced. When the residual population is clustered, monitoring is more likely to result in a false negative result (incorrectly concluding that the mean RTCI is below the threshold) than when the residual population is randomly distributed. This effect is accentuated as the mean RTCI threshold reduces. Thus, if the target for monitoring is solely a mean RTCI, then more monitors will incorrectly pass if the residual possum population is clustered.

The most effective modification of those examined for improving the sensitivity of monitoring to detect levels of RTCI higher than the threshold was to increase spatial coverage (specifically, by halving the length of trap lines and doubling the number of lines). This modification improved the sensitivity of detecting high levels of possum abundance in clustered populations (lowered the curves in Fig. 7), and was associated with a more rapid switch between specificity and sensitivity (steeper negative gradient in Fig. 7). Given that this modification to the protocol had little impact on the monitoring of randomly distributed populations (Fig. 6), we would recommend it as a robust method to improve the monitoring of mean possum abundance.

The use of a line threshold RTCI criterion in combination with an overall mean threshold RTCI criterion represents an attempt to detect clustering in possum monitoring operations. Our analysis of the sensitivity of various line threshold targets in combination with a mean RTCI has concluded that the standard monitoring protocol, as well as various modifications, was unlikely to correctly identify cluster sizes of 12.6 ha or smaller, no matter what line threshold was used. Line thresholds had moderate sensitivity to correctly identify cluster sizes of 50–100 ha, which is consistent with the observed plausible range of clustering needed to generate the RTCI variability in the Wellington/Southland data when mean RTCI is 2% or 5%. Monitoring over 6 nights instead of 3 appeared to correctly identify clusters more consistently than the other monitoring designs explored, and halving the length of trap lines and doubling the number of lines over 6 nights further improved the sensitivity of cluster detection. However, the increase in sensitivity using double the number of shorter lines was more pronounced for a mean RTCI threshold of 1% than it was for higher mean RTCI thresholds with practically no difference when using a mean RTCI threshold of 5%. For mean RTCI thresholds of 1%, the line threshold that most correctly identified clustering was a line RTCI of 10% (i.e. any 5-trap line catching three or more possums over 6 nights). For a mean RTCI threshold of 2%, the line threshold that most correctly identified clustering was a line RTCI of 13.3% (i.e. any 5-trap line catching four or more possums over 6 nights). For a mean RTCI threshold of 5%, the line RTCI threshold that most correctly identified clustering was 15% (i.e. any 10-trap line catching nine or more possums over 6 nights). Alternatively, for a 5-trap line, a line threshold of 20% (any 5-trap line catching 6 or more possums over 6 nights) gave an equivalent sensitivity to the line threshold for the 10-trap line. However, the specificity was slightly lower (92% cw 96%).

The ability to successfully monitor possum populations at very low densities is limited by the inherent variability of trap catch data. In particular, we considered the example of zero trap catch as a monitoring result. Because this result can occur at a range of possum densities, zero trap catch under the existing monitoring protocol provides a poor indication of the probability of local eradication. However, when possum abundance is expressed in terms of %RTCI (where 1% RTCI is equivalent to 0.18 possums/ha, a result of zero trap catch provides strong evidence (0.999 probability) that the “true” RTCI is below 1%, and 0.95 probability that “true” RTCI is below 0.5% RTCI. Therefore, while a single survey using the existing protocol is insufficient to conclude that local eradication has been achieved, a result of zero trap catch can provide the basis for determining the likelihood that population abundance is below 1% RTCI.

7. Conclusions

- Vector managers need to be aware that any method for auditing the outcome of possum control operations represents a compromise between two competing interests: sensitivity (probability of correctly identifying instances that exceed some threshold for population abundance or clustering), and specificity (probability of correctly identifying instances below some population threshold).

- Data on the RTCI monitoring operations obtained from regional councils suggest that possum populations following control are more likely to be clustered than not. Monitoring programmes must, therefore, take clustering into account, in terms of trying to detect clustering within individual operations, and in understanding how clustering affects estimates of mean abundance.

- Monitoring using the standard protocol is more likely to result in a false negative result (incorrectly concluding that the mean RTCI is below the threshold) when the residual population is clustered. This effect is accentuated as the mean RTCI threshold reduces. Increasing the spatial coverage (increasing the number of lines) improved the robustness of monitoring to the effects of clustering.

- The standard monitoring protocol is not well equipped to detect clustering of possum populations. Cluster sizes of 12.6 ha or less are almost impossible to detect. Line thresholds could detect moderately large, high-density clusters (50-ha cluster within a 2500-ha area) with moderate sensitivity (50–80%) and high specificity (>95%) if trapping occurred over 6 nights.

- There is an inherent conflict in the statistical requirements of estimating mean possum abundance and detecting clusters of possums. Whereas estimates of the mean are improved by increasing spatial coverage (i.e. lots of trap lines, but few traps per line), the detection of clusters is improved by increasing the sampling intensity within individual lines (e.g. doubling the number of nights of sampling). If managers are seeking to increase both the precision of mean RTCI estimates and the probability of detecting clusters, their needs to be an increase in the total sampling effort. However, if there is little opportunity to increase total sampling effort, managers will need to decide between these two aims.

- Obtaining zero trap catch as a result of monitoring provides a very poor indication of the probability of local eradication. However, if a zero trap catch is obtained, it provides very strong evidence (P = 0.999) that the actual residual population is less than 1% RTCI and strong evidence (P = 0.95) that the residual population is less than 0.5% RTCI.

8. Recommendations

- The reliability of the threshold mean RTCI criterion as used with current monitoring protocol/intensity should be improved by increasing spatial coverage with more lines, especially when aiming for a low threshold mean RTCI (e.g. 1%). Decreasing the number of traps per line would achieve this most cost effectively.

- The sensitivity of the standard monitoring protocol to identify true clustering in possum populations following control would be increased by modifying the line threshold criterion to allow trapping to be extended over 6 nights. The following 6-night thresholds (for a standard 10-trap line) provide the greatest sensitivity (probability of correctly identifying clustering) over other line thresholds considered, while having a specificity (probability of correctly identifying lack of clustering) of at least 95%:
    - Mean RTCI of 1%: no line over 5%
    - Mean RTCI of 2%: no line over 6.7%
    - Mean RTCI of 5%: no line over 13.3%.

  Alternatively, if a 5-trap line is adopted (and double the usual number of lines is used), the following 6 night thresholds provide the greatest sensitivity and specificity:
    - Mean RTCI of 1%: no line over 6.7%
    - Mean RTCI of 2%: no line over 10%
    - Mean RTCI of 5%: no line over 16.7%

  If 6-night sampling is adopted for cluster detection, it is important that mean RTCI is still based on only the first 3 nights of sampling. Otherwise (i.e. for a mean based on all 6-nights) RTCI will be biased low (and hence contracts more likely to be passed) because removal trapping decreases the per-night catch over time.

- If the objective is to have a high sensitivity of detecting clustering of possum populations following control, then the AHB should consider developing an alternative monitoring system designed specifically for this purpose. Within the constraints of line-based monitoring, we found that the biggest improvement in cluster detection was through extending the trapping period to 6-night samples. Nonetheless, spatial coverage is likely to be an important aspect of cluster detection. Given that increasing the number of monitoring lines is a costly way of improving spatial coverage, we recommend that the AHB explore cluster-detection methods that do not rely on line-based monitoring.

9. Acknowledgements

We would like to thank the AHB for funding this research. We acknowledge Environment Southland (Mark Hunter, Craig Reed, Amy Rush and Lyndon Dynes) and Wellington Regional Council (James Lambie and Wayne O'Donnell) for providing possum monitoring data and associated information on vector management. Thanks to Greg Arnold for invaluable help on the algorithm for generating stochastic capture probabilities in the simulation model. We thank Joanna McKenzie for useful discussions on possum patchiness. Graham Nugent, Bruce Warburton and Guy Forrester gave valuable comments on earlier drafts of this report. Thank also to Christine Bezar for editorial comment and Jemma Callaghan for word processing.

10. References

Buckland,S.T.; Anderson,D.R.; Burnham,K.P.; Laake,J. 1993: Distance sampling: estimating abundance of biological populations.London, Chapman & Hall.

Efford,M.; Dawson,D.K.; Robbins,C.S. 2004: DENSITY: Software for analysing capture-recapture data from passive detector arrays. *Animal Biodiversity & Conservation in press*

Hilborn,R.; Mangel,M. 1997: The Ecological Detective. Confronting models with data.Princeton, NJ., Princeton University Press.

NPCA 2000: Unit 3.1. Trap-catch for monitoring possum populations - protocol for designers. National Possum Control Agencies Contract Report ISBN 0-9583736-7-1

Ramsey,D.; Efford,M.; Ball,S.; Warburton,B.; Nugent,G. 2004: The evaluation of indices of animal abundance using spatial simulation of animal trapping. *Wildlife Research*: submitted

Seber,G.A.F. 1982: The estimation of animal abundance.2nd.edition London, Charles Griffin & Co.

Sokal,R.R.; Rohlf,F.J. 1995: Biometry.3rd.edition New York, W.H. Freeman and Company.

Appendix 1.

Table 3. Estimates of the mean of the parameters g(0) and σ (+se) from the half-normal detection function. Estimates were made using the algorithm of Efford et al (2004) using inverse prediction from mark-recapture data. n = number of closed capture sessions with the exception of Mt Somers where n = No. of individuals.

Site	*g*(0) (+se)	σ (m) (+se)	n
Castlepoint	0.17 (0.015)	35 (0.84)	39
Orongorongo	0.13 (0.007)	31 (0.42)	64
Pigeon Flat	0.19 (0.011)	31 (0.63)	39
Rakiura	0.03 (0.007)	50 (4.51)	8
Waitarere	0.24 (0.055)	50 (3.28)	5
Mt Somers	0.07 (0.007)	55 (4.12)	14
Turitea	0.08 (0.007)	27 (0.63)	23
Overall mean	0.13 (0.0398)	41 (4.08)	192

Based on data in Table 3, a stochastic algorithm to simulate the parameters of the half-normal detection function *g*(0) and σ was developed. Initial analysis of the distribution of σ and log(σ) indicted non-normality. However, there was evidence that a good approximation could be gained by fitting a mixture of two normal distributions to log(σ) using maximum likelihood, giving the estimates;

Parameter	Estimate	se
p	0.781	0.0607
$\overline{x_1}$	3.41	0.0131
σ_1	0.13	0.0105
$\overline{x_2}$	3.84	0.0947
σ_2	0.29	0.0453

which indicated approximately 78% of the mass around $\overline{x_1}$.

The conditional distribution of g(0), given σ, was estimated by fitting a generalised linear model and assuming *g*(0) was inversely related to σ and hence, follows a gamma distribution giving a relationship for 1/g(0) = 0.26 σ – 1.08. Unfortunately, taking this relationship and the estimate of the dispersion parameter (0.35) didn't seem to adequately reproduce the observed inverse relationship. However, changing the rate parameter to 25 gave satisfactory results. Hence g(0) had a gamma distribution with shape = 25/(0.26* σ-1.08) and scale = 1/25. Simulated results are presented in Figure 16.

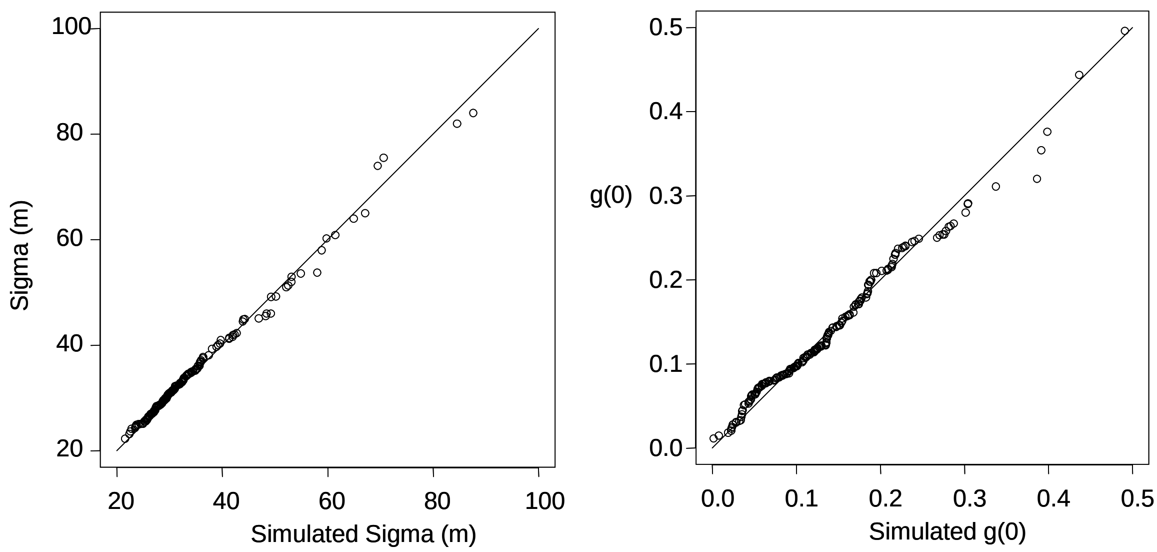

Figure 19. Plot of simulated values of g(0) and σ (x-axis) vs estimates of g(0) and σ from data presented in Figure 1 (y-axis). Random values of σ were drawn from a normal mixture distribution with means of log(σ) of 3.41 (se=0.0131) and 3.84 (se=0.0947) with g(0) taking random values from a gamma distribution with with shape = 25/(0.26*