Residual trap-catch methodology for

low-density possum populations

(Part I – Improving the precision of possum population density estimates)

J. A. Brown[1] and M. D. Thomas[2]

[1]Biomathematics Research Centre
University of Canterbury
Private Bag 4800, Christchurch
New Zealand
Email j.brown@math.canterbury.ac.nz

[2]Pest Control Research
PO Box 7223, Christchurch
New Zealand
Email info@pcr.co.nz

Report to the Animal Health Board, PO Box 3412, Wellington
Research Contract R-80512 (Part 1)

Research report UCDMS 2000/6
Biomathematics Research Centre
Department of Mathematics and Statistics
University of Canterbury

DATE: May 2000

Biomathematics
Research Centre

PEST CONTROL RESEARCH

Contents

1. Summary

1.1 Project and client

The University of Canterbury and Pest Control Research were to investigate the residual-trap-catch (RTC) methodology for monitoring low-density possum populations. The research detailed in this report was carried out for the Animal Health Board between July 1999 and April 2000.

1.2 Objectives

- To improve the precision of possum monitoring estimates by investigating the use of more trap-lines containing fewer traps and comparing their precision and cost-effectiveness at two field sites in the North and South islands.
- To improve the statistical analysis of possum monitoring data by investigating the use of bootstrap methods to provide more precise and accurate estimates of confidence intervals for the residual-trap-catch (RTC) index.
- To identify possible alternative sampling designs that could improve the precision and accuracy of low-density possum monitoring.

1.3 Methods

- Field trials were conducted in the North Island (Tutukau) and South Island (Hohonu) to compare the relative precision of RTC estimates and confidence intervals using 5 trap-lines containing 20 traps, 10 trap-lines containing 10 traps and 20 trap-lines containing 5 traps. In addition times to set and check the traps were recorded to compare the relative cost-efficiency of the three designs.
- Simulated trap-catch data were used to obtain a range of true RTC values characteristic of low-density populations. The simulated data were used to compare the precision and accuracy of confidence intervals calculated using the standard method specified in the NPCA trap-catch protocol with a range of alternative bootstrap methods. Precision and accuracy were determined by comparing mean square errors (MSE), coverage and balance.

1.4 Results

- The estimates of RTC for the three designs ranged from 5% to 7.07% for Tutukau and 3.67% to 6.73% for Hohonu. The most precise RTC estimates were obtained from the 20 lines of 5 traps designs. As expected the cost of the designs increased as the number of trap-lines increased. At Tutukau the design with 20 lines took 27% longer than the design with 5 lines but had a 33% increase in precision. At Hohonu there was a 33% increase in cost but only a 14% improvement in relative precision.
- Monitoring designs that had only five lines of traps gave poor results in terms of the precision, coverage and balance of the confidence intervals.
- The simulated data provided a good fit when compared to the field data. The biased corrected percentile method bootstrap method (BCP) gave the most

precise and accurate calculations of confidence intervals having the lowest MSE's, better coverage and more balance. The results using the standard and BCP methods indicated that the RTC was likely to fall outside the confidence intervals more than 95% of the time (range 81–96%). The RTC estimates falling outside the confidence intervals were more likely to fall above the upper confidence limit 90% of the time when using the standard method and between 52% and 83% of the time when using the BCP method.

1.5 Conclusions

- Designs with a larger number of shorter trap-lines had the highest relative precision, but had the highest field costs. A balance between precision and cost must be met but the importance of having accurate and reliable estimates of residual population size should be understood and appreciated.
- The bias corrected bootstrap method (BCP) had the best performance for estimating confidence intervals among the methods used in this study. This method gives the confidence intervals better coverage and more balance compared to the standard method.
- Confidence limits need to be used with caution because there are always a proportion of the true RTC estimates that fall outside of them. This proportion is more likely to be above rather than below the confidence intervals especially when the standard method is used to calculate confidence intervals. When the true RTC falls above the confidence interval it means there will be more possums in the residual population than the monitoring results indicate.
- Problems of poor accuracy and precision of estimates of RTC and confidence intervals because of small sample sizes is compounded when low-and very low-density populations are monitored because of low animal counts from trap-lines.
- Spatially extensive designs are more likely to provide more precise measures of possum populations. These designs would suit the use of alternative lightweight monitoring devices that will provide larger sample sizes and better spatial coverage.
- Systematic line placement rather than random line placement, as specified in the trap-catch protocol, could provide better estimates of possum density when using trap-catch methods. Systematic sampling should be easier to implement in the field.
- A measure of the proportion of a survey area that contains possums may be a more appropriate method to monitor very low - density possum populations, e.g., those that are targeted for disease eradication.

1.5 Recommendations

- RTC estimates for low-density possum populations calculated using the trap-catch protocol should not be used on their own without reference to their associated confidence intervals. Basing decisions to pay contractors on specified RTC levels without reference to confidence intervals is not recommended.

- Improvements to the trap-catch protocol, (e.g., using more complex techniques to analyse data and estimate confidence intervals, using shorter lines with fewer traps, and systematically locating trap-lines) will not address the underlying problems associated with small sample sizes. Consideration needs to be given to investigating alternative sampling methods that will provide larger sample sizes and to using systematic sampling designs.
- The proportion of an operational area that contains possums is a more informative measure than the proportion of traps that capture possums when measuring very low - density populations. To estimate the proportion of an area that contains possums, the animals do not need to be caught in leg-hold traps and removed. This allows the use of more lightweight monitoring devices that record the presence of possums only, rather than traps that record the presence and numbers of possums.
- Designs with a large number of shorter trap-lines rather than designs with a small number of longer trap-lines are recommended when using the trap-catch protocol as a way to improve precision.
- Monitoring with only five trap-lines is not recommended. Estimates of confidence intervals cannot be considered reliable when they are based on small sample sizes such as five trap-lines.
- Bias-corrected bootstrap methods should be used in preference to the standard method to calculate confidence intervals for estimates of RTC.

2. Introduction

The University of Canterbury and Pest Control Research were contracted to investigate the residual-trap-catch (RTC) methodology for monitoring low-density possum populations. The research detailed in this report was carried out for the Animal Health Board between July 1999 and April 2000.

3. Background

3.1 Trap-catch protocol

The aim of the Animal Health Board (AHB) for possum (*Trichosurus vulpecula* Kerr) control is to reduce populations to levels that prevent the spread of bovine tuberculosis. To do this effectively it is necessary to determine accurately when control operations need to be undertaken, and how successful they have been. Often monitoring the success of the control operation determines whether possum control contractors have achieved the targets specified in their contracts or whether further control work is necessary.

The National Possum Control Authority (NPCA) has developed a standard protocol for monitoring possum population densities (NPCA 2000) in an attempt to achieve these requirements. The protocol defines a method to estimate possum densities using the proportion of leg-hold traps that capture possums over a predetermined number of trap-nights. Most commonly the estimate is calculated after possum control has been undertaken and this is termed the residual trap-catch or RTC. The protocol specifies how many traps are to be used per trap-line, where the traps are to be located in the field and how the capture data is to be analysed. The principal specifications are: the trap-lines are located randomly; the lines contain a specified number of traps; and the number of trap-lines is determined by the size of the area monitored.

Recently there have been concerns that the existing protocol may not be able to measure possum population size precisely enough when possum population densities are low. Specifically the main concern is that the sample size, i.e., the number of individual trap-lines, is not large enough to obtain a precise estimate for meaningful results. To help resolve this issue the protocol has recently been modified (NPCA February, 2000) to increase sample sizes by reducing the number of traps per line and increasing the number of trap-lines.

This project was undertaken to investigate the trap-catch protocol specifically to identify methods to further improve the precision of low-density population estimates. This would give the AHB and contracting agencies more confidence in decisions that specify control targets; control timing and measures of control contractor's success.

Three areas were identified where improvements could be made. These were:

1 Increasing the number of sample units, i.e., the number of trap-lines.
2 Analysing the survey data using different statistical procedures.
3 Investigating alternative sampling designs.

This project was funded by the AHB in two parts and both parts should be read in conjunction. Part II was undertaken by Landcare Research to provide a decision support system for possum control contracting agencies to help them interpret RTC estimates and their associated confidence intervals. It also includes an evaluation of alternative methods to estimate confidence intervals in addition to those investigated in this report.

3.2 Designing protocols for low-density populations

Two of the most frequently asked questions when monitoring population densities are how many sample units should be used and how large should the sample unit be? When using the trap-catch protocol the sample unit is the trap-line and this is used to estimate the sample mean.

To answer this question some background is needed. A design with less effort within sample units (i.e., fewer traps per line) but with many sample units (i.e., more trap-lines) is a spatially extensive design. Survey effort is concentrated among the many sample units. A design with a few sample units, i.e., more traps per line and fewer trap-lines e.g., a few, long trap-lines, is a spatially intensive design. Survey effort is concentrated within sample units.

The spatially extensive design will allow good coverage of the survey area, but will produce less reliable data from each sample unit. The spatially intensive design will have poor coverage of the survey area but will produce more reliable data from each sample unit.

Every animal population is different, but a number of wildlife studies have found that spatially extensive designs provide a more powerful monitoring tool for low-density populations than the intensive design (Roughton and Sweeney 1982, Millard and Lettenmaier 1986, Wilson and Weisburg 1993, , Link et al. 1994, J. Van der Meer 1997, Brown and Miller 1998, Hargreaves 1998). Deciding on the optimal design for monitoring low-density populations is a balance between the within- and among-sample unit efforts. It should also take into account the relative costs of collecting data from within- and among-sample units (Gates 1981).

3.3 Analysis of low-density population monitoring data

Analysis of data from surveys of animal populations can be complicated when there are many zero counts in the data, e.g., the data from a trap-line that catches no possums will be zero. Typically, low-density possum populations will have many zero counts. The same problem occurs with patchily distributed populations, e.g., a few trap-lines may have high catches while the rest may have low, or zero catches. Standard statistical analyses have been developed for normally distributed data but data that contain many zeros are not normally distributed.

Data from trap-lines are often zero when monitoring low-density possum populations so the data becomes skewed to the right. This skewness, combined with small sample sizes, can reduce the accuracy of the confidence interval of the estimated RTC when standard statistical methods are used. There are two options to help overcome this problem. These are:

1 Transform the data so it appears more normal.
2 Use statistical methods that allow for the non-normality, or skewness, in the data.

Option 1 is not generally recommended for data that contain many zeros, because of potential biases that can be introduced, and this was not investigated. We investigated Option 2 and concentrated our investigation on bootstrap methods for estimating confidence intervals for RTC estimates. Further statistical methods are evaluated in Part II.

Bootstrap methods for confidence intervals are computer-intensive techniques that repeatedly sample data. The distribution of the sample estimates is used to estimate the confidence interval. Computationally these techniques are more complex than the method currently used (i.e., $\bar{y} \pm t_{df,1-\alpha/2} \cdot SE$) but the widespread use of computers with suitable software, such as Traplog (McAuliffe 2000), makes them easier to use.

3.4 Confidence intervals and precision

The importance of understanding and using confidence intervals and how to interpret trap-catch data are covered more fully in Part II. However, for completeness, the concept of a confidence interval is briefly reviewed here.

When a population is monitored the survey results are used to estimate the population size. In the case of possum monitoring the estimate is the RTC index. There will always be some uncertainty about how accurately the RTC estimate reflects the true RTC. The degree of this uncertainty is measured in the confidence interval. For example, a 95% confidence interval is a statement that there is a 95% chance that the true RTC falls between the upper and lower confidence limits.

Two ways to improve the confidence, or precision, of the estimate are:

1 *To use large sample sizes.* A monitoring design that uses more trap-lines will give a more precise estimate than a design that uses fewer trap-lines. Consider a simplistic example where 100 traps are to be used to monitor possums. Ignoring for a moment the difference in survey effort, these 100 traps could be either laid out as 20 lines of 5 traps, or 5 lines of 20 traps. The design with 20 lines of 5 traps has a sample size of 20, while the design with 5 lines of 20 traps has a

sample size of 5 or 1/4 the sample size. The design with the larger sample size should have better precision.

2 *To have less sample variation.* Continuing with the above example, if the trap-lines contain too few traps, the comparative estimates from each line will be more variable. This will give high sample variation and will reduce precision.

A balance between the number of lines and the number of traps per line needs to be reached. On one hand a large sample size (i.e., many trap-lines) will improve precision, but if the lines are too short (i.e., too few traps per line) precision will be reduced. A study undertaken by Webster et al. (1999) retrospectively analysed data from Wellington Regional Council and the Department of Conservation. They found that the same precision could be achieved if 5 lines of 20 traps, 15 lines of 5 traps, or 9 lines of 10 traps were used. Therefore, to obtain better precision than five lines of 20 traps, more than 15 lines of 5 traps or 9 lines of 10 traps are required.

To find the optimal balance between lines of traps and traps per line the cost of the different survey design options needs to be considered. The cost of establishing and maintaining 5 lines of 20 traps may not be the same as the cost for 15 lines of 5 traps. Also the cost difference between these two designs may depend on habitat type.

4. Objectives

- To improve the precision of possum monitoring estimates by investigating the use of more trap-lines containing fewer traps and comparing their precision and cost-effectiveness at two field sites in the North and South islands.
- To improve the statistical analysis of possum monitoring data by investigating the use of bootstrap methods to provide more precise and accurate estimates of confidence intervals for the RTC index.
- To identify possible alternative sampling designs that could improve the precision and accuracy of low-density possum monitoring.

5. Methods

5.1 Field trials

Combinations of 5, 10, and 20 traps per line were used in the field trials. These combinations were chosen because these were the combinations recommended in the NPCA protocol and they are multiples of 100 traps. Five trap-lines were used for the lines containing 20 traps, 10 trap-lines for the lines containing 10 traps and 20 trap-lines for the lines containing 5 traps.

Two field trials were conducted: Tutukau, Central North Island in January 2000 and Hohonu, Westland, in March 2000. Tutukau contains low hills covered with mature pine forest and Hohonu is flat and covered with kamahi/podocarp/hardwood forest. These sites were used because they were recorded as having low-density possum populations with RTC levels that are typical for areas that have recently undergone possum control.

At each study site 300 traps were set using the three different line lengths. Because of the large number of trap-lines (i.e., 35), it was not possible to start the lines at random points as specified in the NPCA protocol. Therefore the start points were located no less than 40 m from a road or track and at least 200 m from each other as specified in the protocol for lines that are parallel to each other. The lines were located in five groups of seven trap-lines, i.e., within each group there were 1 line of 20 traps, 2 lines of 10 traps and 4 lines of 5 traps. This layout ensured that the different line lengths were interspersed within a restricted random pattern.

Traps were set and checked for 3 fine nights and the number of possums captured, the number of possum escapes, the number of sprung traps and the number of non-targets captured were recorded as specified by the protocol. In addition estimates of time to move along the line, to move between lines, and to set, check and remove traps were recorded. These estimates were used to calculate the costs of locating, checking and removing lines using the different line lengths.

The relative precision of the three monitoring designs was determined by the ratio of the width of the confidence interval to the RTC. The relative precision was compared among the three sample designs for data from each field trial.

The cost-efficiency data for each line length was calculated using the following formula,

$$Time = ltc_i + ltc_c + lc_a + l(t-1)c_w \,,$$

where there are l lines, t traps, c_i is the cost of installing and removing a trap, c_c is the cost of checking a trap, c_a is the cost of moving among trap-lines, and c_w is the cost of moving within trap-lines. Because the cost of checking a trap will differ between whether a trap catches a possum or not, we assumed the proportion of traps catching a possum was the same among the three designs and a constant, c_c, was used.

5.2 Comparison of methods to estimate confidence intervals

To compare bootstrap methods for estimating confidence intervals, a model to simulate trap-catch data was used to calculate exact confidence limits and RTC's. A negative binomial distribution was used to simulate skewed data that is characteristic of trap-catch results. Comparing it with the actual trap-catch data collected from the field trials checked the suitability of the simulated data.

The model used simulated trap-catches for a range of low-density populations, i.e., 1.67%, 3.33%, 5.00% and 6.67% RTC. The survey designs were the same as those used in the field trial i.e., 5 lines of 20 traps; 10 lines of 10 traps; and 20 lines of 5 traps. To extend the range of the simulations an additional design with 25 lines of 4 traps was also used.

For each combination of survey design and target RTC 1000 sets of trap-catch data were simulated. Estimates of 95% confidence intervals were calculated for each data set using a range of methods. These included the standard method,

$$rtc \pm t_{df,1-\alpha/2} \cdot SE$$

and seven bootstrap methods:

1. The standard bootstrap method (Manly 1997, p. 34),
2. The first-percentile method (Manly 1997, p. 39, Efron 1979),
3. The second-percentile method (Manly 1997, p. 41, Hall 1992),
4. The bias-corrected percentile method (Manly 1997, p. 44, Efron 1981),
5. The accelerated bias-corrected method (Manly 1997, p. 49, Efron 1987, Efron and Tibshirani 1986),
6. The bootstrap-t method (Manly 1997, p. 56), and,
7. Hall's Bootstrap-t transformation method (Manly 1997, p. 59, Fletcher and Webster 1996, Hall 1992).

The suitability of the bootstrap methods was determined by comparing them with each other using four measures. These were:

1. *The bias of the confidence limits.* The bias was calculated as the difference between the true percentile limit and the estimated limit. True percentile limits of the sample mean were calculated using a negative binomial probability generating function.
2. *The mean square error* (MSE). The MSE is the sum of the square of the bias and the variance and can be considered an overall measure of accuracy and precision of the confidence limits. The smaller the MSE, the better the accuracy and precision.
3. *The coverage of the confidence interval.* Coverage is the proportion of the confidence intervals where the true population mean was between the upper and lower confidence limits. With 95% confidence intervals that have ideal coverage, the true population mean should be between the upper and lower limits 95% of the time.
4. *The balance of the confidence interval.* The balance is the proportion of times the true mean falls above the upper limit and below the lower limit. With 95% confidence intervals that have ideal coverage *and* balance, the true population mean will fall above the upper limit 2.5% of the time, and below the lower limit 2.5% of the time.

6. Results

6.1 Field trials

6.1.1 Relative precision

The estimates of RTC for the three monitoring designs at Tutukau and Hohonu were all within the range of what would be considered low-density possum populations. The lowest RTC estimate was 3.67% from the 5 lines of 20 traps at Hohonu and the highest was 7.07% from the 10 lines of 10 traps at Tutukau (Table 1). As expected the relative precision was best (i.e., lowest) at both study sites for the designs that had the largest sample size, i.e., the 20 lines of 5 traps. Conversely the relative precision was worst (i.e., highest) for the designs that had smaller sample sizes, i.e., the five lines of 20 traps at Tutukau, and the ten lines of ten traps at Hohonu (Table 1).

Table 1. *Estimates of* RTC*, their confidence intervals' and relative precision of three monitoring designs at two locations.*

		RTC %	Upper CI	Lower CI	Relative precision
Tutukau	5 lines, 20 traps	5.00	8.87	1.13	1.55
	10 lines, 10 traps	7.07	11.89	2.24	1.37
	20 lines, 5 traps	5.07	7.73	2.40	1.05
Hohonu	5 lines, 20 traps	3.67	5.93	1.40	1.24
	10 lines, 10 traps	6.73	11.61	1.86	1.45
	20 lines, 5 traps	6.53	10.04	3.03	1.07

6.1.2 Cost-efficiency

Times to install and remove traps (c_i) and check traps (c_c) were similar at both study sites (i.e., c_i = 2.00 min and c_c = 0.15 min) when the average of all the time estimates were compared. Times to move among lines (c_a) and within lines (c_w) were longer at Hohonu than at Tutukau. This was attributed to the denser forest understory present at Hohonu. The average times were c_a =5.88 min and c_w = 0.59 min for Tutukau, and c_a = 7.15 min and c_w = 0.71 min for Hohonu.

Total time estimates for the three trap-line designs varied from 300.45 to 379.80 min at Tutukau and 319.15 to 415.60 min at Hohonu. As expected the design with the 5 lines of 20 traps had the least time expended while the design with 20 lines of 5 traps had the most (Table 2).

Table 2. *Cost estimates (time in minutes) for three monitoring designs at two locations.*

		Cost estimate (min)
Tutukau	5 lines, 20 traps	300.45
	10 lines, 10 traps	326.90
	20 lines, 5 traps	379.80
Hohonu	5 lines, 20 traps	319.15
	10 lines, 10 traps	351.30
	20 lines, 5 traps	415.60

At Tutukau the proportional extra time to survey more, shorter lines compared with surveying 5 lines of 20 traps was similar to the gain in precision: the design with 20 lines took 27% longer than the design with 5 lines and there was a 33% improvement in relative precision; and the design with 10 lines took 9% longer than the design with 5 lines and had a 12% improvement in relative precision.

However, at Hohonu the proportional extra time required to survey more lines was not matched with the proportional gain in relative precision. The design with 20 lines took 33% longer than the design with 5 lines but had only a 14% improvement in relative precision. The design with ten lines took 32% longer than the design with five lines and had a 17% loss in relative precision. (Figure 1).

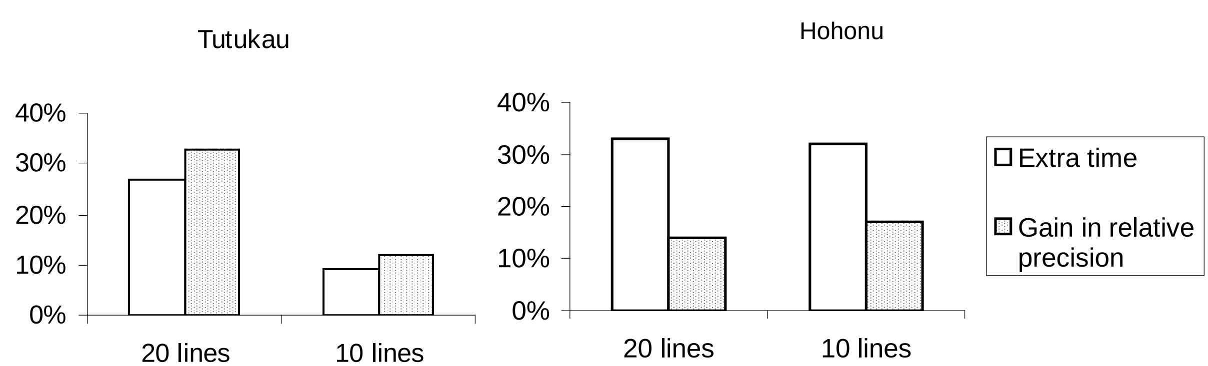

Figure 1. *Proportional extra time to survey more, shorter lines and proportional gain in relative precision for designs with 20 and 10 lines compared to design with 5 lines of 20 traps at the two field sites.*

6.2 Comparison of methods to estimate confidence intervals

6.2.1 Simulation model

There was no evidence that data simulated from the negative binomial model did not provide a good fit when compared to the real trap-catch data from Tutukau and Hohonu. Five of the six sets of field data were well simulated by the negative binomial model. There was no evidence of lack of fit for the three data sets from Tutukau i.e., 5 lines of 20 traps (χ^2 = 1.76, P = 0.63), 10 lines of 10 traps (χ^2 = 2.21, P = 0.53) and 20 lines of 5 traps (χ^2 = 2.69, P = 0.44). The Hohonu data also showed no evidence of lack of fit for the design with 10 lines of 10 traps (χ^2 = 0.23, P = 0.97) and 20 lines of 5 traps (χ^2 = 3.01, P = 0.39). The design with 5 lines of 20 traps at Hohonu had trap-catch data that was different from the data modelled, (χ^2 = 8.73, P = 0.03). This difference was attributed to these lines having no lines of zero possum catch while the model predicted that, on average, 1.5 lines would have a zero catch. While not perfect, the negative binomial model was considered a suitable model to use to simulate low-density possum trap-catch data.

Three parameters were set in order to use the negative binomial model to simulate trap-catch data. The shape parameter *r* was fixed at 1 and the values of the other two parameters *p* and *n* were varied to simulate line trap-catch over 3 nights for the four different population mean *RTC* values and four different numbers of lines (Table 3).

Table 3. Values of the negative binomial parameter p used to simulate line trap-catch data over 3 nights for four RTC *and different survey designs. The sample size is the number of lines. The total number of traps is fixed at 100. The parameter r is fixed at 1.*

	Number of lines			
Mean RTC	5	10	20	25
1.67%	0.50	0.67	0.80	0.83
3.33%	0.33	0.50	0.67	0.71
5.00%	0.25	0.40	0.57	0.63
6.67%	0.20	0.33	0.50	0.56

6.2.2 Comparison of bootstrap methods

The accelerated-bias-corrected percentile (method 5), the bootstrap-t (method 6) and Hall's bootstrap-t transformation (method 7) were not suitable for estimating trap-catch confidence intervals because the elements in the simulated data were often identical and had zero variance. These methods involve dividing by the variance, or average deviation, and this is not possible with a zero variance. One approach to deal with this problem would be to ignore the bootstrap sample that has identical elements. However, when using small sample sizes this occurs frequently and would bias the results.

Of the other bootstrap methods, we found that the bias-corrected percentile method (BCP, method 4) gave better results compared with methods 1, 2 and 3 (see Appendix 1 for how to calculate the BCP). The BCP method gave the lowest MSE, the best coverage and was the most balanced. Therefore we compared the results from this method with the standard method currently used in the trap-catch protocol. Full results of the comparisons using the other methods are given in Appendix 2.

6.2.3 Comparison of MSE

The mean square error (MSE) using the BCP method was smaller compared with the standard method for both the upper (Figure 2) and lower confidence limits (Figure 3) when using the five-trap-line design. The lower confidence limit had the largest difference in MSE between the standard and BCP methods (Figure 3). There was little difference in MSE between the standard and BCP methods for the designs that had 10 trap-lines or greater. The MSE decreased as the number of trap-lines increased especially when increasing from 5 trap-lines to 10 trap-lines (Figures 2 and 3). There was very little reduction in MSE when increasing from 20 to 25 trap-lines. Also the MSE increased as the average RTC increased for both the BCP method and the standard methods.

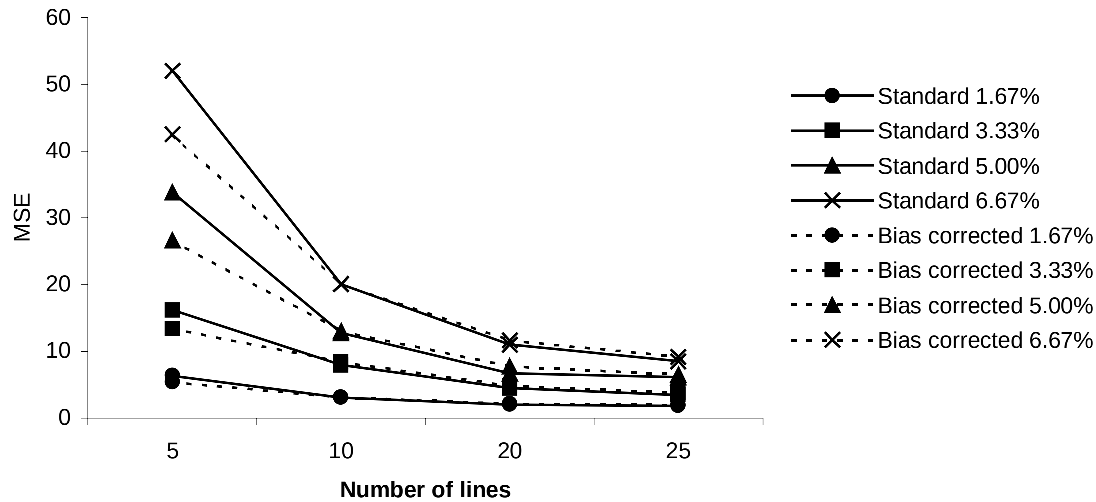

***Figure 2**. MSE of the upper confidence limit with four levels of RTC calculated using the standard and bias-corrected confidence interval methods. The MSE is lower, or better, for the bias-corrected method. As the number of lines increase, the MSE decreases for both methods.*

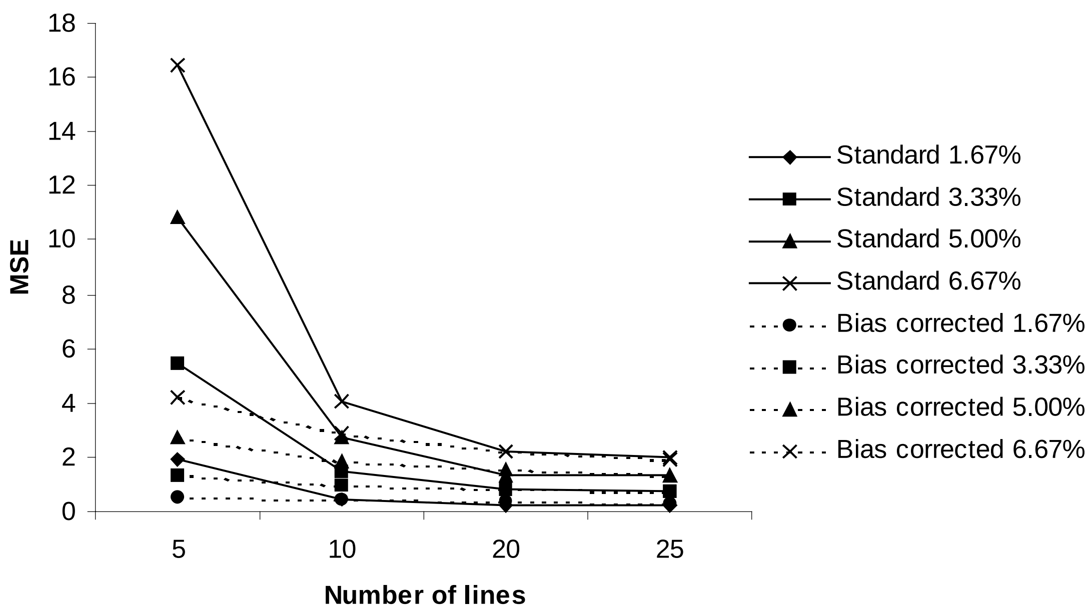

***Figure 3.** MSE of the lower confidence limit with four levels of RTC for the standard and bias corrected confidence intervals.*

6.2.4 Coverage of the confidence intervals

The coverage of the confidence interval (i.e., the proportion of the times the true RTC fell within the confidence intervals) improved as the number of lines increased for both the methods. Coverage was below the ideal 95% for most of the designs (Figure 4). By definition a 95% confidence interval should include the true mean on 95% of occasions. Only the BCP method achieved this and only then with some of the designs that had more than 20 lines (Figure 4). When 5 lines were used, coverage was below 90% for both methods and all RTC levels.

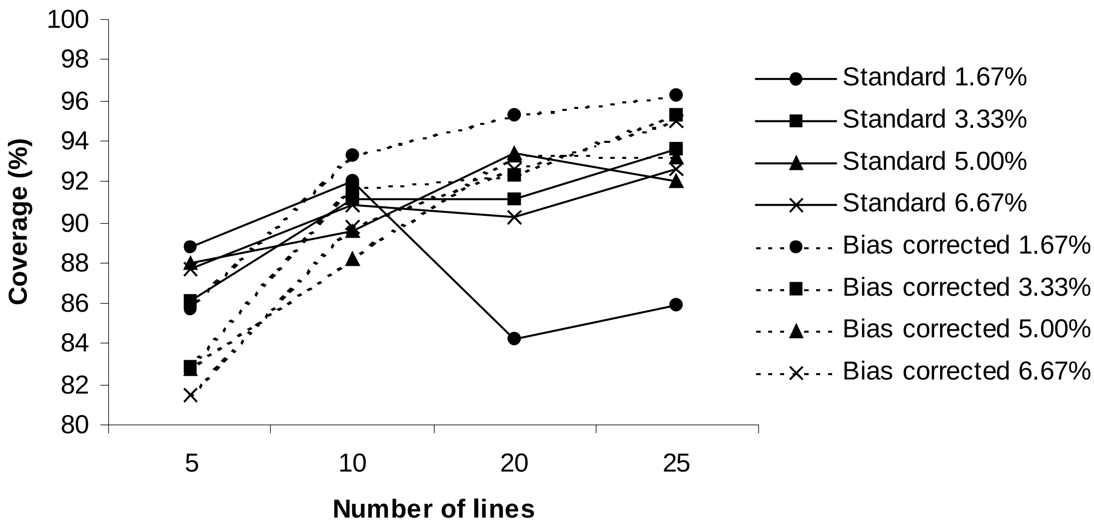

Figure 4. *Total coverage with four different designs, and four levels of* RTC *for the standard and bias-corrected confidence intervals. Total coverage should be 95%, i.e., the true mean should fall within the 95% confidence interval 95% of the time.*

6.2.5 Balance of the confidence intervals

Both the BCP method and standard methods failed to provide balanced coverage of the confidence intervals (i.e., the simulated RTC did not fall above and below the upper and lower confidence intervals the same number of times). When the standard method was used the simulated RTC's that were outside the confidence intervals were almost all (90%) above the upper limit (Figure 5). The confidence intervals from the BCP method were more balanced than the standard method. With this method when the simulated RTC's were outside the confidence interval they fell above the upper limit between 52% and 83% of the time.

The lower limit of the confidence intervals calculated using the BCP method was never negative or zero. However, the confidence intervals calculated using the standard method gave, on average, negative lower limits for the design with 5 lines with an RTC of 1.67%, 3.33%, 5.00% and 6.67% and for the design with 10 lines with an RTC of 1.67%.

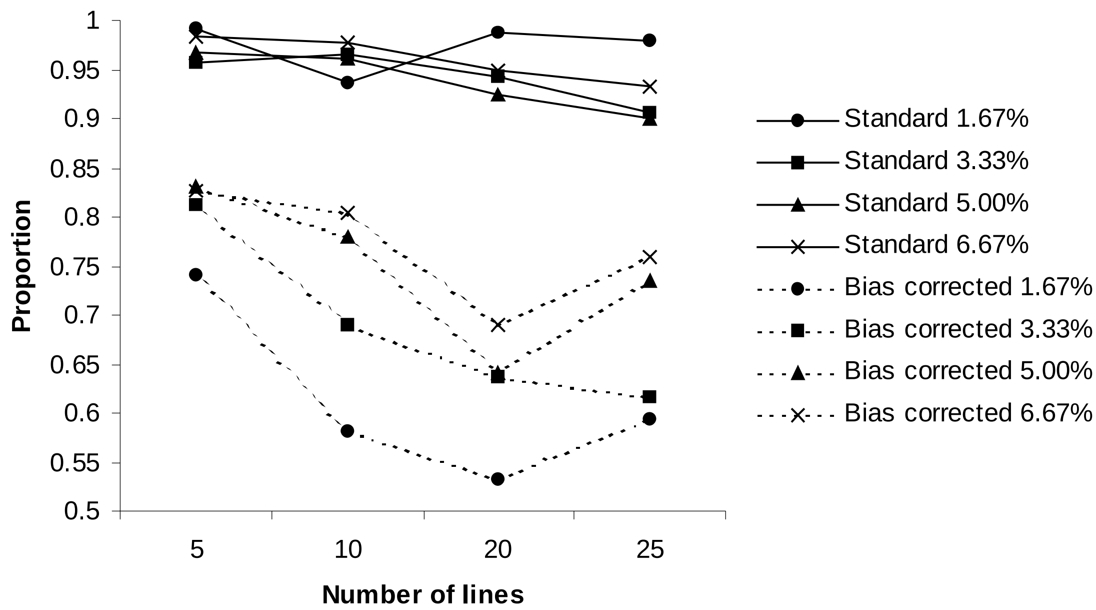

Figure 5. *The proportion of times the true mean fell above the upper confidence limit (when it was outside the confidence interval) using four levels of* RTC. *The proportion should be 0.5.*

7. Discussion and Conclusions

7.1 Field trials

The relative precision of the three monitoring designs i.e., 5, 10, and 20 lines, increased when more, shorter trap-lines were used. As expected, the cost of the surveys increased as the number of lines increased, e.g., the design with the highest relative precision (i.e., 20 lines of 5 traps) also had the highest cost in both field trials.

There are many advantages of having a more precise estimate of RTC, the most important being that the calculated success of the control operation can be more confidently relied on to be the correct result. While it is important to focus on how to improve precision by using more survey lines, this needs to be balanced against the increased cost of using these designs. To achieve an equivalent cost to using the standard 5 lines of 20 traps, 16 lines of 5 traps or 9 lines of 10 traps would have to be used. These designs would have to give more precise estimates otherwise there would be no advantage gained in using them. The importance of having accurate and reliable estimates of RTC needs to be stressed. It may be worth accepting higher monitoring costs if more accurate and reliable estimates of residual population size are to be realised.

These combinations of lines and traps are similar to the recommendations made by Webster et al. (1999) where designs with similar expected precision were considered. When 3 nights of trapping is undertaken Webster et al. recommend that 15 lines of 5 traps or 9 lines of 10 traps have similar precision to 5 lines of 20 traps.

One of the factors in determining the cost of a survey is the distance lines are apart. This is difficult to determine. Each survey is different because of the type of terrain and the random location of the lines. The costs of the surveys undertaken in this study can be used as a guide to the relative costs of these alternative designs.

The differences in the RTC among the three designs used in the field was unexpected, especially those recorded at Hohonu. The trap-lines for each design were interspersed between each other so similar RTC estimates were expected. However, at Hohonu the design using 5 lines of 20 traps recorded an RTC of 3.67% while the design using 10 lines of 10 traps recorded an RTC of 6.73%. Unlike the simulated data used to generate RTC's, the field data could not be used to determine which RTC estimate was the most accurate. The confidence interval for the estimates overlapped and there was no statistical evidence of a difference in RTC among the field designs, but considering that the same population was measured at the same time, this difference is hard to reconcile and the accuracy of RTC should be quetioned.

7.2 Estimation of confidence intervals using bootstrap methods

The three more complex bootstrap methods to estimate confidence intervals, i.e., the accelerated bias-corrected percentile, the bootstrap-t, and the Hall's bootstrap-t transformation were expected to provide more accurate confidence intervals than the other methods used. These methods are not recommended for low-density trap-catch data because of the problems with zero variance. Removing any data sets with zero variance and or adjusting for bias could allow these methods to be used, but this adds a degree of complexity to the analysis.

Of the other bootstrap methods, the BCP performed the best. The BCP method, along with the standard method for estimating confidence intervals had the lowest MSE and achieved the best coverage in terms of the proportion of times the true mean was included in the confidence interval. However, the coverage of the 95% confidence intervals was not 95% for either method.

Theoretically a 95% confidence interval is expected to include the true RTC 95% of the time. In practice we can expect that for every 20 control operations monitored, one (5%) will have the true RTC outside the confidence interval if there was correct coverage. The true RTC that falls outside the limit will either be higher than the upper limit, or lower than the lower limit of the confidence interval. This "error-rate" is what is expected, even if it is not explicitly stated or fully understood by the possum control industry. However, even with the two methods that gave the best coverage, (i.e., the BCP and the standard method), the actual coverage of the 95% confidence intervals was about 90% (Figure 4). Therefore in reality there is a 1/10 chance that the true RTC will fall outside of a confidence interval rather than the theoretical 1/20 chance (Figure 6). On average, for every 10 control operations monitored, one of these is likely to have a true RTC that is either higher than the upper limit, or lower than the lower limit of the confidence interval.

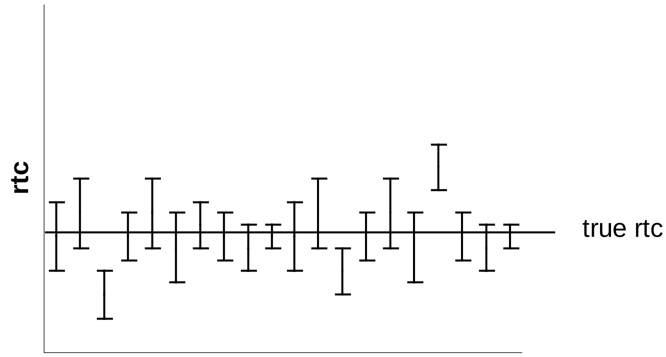

Figure 6. *Theoretical* RTC *estimates from 20 control operations that are monitored with 95% confidence intervals calculated and the true* RTC *is known. Two of these, or 10%, do not include the true* RTC *(i.e, they have a 90% coverage). If the confidence intervals had the correct coverage (i.e.95%) then only one, or 5%, would be expected to not include the true* RTC*.*

The next question is when there is a confidence interval that does not include the true RTC, will the true RTC tend to be on the high side or on the low side of the confidence interval? This question concerns the balance of the confidence interval. For an even balance the true RTC should have an equal chance of falling above or below the confidence interval. Figure 6 shows a balanced example where the true RTC fell above the interval and below the interval. Figure 7 gives an example of an unbalanced confidence interval where the true RTC falls above the upper limit for two of the confidence intervals. In these two cases the residual population levels are higher than the results suggest.

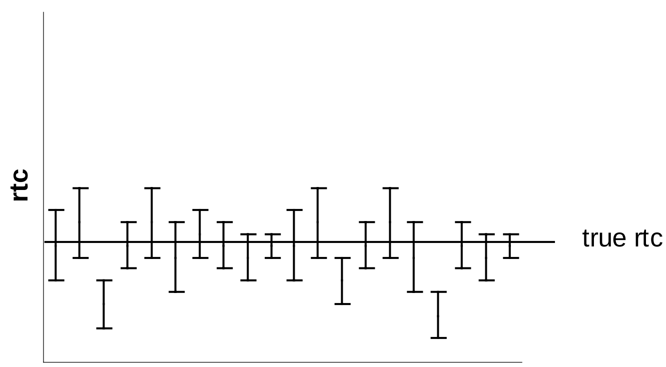

Figure 7. *Theoretical* RTC *estimates from 20 control operations that are monitored with 95% confidence intervals calculated and the true* RTC *is known. Two of these do not include the true RTC. However, this example is unbalance because the true* RTC *falls above both confidence intervals rather than above and below as in Figure 6.*

The standard method for calculating confidence intervals did not produce balanced confidence intervals because for the 10% of the time that the true RTC fell outside of their range it fell above the confidence intervals in 90-100% of the cases (Figure 5). Therefore the 10% of control operations that have "incorrect intervals" have a 90–100% chance of indicating that there is a lower possum population density than is actually the case. Although this may be good for contractors, it may not be so good for reducing Tb levels.

The reason for the occurrence of an imbalance of confidence intervals is because trap-catch data collected from low-density possum populations is commonly skewed. This occurs because there are often lines with either no possums or very few possums captured. The BCP method did not give an ideal balance but it did provide more balanced results than the standard method. When the BCP method was used, the true RTC fell above the confidence interval approximately 70% of the time.

One of the uses of RTC estimates and their associated confidence interval is to decide whether a possum control operation is successful. If it is deemed successful, then further control is not often specified. In cases where the true RTC actually falls outside the confidence interval, it is most likely to occur above the confidence interval, and a repeat control operation may be necessary but would not be identified. Another use of RTC estimates is to set a target of possum density that will prevent the transmission of Tb from possums to domestic stock. If the results from monitoring operations are misleading, i.e., the true RTC is higher than the confidence interval suggests, then management decisions based on RTC target levels need to be more robust by considering this uncertainty.

Another advantage of the BCP method is that the lower limit of the confidence interval is never negative. The standard method gave negative lower limits using the simulated data for the designs using five trap-lines. In practice when the lower limit is negative it can be reported as a 0% RTC. This is a meaningless measure because if at least one possum were caught in a trap, then the true RTC cannot be 0%.

Overall, this study has indicated that the BCP method was the best of the bootstrap methods used in this study for estimating confidence intervals for RTC when using data collected from low-density possum populations. The method was the most likely to include the true RTC in the correct proportion (i.e., 95% of the time). It also produced more balanced confidence intervals and the lower limit cannot be negative. The BCP method for confidence intervals has been incorporated into Traplog (McAuliffe 2000) as a user specified option.

Regardless of the method used, the precision and accuracy of confidence intervals improved if the number of trap-lines was increased. Monitoring with more, shorter lines is preferable to monitoring with fewer longer lines. However no reliable method to estimate confidence intervals was found when only five trap-lines were used. Therefore possum monitoring with only five trap-lines should is not recommended.

7.3 Alternative survey designs

The discussion so far has focused on modifying the existing trap-catch protocol by using different numbers of traps along lines, different numbers of lines, and different methods to estimate confidence intervals. If control agencies are targeting very low possum densities e.g., below 2% RTC, it is appropriate to consider alternative sampling methods. When low-density populations are monitored it is difficult to reliably estimate RTC because the chance of catching a possum is small. The challenge is to design a survey where, if no possums are caught, there is a high probability that there are no possums present rather than there being possums present but not being able to be detected (Brown and Boyce 1996). We discuss two possible design improvements that could improve the chance of detecting possums in very low-density populations using traps.

7.3.1 Line placement

An advantage of using a design that has a larger sample size, (i.e., more trap-lines), is that the sample can be more spatially extensive. This potential advantage can be lost if the lines do not give coverage of the study area but are clustered in one location. The use of random line placement has this potential risk. The recommended minimum 200-m spacing between lines and recommendations for stratifying the area in the NPCA protocol goes some way to avoid possible clustering. However, if the strata are too large, spatial coverage may not be achieved using random line placement.

Better spatial coverage of trap-lines could be achieved using stratification on a finer scale than is currently used and have, for example, one or two lines per stratum. Another method is to systematically locate trap-lines so that there is even spatial coverage over the study area or, if a stratified design is used, even spatial coverage within each stratum (Simmonds and Fryer 1996, Gilbert 1987, p. 89, Ratti and Garton 1980). There are many variations on systematic sampling, but in the simplest design, trap-lines would be more regularly spaced over the study area or stratum.

Locating lines systematically has been discussed within the industry for some time and there has been criticism that the variance of systematic sampling cannot be properly estimated when the sample is analysed as a simple random sample. This would be true if the possum populations have a periodic pattern that matches the interval of the systematic sample, e.g., if there were patches of high possum abundance every 400 m and lines were placed at 400-m intervals, or if there was a population density trend, e.g., a gradient from high to low possum abundance (Lohr 1999). The variance of systematic sampling will then be either under- or over-estimated respectively. If there is no pattern or trend, the use of the simple random sampling formulae will give an unbiased estimate of variance. In practice a number of studies in ecology have found that systematic sampling gives more precise survey estimates than simple random sampling (e.g., Simmonds and Fryer 1996, Skalski et al. 1992, Kraft et al. 1995).

A further advantage of systematic sampling is that it is often logistically easier to undertake in the field (Peshkova 1970, Gilbert 1987, p. 89, Ratti and Garton 1980). With random sampling the field worker has to locate the next start point for the line. They have to know where they are when they finish the preceding line to be able to move to the next-closest line. In systematic sampling there are less decisions to make in the field – the lines are all equally spaced apart. Knowing the exact location of the preceding line is not so critical when the next line is a fixed interval away. When there are fewer decisions there is less chance of making the wrong decision (Ratti and Garton 1980).

Another reason why systematic sampling can have fewer survey errors than simple random sampling is that systematic sampling ensures that no line is too far from any other line. Consider the situation where within a stratum there are six lines to survey and these have been randomly placed. Five of these lines are within 300 m of each other and the sixth is 4 km away. Even with perfect training, human nature is such that chances are the sixth line is not placed where it should be and is placed closer than 4 km away. No one would admit to doing this but it does happen in practice. With systematic sampling this situation would be less likely to occur because no one line would be located a long distance from another.

7.3.2 Alternative monitoring methods

One of the difficulties with monitoring very low density populations is the problem in dealing with discrete numbers. This is best explained by example. Consider two possum populations, one at high and one at low density. The high-density population has an RTC of 20%. Over 3 nights 100 traps caught 60 possums. However, if there had been 61 possums caught, the RTC would have been 20.33% or a 1.67% increase. The low density population has an RTC of 2%. Over 3 nights 100 traps caught six possums. Now consider the effect of catching one more possums. If 7 possums were caught the RTC increases from 2% to 2.33% – a 16.67% increase. Consequently the effect of catching one or two more possums when monitoring low densities is larger than when monitoring high densities. The low-density RTC estimate is more "sensitive" to small differences in trap-catch rates. This raises the question whether counting numbers or catches is the best way to monitor low densities?

An alternative monitoring design for very low-density populations is to determine the proportion of an area that has possums present rather than trying to measure their abundance. Possum trap-catches are spatially aggregated. When one possum is caught in a trap, there is an increased chance of capturing another possum in an adjacent trap (Faddy *et al.* 2000). Therefore catching two possums compared with one may not necessarily mean that the population is twice as large. The count of trapped possums, at very low density, may be less informative than a measure of the proportion of the area where possums are present. There are various ways to design a monitoring scheme to detect such proportions (Brown and Boyce 1996, Thomas and Abery 1995) and these may be viable and effective alternatives to estimating residual populations for targeting very low density populations.

If monitoring were to focus on detecting the proportion of the area where possums are present, rather than estimating possum numbers from trap-catch rates, it may be more effective to use alternative monitoring devices that measure the presence or absence of possums. When monitoring data are presence/absence data rather than count data it is not necessary to mark or remove possums, which makes it less labour-intensive. Also the alternative monitoring devices could be smaller and more compact allowing more to be used in the field, which would provide larger sample sizes and better spatial coverage. In addition the devices would not need to be checked daily, which could substantially reduce monitoring cost.

8. Recommendations

- RTC estimates for low-density possum populations calculated using the trap-catch protocol should not be used on their own without reference to their associated confidence intervals. Basing decisions to pay contractors on specified RTC levels without reference to confidence intervals is not recommended.
- Improvements to the trap-catch protocol, (e.g., using more complex techniques to analyse data and estimate confidence intervals, using shorter lines with fewer traps, and systematically locating trap-lines) will not address the underlying problems associated with small sample sizes. Consideration needs to be given to investigating alternative sampling methods that will provide larger sample sizes and to using systematic sampling designs.
- The proportion of an operational area that contains possums is a more informative measure than the proportion of traps that capture possums when measuring very low - density populations. To estimate the proportion of an area that contains possums, the animals do not need to be caught in leg-hold traps and removed. This allows the use of more lightweight monitoring devices that record the presence of possums only, rather than traps that record the presence and numbers of possums.
- Designs with a large number of shorter trap-lines rather than designs with a small number of longer trap-lines are recommended when using the trap-catch protocol as a way to improve precision.
- Monitoring with only five trap-lines is not recommended. Estimates of confidence intervals cannot be considered reliable when they are based on small sample sizes such as five trap-lines.
- Bias-corrected bootstrap methods should be used in preference to the standard method to calculate confidence intervals for estimates of RTC.

9. Acknowledgements

We thank Phillip Commins and John Hughes for assisting with fieldwork and providing helpful suggestions. Thank you to Ray Webster (Landcare Research) and Nigel Sinclair (University of Canterbury) for their assistance to investigate alternative methods to calculate confidence intervals and to Ray for his useful comments on an earlier draft. Christine Bezar provided editorial assistance.

10. References

Brown, J.A. and Boyce, M.S. (1996), Monitoring of Karner blue butterflies (*Lycaeides melissa samuelis*) for the proposed habitat conservation plan, Wisconsin. Report to the US National Fish and Wildlife Foundation.

Brown, J.A. and Miller, C.M. (1998), Monitoring stoat *Mustela erminea* control operations: power analysis and design. *Science for Conservation*: 96, Department of Conservation, Wellington

Efron, B. (1979), Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7, 1-26.

Efron, B. (1981), Nonparametric standard errors and confidence intervals. *Canadian Journal of Statistics*, 9, 139-72.

Efron, B. (1987), Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82, 171-85.

Efron, B and Tibshirani, R. (1986), Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1, 54-77.

Faddy, M.J., Brown, J.A. and Commins, P., (unpublished), Likelihood based modelling of possum trapping data, submitted to *Journal of Agricultural, Biological, and Environmental Statistics.*

Fletcher D. and Webster R. (1996), Skewness-adjusted confidence intervals in stratified biological surveys, *Journal of Agricultural, Biological, and Environmental Statistics*, 1**,** 120-130.

Gates, C.E. (1981), Optimizing sampling frequency and numbers of transects and stations. *Studies in Avian Biology,* 6; 399-404.

Gilbert, R.O. (1987), *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.

Hargreaves, L. (1998), Power analysis of possum monitoring techniques. Department of Mathematics and Statistics, University of Canterbury.

Kraft, K.M., Johnson, D.H., Samuelson, J.M, and Allen, S.H. (1995), Using known populations of pronghorn to evaluate sampling plans and estimators. *Journal of Wildlife Management*, 59; 129-137.

Link, W.A., Barker, R.J., Sauer, J.R. and Droege, S. (1994). Within-site variability in surveys of wildlife populations. *Ecology*, 75;1097-1108.

Lohr, S. L. (1999), *Sampling: Design and Analysis*. Duxbury Press, Pacific Grove.

McAuliffe, R.J. (2000), Traplog. Russel McAuliffe, "The Grotto", Days Rd, RD 4, Chch.

Manly B. F. J. (1997), *Randomization, Bootstrap and Monte Carlo Methods in Biology*, Chapman and Hall, London.

Millard, S.P and Lettenmaier, D.P. 1986. Optimal design of biological sampling programs using analysis of variance. *Estuarine, Coastal and Shelf Science,* 22; 637-656.

National Possum Control Authority, (2000), *Trap-catch for Monitoring Possum Populations*, NPCA, Wellington, New Zealand.

Peshkova, N.V. (1970), Comparison of results of determinations of the density of pant populations by systematic and by restricted random sampling procedures. *Ekologiia,* 2; 180-181.

Ratti, J. T. and Garton, E.O (1980), Research and experimental design. In: *Research and Management Techniques for Wildlife and Habitats*, eds, T.A. Bookhout, The Wildlife Society USA., pp. 1-23.

Roughton, R.D. and Sweeney, M.W. (1982), Refinements in scent-station methodology for assessing trends in carnivore populations. *Journal of Wildlife Management,* 46; 217-229.

Simmonds, E.J. and Fryer, R.J. (1996), Which are better, random or systematic acoustic surveys? A simulation using North Sea herring as an example. *ICES Journal of Marine Science,* 53; 39-50.

Skalski, J.R., Hoffmann, A., Ransom, B.H., and Steig, T.W. (1992), Fixed-location hydroacoustic monitoring designs for estimating fish passage using stratified random and systematic sampling. *Canadian Journal of Fisheries and Aquatic Sciences*, 50; 1208-1221.

Thomas, C.D. and Abery, J.C.G. (1995), Estimating rates of butterfly decline from distribution maps: the effect of scale. *Biological Conservation,* 73; 59-65.

Van der Meer, J. (1997), Sampling design of monitoring programmes for marine benthos: a comparison between fixed versus random selected stations. Journal of Sea Research, 37; 167-179.

Webster, R.A., Ruscoe, W.A. and Warburton, B. (1999), Optimising the trap-catch method for monitoring possum populations. Landcare Research Contract Report: LC9899/141. Report to the Animal Health Board, Unpublished.

Wilson, H.T. and Weisburg, S.B. (1993), Design considerations for beach seine surveys of striped bass. *North American Journal of Fisheries Management*, 13; 376-382.

11. Appendices

11.1 Calculation of the Bias-Corrected Percentile (BCP) Method

Bias corrected percentile method (Manly 1997, p. 44, Efron 1981). This method involves calculating a value *p*, which is the proportion of bootstrap estimates where $RTC_b > RTC$. The lower confidence limits is the $100(\phi(2z_0 - z_{\alpha/2}))^{th}$ percentile of the bootstrap distribution of estimates, RTC_b and the $100(\phi(2z_0 + z_{\alpha/2}))^{th}$ percentile is the upper limit, where z_0, the value from the standard normal distribution that is exceeded with probability *p* and $\phi(2z_0 - z_{\alpha/2})$, the proportion of the standard normal distribution less than $2z_0 - z_{\alpha/2}$.

11.2 Comparisons of bootstrap methods

			Average CI		Standard Deviation		MSE		Coverage (%)		
Method	E[RTC]	# Lines	Lower	Upper	Lower	Upper	Lower	Upper	Lower (2.5)	Upper (2.5)	Total (95)
Standard	1.67%	5	-0.90	4.22	0.962	2.512	1.918	6.312	0.1	11.1	88.8
		10	-0.19	3.55	0.545	1.736	0.457	3.058	0.5	7.5	92.0
		20	0.03	3.29	0.432	1.392	0.242	1.995	0.2	15.6	84.2
		25	0.07	3.21	0.422	1.326	0.222	1.841	0.3	13.8	85.9
	3.33%	5	-1.10	7.66	1.620	4.021	5.487	16.170	0.6	13.3	86.1
		10	0.18	6.58	0.958	2.810	1.445	7.919	0.3	8.6	91.1
		20	0.74	5.93	0.795	2.090	0.793	4.442	0.5	8.4	91.1
		25	0.88	5.86	0.788	1.860	0.717	3.472	0.6	5.8	93.6
	5.00%	5	-1.28	11.36	2.260	5.805	10.828	33.809	0.4	11.6	88.0
		10	0.72	9.14	1.351	3.540	2.706	12.710	0.4	10.0	89.6
		20	1.62	8.62	1.092	2.581	1.364	6.666	0.5	6.1	93.4
		25	1.76	8.29	1.081	2.457	1.332	6.083	0.8	7.2	92.0
	6.67%	5	-1.31	14.56	2.784	7.215	16.439	52.050	0.2	12.1	87.7
		10	1.32	12.04	1.677	4.452	4.059	20.035	0.2	9.0	90.8
		20	2.45	10.99	1.379	3.296	2.222	10.983	0.5	9.3	90.2
		25	2.65	10.63	1.311	2.891	1.994	8.468	0.5	6.9	92.6
Bias Corrected	1.67%	5	0.58	3.90	0.523	2.299	0.501	5.351	3.7	10.6	85.7
		10	0.60	3.85	0.507	1.756	0.409	3.092	2.8	3.9	93.3
		20	0.64	3.82	0.464	1.428	0.357	2.119	2.2	2.5	95.3
		25	0.64	3.77	0.442	1.352	0.330	1.901	1.5	2.2	96.3
	3.33%	5	1.23	6.82	0.967	3.568	1.349	13.403	3.2	13.9	82.9
		10	1.36	6.79	0.886	2.896	0.993	8.393	2.6	5.8	91.6
		20	1.54	6.43	0.810	2.173	0.814	4.778	2.8	4.9	92.3
		25	1.57	6.35	0.779	1.895	0.750	3.724	1.8	2.9	95.3
	5.00%	5	1.96	9.96	1.416	5.050	2.724	26.613	2.9	14.3	82.8
		10	2.20	9.28	1.245	3.592	1.851	12.983	2.6	9.2	88.2
		20	2.60	9.14	1.122	2.734	1.577	7.706	2.4	4.3	93.3
		25	2.56	8.76	1.088	2.533	1.337	6.486	1.8	5.0	93.2
	6.67%	5	2.66	12.68	1.783	6.244	4.220	42.502	3.2	15.3	81.5
		10	3.10	12.06	1.561	4.456	2.881	20.050	2.0	8.2	89.8
		20	3.57	11.52	1.391	3.403	2.244	11.616	2.3	5.1	92.6
		25	3.55	11.10	1.324	3.022	1.891	9.151	1.2	3.8	95.0
Standard Bootstrap	1.67%	5	0.05	3.27	0.622	1.943	0.390	4.561	1.7	18.8	79.5
		10	0.15	3.22	0.532	1.588	0.287	2.821	1.0	17.7	81.3
		20	0.17	3.15	0.450	1.343	0.211	1.950	0.5	15.6	83.9
		25	0.18	3.10	0.439	1.289	0.203	1.822	0.7	13.8	85.5
	3.33%	5	0.51	6.05	1.134	3.135	1.291	12.371	2.2	19.9	77.9
		10	0.75	6.00	0.935	2.567	0.897	7.113	1.3	12.9	85.8
		20	0.96	5.70	0.820	2.020	0.704	4.328	0.8	12.3	86.9
		25	1.05	5.69	0.804	1.815	0.667	3.379	1.1	7.4	91.5
	5.00%	5	1.05	9.03	1.636	4.549	2.682	24.624	1.7	19.6	78.7
		10	1.47	8.39	1.341	3.248	1.832	11.935	1.9	13.8	84.3
		20	1.93	8.31	1.116	2.499	1.257	6.365	1.3	7.2	91.5
		25	1.99	8.06	1.108	2.400	1.257	5.954	1.1	8.2	90.7
	6.67%	5	1.62	11.63	2.055	5.666	4.224	40.689	2.5	19.5	78.0
		10	2.27	11.08	1.662	4.082	2.787	18.691	1.0	13.3	85.7
		20	2.83	10.61	1.420	3.189	2.052	10.691	1.3	10.0	88.7
		25	2.93	10.35	1.338	2.826	1.851	8.357	1.2	8.2	90.6

First Bootstrap Percentile	1.67%	5	0.32	3.38	0.469	2.054	0.271	4.819	1.9	18.8	79.3
		10	0.35	3.33	0.450	1.637	0.222	2.866	1.0	8.5	90.5
		20	0.38	3.27	0.429	1.391	0.197	2.003	0.7	6.2	93.1
		25	0.38	3.23	0.412	1.336	0.181	1.857	0.7	5.9	93.4
	3.33%	5	0.90	6.23	0.932	3.295	0.967	12.858	2.4	17.6	80.0
		10	1.06	6.21	0.860	2.685	0.763	7.474	1.4	11.7	86.9
		20	1.22	5.89	0.795	2.090	0.639	4.463	1.0	8.5	90.5
		25	1.30	5.88	0.771	1.876	0.606	3.529	1.5	5.5	93.0
	5.00%	5	1.61	9.31	1.414	4.792	2.241	25.867	2.3	18.1	79.6
		10	1.86	8.65	1.226	3.405	1.544	12.430	1.8	12.2	86.0
		20	2.24	8.55	1.078	2.593	1.205	6.737	1.1	7.3	91.6
		25	2.28	8.29	1.084	2.466	1.187	6.126	1.1	7.2	91.7
	6.67%	5	2.26	11.95	1.757	5.915	3.475	41.806	2.7	19.0	78.3
		10	2.74	11.40	1.547	4.247	2.487	19.261	1.5	10.8	87.7
		20	3.19	10.92	1.392	3.316	1.967	11.173	1.5	8.3	90.2
		25	3.26	10.62	1.309	2.921	1.719	8.655	1.0	6.3	92.7
Bootstrap Percentile	1.67%	5	-0.06	3.00	0.743	1.789	0.578	4.551	2.2	22.2	75.6
		10	0.03	3.01	0.596	1.526	0.387	2.894	1.3	19.1	79.6
		20	0.04	2.94	0.488	1.316	0.289	2.084	0.7	16.3	83.0
		25	0.05	2.90	0.471	1.265	0.275	1.963	0.6	15.0	84.4
	3.33%	5	0.34	5.66	1.324	2.934	1.815	12.518	2.3	20.9	76.8
		10	0.55	5.69	1.023	2.405	1.173	6.842	1.3	15.1	83.6
		20	0.78	5.43	0.868	1.952	0.879	4.400	0.8	14.7	84.5
		25	0.86	5.44	0.849	1.761	0.829	3.396	0.6	11.4	88.0
	5.00%	5	0.79	8.48	1.871	4.198	3.602	24.045	1.9	22.4	75.7
		10	1.21	7.99	1.449	3.065	2.294	11.874	1.3	17.1	81.6
		20	1.68	7.99	1.182	2.402	1.524	6.218	0.8	9.4	89.8
		25	1.77	7.77	1.160	2.328	1.508	5.948	0.6	10.7	88.7
	6.67%	5	1.31	10.98	2.369	5.296	5.723	40.832	2.5	20.5	77.0
		10	1.95	10.61	1.791	3.884	3.444	18.681	1.0	14.2	84.8
		20	2.53	10.25	1.472	3.079	2.402	10.668	0.8	11.6	87.6
		25	2.68	10.26	1.383	2.739	2.165	8.003	0.9	9.9	89.2