

Estimating and Correcting for Bias in Population Assessments of Sooty Shearwaters

S. MCKECHNIE,¹ *Zoology Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

D. FLETCHER, *Mathematics and Statistics Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

H. MOLLER, *Zoology Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

D. S. SCOTT, *Zoology Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

J. NEWMAN, *Zoology Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

C. BRAGG, *Zoology Department, University of Otago, P.O. Box 56, Dunedin, New Zealand*

ABSTRACT We investigated the precision and accuracy of an infrared burrowscope for detecting sooty shearwater (*Puffinus griseus*) chicks at 13 plots from 3 islands in southern New Zealand in 2003. We partially excavated burrows systems to reveal the entire burrow contents after 2 teams of observers had prospected all burrow entrances. Accuracy was similar between islands and observer teams at approximately 85%. The majority of the inaccuracy stemmed from failure to detect some chicks. Logistic regression modeling identified 4 burrow characteristics occurring between the entrance and the nest-site that influenced detection of burrow occupants. Detection was lower at nest-sites further from burrow entrances, in burrows with a high rate of burrow division, and in burrows with a high level of curvature. There was a positive relationship between the interaction of rate of division and curvature and detection of chicks. Distance from the burrow entrance was the only parameter that could be reliably used as a predictor of detection rate, so a reduced model containing only this variable was constructed to correct for burrowscope bias. The correction factor performed well on The Snares and Bench Island where predicted bias was very similar to observed levels (within 5%), but bias was overestimated on Putauhinu by up to 19.1%. Consistent bias, lack of damage to burrows from excavation, and the successful application of a correction factor all indicate the value of further testing burrowscope accuracy on other burrow-nesting seabird species. (JOURNAL OF WILDLIFE MANAGEMENT 71(4):1325–1335; 2007)

DOI: 10.2193/2006-018

KEY WORDS bias, burrowscope, correction factor, harvest, infrared, New Zealand, population monitoring, *Puffinus griseus*, sooty shearwater.

Adequate monitoring of a wildlife population relies on the robust estimation of population size. For many species, obtaining estimates is difficult or impossible due to rarity or aspects of their biology. Assessing population size and studying the breeding biology of many seabirds can be difficult due to the frequency of burrow-nesting behavior (Warham 1990). Accurate methods for detecting occupants and identifying breeding status in burrows that are often long and convoluted are scarce (Hamilton 1998a). Traditional methods include the use of smell or signs at the burrow entrance, chick cheeping responses, barricades at burrow entrances, and placement of inspection hatches over nesting chambers. These techniques are either inaccurate, imprecise, subjective, create unacceptable disturbance, require knowledge of the size of the nonbreeding population, or a combination (Hamilton 1998a, b; Cuthbert and Davis 2002; Moller et al. 2003a, b).

Some of these difficulties have been overcome by the recent development of infrared burrowscopes. Burrowscopes were first described by Dyer and Hill (1991) as a method of determining the occupancy of procellariiform burrows. A burrowscope consists of a miniature videocamera mounted on the end of a long tube that is maneuvered down the burrow. Burrow occupants are illuminated by infrared lights mounted around the camera, and the image is transmitted by cable to a monitor at the surface. Advantages over traditional methods include visual recognition of occupants, minimal

disturbance, and relative speed of use, making them suitable for both intensive and extensive monitoring approaches. Use of the equipment internationally, on a range of species, has made the burrowscope an important tool for monitoring seabird populations (Seto and Jansen 1997, Lyver et al. 1998, Cuthbert and Davis 2002, Dunlop et al. 2002).

Sooty shearwaters (*Puffinus griseus*) are a medium-sized procellariiform that breed on islands around New Zealand, South America, and southern Australia (Warham and Wilson 1982). Between mid-November and early December, breeding pairs lay a single egg in burrows that are often extremely long and convoluted (Hamilton 2000). The current monitoring program of sooty shearwaters in southern New Zealand aims to assess the sustainability of a cultural harvest of chicks and fledglings by Rakiura Māori (Moller 1996). The project incorporates a combination of monitoring approaches involving the use of burrowscopes. Intensive monitoring of abundance occurs at long-term study sites as an independent validation of demographic model predictions. An extensive approach involving single visits to other sites provides replication for a spatial comparison of abundance between harvested and non-harvested islands (Moller et al. 1999). Accurate, absolute measures of burrow occupancy are necessary for estimating breeding success, total population size, and the proportion of chicks harvested. However, relative burrow occupancy measures may be adequate for detecting trends in abundance at intensively studied sites, providing detection probability remains constant through time.

¹ E-mail: mcksa534@student.otago.ac.nz

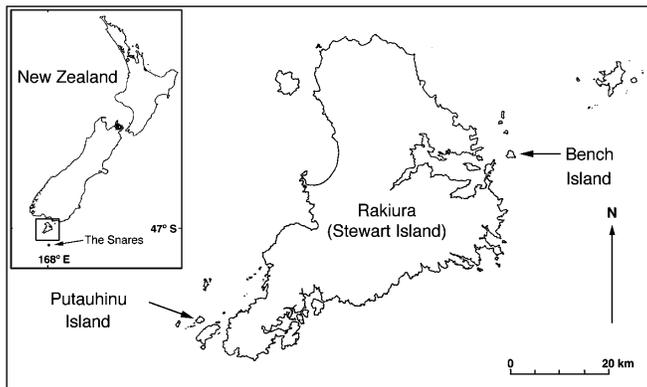


Figure 1. Map of Rakiura (Stewart Island) showing the location of the 3 New Zealand study islands visited in 2003.

Although burrowscopes are a significant advance on alternative techniques, several studies have indicated that a proportion of occupants are consistently missed (Seto and Jansen 1997; Hamilton 1998a, 2000; Cuthbert and Davis 2002). Failure to detect occupants in several of these studies has been attributed to unfavorable burrow geometry, with factors such as corners, loose substrate, and burrow length preventing passage of the burrowscope (Lyver et al. 1998, Hamilton 2000).

Published validation of burrowscope accuracy is limited to Hamilton (2000) and Jones et al. (2003) for sooty shearwaters, and a study by Cuthbert and Davis (2002) on Hutton's shearwater (*Puffinus huttoni*). However, no robust and replicated investigation has currently been undertaken. Our study builds on the pilot study undertaken by Hamilton (2000) by increasing sample sizes and replication, with the aims of 1) describing burrow structure and complexity, 2) estimating the accuracy and precision of the burrowscope, 3) identifying the burrow characteristics that prevent the correct detection of burrow occupants, and 4) building a predictive model that allows burrowscope data to be corrected.

STUDY AREA

We partially excavated 10 plots: 5 on Northeast Island, The Snares (48°01'S, 166°36'E) between 2 February and 10 February 2003, 2 on Bench Island (46°55'S, 168°14'W) between 2 March and 5 March 2003, and 3 on Putauhinu (47°13'S, 167°23'W), New Zealand (Fig. 1). We fully excavated an additional 3 plots at sites on Putauhinu where huts were being constructed by muttonbirders. We established 4 of the plots on Putauhinu at sites where chicks are harvested by muttonbirders, whereas the remaining 2 were situated within a Rāhui (restricted) area, where harvesting has not taken place for about 30–40 years (Moller et al. 1999). The study on Putauhinu took place between 22 March and 5 April 2003.

METHODS

Burrowscope Design

The burrowscope model used in our study was identical to Prototype C in Lyver et al. (1998) and Hamilton (2000).

The camera head consisted of a 6-mm printed circuit board and encapsulated charge-coupled device black-and-white standard security camera (KPC-500B, Sony, Tokyo, Japan) with 510 (horizontal) × 492 (vertical) effective pixels, housed in a solid aluminum cylinder of 46-mm diameter. The camera was surrounded and lit by 6 infrared light-emitting diodes (LEDs). The camera used the emitted light from the diodes to record the burrow contents. We attached a flexible, corrugated plastic drainage hose, with a diameter of 48 mm and a length of 3.3 m, to the camera head. A cylindrical control handle was attached to the opposite end of the hose, and it was connected to a ball joint pivot at the base of the camera head with motorcycle control cables running along the inside of the hosing. By manipulating the control handle, we could achieve bidirectional movement of the camera head (45° either side of neutral on one axis). The camera extension cord was contained within the hosing and connected to a standard security monitor (105 × 88-mm screen) by electrical connectors. This monitor was contained within waterproof housing and a 12-V battery powered the unit.

Burrowscoping Procedure

We randomly selected plots after stratifying areas based on burrow entrance density to ensure that we sampled a wide range of densities. We located the 3 plots excavated at hut building sites randomly within the building areas designated by the muttonbirders. After a plot was marked out and all entrances within the boundaries were identified, 2 burrowscoping teams prospected each burrow entrance to determine its occupancy status. Burrowscope teams consist of 2 members, one member operating the hose section and the second member manipulating the camera head by using the control handle. The skill and efficiency of the hose operator largely determines the outcome of burrow prospecting, keeping the operators of the hose section consistent throughout the study. Logistical constraints prevented the consistency of the monitor operator in the second team (we used different monitor operators on each island). The influence of the monitor operator is generally thought to be minimal, but it could potentially affect efficacy through interpreting the progress of the burrowscope from the monitor picture. Both hose operators were relatively experienced in their use (>1000 burrow entrances previously prospected), and they were selected to represent the skill level that is maintained during monitoring of sooty shearwater populations.

Our burrowscoping procedure was identical to the usual monitoring practice. We prospected each viable burrow entrance by inserting the camera head into the entrance and manipulating the hose and control handle to achieve passage down the burrow. By rotating the hose in conjunction with movement of the camera head, we achieved systematic coverage of the tunnel and negotiation of corners and tunnel divisions. The infrared LEDs provided the illumination necessary to identify tunnel contents such as divisions, adults, chicks, and obstructions, on the monitor above ground. We considered burrow entrances viable if the

burrow extended ≥ 30 cm underground (Hunter 2001). Similarly, when burrows divided, we considered any branch that extended a further 30 cm from the point of divergence an individual burrow. We attempted to prospect each individual burrow beyond the divergence. The presence of burrow occupants often prevents passage of the burrowscope further down the burrow. If we did not reach the perceived end of the burrow due to any number of factors other than the presence of a chick, we considered the burrow unsolved and noted the reason we did not reach the end. We measured the distance to burrow occupants and end walls by using a graduated scale marked on the burrowscope hose.

The second team started prospecting burrows roughly >15 min after the first team prospected them. The time we spent prospecting each individual entrance was dependent on the complexity of the burrow system diverging from it, although typically this time would be between 1 minute and 10 minutes. We kept the order with which the 2 teams prospected the burrows consistent throughout the study. The results of burrow status remained confidential within the teams until both had finished prospecting the plot.

Burrow Curvature

The first team of observers to burrowscope each plot also recorded orientation of the tunnel at the entrance, at the depth of 40 cm (the point to which the burrow can be easily investigated with an arm from the entrance), and perceived orientation every 30 cm beyond, including all diverging tunnels, with a compass.

Partial Excavations

After each burrow entrance in the plot had been prospected, we constructed a grid system of 1×1 -m squares on the plot surface by using string and metal pegs. For each burrow entrance, we inserted inspection holes at arm-length intervals along the burrow system until the entire system had been prospected by hand to reveal all chicks, burrow divisions, and side chambers. We prospected tunnels extending outside the plot until either we reached the burrow end or until we encountered a burrow entrance outside the plot. We mapped out the architecture of the burrow system on waterproof paper by using the grid system as scale.

We measured burrow depth at each inspection hole as the distance between the ground surface and the floor of the burrow. We calculated burrow height by subtracting the distance between the ground surface and the burrow roof, and total burrow depth. We measured distances between inspection holes by using 500-mm rulers, and we measured burrow angles by using a compass. This allowed us to calculate the curvature (cumulative change in angle of the burrow), distance to chicks, divisions and end-walls, and the angle of corners.

Once we revealed and measured the entire burrow system, we plugged all inspection holes by using a lattice of sticks pushed into the walls of the hole, covered with a layer of leaves, and finally a compacted layer of soil.

Full Excavations

The methods we used in the 3 fully excavated plots on Putauhinu were identical to those used in partial excavations, except that we removed the entire burrow roof to reveal burrow structure instead of only placing inspection holes. Any chicks encountered had their position in the burrow marked before we removed and euthanized them. After measurements of burrow geometry were completed, we filled the excavated burrows by using the originally displaced soil.

Measures of Burrow Occupancy

We classified individual burrows as solved if either a chick was detected or the end of the burrow was reached. Unsolved burrows were those burrows in which we failed to reach the end with the burrowscope due to unfavorable burrow geometry and obstructions such as pooled water, roots, and adult birds (McKechnie 2004). Solved entrances were those for which all diverging burrows were solved with the burrowscope.

We calculated 2 measures of occupancy per entrance. We measured the improved occupancy per entrance (*IO*) to take into account the potential for occupants being present beyond the point to which the burrowscope prospected. We calculate *IO* as

$$IO = \frac{SB}{SE} \times \frac{C}{S}$$

where *SB* is the number of solved burrows diverging from solved entrances (*SE*), *C* is the total number of chicks detected, and *S* is the number of solved burrows (irrespective of whether they came from solved or unsolved entrances).

Simple occupancy (*SO*) is the most common method of calculating occupancy in other studies, which we calculated as

$$SO = \frac{C}{AE}$$

where *AE* is the total number of entrances prospected (solved and unsolved).

We calculated actual occupancy as the number of chicks detected by partial or complete excavation divided by the number of entrances in the plot. We excluded chicks that were closer to an entrance outside the plot (compared with an entrance inside the plot) from our calculations.

Measuring Accuracy

We defined accuracy as the proportion of chicks present that were correctly detected by the burrowscope operators. We achieved this accuracy by comparing the positions of chicks on the burrow system maps with the results of the burrowscoping procedure. We considered occupants to be incorrectly detected if they were either entirely missed by the burrowscope operators, or they were unknowingly counted from 2 or more burrow entrances.

The observed level of double counting of chicks is an underestimate. Burrow systems within plots were often connected to burrow entrances outside the plot boundaries

that we did not prospect. The probability of double counting chicks in these systems was therefore reduced. We mitigated this potential bias by calculating the proportion by only including those chicks in burrow systems for which all possible connecting burrow entrances were prospected with the burrowscope.

Although we investigated the reliability of the burrowscope in square plots of burrows, much of the monitoring of sooty shearwaters relies on the estimation of occupancy from burrows within 2-m-wide transects. If occupants can be detected from more than one burrow entrance, there will be a greater probability that the additional entrance will be outside the area of a transect than a plot; hence, this entrance will not be prospected. We estimated this potential difference in probabilities empirically by placing a pair of simulated transects, of the same length used in regular monitoring, over the maps, running vertically and horizontally, respectively, and we noted whether both entrances from which double-counted chicks were detected fell within the transect bounds.

Statistical Analyses

We used logistic regression to assess which burrow variables affected detection of chicks. We used a binary response variable of whether a chick was detected (1) or remained undetected (0). We used the following predictor variables: 1) distance from the chick to the nearest burrow entrance, 2) the number of burrow divisions (where the burrow branched into >1 tunnel) per meter between the chick and the nearest burrow entrance, 3) the curvature (cumulative change in angle) per meter of the burrow between the chick and the nearest burrow entrance, 4) the depth of the nest site, and 5) all relevant pairwise interactions between predictor variables. In addition to the predictor variables based on burrow geometry, we included a categorical variable that identified the plot, observer, and island combination (e.g., 132 would indicate chicks on island 1, plot 3 for observer 2), to control for variation between these combinations. In no cases were large Pearson's correlation coefficients observed between predictor variables. We undertook model selection by using backwards elimination after fitting a model with all predictor variables included. We manually removed the predictor variable with a highest P -value >0.05 from the model, and the model was rerun until only predictor variables with a P -value <0.05 were included. We deemed sample sizes of <10 chicks in a plot too small for the modeling procedure, so we excluded chicks from these plots.

We performed analyses of variance on several response variables by using a model incorporating the factors island, plot, and observer. The latter 2 factors were nested within island. We considered the observer factor random due to the variation in observer combinations in observer team 2 between the 3 islands. We considered the island and plot factors fixed and random, respectively. All tests were 2-tailed, and differences were considered significant at the P < 0.05 level. We undertook procedures using the MINITAB

14 (Minitab, Inc., State College, PA) and SPSS 11.0 (SPSS Inc., Chicago, IL) statistical software programs.

Developing a Correction Factor to Improve Burrow Occupancy Estimates

The distance to a chick is the only potential predictor variable included in the logistic regression model that can also be accurately measured when burrowscoping, so we developed correction factors by using detection probabilities derived from models that only included this variable. We used 2 techniques to correct occupancy measures estimated from burrowscoping data. The incremental method incorporated the detection probability p estimated using the logistic regression procedure

$$p = \frac{1}{1 + \exp\{-(a + bx)\}}$$

where x is the distance of the chick from the burrow entrance, a is the intercept term, and b is the slope term. We derived the intercept and slope terms from the various detection function models developed. The total number of chicks (N) is then

$$\hat{N} = \sum_x \frac{n_x}{\hat{g}(x)}$$

where n_x is the count of chicks at each distance increment (10-cm intervals) and $g(x)$ is the probability that we detected a chick x meters from the burrow entrance, given that it is present at that distance; we derived the estimate $\hat{g}(x)$ from the logistic regression analysis.

We calculated the average method by using the actual distribution of all chicks in the excavation plots and the estimated detection functions. Suppose p is the probability of detecting a chick in a burrow. For simplicity, we assume that p is the same for all burrows at a site. If we detect n chicks from a total of B burrows, we can estimate p by using the observed proportion of burrows in which we detected a chick, that is,

$$\hat{p} = \frac{n}{B}$$

Now p can be written as

$$p = p_1 p_2$$

where p_1 is the probability that a chick is detected in a burrow given that a chick is present in that burrow; and p_2 is the probability that a chick is present in a burrow, that is, the chick density at that site. The total number of chicks (N) is given by $N = p_2 B$. Using the aforementioned equations, we can estimate N by writing

$$\hat{p} = \hat{p}_1 \hat{p}_2 = \frac{n}{B},$$

which leads to

$$\hat{N} = \hat{p}_2 B = \frac{n}{\hat{p}_1}$$

Table 1. Precision and accuracy of the infrared burrowscope for detecting occupants of sooty shearwater burrows on 3 New Zealand breeding islands in 2003. All values are percentages.

Island	Observer	<i>n</i>	Chicks detected correctly		Chicks missed		Chicks double counted		Chicks correctly detected by both observers		Chicks missed by both observers	
			\bar{x}	SE	\bar{x}	SE	\bar{x}	SE	\bar{x}	SE	\bar{x}	SE
Snares	1	5	85.5	7.8	11.5	5.5	4.0	2.5	77.6	6.2	7.4	3.1
	2	5	80.4	5.2	15.6	4.3	4.0	2.5				
Bench	1	2	86.7	13.3	13.3	13.3	0		76.7	23.3	6.7	6.7
	2	2	83.3	16.7	16.7	16.7	0					
Putauhinu	1	6	85.2	3.3	11.9	2.7	2.9	1.8	75.0	5.9	4.0	2.7
	2	6	86.1	5.4	6.5	3.2	7.3	2.6				

Finally, we can see how to estimate p_1 by rewriting it as

$$p_1 = \sum_x g(x)b(x)$$

where $b(x)$ is the probability that a chick is present at that distance given that a chick is present somewhere in the burrow. This leads to

$$\hat{N} = \frac{n}{\sum_x \hat{g}(x)\hat{b}(x)}$$

where $b(x)$ is estimated from the observed distribution of chicks (with respect to distance from the burrow entrance) on each island.

RESULTS

Accuracy and Precision

The burrowscope correctly detected 82.4% of chicks on The Snares, 85.0% on Bench Island, and 85.7% on Putauhinu, when data from both sets of observers were combined (Table 1). Differences in accuracy among islands were not statistically significant ($F=0.11$, $df=2$, 25 , $P=0.90$). The major source of inaccuracy was due to chicks being missed ($\bar{x}=6.5$ – 16.7%), although a proportion ($<7.3\%$) of the chicks on The Snares and Putauhinu were double counted, after being detected from >1 burrow entrance (Table 1). Missed chicks were the only source of inaccuracy on Bench Island. We detected some chicks from up to 3 burrow entrances; however, burrowscope operators could often identify that the same chick was being detected, so they did not record each detection as a new individual. This prevented us from double counting some chicks, and no chicks were triple counted. The proportions of chicks missed and the proportions double counted were not significantly different among islands (missed: $F=0.49$, $df=2$, 25 , $P=0.64$; double counted: $F=1.15$, $df=2$, 25 , $P=0.36$; Table 1).

The probability of double counting on transects was estimated to be 73% of the level observed in the plots. Applying this correction to double counting measures resulted in values of 2.9% on The Snares for both observers, and 2.1% and 5.3% on Putauhinu for observers 1 and 2, respectively.

Very consistent results were observed between observers in the average probability of correctly detecting chicks ($<6\%$

points difference on each island; $F=0.28$, $df=3$, 25 , $P=0.84$), the average proportion missed ($<6\%$ points difference on each island; $F=0.68$, $df=3$, 25 , $P=0.58$; Table 1), and the average proportion double counted ($<5\%$ points difference on each island; $F=3.32$, $df=3$, 25 , $P=0.07$; Table 1). Although both teams of observers correctly detected a similar proportion of chicks, often different chicks were missed or double counted. The average proportion of chicks correctly detected by both observers ranged between 75% and 78%, whereas both observers (Table 1) missed small proportions. The proportion of chicks missed by both observers was higher than expected by chance, assuming independence of observers ($F=8.13$, $df=1$, 25 , $P=0.02$), which suggests that some chicks were particularly hard to detect by both observer teams. The end of the burrow was mistakenly classified as having been reached on 31.5% ($n=127$) and 29.0% ($n=114$) of occasions by observer 1 and observer 2, respectively.

Comparison of Occupancy Measures

The simple occupancy per entrance measure consistently underestimated true occupancy, despite double counting (Table 2; Fig. 2A). The only case we overestimated occupancy was by observer 2 on Putauhinu, when the high rate of double counting overcompensated for the low proportion of chicks missed. Linear regressions displayed very strong relationships between simple occupancy and actual occupancy, for both observers (observer 1: $R^2=0.94$, $P<0.001$; observer 2: $R^2=0.86$, $P<0.001$). When simple occupancy measures were compared with actual occupancy (Fig. 2A), the slope of the regression line was significantly different from 1 for observer 1 ($t=3.11$, $df=11$, $P=0.01$) but not for observer 2 ($t=1.57$, $df=11$, $P=0.15$).

Improved occupancy per entrance was a better estimate of actual occupancy than simple occupancy (Table 2; Fig. 2B). Occupancy was consistently overestimated using this measure, although the error was generally not substantial. Improved occupancy measures estimated for observer 1 on Putauhinu were the only case of underestimation, although this difference was minimal. Linear regressions exhibited very strong relationships between improved occupancy and actual occupancy (observer 1: $R^2=0.92$, $P<0.001$; observer 2: $R^2=0.91$, $P<0.001$). Slope of the linear regression equation was not significantly different from 1, for both

Table 2. Comparison of estimated occupancy of sooty shearwater burrow entrances and actual occupancy levels on 3 New Zealand breeding islands in 2003. Values are estimates and standard errors.

Island	Observer	Actual occupancy		Simple estimate of occupancy		Relative error (%)		Improved estimate of occupancy		Relative error (%)	
		\bar{x}	SE	\bar{x}	SE	\bar{x}	SE	\bar{x}	SE	\bar{x}	SE
Snares ^a	1	0.45	0.13	0.40	0.11	-7.4	3.1	0.48	0.12	12.8	14.2
	2			0.39	0.11	-12.9	4.5	0.46	0.13	4.4	7.8
Bench ^a	1	0.44	0.39	0.33	0.28	-13.3	13.3	0.40	0.34	6.5	18.5
	2			0.30	0.25	-16.7	16.7	0.42	0.36	4.0	10.6
Putauhinu ^a	1	0.60	0.02	0.55	0.02	-7.9	1.6	0.59	0.02	-0.8	4.8
	2			0.61	0.03	2.6	2.5	0.65	0.03	8.5	3.5

^a Sample sizes of plots on The Snares, Bench Island, and Putauhinu were 5, 2, and 6, respectively.

observers (observer 1; $t = 1.73$, $df = 11$, $P = 0.11$; observer 2: $t = 0.50$, $df = 11$, $P = 0.62$).

Variation in Burrow Complexity

Measures of burrow complexity differed between islands (Table 3). Burrow density and the number of connections per burrow entrance were considerably higher on The

Snares and Putauhinu than on Bench Island. The number of burrows per entrance, excluding burrows connecting to other entrances, was 0.36 (standard error of the difference [SED] = 0.19) and 0.44 (SED = 0.25) burrows per entrance higher on Putauhinu than on The Snares and Bench Island respectively. The total number of burrows per entrance was only 0.04 (SED = 0.28) burrows per entrance higher on Putauhinu than The Snares, but it was 1.05 (SED = 0.38) burrows per entrance higher than on Bench Island (Table 3). The average proportion of burrows that were simple (self-contained that do not connect to other burrow entrances) was highest on Bench Island (0.76 ± 0.13 , $n = 2$) followed by Putauhinu (0.45 ± 0.06 , $n = 6$), and it was lowest on The Snares (0.30 ± 0.06 , $n = 5$). Average burrow length on Putauhinu (133.25 ± 3.47 cm, $n = 6$) and Bench Island (132.31 ± 5.81 cm, $n = 2$) was longer than that on The Snares (102.04 ± 1.31 cm, $n = 5$; Table 3). The number of divisions per meter of burrow was highest on The Snares (1.04 ± 0.09 , $n = 5$) followed by Putauhinu (0.67 ± 0.06 , $n = 6$), with Bench Island showing the lowest rate of division (0.47 ± 0.07 , $n = 2$). We observed very consistent results between islands in the degree of burrow curvature per unit burrow length ($<4^\circ/\text{m}$ difference).

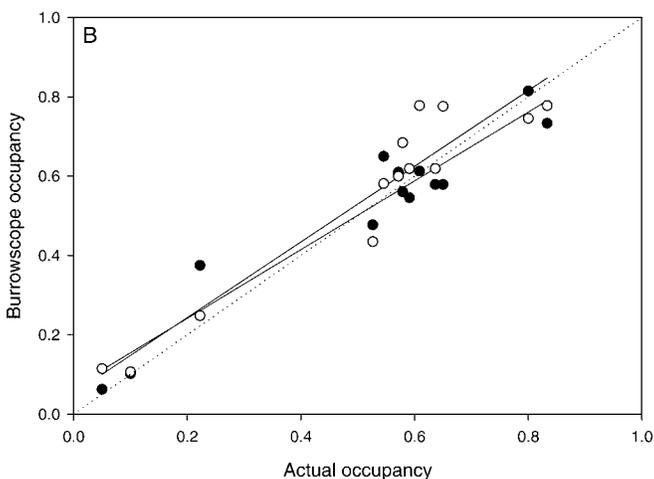
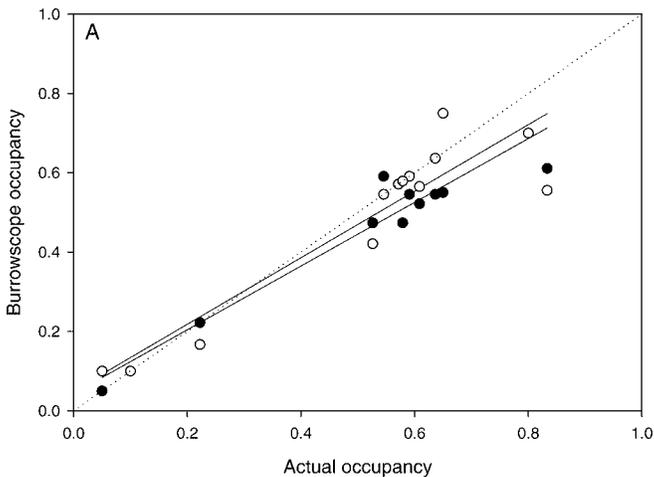


Figure 2. Relationship between actual sooty shearwater occupancy and (A) simple occupancy and (B) improved occupancy, measured with an infrared burrowscope, for observers 1 (closed circles) and 2 (open circles). Data are from 3 New Zealand islands in 2003 combined.

Modeling the Detection of Burrow Occupants

Distance to the chick, the number of divisions, the degree of curvature, and an interaction between the number of divisions and curvature were the predictor variables found to significantly influence detectability of chicks at the 5% level (Table 4). The model predicted 85.6% of the responses correctly.

Chick detection probability decreased with distance from the burrow entrance, and we predicted very few chicks detected beyond 250 cm (Fig. 3A). Estimated detection probabilities at the mean distance to chicks on The Snares (82 cm), Bench Island (110 cm), and Putauhinu (101 cm) were 0.92 (CI = 0.73–0.98), 0.83 (CI = 0.54–0.95), and 0.87 (CI = 0.60–0.96), respectively. The relationships between detection probability and the degree of curvature and divisions were also negative. We predicted $<50\%$ of chicks to be detected when the number of divisions reaches 2 chicks/m (Fig. 3B). Burrows displaying a high level of curvature prevented detection of chicks, and we predicted $<50\%$ detected when curvature reaches $150^\circ/\text{m}$ (Fig. 3C).

Table 3. Comparison of sooty shearwater burrow complexity measures between plots on 3 New Zealand breeding islands in 2003.

Measure	Snares ^a		Bench ^a		Putauhinu ^a		F	P
	\bar{x}	SE	\bar{x}	SE	\bar{x}	SE		
Burrow density (burrows/m ²)	0.98	0.13	0.33	0.14	0.86	0.09	7.14	0.01
Connections/entrance	1.94	0.40	0.24	0.13	1.05	0.22	4.98	0.03
Burrows/entrance (connections excluded)	1.39	0.09	1.32	0.12	1.75	0.10	2.57	0.13
Burrows/entrance (connections included)	2.54	0.17	1.53	0.14	2.58	0.18	4.12	0.05
Proportion of burrows simple ^b	0.30	0.06	0.76	0.13	0.45	0.06	5.10	0.03
Divisions/m	1.04	0.09	0.47	0.07	0.67	0.06	10.71	<0.01
Curvature/m	62.54	2.90	61.92	0.17	65.76	2.85	0.46	0.65
Burrow length (cm)	102.04	1.31	132.31	5.81	133.25	3.47	31.85	<0.001

^a Sample sizes of plots on The Snares, Bench Island, and Putauhinu were 5, 2, and 6 respectively.

^b Proportion of burrows that were self-contained and did not connect to other burrow entrances.

Although 3 burrow characteristics and one interaction term were found to be important by the modeling procedure, these characteristics must be able to be measured efficiently and accurately when burrowscoping for successful correction of occupancy. Although a regression of curvature of the burrow before reaching a chick (measured with the burrowscope), against true curvature, was significant ($R^2 = 0.22$, $P < 0.001$; Fig. 4), we observed substantial variation between the measures. Furthermore, the ability of the burrowscope operator at detecting changes in the direction of the burrow seems to be significantly reduced at greater distances down the burrow (Fig. 5).

Detection of burrow divisions was affected by distance from the burrow entrance in a similar way to detection of chicks, although the effects of distance were even more pronounced (coeff. = -0.05 , SE = 0.01, $z = -8.63$, $P < 0.001$). We detected very few divisions >80 cm.

Reduced Model Only Incorporating the Distance to Chick

We repeated the modeling procedure incorporating the distance to chicks as the only predictor variable. Although this model is inferior, the model correctly predicted 80.5% of the responses, so elimination of the other variables resulted in a reduction of only 5.1 percentage points in the proportion assigned correctly.

In fitting this simpler model, we also assessed its generality across islands by pooling data from all plots and observers within islands. Predictor variables initially included were distance to the chick, island, and an interaction between distance and island. We removed the interaction term and the island variable sequentially, with the model refitted after each removal. Removing the interaction term resulted in a

change in deviance of the model that approached formal significance (Table 5), suggesting that the slope of the detection functions may differ between islands. Slope seemed to be particularly pronounced on Bench Island relative to The Snares and Putauhinu, with very few chicks predicted to be detected >200 cm (Fig. 6). However, very small sample sizes may have influenced the form of the detection function on this island. We detected a highly significant change in deviance of the model when the island variable was removed (Table 5), which is evidence that intercept terms of the detection function varied between islands. Detection functions incorporating island-specific intercept terms remained relatively consistent between islands (Fig. 6). We developed a simple model containing only the distance predictor to test whether a general nonisland-specific detection function could have utility for future corrections of occupancy estimates (Fig. 7).

Using the Correction Factor for Predicting Bias Levels

We compared the different detection function models to test the utility of a correction factor based on distance to burrow occupants (Table 6). By using the incremental correction and incorporating the detection function derived from the simple model, our predicted levels of bias, for both observers, showed a tight relationship (within 5 percentage points) with the observed levels of bias at The Snares and Bench Island (Table 6). The performance of the correction factor for estimating bias on Putauhinu was substantially decreased. We overestimated negative bias by 7.2% and 19.1% for observers 1 and 2, respectively. The accuracy of the average correction method produced very similar results, with our predictions again closely approximating reality on The Snares and Bench Island but overestimating bias on Putauhinu. Overall, our predictions of bias when data from all islands were pooled overestimated actual bias, presumably due to the influence of the large sample size at Putauhinu. We observed variable results when the island-specific intercept models and the island-specific slope and intercept models were used (Table 6). We observed considerable overestimation of bias on The Snares by using both models, although more accurate correction occurred on Putauhinu than for the simple model, especially when the island-specific intercept model was used. The accuracy of the

Table 4. Logistic regression results for the influence of burrow characteristics on the detection of sooty shearwater chicks on 3 New Zealand breeding islands in 2003.

Variable	Coeff.	SE	z	df	P
Constant	7.34	1.13	5.95	1	<0.001
Distance	-0.03	0.01	-6.32	1	<0.001
Division	-1.12	0.45	-2.47	1	0.01
Curvature	-0.03	0.01	-4.12	1	<0.001
Division \times curvature	0.01	0.00	2.30	1	0.02

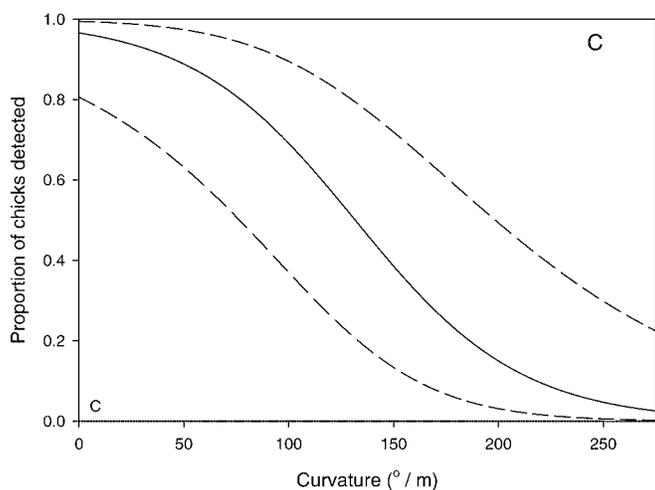
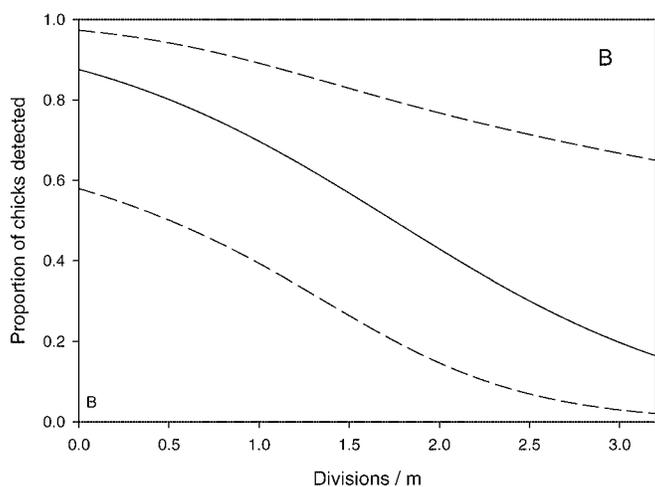
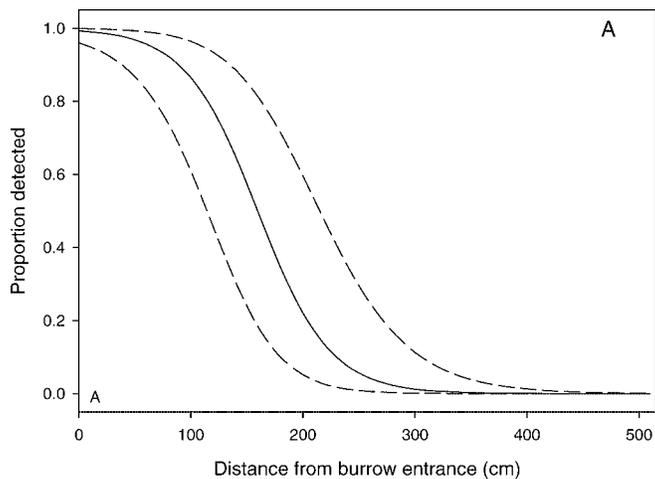


Figure 3. Detection functions of sooty shearwater chicks on New Zealand islands in 2003 with (A) distance from burrow entrances, (B) divisions per meter, and (C) curvature per meter. Equations are back-transformed logistic regressions from the model containing all important predictor variables. Functions are expressed while keeping the other predictor variables at their mean values. Dashed lines represent 95% confidence intervals.

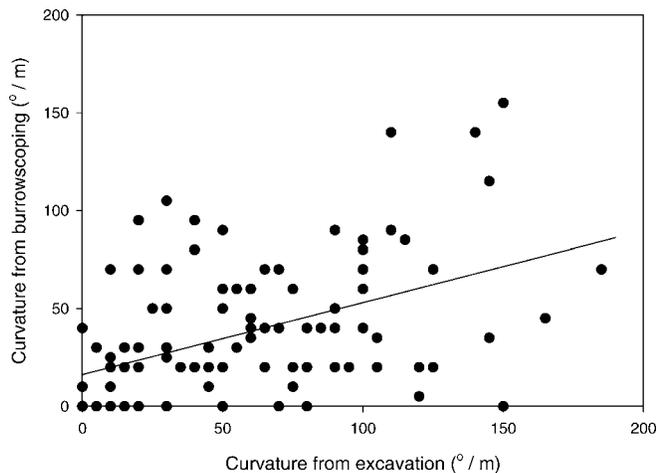


Figure 4. Comparison of curvature of sooty shearwater burrows (measured as the cumulative change in angle) estimated by observers with an infrared burrowscope, with actual curvature measured during excavation of burrow systems on 3 New Zealand islands in 2003.

island-specific slope and intercept model was poor on Bench Island with bias overestimated for observer 1 by almost 15%. The island-specific intercept model was reasonably accurate on Bench Island, with bias underestimated by 7.3% and 4.0% for observers 1 and 2, respectively.

DISCUSSION

Comparison of Accuracy to Previous Studies

The burrowscope accuracy observed in our study was considerably higher than observed in a pilot study undertaken on The Snares in 1996 (Hamilton 2000). This is the only other published study that achieved complete enumeration of occupants after burrowscoping a nonbiased selection of burrow entrances. Both studies used the “prototype C” burrowscope model of Lyver et al. (1998). Converting Hamilton’s results into our measure of accuracy corresponds to 60%, 49%, and 59%, for the three burrowscope checks.

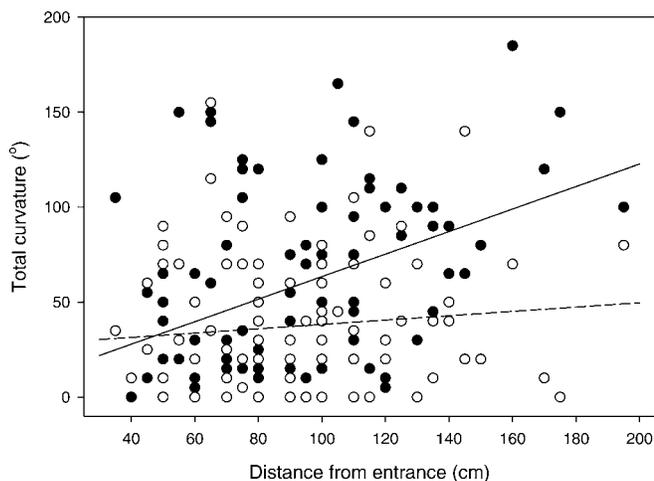


Figure 5. Change in curvature (cumulative change in angle of the burrow) with distance from the burrow entrance, for actual curvature, as measured during excavation of sooty shearwater burrow systems on 3 New Zealand islands in 2003 (solid circles, solid line), and curvature estimated using an infrared burrowscope (open circles, dashed line).

Table 5. Analysis of deviance table comparing the fit of the 3 forms of the detection function model to data on the detection of sooty shearwater chicks on 3 New Zealand breeding islands in 2003.

Model	Parameter	Deviance	Δ Deviance	Δ df	P
Island + distance					
+ interaction	6	267.43			
Island + distance	4	272.65	5.22	2	0.07
Distance	2	295.91	23.26	2	<0.001

The percentage of eggs missed during the three checks was 33%, 34%, and 20%, respectively. This accuracy seems to be relatively low compared with our study, although the unreplicated nature of the pilot investigation, different breeding stages investigated, high burrow complexity observed in the pilot study (only 5% of entrances led to self-contained burrows not connected to other entrances, by comparison with an average of 29.8% for our plots on The

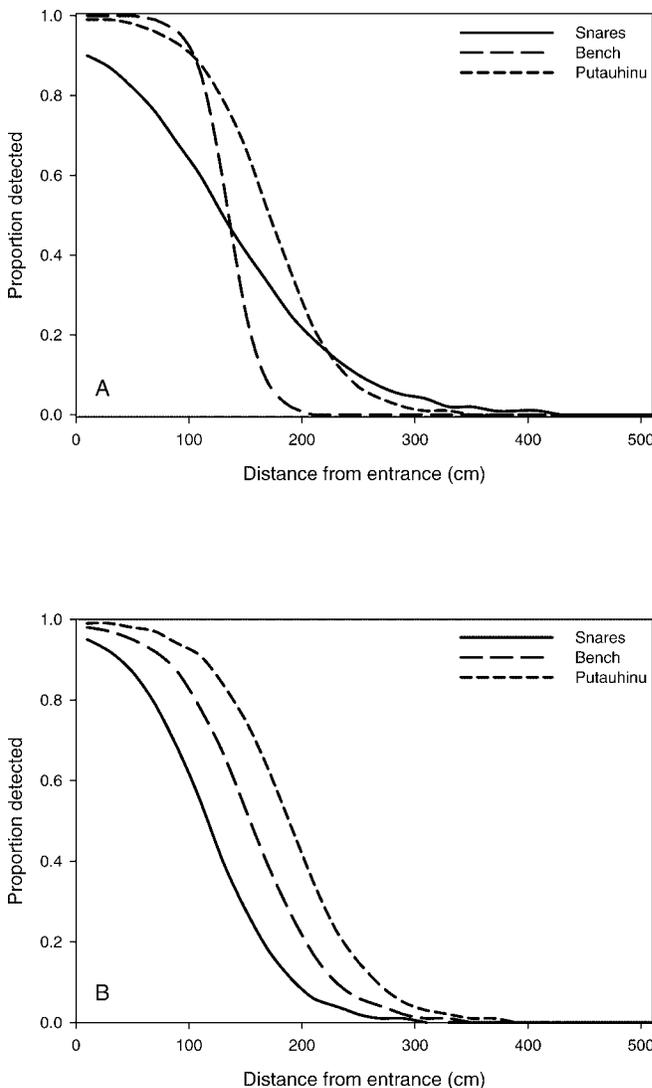


Figure 6. Mean detection rate of sooty shearwater chicks measured on 3 New Zealand islands in 2003 versus distance from burrow entrances. Regressions are backtransformed logistic equations, based on models with (A) island-specific slopes and intercepts and (B) island-specific intercept terms, and a common slope.

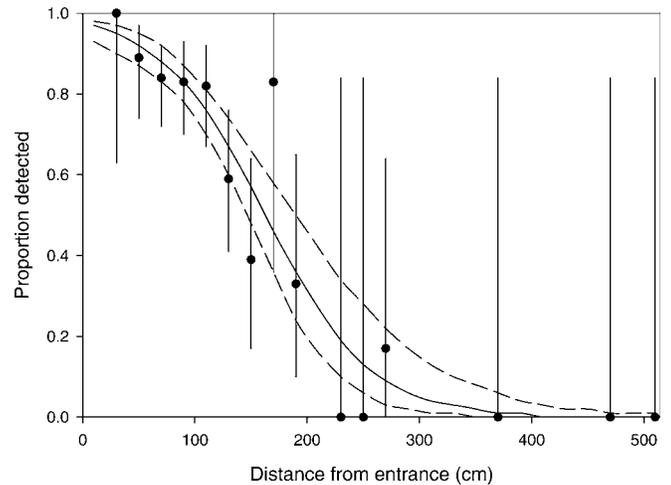


Figure 7. Mean detection rate of sooty shearwater chicks measured on 3 New Zealand islands in 2003 versus distance from burrow entrances. The regression is a backtransformed logistic equation (solid line), with 95% confidence intervals (dashed line), based on the simple model in which it is assumed that there is a common intercept and common slope across all islands. Circles are observed detection rates with 95% binomial confidence intervals.

Snares), and the potential for different observer effects allow for little reliable comparison with our study.

Factors Influencing Chick Detection

Our analyses identified the negative influence of burrow characteristics associated with elevated burrow complexity; increased distances to the nest site, the presence of divisions and high burrow curvature. In general, complexity of burrow systems was higher at The Snares and Putauhinu than at Bench Island. Presumably, these differences can be attributed to the very low burrow densities encountered on the latter, and hence to the reduced levels of burrow bifurcation and connections between systems that result from burrow sharing and the proximity of entrances. The lack of connecting burrow entrances prevented double counting of chicks on Bench Island. Interestingly, despite differences in the burrow characteristics affecting chick detection (those identified by logistic regression) existing among the 3 islands, the mean accuracy of the burrowscope was reasonably consistent at between 80% and 87%. Sample sizes for plots on all islands were small, and variation between plots, especially on Bench Island, was high. This presumably resulted in low power to detect differences. However, the observed consistency remains encouraging, and it suggests that the differences in habitat, especially burrow characteristics, between islands may be low enough to prevent substantial variation in bias between islands. Alternatively, differences in detectability caused by variation in one burrow characteristic may be offset by the opposite influence of variation in another characteristic.

Accuracy of Occupancy Measures

Simple occupancy consistently underestimated true occupancy, whereas improved occupancy exhibited positive bias, although the latter deviated less from true occupancy.

Table 6. Comparison of mean and 95% confidence interval predicted bias levels when detecting sooty shearwater chicks on 3 New Zealand breeding islands in 2003, estimated using 2 correction factor methods, based on detection functions derived from logistic regression modeling, with actual bias levels for both observer teams. Detection functions used were based on the simple model, models with island-specific intercept terms, and models with island-specific slope and intercept terms. All values are percentages.

Island	Obs	Actual proportion missed ^a	Simple model				Intercept model				Slope + intercept model			
			Avg method		Increment		Avg method		Increment		Avg method		Increment	
			\bar{x}	CI	\bar{x}	CI	\bar{x}	CI	\bar{x}	CI	\bar{x}	CI	\bar{x}	CI
Snares	1	17.7	16.4	12.1–22.0	15.9	11.6–21.6	34.5	25.3–44.8	31.7	21.5–44.1	33.5	22.6–45.6	30.1	19.7–43.4
	2	17.7	16.4	12.1–22.0	19.8	14.8–26.0	34.5	25.3–44.8	39.8	27.6–53.7	33.5	22.6–45.6	34.9	22.7–50.1
Bench	1	25.0	26.3	20.7–32.6	20.4	15.8–26.0	25.1	12.4–43.6	17.7	7.9–34.9	23.3	11.0–51.9	10.6	3.1–36.2
	2	31.3	26.3	20.7–32.6	28.1	22.4–34.6	25.1	12.4–43.6	27.3	13.2–48.2	23.3	11.0–51.9	31.5	12.4–61.0
Putauhinu	1	12.0	22.9	17.7–29.1	19.2	14.3–25.4	10.9	6.5–17.6	7.4	3.9–13.6	20.4	8.9–20.7	9.7	2.8–10.4
	2	6.5	22.9	17.7–29.1	25.6	18.9–34.1	10.9	6.5–17.6	12.3	6.7–21.8	20.4	8.9–20.7	17.6	9.8–47.9
Total	1	15.4	21.1	16.1–27.0	18.2	13.5–24.2	21.4	15.7–28.4	18.5	12.7–26.3	18.7	11.6–28.5	15.9	8.3–30.7
	2	13.2	21.1	16.1–27.0	24.4	18.2–32.1	21.4	15.7–28.4	28.1	19.2–40.0	18.7	11.6–28.5	32.8	15.4–67.2

^a Data from all plots within islands were pooled to give an overall the proportion of chicks missed, because correction factors were applied to pooled data (due to small sample sizes). Hence, this figure differs slightly from the average proportion missed. Sample sizes of chicks on The Snares, Bench Island, and Putauhinu were 46, 16, and 75, respectively.

Assessment of sooty shearwater population densities relies largely on determining occupancy on strip transects, where the incidence of double counting is reduced. The resulting reduction in positive bias for both measures would increase the accuracy of the improved measure relative to simple occupancy. However, caution is necessary when interpreting this result, because although improved occupancy aims to take into account burrows in which the end is not reached, this requires accurate assessment of whether the burrow was fully prospected (classified as solved). This is not the case, because we observed high rates of misclassification. We, therefore, do not promote the use of this measurement in the future.

Consequences of Negative Bias in Occupancy Measures

For comparisons of abundance and detecting trends in population size, a relative abundance index may be adequate provided that detection probabilities remain constant over time and space. Because this condition is seldom met in the field (Anderson 2001, 2003), we advise caution when relying on raw occupancy measures, despite the strong linear relationships observed between both simple and improved occupancy, and actual occupancy. Moreover, estimates of absolute occupancy are required when estimating population size and harvest intensity (by comparing off-take rates with local population size). Furthermore, the number of experimental plots was small, and the islands on which the study was conducted do not cover all habitat types occupied by sooty shearwaters, thereby limiting inferences from this study. The structure of this relationship outside the range of abundance levels encountered in the experimental plots remains unknown. Our study confirms that some burrow characteristics influencing detection of occupants vary between islands, although we examined too few sites to assess whether systematic variation occurs, especially with respect to the presence of human harvesting.

Development of Correction Model

The correction factors developed in this study showed potential for estimating absolute population sizes, especially

on The Snares and Bench Island. Although we observed similar accuracy when the improved occupancy measure was calculated, performance of this raw occupancy measure will be reduced at sites where chicks are distributed at greater distances from burrow entrances. Overall, correction factors incorporating the simple detection function outperformed those that used alternative functions.

The incremental correction method is more suitable for estimating burrowscope bias than the average correction factor, which requires knowledge of the actual distribution of chicks within burrow systems. This information is seldom available in burrowscoping studies, thereby restricting its use for future surveys. However, this method allowed the reliability of the incremental method to be assessed. The similar results observed between methods suggest that the methodology of the incremental approach is valid.

Emphasis must now be placed on determining the utility of applying a general model to subsequent sites monitored. The reliability of model predictions can only be tested with independent data over the full range of conditions likely to be encountered when monitoring sooty shearwaters. We observed uncertainty over the form of the detection function at greater distances from burrow entrances, due to the relatively shallow distributions of chicks on the study islands, which prevented adequate modeling at long distances. Ideally, validation would occur at some sites where chicks are distributed further away from burrow entrances to overcome the current uncertainty in the detection function at those distances.

Further research also could validate the structure of the model, although the full detection model tested here seems to encompass most of the important features that prevent correct detection. Many of the minor influential characteristics such as substrate, water, and roots are probably controlled for by incorporating distance as a predictor.

Small sample sizes of double-counted chicks prevented adequate modeling of their occurrence based on burrow features. The presence of double counting will inflate

corrected occupancy estimates, although in most cases, this effect seems to be relatively minor, especially where sampling occurs on transects.

MANAGEMENT IMPLICATIONS

We recommend using our simple detection function model to estimate and incorporate detection probabilities into future population estimates of sooty shearwaters, especially after model validation and refinement has been carried out using independent data. Any current attempts to assess the impacts of human harvesting will contain measurement error of an unknown magnitude; so, accounting for detection probability by measuring and predicting the effects of measurable burrow characteristics will help remove a further source of uncertainty.

ACKNOWLEDGMENTS

We thank the Rakiura Titi Islands Administering Body for guidance throughout the study and the New Zealand Department of Conservation for permission to undertake the experiment on The Snares and Bench Island. We are grateful to the Muttonbirders on Putauhinu for hospitality and for allowing the research to be undertaken on the island. S. Pilkington, H. Bryant, D. Vercoe, G. Pikerell, and B. McKinley assisted with fieldwork. Financial assistance was provided by the Foundation for Research, Science and Technology, University of Otago, and New Zealand Aluminum Smelters Ltd. The manuscript was improved greatly by the comments of the Associate Editor and 2 anonymous reviewers.

LITERATURE CITED

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin* 29:1294–1297.

Anderson, D. R. 2003. Response to Engeman: index values rarely constitute reliable information. *Wildlife Society Bulletin* 31:288–291.

Cuthbert, R., and L. S. Davis. 2002. Adult survival and productivity of Hutton's shearwaters. *Ibis* 144:423–432.

Dunlop, J. N., P. Long, I. Stejskal, and C. Surman. 2002. Inter-annual variation in breeding participation at four Western Australian colonies of the Wedge-tailed Shearwater *Puffinus pacificus*. *Marine Ornithology* 30: 13–18.

Dyer, P. K., and G. J. E. Hill. 1991. A solution to the problem of

determining the occupancy status of wedge-tailed shearwater *Puffinus pacificus* burrows. *Emu* 91:20–25.

Hamilton, S. 1998a. Determining burrow occupancy, fledging success and land-based threats to mainland and near-shore island sooty shearwater (*Puffinus griseus*) colonies. *New Zealand Journal of Zoology* 25:443–453.

Hamilton, S. 1998b. A test of burrow occupancy of Sooty Shearwaters (*Puffinus griseus*) using chick response to sound. *Notornis* 45:64–66.

Hamilton, S. 2000. How precise and accurate are data obtained using an infra-red scope on burrow-nesting sooty shearwaters *Puffinus griseus*? *Marine Ornithology* 28:1–6.

Hunter, C. M. 2001. Demography of procellariids: model complexity, chick quality, and harvesting. Dissertation, University of Otago, Dunedin, New Zealand.

Jones, C., S. Bettany, H. Moller, D. Fletcher, P. Lyver, and J. de Cruz. 2003. Burrow occupancy and productivity at coastal sooty shearwater (*Puffinus griseus*) breeding colonies, South Island, New Zealand: can mark-recapture be used to estimate burrowscope accuracy. *Wildlife Research* 30:377–388.

Lyver, P., S. Hamilton, M. McKenzie, I. Dickson, M. Doohar, T. Broad, and H. Moller. 1998. A burrowscope for examining petrel nests in burrows. Conservation Advisory Science Notes No. 209. Department of Conservation, Wellington, New Zealand.

McKechnie, S. 2004. Sooty Shearwater burrow systems: detection of occupants, chick quality and nest-site selection. Thesis, University of Otago, Dunedin, New Zealand.

Moller, H. 1996. Customary use of indigenous wildlife: towards a bicultural approach to conserving New Zealand's biodiversity. Pages 89–125 in B. McFagen and P. Simpson, compilers. Biodiversity: papers from a seminar series on biodiversity, hosted by Science & Research Unit, Department of Conservation, Wellington, New Zealand.

Moller, H., J. De Cruz, D. Fletcher, K. Garrett, C. Hunter, C. J. Jones, J. Kitson, P. Lyver, J. Newman, B. Russell, P. Scofield, and D. Scott. 1999. Kia Mau Te Titi Mo Ake Tōnu Atu: goals, design and methods. Wildlife Management Report 117, University of Otago, Dunedin, New Zealand.

Moller, H., M. Charteris, B. Russell, and J. Newman. 2003a. Detection of sooty shearwater chicks (*Puffinus griseus*) by response to sound. DOC Science Internal Series 114, Department of Conservation, Wellington, New Zealand.

Moller, H., C. Hunter, M. Howard, P. McSweeney, P. Scofield, D. Scott, and S. Uren. 2003b. Automated system for measuring breeding burrows entry and exit by sooty shearwaters (*Puffinus griseus*). DOC Science Internal Series 129, Department of Conservation, Wellington, New Zealand.

Seto, N. W. H., and P. Jansen. 1997. A miniature camera system for examining petrel burrows. *Journal of Field Ornithology* 68:530–536.

Warham, J. 1990. The petrels—their ecology and breeding systems. Academic Press, London, England.

Warham, J., and G. J. Wilson. 1982. The size of the sooty shearwater population at the Snares Islands, New Zealand. *Notornis* 29:23–30.

Associate Editor: Rodewald.