

**ON THE POWER TO DETECT CHANGES USING THE STANDARD METHODS FOR MONITORING PARAMETERS OF PREDATORY SPECIES**

P. Boveng and J.L. Bengtson

**Abstract**

The CCAMLR guidelines (CCAMLR, 1989) for analyzing the statistical power to detect changes are applied using variance estimates from studies at the South Shetland Islands of lengths of first incubation shifts by penguins (Standard Method A2.1), duration of penguin foraging trips (A5.1), penguin breeding success (A6.1), penguin chick diet (A8.1), and fur seal foraging trip duration (C2.0). The other Standard Methods are not evaluated because the power analysis in the CCAMLR guidelines is inappropriate (fur seal pup growth rate C1.0), and/or because data from the studies considered here are insufficient for conducting power analyses (penguin arrival weight A1.1, trend in penguin population size A3.1, penguin demography A4.1, penguin fledging weight A7.1). Alternative forms for analyses of the power of Standard Methods A3.1, A5.1, C1.0, and C2.0 are proposed.

**Résumé**

Les lignes directrices de la CCAMLR (CCAMLR, 1989) concernant l'analyse de la puissance statistique de détection des changements sont mises en application en utilisant des estimations de variance provenant d'études faites aux îles Shetland du Sud. Celles-ci portent sur la durée du premier tour d'incubation des manchots (Méthode standard A2.1), la durée des sorties alimentaires des manchots (A5.1), la réussite de la reproduction des manchots (A6.1), le régime alimentaire des jeunes manchots (A8.1) et la durée des sorties alimentaires des otaries (C2.0). Les autres Méthodes standard ne sont pas évaluées, l'analyse de la puissance dans les lignes directrices de la CCAMLR étant impropre (taux de croissance de jeunes otaries C1.0), et/ou les données provenant des études prises en considération dans ce document étant insuffisantes pour effectuer des analyses de la puissance (poids du manchot à l'arrivée A1.1, tendances dans la taille de la population de manchots A3.1, démographie des manchots A4.1, poids des jeunes à la première mue A7.1). D'autres formes d'analyses de la puissance des Méthodes standard A3.1, A5.1, C1.0 et C2.0 sont proposées.

**Резюме**

Руководства АНТКОМа (CCAMLR, 1989) по анализу статистической мощности при выявлении изменений применяются с использованием оценок продолжительности первой инкубационной смены пингвинов (Стандартный метод А2.1), продолжительности периодов кормления пингвинов в море (Стандартный метод А5.1) репродуктивного успеха пингвинов (Стандартный метод А6.1), рациона птенцов пингвина (Стандартный

метод А8.1) и продолжительности периодов кормления морского котика (Стандартный метод С2.0), полученные в ходе исследований на Южных Шетландских островах. Остальные Стандартные методы не оцениваются в связи с тем, что приведенный в руководствах АНТКОМа анализ мощности неприменим (темп роста щенков морского котика С1.0), и/или в связи с тем, что данных, полученных в результате рассматриваемых исследований, недостаточно для выполнения анализа мощности (вес пингвинов по прибытии в колонию А1.1, тенденции изменения размера популяции пингвинов А3.1, демография пингвинов А4.1, вес пингвинов при оперении А7.1). Предлагаются альтернативные формы анализа мощности Статистических методов А3.1, А5.1, С1.0 и С2.0.

### Resumen

Las directivas de la CCRVMA (CCAMLR, 1989) para analizar la capacidad estadística para detectar cambios, se aplican utilizando estimaciones de variación procedente de los estudios realizados en las Islas Shetland del Sur sobre las duraciones de los primeros turnos de incubación de pingüinos (Método Estándar A2.1), duración de los viajes de alimentación de los pingüinos (A5.1), éxito en la reproducción de pingüinos (A6.1), dieta de los polluelos de pingüinos (A8.1), y duración de los viajes de alimentación de los lobos finos (C2.0). Los otros Métodos Estándar no están evaluados porque la capacidad de análisis de las directivas de la CCRVMA no es apropiada (índice de crecimiento de los cachorros de lobo fino C1.0), y/o porque datos de los estudios considerados aquí son insuficientes para conducir análisis de capacidad (peso de llegada de los pingüinos A1.1, tendencias en el tamaño de la población de pingüinos A3.1, demografía de pingüinos A4.1, peso de pingüinos al emplumaje A7.1). Se proponen maneras alternativas para los análisis de la capacidad de los Métodos Estándar A3.1, A5.1, C1.0 y C2.0.

## 1. INTRODUCTION

Members of the Working Group for the CCAMLR Ecosystem Monitoring Program (CEMP) have been asked to conduct "sensitivity analyses" of existing data on predator parameters. Those analyses, perhaps more properly called "power analyses", should provide information about the power of the standard methods to detect changes in parameters monitored for predators.

Guidelines for conducting one type of power analysis were compiled by the CCAMLR Secretariat (CCAMLR, 1989). That paper described the method for estimating the sample size required to detect a specified difference ( $\delta$ ) between or among years, subject to acceptable probabilities of rejecting a true null hypothesis ( $\alpha$ ) or accepting a false null hypothesis ( $\beta$ ). The required input for the analysis is an estimate of the standard deviation ( $s$ ) if  $\delta$  is in the same units as the mean, or the coefficient of variation ( $cv$ ) if  $\delta$  is expressed as a percentage of the typical mean value.

The analysis described in CCAMLR (1989) provides information about the power to distinguish between the mean values of parameters obtained in two or more years. We may also wish to know what sample size is required in an annual monitoring effort to detect a linear or log-linear trend in a parameter. Methods for such analysis were presented by Gerrodette (1987). When the parameter values are proportions, the sample sizes required to detect a specified slope can be obtained as described by Nam (1987). The quantities required in a power analysis for detecting trends are essentially the same as those required for the previous analysis, though the formulae are slightly different. The quantity  $\delta$  must be expressed as the minimum detectable slope of a regression.

For some parameters the monitoring design is such that neither analysis mentioned above is strictly appropriate. Because of the complexity of some of the designs, there may be no general formula for evaluating power or determining a required sample size. For those parameters we advocate using numerical simulations to assess the ability to detect changes.

In the remainder of this paper, we first suggest several changes to the guidelines suggested in CCAMLR (1989). Then we consider each of the standard methods for monitoring parameters of predatory species. For each method we discuss appropriate techniques for evaluating the power to detect changes or for determining required sample sizes. For some of the parameters, we apply those techniques to existing estimates of the uncertainty in the parameter. We consider unpublished data from the US AMLR monitoring program at Seal Island, Elephant Island (60°59.5'S, 55°24.5'W), and from the study by Trivelpiece et al. (1989) at Admiralty Bay, King George Island (62°10'S, 58°30'W).

## 2. COMMENTS ON CCAMLR (1989)

First, we note that there was a typographical error in equation (1) of CCAMLR (1989) by the omission of an exponent on the sum of the t-values. The corrected form of the equation is reproduced here as

$$n \geq 2(s/\delta)^2 \{t_{\alpha, [v]} + t_{2(1-P), [v]}\}^2 \quad (1)$$

where  $n$  = the required sample size,  
 $s$  = the standard deviation computed from the raw or transformed data,  
 $\delta$  = the smallest true difference that it is desired to detect,  
 $v$  = degrees of freedom for error in the ANOVA or t-test [ $v=r(n-1)$  where  $r$  is the number of years or replicates],  
 $P$  = desired probability that a difference will be found if it is as large as  $\delta$  ( $P$  is the statistical power),

- $t_{\alpha,[v]}$  = value from a two-tailed t-table with  $v$  degrees of freedom and corresponding to probability  $\alpha$ , and
- $t_{2(1-P),[v]}$  = value from a two-tailed t-table with  $v$  degrees of freedom and corresponding to probability  $2(1-P)$ .

We also suggest that an additional column be added to the table described in paragraph 19 of CCAMLR (1989). The additional column should contain the sample sizes required when  $\alpha = \beta = .05$ . For many of the parameters, these conditions will not be attainable due to logistic impracticalities, so  $\alpha$  and  $\beta$  may need to be reduced. However, unlike situations typically presented in statistical textbooks, we suggest that in a conservation context, type 2 errors ( $\beta$  errors) may be at least as important as type 1 errors ( $\alpha$  errors). For example, failure to detect a decrease in penguin reproductive success that actually occurred (type 2 error) may be equally or perhaps more serious than detecting an apparent, but false, decrease (type 1 error).

The probabilities of committing the two types of errors have a direct bearing on the circumstances under which decisions are taken to impose conservation measures. If one were required to choose between the types of error that one would be willing to accept, a conservative approach might suggest that it would be preferable to take protective action when there was an apparent but false need rather than failing to take protective action when there was a true but undetected need for conservation measures.

Finally, the suggestion that members compute estimates of required sample sizes for both two and five years of monitoring seems unnecessary. The question addressed by equation (1) simply uses the estimated variance of a parameter to determine how small a change ( $\delta$ ) could be detected between or among years with a given sample size. The number of years monitored will have little effect on the size of the change that is detectable. Therefore, the required sample sizes will differ by very little, if at all, for estimates based on two years or five years of monitoring (though for detecting a trend, the number of years will be very important). This can be seen in the table presented in paragraph 19 of CCAMLR (1989). In the following analyses, we present only estimates of sample sizes required assuming two years of monitoring.

### 3. METHODS FOR PENGUINS

#### 3.1 Weight on Arrival at Breeding Colonies (A1.1)

As shown in CCAMLR (1989), the standard method for estimating the population mean weight on arrival at the rookery is incomplete because it does not account for seasonal variation in rates of arrival. No data were available, from the sources mentioned above, to evaluate the variability in this parameter. Croxall et al. (1988) evaluated the power to detect changes in this parameter for macaroni and gentoo penguins at Bird Island and for Adélie and chinstrap penguins at Signy Island.

#### 3.2 Length of First Incubation Shift (A2.1)

Trivelpiece et al. (1989) measured the durations of first incubation shifts of male and female Adélie penguins for five years and of male and female chinstrap penguins for four years at King George Island. Although a separate analysis could be conducted to determine the statistical power which could be achieved by routine monitoring at the actual sampling intensity attained in each year, a more useful approach would combine the standard deviations from all five years (four years for chinstraps) into a single statistic that would

be more representative of the variation in this parameter. Then the analysis suggested in CCAMLR (1989) would be performed to investigate the trade-off between sample size and the power to detect changes in the parameter.

One appropriate statistic for estimating the standard deviation of a parameter when several years' data are available is the "mean-squared-error" (MSE) term from a one-way analysis of variance with years representing treatments. The square root of the MSE term is a "pooled" estimate of the quantity  $s$  of equation (1). Given the resulting value of  $s$  and specified levels of acceptable type 1 and 2 errors ( $\alpha$  and  $\beta$  respectively), the minimum detectable difference  $\delta$  can be expressed as a function of sample size, or vice-versa.

The MSE within years for first incubation shifts by male Adélie penguins was 6.92 and for female Adélies was 13.07. Using the square-root of the MSE as an estimate of the typical standard deviation in this parameter, we used equation (1) to compute the requisite sample sizes for detecting changes of 10%, 20% and 30% of the mean value (males, 12.4 days; females, 10.0 days) for the 5 years studied (Table 1).

For chinstrap penguins, the MSE for length of first incubations by males was 9.0 and for first incubations by females was 5.5. The mean length of first incubations by males was 5.8 days and by females was 3.9 days. The requisite sample sizes for detecting changes of 10%, 20% and 30% of the mean values are shown in Table 2.

No data on this parameter were available from the study at Seal Island.

### 3.3 Trends in Breeding Population Size (A3.1)

There were no suitable estimates of variance for the data on breeding population size from the two studies considered here. This will likely be a problem for other studies as well. Some techniques for abundance estimation result in particular and known relations between the mean and variance of the estimate (Gerrodette, 1987). For studies utilizing such techniques, a range of reasonable guesses about the variance in the estimates could be applied to the formulae presented by Gerrodette (1987) to make a preliminary assessment of the power to detect trends. As an initial upper bound to the range of variance, one might simply use the variance between several recently obtained annual censuses. We have not done such an analysis for the data considered here. Alternatively, replicate counts (at least three replicates) of colonies in future seasons would allow estimating the variance.

### 3.4 Demography (A4.1)

The parameters to be monitored by method A4.1 are annual survivorship, age at first reproduction, and cohort strength. The banding program in "Method A" will provide estimates of the average survival of adult breeding birds. "Method B" will provide age-specific survival estimates as well as estimates of the age at first reproduction. Cohort strength is not well defined in the method's description but presumably involves some measure of the size of a cohort relative to that of adjacent cohorts.

Estimates for any of the three parameters will be available only after several years of monitoring. Estimating the variability of the parameters may require an even longer period, though some of the methods for estimating survival rates and age at first reproduction in a mark-resight study are associated with theoretically derived analytic estimates of variance. Sufficient data for evaluating the variability in any of the three demographic parameters were not available from either of the two studies considered here.

### 3.5 Duration of Foraging Trips (A5.1)

The design for monitoring the duration of foraging trips is a nested, or hierarchical, experiment. That is, for any particular year, several breeding birds are chosen (at random) for monitoring and the lengths of foraging trips by each bird are measured. In general, the sample of birds monitored one year is not the same as the sample monitored the next year. Thus, each sample of birds is said to be "nested" within years when two or more years are being compared. The analysis must take into account that some of the variation between or among years is simply due to the fact that the birds used were not the same every year.

In addition to the complication of a nested design, the standard method for monitoring foraging trip duration may be complicated by an inverse relationship between trip duration and number of trips. That is, if a bird with a propensity to make long trips tends to make fewer of them than a bird which makes short trips, the usual statistical tests will not be reliable. The usual methods for unbalanced (unequal sample size) nested ANOVA assume that sample size and magnitudes of the items in the sample are independent.

The consequence of the problems mentioned above is that there is no simple or general formula analogous to equation (1) to produce the required sample size for effective monitoring of foraging trip duration. Existing estimates of the variance of trip duration for an individual bird, the covariance of trip duration with number of trips, and the variance between birds could be used in numerical simulations to assess the power of the method or to choose reasonable sample sizes. Such an exercise is beyond the scope of this paper. Until such time as those simulations are conducted, we advocate the following as a first approximation to a power analysis of the standard method for foraging trip duration.

The problem of correlation between trip duration and numbers of trips can be avoided by computing a mean trip duration separately for each bird and giving equal weight to the mean for each bird when computing the annual mean. Then the sample size is simply the number of birds, without regard to the number of trips made by individual birds. Likewise, the variance is the variance between the means of individual birds, without regard for the variance of the lengths of trips made by individual birds. Applying existing estimates of the between-bird variance to the formula in equation (1), yields the approximate required sample sizes, in terms of numbers of birds.

Trivelpiece et al. (1989) used visual monitoring of pairs of Adélie and chinstrap penguins to estimate average foraging trip durations for each species at King George Island in the 1981/82 and 1982/83 seasons. The mean trip duration for 37 pairs of Adélie penguins in the two years was 24.3 hr. The mean duration of trips made by 47 pairs of chinstrap penguins was 16.7 hr. The MSE for duration of trips was 20.5 for Adélies and 40.7 for chinstraps. Sample sizes, estimated by the approximation method described in the previous paragraph, for detecting changes of 10%, 20%, and 30% of the mean values are shown for Adélie and chinstrap penguins in Table 3.

The data on durations of chinstrap and macaroni penguin foraging trips at Seal Island, monitored with radio-transmitters, were not summarized in time for analysis in this report. However, a comparison of data obtained using visual and radio-transmitter methods would be useful in determining whether sample sizes required for these techniques differ.

Finally, we note that the scheme presented in CCAMLR (1989) for computing an annual mean trip duration (paragraphs 28 and 29) ignores the problems of nested design and correlation between trip duration and number of trips. That scheme gives equal weight to each foraging trip made during a particular year. The simplification we presented above partially avoids that problem, but ignores the five-day stratification scheme which may be necessary to account for systematic changes in trip duration within a season.

### 3.6 Breeding Success (A6.1)

The reproductive success of 150 pairs of chinstrap penguins was monitored at Seal Island in 1987/88, though the field work was begun on 25 December, well after the first egg was laid. Assuming no egg loss occurred prior to the study, the mean number of chicks raised to creche age per pair was 1.26, with a standard deviation of 0.87. If those numbers are typical, the minimum number of nests which would have to be monitored to detect changes of 10% to 30% of the mean value are shown in Table 4\*.

Breeding success was monitored in 33 pairs of macaroni penguins at Seal Island in 1987/88. Because the number of macaroni chicks raised per pair can only equal zero or one, the data can be treated as binomial proportions. Fleiss (1981, pp. 41-42) gives the appropriate formulae for the sample sizes required to detect a specified difference between two proportions. Table 5 presents estimated minimum sample sizes to detect a change from the observed mean of 0.91 chicks per pair to each of several new values for the reproductive success parameter. Note that the sample sizes in Table 5 are generally much larger than the 33 pairs used to estimate the mean for the analysis. Therefore, this analysis is presented more as an example of the method than as a true power analysis.

The reproductive success of Adélie penguins, measured as the number of chicks fledged per breeding pair was monitored by Trivelpiece et al. (1989) at Admiralty Bay, King George Island. During each of six years (not all consecutive), several hundred breeding pairs were observed. Therefore, a mean and standard deviation of the number of chicks fledged per pair is available for each of the six breeding seasons. However, at the time of this writing we were unable to adequately interpret those data because the quoted standard deviations were larger than is possible for a variable which can only take on values of zero, one, or two.

### 3.7 Chick Fledging Weight (A7.1)

As shown in paragraphs 9 and 10 of CCAMLR (1989), the standard method for estimating the population mean fledging weight may not yield a reliable annual index if the true average fledging weight varies between five-day periods. The average weights of chinstrap penguins fledging in consecutive five-day periods differed significantly at Seal Island in 1989. If this result is confirmed for other species, sites, and years, the standard method will need to be revised to include a strategy for weighting the mean from each five-day period by the number of chicks fledged in that period.

One potential method for obtaining an index of the numbers of chicks fledged in each 5-day period would be to keep daily counts of the numbers of chicks (unfledged) remaining in several discrete, small or peripheral colonies. Adding up the daily decrements in chick numbers might yield a satisfactory estimate of the relative numbers fledging in the several 5-day periods. The daily decrements will include mortality, but if the rate of mortality is small relative to the rate of fledging and if the mortality rate is not dramatically different between different five-day periods, the estimates should still be reliable.

### 3.8 Chick Diet (A8.1)

The data from the standard methods for monitoring penguin chick diet will consist of total wet weights of the contents of individual stomachs or of more detailed information such

---

\* E. Marschoff (personal communication) has suggested that changes in the individual proportions of pairs raising zero, one or two chicks might be detectable using sample sizes much smaller than those apparently required to detect changes in the mean number of chicks raised.

as the proportion of the total weight comprising a particular prey taxon. We consider here only the data on total weight, deferring analysis of the proportion data to another paper evaluating the different types of analysis appropriate for proportions.

Trivelpiece et al. (1989) measured meal sizes of Adélie and chinstrap penguin chicks during three years at King George Island. The average meal size for Adélie penguins was 533 g and the average for chinstraps was 443 g. Pooling the variances as described above, the square-root of the MSE for meal size of Adélie penguins was 191.8 and for chinstrap penguins the square-root of the MSE was 219.8. Using the square-root of the MSE as the typical standard deviation, we estimated the sample sizes required to detect changes of 10%, 20%, and 30% of the mean values for Adélie and chinstrap meal size (Table 6). Because of the relatively high variability in this parameter, the samples required to detect a modest change in the mean are very large.

#### 4. METHODS FOR FUR SEALS

##### 4.1 Fur Seal Pup Growth (C1.0)

The data from this method consist of periodic weights of a marked sample of pups or of random samples of pups. The resulting index will be the value of the slope of a regression of pup weight versus date. The techniques for comparing the slopes of two or more regressions are well established, and found in most introductory statistics texts. However, several aspects of this technique should be considered in more detail before a formal power analysis is conducted.

As discussed in CCAMLR (1989), the assumptions for linear regression analysis (homogeneity of variance, fit to linear model) should be investigated. Assuming that those investigations reveal that linear regression is appropriate, a power analysis will still be more difficult than one which relies on equation (1) for the following reason.

There are two measures of sample size which must be considered in designing a test for differences between pup growth rates among several years. The first is the number of pups to be weighed on a particular date. The second is the number of dates within the pup-rearing season on which pup weights are to be sampled. The issue of how many times samples of pups should be weighed is further complicated by consideration of the time intervals between successive weighings. Weights taken on each of four consecutive days will not provide the "contrast" afforded by weights collected at four one-week intervals. A thorough analysis of the power to detect changes in pup growth rates will likely require numerical simulations in which the response of the parameter to systematic changes in numbers of pups and in numbers and spacing of weighing dates is explored. Because good estimates of the "within-date" variance of pup weights are available from two years at Seal Island, we intend to pursue this topic in the near future.

##### 4.2 Foraging/Attendance Duration (C2.0)

As we discussed above in the section on penguin foraging trip duration, this monitoring scheme is a nested design with the complication that trip duration and number of trips made by an individual are inversely related (Figure 1). One approach, to prevent undue weighting of females which tend to make short trips, would be to weight the mean trip duration for each female equally when computing the overall mean trip duration for the year. Using that method on data from Seal Island, the following results were obtained.

The average of the mean trip durations of 19 fur seal females at Seal Island in 1986/87 was 59.7 hr, with a standard deviation of 43.6 hr. In 1988/89, the average from 13 females was 52.4 hr, with a standard deviation of 44.4 hr. With coefficients of variation



of about 80%, it would clearly be very difficult to detect differences between years using this type of analysis. This is shown in Table 7, by the very large sample sizes which would be required (again calculated by using a pooled estimate of variance and equation (1)).

The data shown in Figure 2 suggest an explanation for the low statistical power obtained by the previous method: the lengths of foraging trips tend to decrease as the season progresses. Therefore, some of the "within-year" variation may be due to differences between the dates on which different females' trips were made. Modifications to the standard method, such as stratifying foraging data by calendar date or using analysis of covariance to remove the effect of date, will likely reduce the within-year variation, thereby reducing the size of the sample of females required to detect a specified difference between years.

Other modifications, addressing the problems we described above, to the standard method for fur seal foraging trip length are also possible. For example, another way to give equal weight to the mean of each female's trip durations would be to use only a specified number, say the first 5, of her postpartum foraging trips. That strategy may also partially remove the "date effect", by reducing the portion of the season during which trips are measured. In any case, a thorough investigation of the power of standard method C2.0, with the modifications we described above, is beyond the scope of this paper. We will use the data from Seal Island to conduct such an analysis, making the results available to the Working Group as soon as they are completed.

#### REFERENCES

- CCAMLR. 1989. Instructions for the preparation of sensitivity analysis. WG-CEMP-89/13.
- CROXALL, J.P., T.S. MCCANN, P.A. PRINCE, and P. ROTHERY. 1988. Reproductive performance of seabirds and seals at South Georgia and Signy Island, South Orkney Islands, 1976-1987: Implications for Southern Ocean monitoring studies. *In*: SAHRHAGE, D. (Ed.). *Antarctic Ocean and Resources Variability*. Berlin: Springer-Verlag.
- FLEISS, J.L.. 1981. *Statistical Methods for Rates and Proportions*. New York: Wiley. 321 pp.
- FRASER, W.R., and D.G. AINLEY. 1989. United States seabird research undertaken as part of the CCAMLR Ecosystem Monitoring Program at Palmer Station, 1988-1989. Report submitted to National Marine Fisheries Service, Northwest and Alaska Fisheries Center, National Marine Mammal Laboratory, 7600 Sand Point Way N.E., Seattle, WA 98115.
- GERRODETTE, T. 1987. A power analysis for detecting trends. *Ecology* 68(5): 1364-1372.
- NAM, J. 1987. A simple approximation for calculating sample sizes for detecting linear trend in proportions. *Biometrics* 43:701-705.
- SOKAL, R.R. and F.J. ROHLF. 1981. *Biometry*. San Francisco: Freeman. pp. 859.
- TRIVELPIECE, W.Z., S.G. TRIVELPIECE and J. KJELMYR. 1989. Evaluation of annual variation in life history parameters of Adélie and chinstrap penguins: Their potential for monitoring of fishery impacts on the Southern Ocean Marine Ecosystem. Report submitted to National Marine Fisheries Service, Northwest and Alaska Fisheries Center, National Marine Mammal Laboratory, 7600 Sand Point Way N.E., Seattle, WA 98115.

Table 1: Estimates of the minimum sample sizes (number of birds) required to detect a change of magnitude equal to  $\delta$  in the mean lengths (days) of Adélie penguin first incubation shifts are presented for a range of values for desired statistical power. All estimates are for tests conducted after two years of monitoring. Data are from Trivelpiece et al. (1989) Table 11.

Incubation shifts by males:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
1.24 (10%)	119	96	72	46
2.48 (20%)	31	25	19	13
3.72 (30%)	15	12	9	7
For $\alpha = 0.10$ :				
1.24 (10%)	99	79	57	34
2.48 (20%)	26	21	15	10
3.72 (30%)	12	10	8	5

Incubation shifts by females:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
1.00 (10%)	341	276	207	130
2.00 (20%)	86	70	53	34
3.00 (30%)	39	32	24	16
For $\alpha = 0.10$ :				
1.00 (10%)	284	225	163	95
2.00 (20%)	72	57	42	25
3.00 (30%)	33	26	19	12

Table 2: Estimates of the minimum sample sizes (number of birds) required to detect a change of magnitude equal to  $\delta$  in the mean lengths (days) of chinstrap penguin first incubation shifts are presented for a range of values for desired statistical power. All estimates are for tests conducted after two years of monitoring. Data are from Trivelpiece et al. (1989), Table 12.

Incubation shifts by males:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
0.58 (10%)	698	564	422	264
1.16 (20%)	176	142	107	67
1.74 (30%)	79	64	48	31
For $\alpha = 0.10$ :				
0.58 (10%)	581	460	332	194
1.16 (20%)	146	116	84	50
1.74 (30%)	66	52	38	23

Incubation shifts by females:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
0.39 (10%)	953	771	576	360
0.78 (20%)	239	194	145	91
1.17 (30%)	107	87	65	41
For $\alpha = 0.10$ :				
0.39 (10%)	794	629	454	265
0.78 (20%)	200	158	114	67
1.17 (30%)	89	71	52	31

Table 3: Estimates of the minimum sample sizes (number of birds) required to detect a change of magnitude equal to  $\delta$  in the mean duration (hours) of chinstrap and Adélie penguin foraging trips are presented for a range of values for desired statistical power. All estimates are for tests conducted after two years of monitoring. Data are from Trivelpiece et al. (1989), Table 9.

Adélie penguins:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
2.43 (10%)	92	74	56	36
4.86 (20%)	24	20	15	10
7.29 (30%)	12	10	8	6
For $\alpha = 0.10$ :				
2.43 (10%)	77	61	44	26
4.86 (20%)	20	16	12	8
7.29 (30%)	10	8	6	4

Chinstrap penguins:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
1.67 (10%)	382	309	231	145
3.34 (20%)	97	78	59	37
5.07 (30%)	44	36	27	18
For $\alpha = 0.10$ :				
1.67 (10%)	318	252	182	107
3.34 (20%)	81	64	47	28
5.07 (30%)	37	29	21	13

Table 4: Estimates of the minimum sample size (number of pairs) required to detect a change of magnitude equal to  $\delta$  in the mean number of chinstrap chicks raised to creche age per breeding pair are presented for a range of values for desired statistical power. All estimates are for tests conducted after two years of monitoring. Data are from Seal Island, 1987/88.

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
0.129 (10%)	1185	958	716	448
0.258 (20%)	297	241	180	113
0.387 (30%)	133	108	81	51
For $\alpha = 0.10$ :				
0.129 (10%)	986	782	564	329
0.258 (20%)	248	196	142	83
0.387 (30%)	111	88	64	38

Table 5: Estimates of the minimum sample sizes (number of pairs) required to detect a change in macaroni penguin reproductive success from a value of 0.91 chicks raised per pair to each of the new values (P2) listed below. The usual quantity  $\delta$  would be the difference between P2 and 0.91. All estimates are for tests conducted after two years of monitoring. Data are from Seal Island, 1987/88.

P2	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
0.95	1104	902	686	446
0.85	792	647	492	319
0.80	281	231	177	117
0.75	152	125	97	65
0.70	98	81	63	43
For $\alpha = 0.10$ :				
0.95	927	744	551	341
0.85	665	533	394	243
0.80	236	191	143	90
0.75	128	104	79	51
0.70	83	68	52	34

Table 6: Estimates of the minimum sample sizes required to detect a change of magnitude equal to  $\delta$  in the average wet weight of Adélie and chinstrap penguin chick meals (grams) are presented for a range of values for desired statistical power. All estimates are for tests conducted after two years of monitoring. Data are from Trivelpiece et al. (1989), Table 6.

Adélie penguins:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
53.3 (10%)	338	273	205	128
106.6 (20%)	86	69	52	33
159.9 (30%)	39	32	24	16
For $\alpha = 0.10$ :				
53.3 (10%)	281	223	161	95
106.6 (20%)	71	57	41	25
159.9 (30%)	33	26	19	12

Chinstrap penguins:

$\delta$	Statistical Power (1- $\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
44.3 (10%)	641	519	388	243
88.6 (20%)	162	131	98	62
132.9 (30%)	73	59	44	28
For $\alpha = 0.10$ :				
44.3 (10%)	534	423	306	179
88.6 (20%)	135	107	77	46
132.9 (30%)	61	48	35	21

Table 7: Estimates of the minimum sample sizes (number of females) required to detect a change of magnitude equal to  $\delta$  in the mean duration (hours) of fur seal foraging trips are presented for a range of values for desired statistical power. These estimates are likely to be too high because they do not account for the "within-year" variation of the type shown in Figure 2. All estimates are for tests conducted after two years of monitoring. Data are from Seal Island, 1986/87 and 1988/89 seasons.

$\delta$	Statistical Power ( $1-\beta$ )			
	0.95	0.90	0.80	0.60
For $\alpha = 0.05$ :				
5.67 (10%)	1558	1261	942	589
11.35 (20%)	391	316	237	148
17.02 (30%)	175	141	106	67
For $\alpha = 0.10$ :				
5.67 (10%)	1298	1029	742	433
11.35 (20%)	326	258	186	109
17.02 (30%)	146	115	84	49

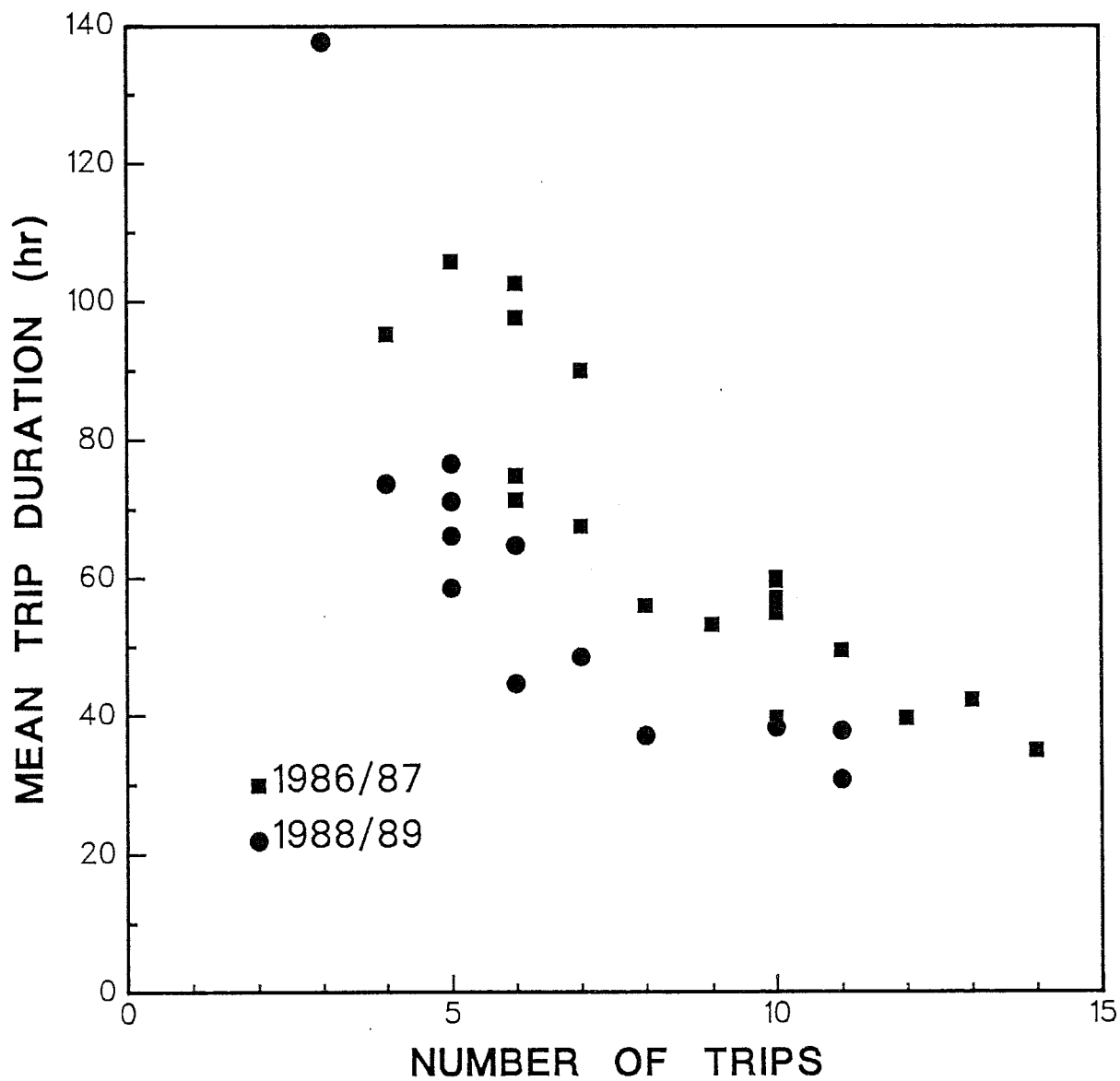


Figure 1: Relationship of individual duration and total number of foraging trips to sea made by female Antarctic fur seals at Seal Island, South Shetland Islands, Antarctica. Data from the 1986/87 season sampled trips from late December to mid-February; 1988/89 data sampled trips from mid January to late February.



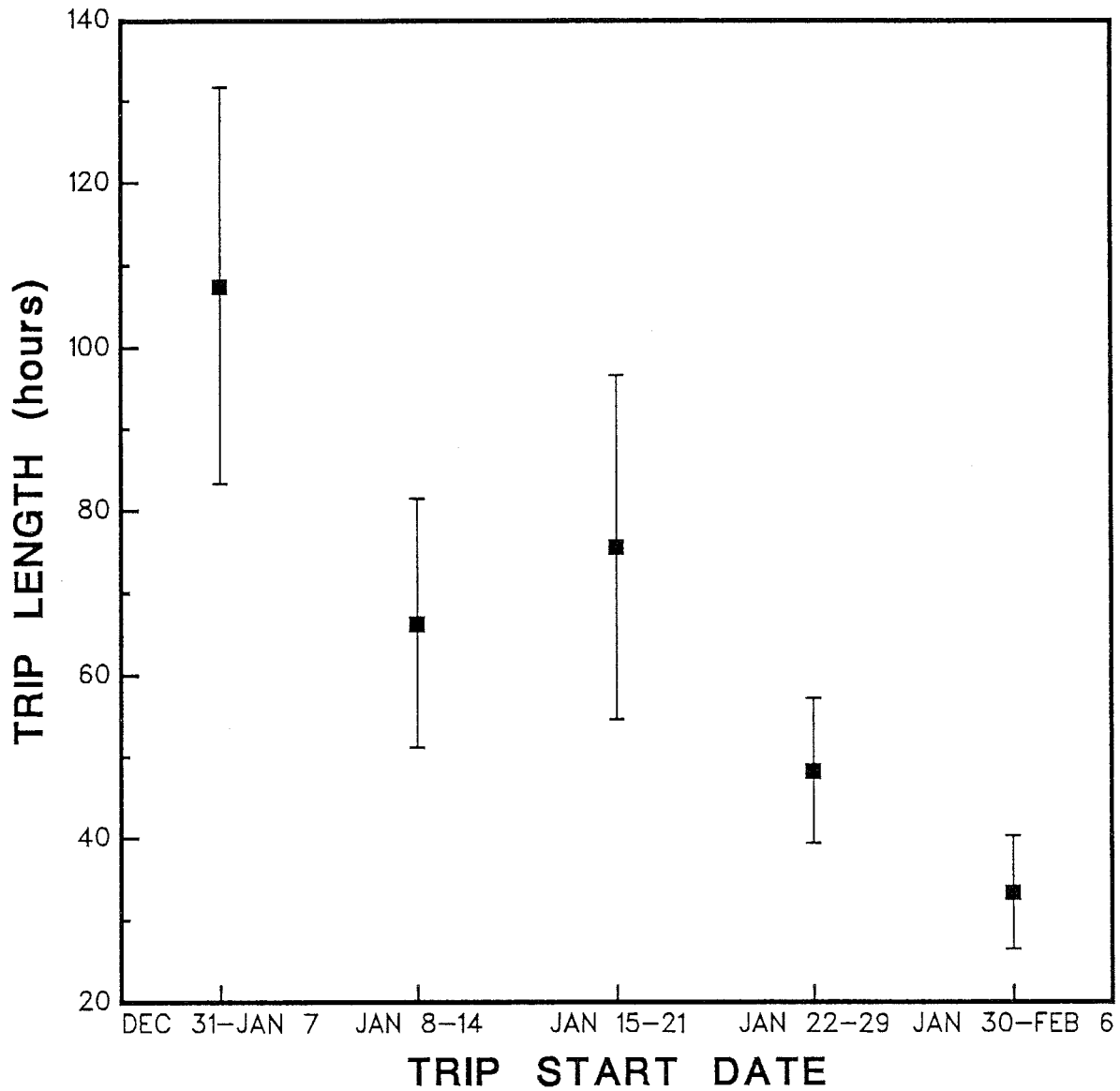


Figure 2: Within-season changes in the duration of foraging trips to sea made by female Antarctic fur seals at Seal Island, South Shetland Islands, Antarctica. Data are from the 1986/87 season.

## Liste des tableaux

- Tableau 1: Les estimations de taille minimum d'échantillons (nombre d'oiseaux) nécessaires pour détecter un changement d'ampleur égale à  $\delta$  dans la durée moyenne (jours) du premier tour d'incubation du manchot Adélie sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données sont fournies par Trivelpiece et al. (1989), Tableau 11.
- Tableau 2: Les estimations de taille minimum d'échantillons (nombre d'oiseaux) nécessaires pour détecter un changement d'ampleur égale à  $\delta$ , dans la durée moyenne (jours) du premier tour d'incubation du manchot à jugulaire, sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données sont fournies par Trivelpiece et al. (1989), Tableau 12.
- Tableau 3: Les estimations de taille minimum d'échantillons (nombre d'oiseaux) nécessaires pour détecter un changement d'ampleur égale à  $\delta$  dans la durée moyenne (heures) des sorties alimentaires des manchots à jugulaire et Adélie; elles sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données sont fournies par Trivelpiece et al. (1989), Tableau 9.
- Tableau 4: Les estimations de taille minimum d'échantillons (nombre de couples) nécessaires pour détecter un changement d'ampleur égale à  $\delta$  dans le nombre moyen de jeunes manchots à jugulaire par couple reproducteur, élevés jusqu'à l'âge d'entrée en crèche, sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données ont été recueillies à l'île Seal, 1987/88.
- Tableau 5: Les estimations de taille minimum d'échantillons (nombre de couples) nécessaires pour détecter un changement dans la réussite de la reproduction à partir d'une valeur de 0.91 jeunes élevés par couple, à chacune des nouvelles valeurs (P2) figurant sur la liste ci-dessous. La quantité normale de  $\delta$  serait la différence entre P2 et 0.91. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données ont été recueillies à l'île Seal, 1987/88.
- Tableau 6: Les estimations de taille minimum d'échantillons nécessaires pour détecter un changement d'ampleur égale à  $\delta$  dans le poids moyen mouillé (grammes) des repas des jeunes manchots à jugulaire et Adélie sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données sont fournies par Trivelpiece et al. (1989), Tableau 6.
- Tableau 7: Les estimations de taille minimum d'échantillons (nombre de femelles) nécessaires pour détecter un changement d'ampleur égale à  $\delta$  dans la durée moyenne (heures) des voyages alimentaires des otaries de Kerguelen sont présentées pour une fourchette de valeurs de la puissance statistique voulue. Ces estimations risquent d'être trop élevées parce qu'elles ne tiennent pas compte de la variation "dans le courant d'une année" du genre figurant à la Figure 2. Toutes les estimations proviennent de tests conduits après 2 ans de contrôle. Les données ont été recueillies à l'île Seal, saisons 1986/87 et 1988/89.

### Liste des figures

- Figure 1: Relation entre la durée d'un voyage et le nombre total de voyages alimentaires en mer entrepris par les otaries femelles de Kerguelen à l'île Seal, aux îles Shetland du Sud, en Antarctique. Les données de la saison 1986/87 proviennent d'échantillons prélevés lors de voyages effectués de fin décembre à mi-février; les données de 1988/89 proviennent d'échantillons prélevés lors de voyages effectués de mi-janvier à fin février.
- Figure 2: Changements, au cours d'une saison, de la durée des voyages alimentaires en mer entrepris par les otaries de Kerguelen femelles à l'île Seal, aux îles Shetland du Sud, en Antarctique. Les données concernent la saison 1986/87.

### Список таблиц

- Таблица 1: Предположительные минимальные размеры проб (количество особей), необходимые для выявления изменений порядка, равного  $\delta$ , средней продолжительности (дни) первой инкубационной смены пингвинов Адели представлены для ряда уровней статистической мощности. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. По данным Тривелписа и др. (Trivelpiece et al., 1989) б Таблица 11.
- Таблица 2: Предположительные минимальные размеры проб (количество особей), необходимые для выявления изменений порядка, равного  $v$ , средней продолжительности (дни) первой инкубационной смены пингвинов чинстрап представлены для ряда уровней статистической мощности. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. По данным Тривелписа и др. (Trivelpiece et al., 1989) б Таблица 12.
- Таблица 3: Предположительные минимальные размеры проб (количество особей), необходимые для выявления изменений порядка, равного  $\delta$ , средней продолжительности (часы) периода добывания пищи пингвинами чинстрап и Адели представлены для ряда уровней статистической мощности. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. По данным Тривелписа и др. (Trivelpiece et al., 1989) б Таблица 9.
- Таблица 4: Предположительные минимальные размеры проб (количество пар), необходимые для выявления изменений порядка, равного  $\delta$ , среднего количества выживших до "ясельного" возраста птенцов пингвина чинстрап на родительскую пару, представлены для ряда уровней статистической мощности. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. Данные были получены на острове Сил, в 1987/88 гг.
- Таблица 5: Предположительные минимальные размеры проб (количество пар), необходимые для выявления изменения репродуктивного успеха золотоволосых пингвинов от 0,91 выжившего птенца на пару до уровней (P2), указанных ниже. Обычно значение  $\delta$  будет

равно разнице между величинами  $P_2$  и 0,91. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. Данные были получены на острове Сил в 1987/88 гг.

Таблица 6: Предположительные минимальные размеры проб, необходимые для выявления изменения порядка, равного  $\delta$ , среднего сырого веса (граммы) пищи, потребляемой птенцами пингвинов Адели и чинстрап за один прием, представлены для ряда уровней статистической мощности. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. По данным Тривелписа и др. (Trivelpiece et al., 1989) б Таблица 6.

Таблица 7: Предположительные минимальные размеры проб (количество самок), необходимые для выявления изменений порядка, равного  $\delta$ , средней продолжительности (часы) периода добывания пищи морскими котиками представлены для ряда уровней статистической мощности. Вероятно, что эти оценки завышены, поскольку они не принимают во внимание "внутригодовые" колебания показанного на Рисунке 2 типа. Все оценочные величины были получены в результате экспериментов, проведенных через два года после начала мониторинга. Данные были получены на острове Сил в течение сезонов 1986/87 и 1988/89 гг.

#### Список рисунков

Рисунок 1: Зависимость продолжительности отдельных периодов поиска пищи и общего количества периодов добывания пищи в море самками южного антарктического котика на острове Сил, Южные Шетландские острова, Антарктика. Данные за сезон 1986/87 г. по периодам добывания пищи в море с конца декабря до середины февраля; данные за 1988/89 г. по периодам добывания пищи в море с середины января до конца февраля.

Рисунок 2: Внутрисезонные изменения продолжительности периодов добывания пищи в море самками южного морского котика на острове Сил, Южные Шетландские острова, Антарктика. Данные были получены в течение сезона 1986/87 г.

#### Lista de las tablas

Tabla 1: Se presentan las estimaciones de los tamaños menores de las muestras (números de aves) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en la duraciones medias (días) de los primeros turnos de incubación de los pingüinos Adelia, para un rango de valores para la habilidad estadística deseada. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos son de Trivelpiece et al. (1989), Tabla 11.

Tabla 2: Se presentan las estimaciones de los tamaños menores de las muestras (números de aves) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en la duraciones medias (días) de los primeros turnos de incubación de los

pingüinos Barbijo para un rango de valores para la habilidad estadística deseada. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos son de Trivelpiece et al. (1989), Tabla 12.

- Tabla 3: Se presentan las estimaciones de los tamaños menores de las muestras (números de aves) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en la duración media (días) de los viajes de alimentación de los pingüinos Barbijo y Adelia. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos son de Trivelpiece et al. (1989), Tabla 9.
- Tabla 4: Se presentan las estimaciones de los tamaños menores de las muestras (números de aves) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en el número medio de polluelos Barbijo que llegan a la edad de crêche por cada pareja reproductiva para un rango de valores para la habilidad estadística deseada. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos provienen de la Isla Seal, 1987/88.
- Tabla 5: Estimaciones de los tamaños menores de las muestras (números de pares) necesarias para descubrir un cambio en el éxito reproductivo de los pingüinos Macaroni de un valor de 0.91 de polluelos incubados por cada pareja a un nuevo valor indicado (P2) aparecen en la tabla siguiente. La cantidad normal  $\delta$  debería ser la diferencia entre P2 y 0.91. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos provienen de la Isla Seal, 1987/88.
- Tabla 6: Se presentan las estimaciones de los tamaños menores de las muestras (números de aves) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en el peso húmedo promedio (gramos) del alimento de los polluelos de los pingüinos Adelia y Barbijo para un rango de valores para la habilidad estadística deseada. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos son de Trivelpiece et al. (1989), Tabla 6.
- Tabla 7: Se presentan las estimaciones de los tamaños menores de las muestras (números de hembras) necesarias para descubrir un cambio de magnitud igual a  $\delta$  en la duración media (horas) de los viajes de alimentación de los lobos finos para un rango de valores para la capacidad estadística deseada. Es posible que estas estimaciones sean demasiado altas porque no toman en consideración la variación "dentro del año" del tipo que se muestra en la Figura 2. Todas las estimaciones son para los estudios conducidos después de 2 años de monitoreo. Datos provienen de la Isla Seal, durante las temporadas 1986/87 y 1988/89.

#### Lista de las figuras

- Figura 1: Relación entre la duración individual y el número total de los viajes de alimentación hechos al mar por la hembra de los lobos finos antárticos en la Isla Seal, Islas Shetland del Sur, Antártica. Datos de los viajes observados desde finales de diciembre hasta mediados de febrero durante la temporada 1986/87; datos de los viajes observados desde mediados de enero hasta finales de febrero durante la temporada 1988/89.
- Figura 2: Cambios dentro de la temporada en la duración de los viajes de alimentación hechos al mar por la hembra de los lobos finos antárticos en la Isla Seal, Islas Shetland del Sur, Antártica. Datos son de la temporada 1986/87.