STATISTICAL POWER TO DETECT CHANGES IN GROWTH RATES OF ANTARCTIC FUR SEAL PUPS

P. Boveng, J. L. Bengtson and M. E. Goebel

Abstract

Numerical simulations were used to investigate the power of methods for detecting changes (between years) in pup growth rates, the loss of power associated with obtaining weight samples on only three dates in a season, and potential means of increasing the power of the CCAMLR Ecosystem Monitoring Program Standard Method C1 (B). Using estimates of variance in pup weights obtained at Seal Island (60°59.5'S, 55°24.5'W), the simulations suggest that following the protocol of the Standard Method (50 pups of each sex, weighed on four dates at 28-day intervals) would result in the ability to detect a 17% change in the growth rate with a 10% chance of committing Type 1 or Type 2 errors. If only three samples (50 pups, 28-day intervals) are obtained, the detectable change increases to about 34% under the same conditions. Slight gains in power can be obtained in some cases, without increasing the total number of pups weighed in a season, by decreasing the time between weighings.

Résumé

Des simulations numériques ont été utilisées pour examiner l'efficacité des méthodes de détection des changements (entre années) dans les taux de croissances des jeunes, la perte d'efficacité liée au fait que les échantillons de poids n'ont été obtenus que trois fois par saison, et les moyens potentiels d'augmenter l'efficacité des Méthodes standard du Programme de contrôle de l'écosystème de la CCAMLR C1 (B). En se basant sur les estimations de variance de poids des jeunes, obtenues à l'île Seal (60°59,5'S, 55°24,5'W), les simulations suggèrent qu'en suivant le protocole de la Méthode standard (50 jeunes de chaque sexe, pesés à quatre dates espacées de 28 jours), l'on pourrait déceler un changement de 17% dans le taux de croissance, avec 10% de chances de commettre des erreurs de type 1 ou 2. Si l'on n'obtient que trois échantillons (50 jeunes, à intervalle de 28 jours), les changements décelables augmentent d'environ 34%, dans les mêmes conditions. Dans certains cas, l'efficacité peut être légèrement augmentée sans augmentation du nombre total de jeunes pesés en une saison, en diminuant le temps passé entre les pesées.

Резюме

При исследовании эффективности методов в выявлении (межгодовых) изменений темпа роста щенков, рассмотрении снижении эффективности вследствие получения данных по весу только за три дня в течение сезона и потенциальных путей повышения эффективности Стандартного метода C1 (B), являющегося частью Программы АНТКОМа по мониторингу экосистемы, было использовано математическое моделирование. В результате применения при моделировании полученных на острове Сил (60°59,5'ю.ш., 55°24,5'з.д.) оценочных величин изменчивости веса щенков, было сделано заключение о том, что при следовании процедуры Стандартного метода (50 щенков каждого пола взвешиваются четыре раза с интервалом в 28 дней) можно выявить 17%-ное изменение темпа роста, при этом существует 10%-ная вероятность допущения погрешности типа 1 или 2. При получении лишь 3 проб (50 щенков в каждой с интервалом в 28 дней) при таких же условиях можно выявить более значительное, 34%-ное изменение. В некоторых случаях эффективность может быть повышена в некоторой степени при отсутствии увеличения количества щенков, взвешенных в течение одного сезона за счет сокращения интервала между взвешиваниями.

Resumen

Simulaciones numéricas se utilizaron para investigar la capacidad de los métodos para detectar los cambios (entre años) en el índice de crecimiento de los cachorros, la incapacidad asociada con la obtención de muestras de peso en sólo tres fechas en una temporada, y maneras potenciales de aumentar la capacidad del Método Estándar C1 (B) del Programa de Seguimiento del Ecosistema de la CCRVMA. Usando estimaciones de variación en el peso de los cachorros obtenidas en la Isla Seal (60°59,5'S, 55°24.5'W), las simulaciones sugieren que siguiendo el protocolo del Método Estándar (50 cachorros de cada sexo, pesados en cuatro fechas a intervalos de 28 días) resultaría en la habilidad de detectar un 17% en el índice de crecimiento, con una posibilidad de 10% de que se cometan errores del Tipo 1 o Tipo 2. Si sólo se obtienen tres muestras (50 cachorros, intervalos de 28 días), el cambio detectable aumenta aproximadamente 34% bajo las mismas condiciones. Leves aumentos en veracidad se pueden obtener en ciertos casos, sin aumentar el número total de cachorros pesados en una temporada, disminuyendo el tiempo entre cada pesaje.

1. INTRODUCTION

The CCAMLR Ecosystem Monitoring Program (CEMP) Standard Method C1 describes two techniques for estimating growth rates of Antarctic fur seal (*Arctocephalus gazella*) pups. Method (A) requires weighing each member of a sample of individually marked pups several times between birth and weaning. Method (B) requires weighing samples of 50 male and 50 female pups at monthly intervals beginning about 30 days after mean pupping date and concluding prior to weaning, when pups are about 110 days old. The analysis presented here pertains to the statistical power to detect changes in pup growth rate using Method (B), deferring a similar treatment of Method (A).

Boveng and Bengtson (1989) suggested that the effective sample size for estimating power to detect changes in growth rates depends not only on the total number of pups weighed, but on the number of weighing dates and the time intervals between them. Method B does not specify clearly the number of dates on which pups should be weighed. If the first sample is weighed 30 days after mean pupping date, three additional samples at monthly (say 28-day) intervals would conclude on the 114th day after mean pupping. Thus, it seems that no more than four samples of pups will be weighed following this method and that possibly three or fewer samples might be obtained.

A monitoring program in which pups are weighed on only three dates may not attain acceptable statistical power to detect changes in pup growth rate. Furthermore, the ability to detect non-linearity in the growth curve is minimized when only three samples are available. This paper presents an investigation into the power of the standard method to detect changes in pup growth rates, the loss of power associated with obtaining only three weight samples, and possible means of increasing the power of the standard method.

2. METHODS

Antarctic fur seal (*Arctocephalus gazella*) pups were weighed on four dates in 1988 and three dates in 1989, at Seal Island, Elephant Island (60°59.5'S, 55°24.5'W). Linear regressions (Sokal and Rohlf, 1981, pp. 480-482) of pup weight versus date, were computed separately for each sex and year. Bartlett's test was used to evaluate heterogeneity of variances. For each sex, pooled (i.e. weighted average) estimates of the slope and of the mean squared error (MSE) from both years were formed for use in a power analysis.

Monte Carlo simulations were used to investigate the effects of number and spacing of sampling dates, and number of pups weighed per date, on the power to detect changes of 5 to 50% of the observed mean pup growth rates. The simulations proceeded as follows:

(i) A simulated data set, with a specified number and spacing of sampling dates and number of pups weighed per date, was drawn at random from normal distributions with means determined by our observed growth rate and variances equal to our observed MSE (the "null" data);

(ii) A second data set was drawn under similar conditions to the first, except that the underlying growth rate differed by a specified amount from the first (the "alternative" data);

(iii) Linear regressions were computed for each of the data sets ("null" and "alternative"), and the regression coefficients compared (Sokal and Rohlf 1981, pp. 499-506);

(iv) Steps (i) to (iii) were repeated 500 times with the same input parameters, recording the number of times that the growth rates in the two data sets were significantly different. The number of significantly different results divided by 500 is an estimate of the power to detect a change in the growth rate from the null value to the alternative value.

(v) Steps (i) to (iv) were repeated for several alternative growth rates, ranging from 1.0 to 1.5 times the null rate.

Simulations were conducted to estimate the effects of using three versus four weighing dates at 28-day intervals, of sampling smaller numbers of pups at more frequent intervals, and of the choice of α for comparisons of regression coefficients.

3. RESULTS AND DISCUSSION

The weighing dates, numbers of pups weighed, mean weights, and standard deviations of weights from Seal Island are shown for each sex and year in Table 1. There were no significant differences in variances of pup weights among dates, for either sex in either year ($P > 0.13$ in all cases). Table 2 shows estimates of growth rates from linear regressions of Seal Island pup weights versus date. Male pups grew faster and had more variable weights than females (Table 2). Tests for departures from linearity were all non-significant ($P > 0.25$ in all cases). Because the weighing periods at Seal Island were shorter than specified in the standard method, this result is not a confirmation of the pup growth curve being generally linear.

Figure 1 shows the statistical power achieved in a simulated monitoring program following Standard Method C1 (B), and assuming that the typical growth rate and within-date variance in pup weights are equal to the weighted averages of those observed in females at Seal Island (Table 2). If four samples of 50 pups each are weighed at 28-day intervals, there would be a 90% chance of detecting a 17% change in the true growth rate (using α equal to 0.10 for comparing rates). Figure 1 also shows the reduction in statistical power that would occur if samples of 50 female pups were weighed each year on three rather than four dates, again at 28-day intervals. The lower power obtained in a program sampling on only three dates is a result of the smaller total number of pups weighed (150 versus 200), as well as the shorter duration of the weighing period.

Figure 2 shows that if the duration of the weighing period is limited (e.g. by logistic or other constraints), weighing smaller samples of pups more often can result in higher statistical power, even though the total number of pups weighed is not increased. No such increase in power, however, was observed in simulations of a monitoring program in which pups are weighed at intervals shorter than 28 days for the entire available weighing period. For example, if samples of 28 pups are weighed on seven dates at 14-day intervals, so that total sample size and duration of the weighing period are nearly identical to those in the standard method, there is no increase in the power to detect changes. Therefore, the power of a monitoring program that utilizes the entire available weighing period can be increased only by increasing the total number of pups weighed or by accepting a lower value of α for comparing growth rates.

Figure 3 shows the effect of the choice of α, the acceptable risk of a Type 1 error, on the power to detect a change in growth rate. As expected, the simulations show that a monitoring program using a relatively high α-level for comparing regression coefficients (growth rates) will be more sensitive (i.e. will have higher power to detect a significant departure from the typical growth rate) than a program using a lower α-level. Boveng and Bengtson (1989) suggested that α and β (the risk of a Type 2 error) probabilities should be equal in some resource conservation contexts. The magnitude of change detectable with α and β probabilites equal, are shown for all simulations in Figures 1 to 3.

If the relationships between the growth rates for males and females and between the variance in weights of males and females observed at Seal Island are typical, the power to detect a proportional change in growth rate will be similar for both sexes. This is because the ratio of the standard deviation of weights to the growth rate is similar for both sexes. Though simulations were conducted using data from both sexes, only the results from female pup weights have been presented here.

The simulations were used to verify that the power curves are symmetric for positive and negative deviations from the typical growth rate. Therefore, the statistical power shown in Figures 1 to 3 pertains to increases or decreases in the growth rate.

LITERATURE CITED

BOVENG, P. and J.L. BENGTSON. 1989. On the power to detect changes using the standard methods for monitoring parameters of predatory species. CCAMLR, WG-CEMP-89/6.

SOKAL, R. R. AND F. J. ROHLF. 1981. *Biometry*. San Francisco: Freeman. pp. 859.

Table 1: Dates, sample sizes, means, and standard deviations of male and female fur seal pup weights measured at Seal Island in 1988 and 1989.

	Males			Females		
Date	n	mean weight (kg)	s.d.	n	mean weight (kg)	s.d.
1988:						
2 Jan	30	9.91	1.60	20	8.78	1.13
15 Jan	40	11.28	1.64	26	9.50	1.49
27 Jan	41	12.91	1.64	20	10.76	1.15
12 Feb	36	15.91	2.22	35	12.38	1.66
1989:						
17 Jan	21	10.65	1.94	29	9.19	1.81
31 Jan	26	12.81	1.88	24	10.62	1.70
13 Feb	30	14.42	1.77	35	12.61	1.42

Table 2: Results of linear regressions of pup weight versus data. Growth rates, standrd errors and estimated "within-date" variance in pup weights (MSE) are shown separately for each year and as weighted averages of both years.

		Growth Rate (kg/d)	Standard Error of Growth Rate	MSE
Males,	1988:	0.147	0.0147	3.21
Males,	1989:	0.139	0.0089	3.45
Males,	wt. ave.:	0.145	- -	3.29
Females,	1988:	0.092	0.0077	2.04
Females,	1989:	0.127	0.0130	2.67
Females,	wt. ave.:	0.108	- -	2.34

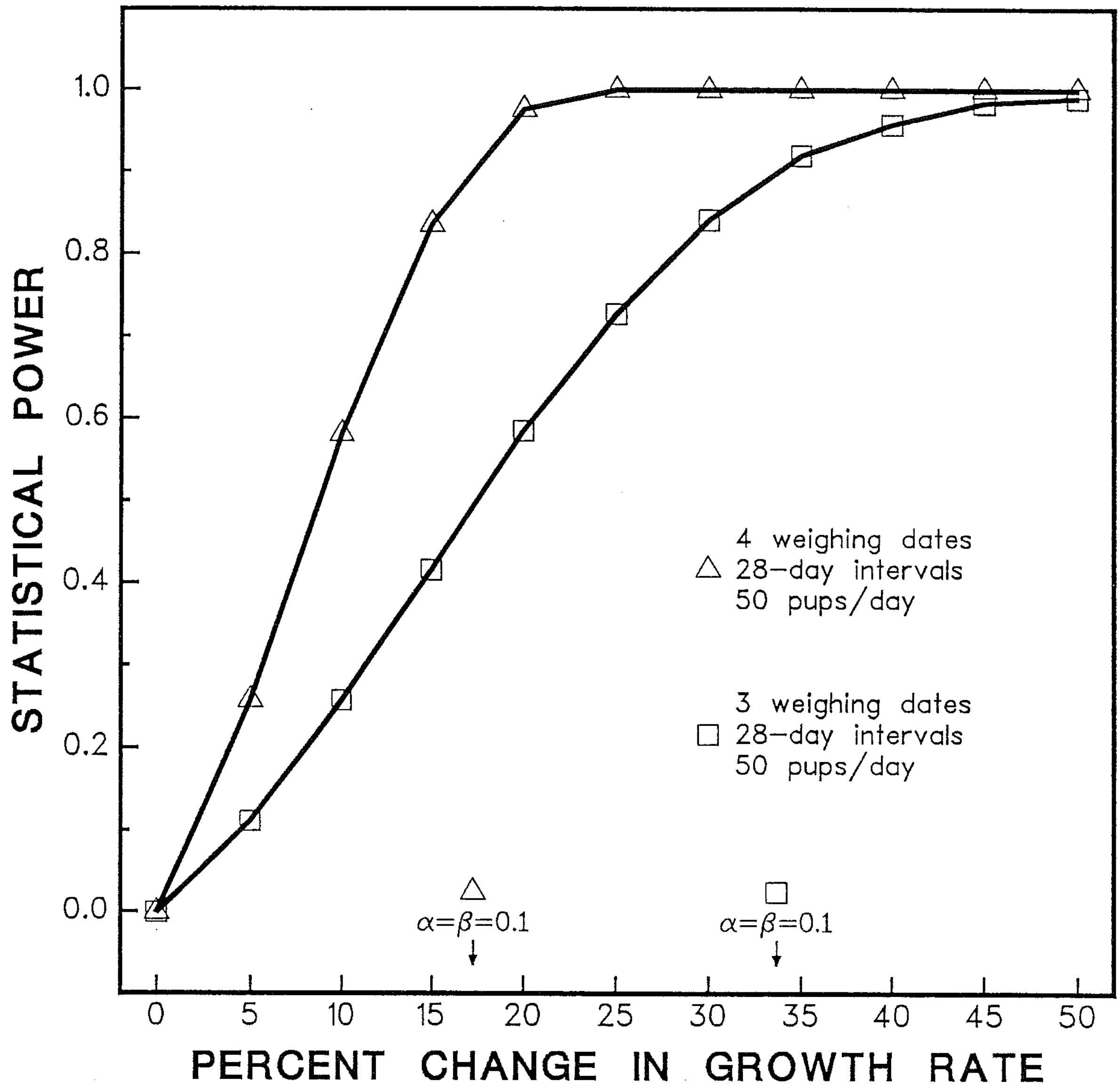

Figure 1: Difference in power to detect changes in fur seal pup growth rate using Standard Method C1 (B) when three or four samples of pups are weighed. The typical growth rate, or "null" rate, used in the simulation was 0.108 kg/d, the average rate observed for females at Seal Island in 1988 and 1989. The mean-squared-error for pup weights was 2.336, the average for females at Seal Island.

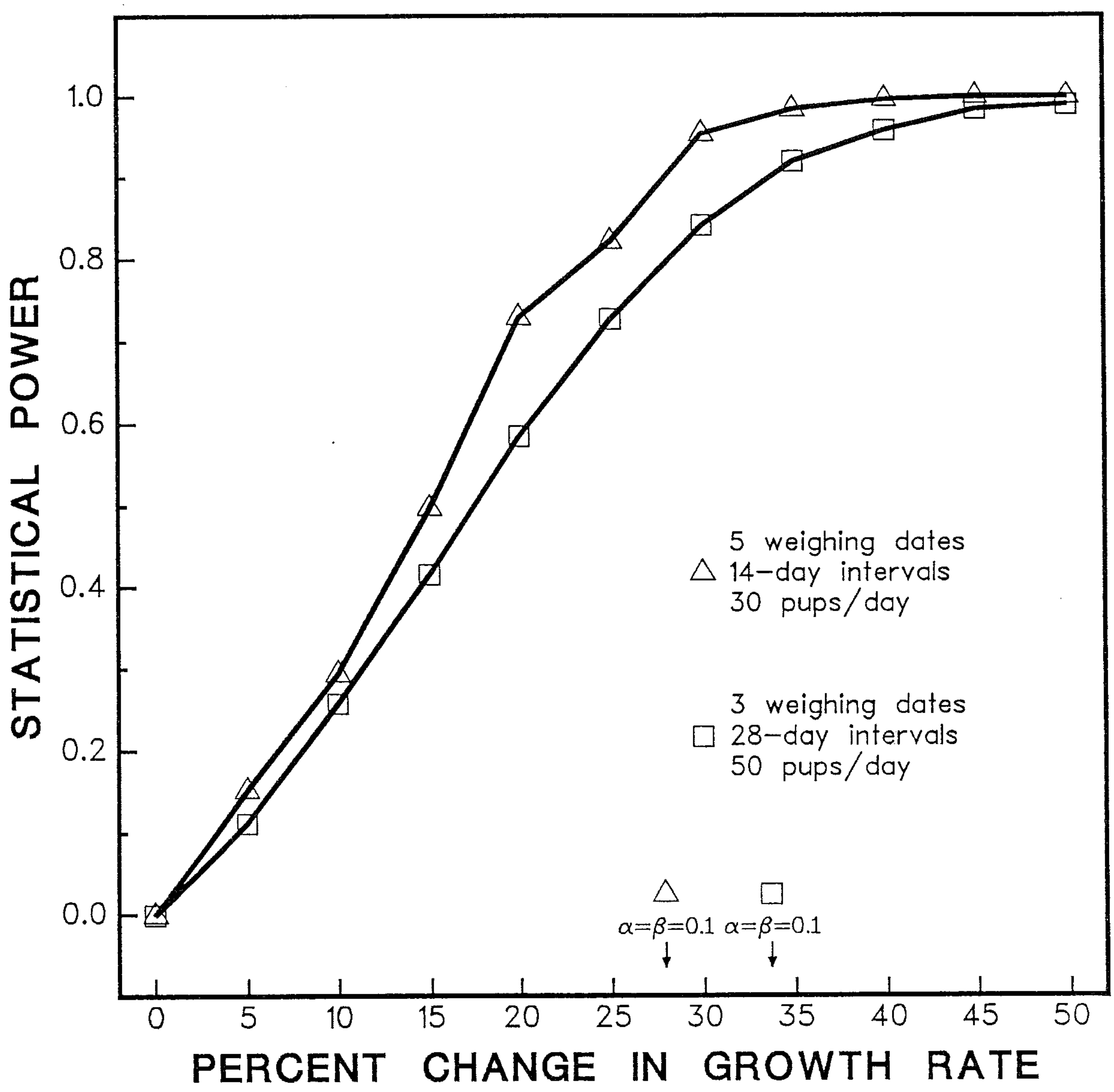

Figure 2: If the weighing period is limited to approximately 60 days, a slight increase in power can be obtained by weighing smaller samples of pups at more frequent intervals than specified in the Standard Method. As in Figure 1, null growth rate equals 0.108 kg/d, mean-squared-error equals 2.336.

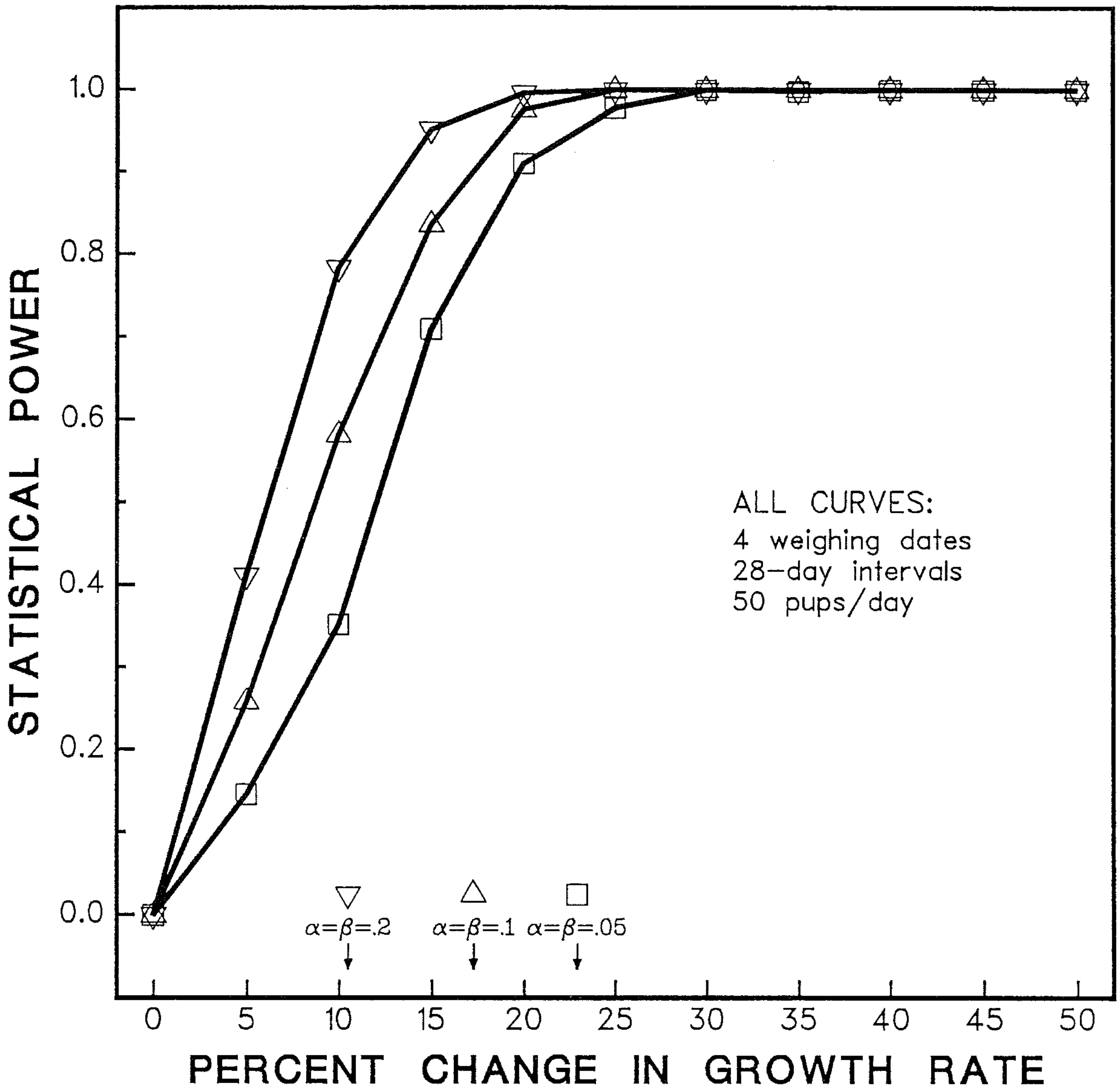

Figure 3: The effect of the choice of α for comparing regression coefficients (growth rates), on the power to detect changes. As in Figure 1, null growth rate equals 0.108 kg/d, mean-squared-error equals 2.336.

Liste des tableaux

Liste des figures

Список таблиц

Список рисунков

0,108 кг/день и равнялся среднему темпу роста самок на острове Сил в 1988 и 1989 гг. Среднее квадратическое отклонение веса щенков равнялось 2,336 - среднее значение для самок на острове Сил.

Рисунок 2: Если период взвешивания не превышает приблизительно 60 дней, некоторое повышение эффективности может быть достигнуто за счет взвешивания менее многочисленных выборок через более короткие интервалы, чем рекомендуемые стандартным методом. Также, как и на Рисунке 1, нулевой темп роста равняется 0,108 кг/день, среднее квадратическое отклонение равняется 2,336.

Рисунок 3: Влияние выбора величин α для сравнения коэффициентов регрессии (темпов роста) на эффективность выявления изменений. Также, как и на Рисунке 1, нулевой темп роста равняется 0,108 кг/день, среднее квадратическое отклонение равняется 2,336.

Lista de las tablas

Tabla 1: Fechas, tamaños de las muestras, medias y desviaciones cuadráticas medias de los pesos de los cachorros macho y hembra de lobo fino antártico obtenidos en la Isla Seal en 1988 y 1989.

Tabla 2: Resultados de las regresiones lineales del peso del cachorro versus datos. Indices de crecimiento, errores estándar y variación estimada "dentro de la fecha" en los pesos de los cachorros (MSE) se muestran separadamente para cada año y como pesos promedios para ambos años.

Lista de las figuras

Figura 1: Diferencia en habilidad para detectar cambios en el índice de crecimiento de los cachorros de lobo fino antártico usando el Método Estándar C1 (B) cuando se pesan 3 ó 4 muestras de cachorros. El índice típico de crecimiento, o índice de nulidad, usado en la simulación fue 0.108 kg/d, el índice promedio observado para las hembras en la Isla Seal en 1988 y 1989. El error cuadrado medio para los pesos de los cachorros fue 2.336, el promedio para las hembras en la Isla Seal.

Figura 2: Sí el período para pesar se limita a aproximadamente 60 días, un leve aumento en veracidad se puede obtener pesando muestras mas pequeñas a intervalos más frecuentes que lo especificado en el Método Estándar. Según la Figura 1, el índice de nulidad de crecimiento es igual a 0.108 kg/d, el error cuadrado medio es igual a 2.336.

Figura 3: El efecto de la elección de α para comparar los coeficientes de regresión (índices de crecimiento), en la habilidad para detectar los cambios. Al igual que Figura 1, el índice de nulidad de crecimiento es igual a 0.108 kg/d, el error cuadrado medio es igual a 2.336.