Sample Size Requirements for Estimating the Mean Length of Penaeid Shrimp Within a Trawl Sample
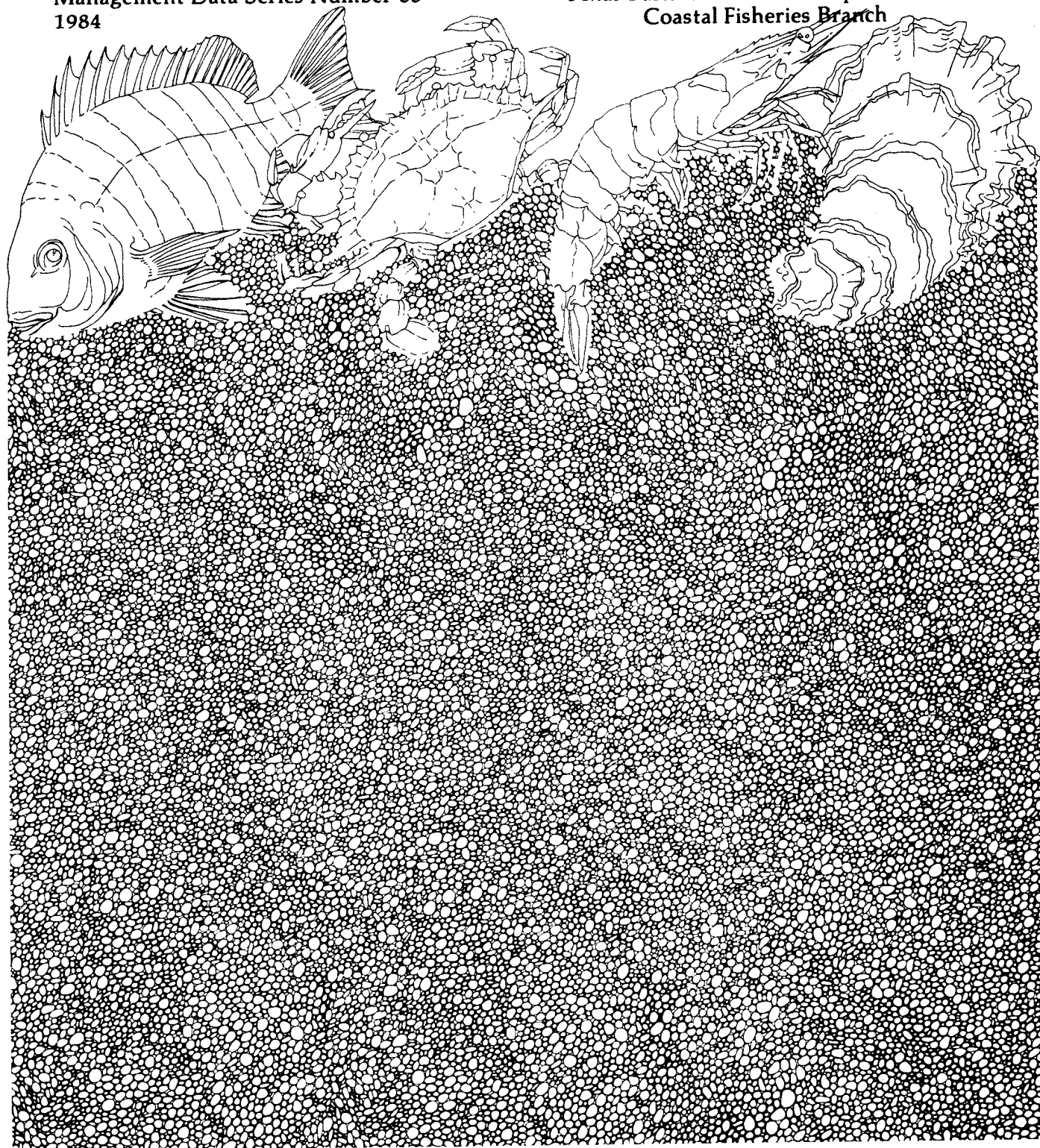
by T.J. Cody

SAMPLE SIZE REQUIREMENTS FOR ESTIMATING
THE MEAN LENGTH OF PENAEID SHRIMP WITHIN A TRAWL SAMPLE

By

Terry J. Cody

MANAGEMENT DATA SERIES
NO. 65

Texas Parks and Wildlife Department
Coastal Fisheries Branch
4200 Smith School Road
Austin, Texas 78744

ACKNOWLEDGEMENTS

ABSTRACT

The Texas Parks and Wildlife Department collected 34 samples along the central Texas coast during June and July 1982 as part of the Southeast Area Monitoring and Assessment Program (SEAMAP). An analysis was conducted on 24 of these samples with catches >25/tow to determine the optimum number of penaeid shrimp to measure to determine differences in mean length with a known precision. Sample mean lengths ranged from 96 to 134 mm total length and coefficients of variation ranged from 8.3 to 22.3%. There were significant differences among the sample mean lengths ($F_s = 167.38$; $P < 0.001$). There were no significant correlations between sample mean lengths and variances, or numbers caught and variances ($P \geq 0.05$).

Sample sizes required to detect a 10% difference among means ranged from 24 to 52 with a CV of 12% and 75 to 140 measurements for a CV of 23.3%. These analyses should help define the variability associated with the 1982 SEAMAP shrimp data and suggests that measuring 50 shrimp should be adequate to determine the mean length of Penaeus aztecus in a sample off Texas during June-July.

INTRODUCTION

Size measurements are essential for estimating growth, one of the basic elements for understanding the population dynamics of a species. Most literature on sampling design focuses on the reduction of bias and variance in the sampling program (Gulland 1966; Cochran 1977; Green 1979). Very little literature beyond statistical textbooks focuses on the optimization of sampling. Kimura (1984) noted that appropriate sampling levels are often left to logistical and budgetary contraints which can lead to either inadequate or excessive sampling. Some recent literature has addressed the problem of combining sampling efficiency and economic efficiency into a decision-making model that can be used to minimize, or at least define, the precision of the estimates at a given sample size (Thornton et al. 1982; Kimura 1984).

For the past two years the Texas Parks and Wildlife Department has participated in the Southeast Area Monitoring and Assessment Program (SEAMAP) which provides comprehensive biological and environmental data from the Gulf of Mexico. One important part of the program is the collection of trawl data on the abundance and size of penaeid shrimp during the Texas Closure--a closed season for gulf shrimping of at least 45 days during June and July established by the Gulf of Mexico Fishery Management Council to complement the traditional closure of Texas waters. The purpose of the closure is to increase the yield of shrimp and to eliminate waste by discard of undersized brown shrimp (Penaeus aztecus) in the Fishery Conservation Zone (FCZ) off Texas (Center for Wetland Resources 1980).

A major goal of the SEAMAP program was to standardize the sampling gear and techniques used by agencies conducting research in the northern Gulf of Mexico. Workshops were held and operations manuals prepared to help standardize data collected in the program (Watson 1984).

During 1982 the Texas Parks and Wildlife Department (TPWD) and the National Marine Fisheries Service (NMFS) collected 34 samples along the central Texas coast during the SEAMAP program. In order to maximize the sampling efficiency and possibly to reduce the workload by not having to measure excessive numbers of shrimp, an analysis was conducted to determine the optimum number of penaeid shrimp to measure from the trawl samples to determine differences in mean length with a known precision.

MATERIALS AND METHODS

Data used in these analyses were collected during SEAMAP sampling off the coast of Texas between 23 June and 12 July, 1982 (Matthews 1982). Stations were chosen at random by NMFS at depths ranging from 9-48 m. Each sample consisted of the catch from a 40-foot semiballoon shrimp trawl towed 15-30 minutes.

At each station, all Penaeus shrimps were culled from the total catch and weighed. Penaeids were sorted by species and total lengths from the

tip of the rostrum to the tip of the telson were measured in millimeters. Only P. aztecus were used in the report to simplify the analyses. Total numbers caught ranged from 2 to 2307 and the number of P. aztecus measured in each sample ranged from 2 to 396. Samples with less than 25 shrimp were eliminated from the analyses to minimize the effect of extremely small sample sizes on the overall variation.

Sample mean lengths and variances were calculated for untransformed length data according to standard procedures (Sokal and Rohlf 1981; Snedecor and Cochran 1980). A single classification analysis of variance (ANOVA) was done using each sample as a group. The correlations between mean length and variance and sample size and variance were tested to see if variability increased with mean length or sample sizes. Sensitivity curves which estimated the number of measurements required to detect a given difference between two means within specified confidence limits were calculated using the formula (from Sokal and Rohlf 1981):

where   n =   number of replications

σ =   true standard deviation

δ =   the smallest true difference that it is desired to detect. (Note: it is necessary to know only the ratio of σ to δ , not their actual values)

v =   degrees of freedom of the sample standard deviation

($\sqrt{MS_{within}}$) with a groups and n replications per group

α =   significance level (such as 0.05)

P =   desired probability that a difference will be found to be significant (if it is as small as δ). This is the intended power of the test.

$t_{\alpha(v)}$ and $t_{2(1-P)}$ (v)   = values from a two tailed t-table with v degrees of freedom and corresponding to probabilities of α and 2(1 - P), respectively. Note: if P = ½ then t_1 = 0.

The coefficient of variation (CV) used in the formula was calculated by using

$$CV = \frac{\sqrt{variance}}{mean} \times 100\%.$$

RESULTS

Of 34 stations sampled during the 1982 SEAMAP program, 24 samples had $\geq$ 25 $\underline{P}$. $\underline{aztecus}$. Mean lengths, variances and coefficients of variation for these samples were calculated (Table 1). Mean lengths ranged from 96 to 134 mm and coefficients of variation ranged from 8.3 to 22.3%.

There were significant differences among the sample (group) mean lengths (Table 2; F_s = 167.38; P < 0.001). The correlation between mean lengths and variances was not significant (r = 0.09; P $\geq$ 0.05). The correlation between the numbers caught and variances was not significant (r = -0.48; P $\geq$ 0.05). Since there were no significant correlations, the within mean square was selected as the best estimate of the variance for all groups (samples). This variance and the unweighted mean of all samples (112 mm) results in a CV of 11.55%.

Sample size determinations were made using a CV of 12% (11.55% rounded upward) and 22.3% (a worst case - highest CV encountered in a sample). The sample size determinations were also made using the 0.01 and 0.05 levels of significance and the 0.95 and 0.80 probabilities of being detected (Table 3).

Sample sizes required to detect a 10% difference in means ranged from 24 to 52 with a CV of 12% (Figure 2). A CV of 22.3% required sample sizes of 75 to 140 measurements to achieve the same level of precision (Figure 3).

DISCUSSION

With increasing demands on limited living marine resources, resource agencies are often called upon to accomplish more without additional equipment, manpower or money. Under these conditions, it is imperative that monitoring programs be designed to obtain the greatest amount of reliable data for the least effort and cost (Hegen et al. 1983). In order to do this properly, programs need to be continually evaluated and refined as the accumulation of data permits.

This analysis of the 1982 SEAMAP shrimp data indicates rather low variation in the lengths of $\underline{P}$. $\underline{aztecus}$ in individual samples. Much more of the overall variation in the shrimp data came from the differences among samples, which is the situation often found in trawl data and fishery landings data (Gulland 1966).

In a program like SEAMAP, several different agencies are working toward a common goal and, at the same time, trying to maintain their respective programs. The Texas Parks and Wildlife Department has traditionally used sample sizes of 50 to estimate the mean length of each penaeid shrimp species in trawl samples. SEAMAP sampling required measurements to be taken on at least 200 shrimp of each species from each trawl sample. The sensitivity curves developed in this report point out the extra effort required to increase the precision of the data. Quadrupling

the effort reduces the detectable percent difference between means by
only one half. At a conservative level of significance ($P \leq 0.01$), a
sample size of n = 50 could detect a 10% difference in means 95% of the
time, while n = 200 should detect a 5% difference in means.

Good experimental design should minimize uncertainty of estimates
made with data, while keeping costs to a minimum. Even in the best
designed programs the practical situation often limits the size of
the sample (Gulland 1966; Hegen et al. 1983). While a loss of
precision may not be desirable, analyses that help define the uncertainty
associated with the data can be useful in comparing and discussing
the data (Thornton et al. 1982).

The analyses in this report should help define the variability
associated with the 1982 SEAMAP shrimp data and can help in the design
of future sampling programs. The data indicated that measuring 50
shrimp of each species from a sample should be adequate to determine
the mean length of penaeid shrimp in a sample off Texas during the
SEAMAP time period (June-July).

LITERATURE CITED

Center for Wetland Resources. 1980. Management plan and final environmental impact statement for the shrimp fishery of the Gulf of Mexico, United States waters. Louisiana State Univ., Baton Rouge. 185 p.

Cochran, W. G. 1977. Sampling techniques, 3rd. ed., John Wiley & Sons, Inc., New York. 428 p.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, Inc. New York. 257 p.

Gulland, J. A. 1966. Manual of sampling and statistical methods for fisheries biology. Part I. Sampling methods. FAO, Rome. 96 p.

Hegen, H. E., G. C. Matlock and A. W. Green. 1983. Evaluation of gill and trammel net sampling strategies for monitoring finfish availability in Texas bays. Tx. Pks. Wildl. Dept., Coastal Fish. Br., Tech. Ser. No. 33. 24 p.

Kimura, D. K. 1984. Determining appropriate sampling levels for estimating age distribution from trawl landings. Wash. Dept. of Fish., Tech. Rept. No. 80. 18 p.

Matthews, G. A. 1982. Abundance and size distribution of _Penaeus_ spp. shrimps in the northern and northwestern Gulf of Mexico during the 1982 closure period. NOAA Technical Memorandum, NMFS-SEFC-109. 58 p.

Snedecor, G. W. and W. G. Cochran. 1980. Statistical Methods. 7th Ed. Iowa State Univ. Press., Ames. 507 p.

Sokal, R. A. and F. J. Rohlf. 1981. Biometry: The principles and practices of statistics in biological research. 2nd Ed., W. H. Freeman and Co., San Francisco. 859 p.

Thornton, K. W., R. H. Kennedy, A. D. Magoun and G. E. Saul. 1982. Reservoir water quality sampling design. Wat. Resource Bull. 18:471-480.

Watson, J. W., Jr. 1984. Proceedings of SEAMAP gear workshop. Gulf States Marine Fisheries Commission (In Press).

Table 1. Sample statistics for P. aztecus collected by Texas Parks and Wildlife during 1982 Southeast Area Monitoring and Assessment Program (SEAMAP) sampling off the central Texas coast.

Station	Mean length (mm)	Variance	Standard error	Number	Coefficient of variation (%)
T 84	134	369	2.10	84	14.4
T 83	99	249	1.31	144	15.9
T 81	96	216	1.07	190	15.3
T 79	108	581	2.88	70	22.3
T 78	123	460	1.81	141	17.4
T 77	100	258	1.17	190	16.0
T 74	112	246	1.11	201	14.1
T 72	123	118	1.32	68	8.8
T 71	124	142	0.84	199	9.6
T 69	118	121	1.02	116	9.3
T 68	118	144	0.85	200	10.2
T 67	125	130	0.83	189	9.1
T 66	122	101	0.72	197	8.3
T 65	122	141	0.84	202	9.7
T 63	110	138	0.63	354	10.7
T 61	101	158	0.63	396	12.5
T 60	108	112	0.53	394	9.7
T 55	109	97	0.51	376	9.0
T 30	108	143	1.04	131	11.0
T 29	108	170	1.41	86	12.0
T 27	110	209	1.04	193	13.2
T 19	96	114	0.55	379	11.1
T 09	110	158	0.69	328	11.4
T 08	101	120	0.60	333	10.8

Table 2. Single classification analysis of variance testing the equality of mean total lengths among trawl samples of Penaeus aztecus collected off the central Texas coast by Texas Parks and Wildlife Department during 1982 Southeast Area Monitoring and Assessment Program (SEAMAP) sampling.

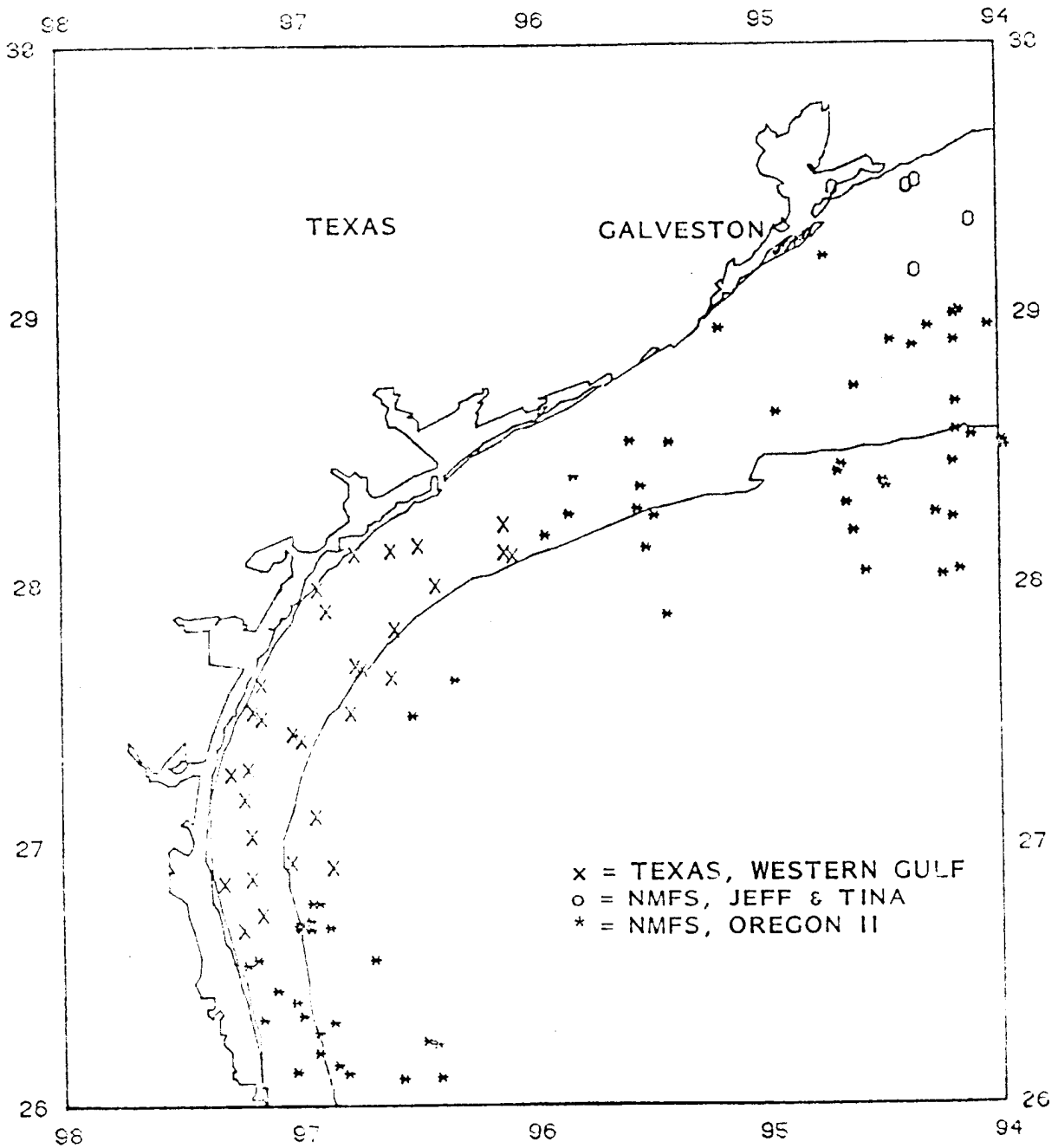| Source of Variation	df	SS	MS	F_s
Among samples	23	445906.29	19387.23	115.83 ***
Within samples	5137	859834.71	167.38	
Total	5160	1305741.00		

*** P 0.001

Figure 1. Trawl stations off the Texas coast sampled by Southeast Area Monitoring and Assessment Program (SEAMAP) vessels during June-July 1982. (X = samples used in this report).

Figure 2. Number of <u>Penaeus aztecus</u> measurements required to detect
X % difference between trawl means at each precision level using a
CV = 12%.

A   P = 80 %   α = 0.05
B   P = 80 %   α = 0.01
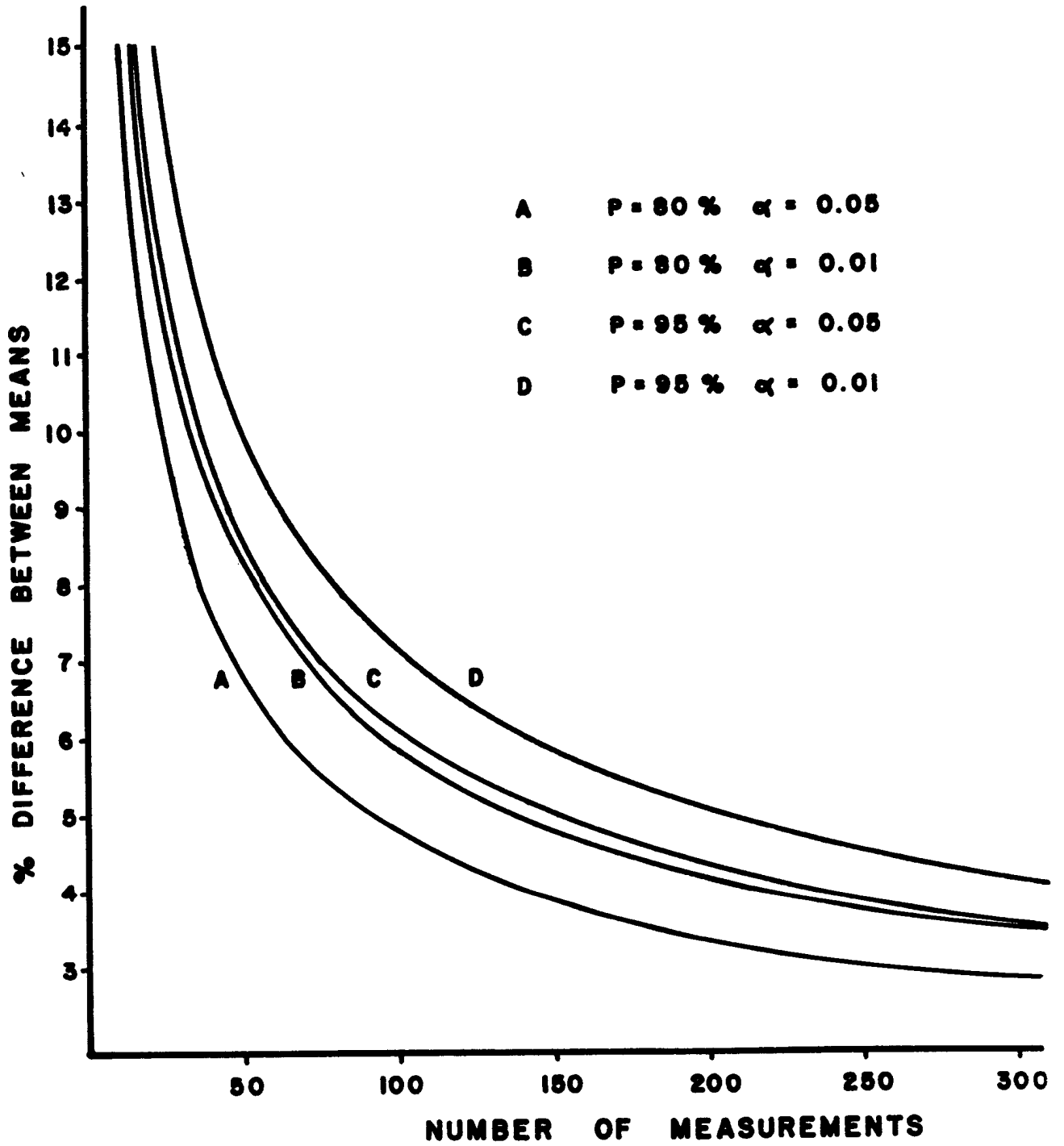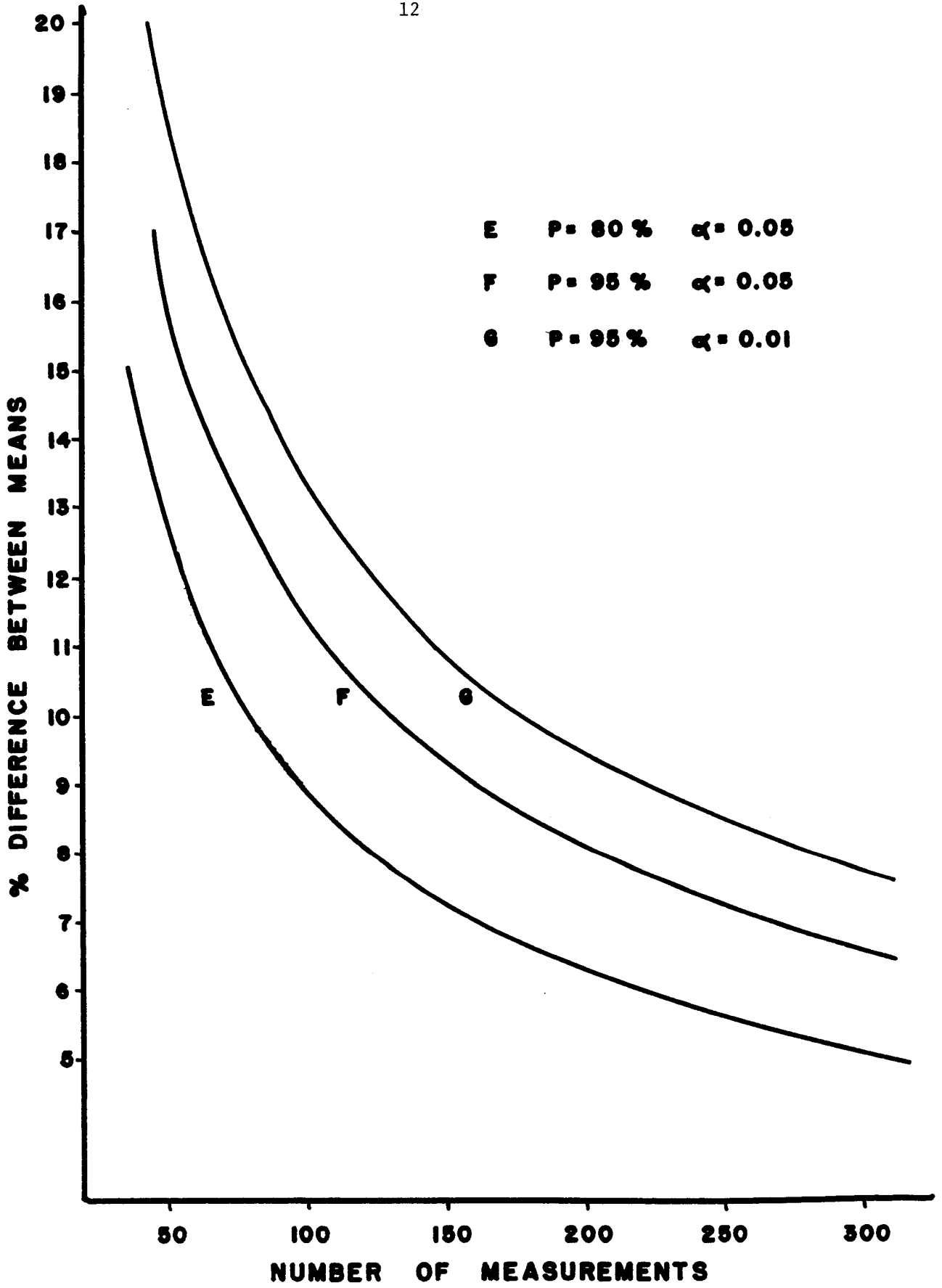C   P = 95 %   α = 0.05
D   P = 95 %   α = 0.01

Figure 3. Number of Penaeus aztecus measurements required to detect
X % difference between trawl means at each precision level using a
CV = 22.3%.

E   P = 80 % α = 0.05

F   P = 95 % α = 0.05

G   P = 95 % α = 0.01

% DIFFERENCE BETWEEN MEANS

NUMBER OF MEASUREMENTS

PWD Report 3000-182
August 1984