ICES 1989 PAPER C.M. 1989/D:22

Ref. K

Optimum Size of Sampling Unit for Estimating the Density of Marine Populations.

Michael Pennington

National Marine Fisheries Service, Woods Hole Laboratory

Woods Hole, Massachusetts 02543, U.S.A.

and

Jon H. Vølstad

Institute of Marine Research

5024 Bergen, Norway

Abstract

Data from several trawling experiments and from some scallop dredge surveys indicate that, within limits, a smaller sampling unit can be more effective than a larger unit for marine abundance surveys. Taking into account survey costs and sampling variability, the unit size is found which produces the most precise density estimate given a fixed amount of survey resources or, if a certain level of precision is required, the size of sampling unit which minimizes the total cost of the survey. As an illustration, the optimum sampling unit sizes are derived for surveys of some fish populations and for a scallop stock on Georges Bank.

Key words: Optimum sampling unit size; Variance function; Minimizing survey costs; Marine density estimates.

1. Introduction

Marine surveys are conducted throughout the world to assess the abundance or to track the relative abundance of many commercially important fish and shellfish stocks. These surveys are expensive and resources are often limited. Thus it is essential that all aspects of a survey's design are cost efficient. The design element considered here is the appropriate size of the sampling unit.

For a particular survey, a trawl or dredge, and an associated towing speed, are selected which are suitable for sampling the target species in the region of interest; hence tow duration determines the sampling unit size. Currently, a towing time of from thirty minutes to as long as two hours is standard for many groundfish surveys, while for shellfish surveys a dredge is usually towed for ten to fifteen minutes. Within limits set by gear saturation effects, the mean catch per tow is a linear function of tow duration, but the relationship between the sampling variance and unit size depends, for the most part, on the species' spatial distribution. In general, the distribution of marine animals is highly patchy and the coefficient of variation (cv) usually is fairly large (see, e.g. Seber, 1986).

One way to compare the effectiveness of different sampling unit sizes for a particular survey is to calculate the cv's of the resulting density estimates. Then the optimum unit size may be thought of as the one which produces the lowest cv for a given cost. That is, for each unit size, t, a number of samples, n_t, can be collected with a fixed amount of resources. If μ_t is the mean number of animals per unit and σ_t^2 is the sampling variance, then the optimum unit size may be defined as the one for which the value of $(\sigma_t/\mu_t)/\sqrt{n_t}$ is minimized.

Considerable information exists on the empirical relationship between the sampling variance and the mean for different population levels sampled with a fixed unit size (Taylor, Woiwood, and Perry, 1979). However, little seems to be available on the form of the variance function if the same population is sampled but with different unit sizes.

(Table 1) Estimates of the population cv (σ_t/μ_t) for various sampling unit sizes from several scallop surveys are given in Table 1. Dredge surveys for sea scallops (Placopecten magellanicus) on northeastern Georges Bank were conducted independently by the United States and Canada in 1982, 1983 and 1984 (Serchuck and Wigley, 1986). The Canadian sampling unit was one-fourth smaller than that of the United States, but the estimated cv's are nearly identical. Results from experimental sampling of the South Channel scallop population are similar (Table 1).

The same pattern appears to hold for some fish populations. In January, 1965 experimental tows of four durations were made on Georges Bank with a standard otter trawl by the Northeast Fisheries Center, Woods Hole, and again, (Table 2) the cv's are approximately constant (Table 2). In Table 3 are the results (Table 3) from two trawling experiments in the Barents Sea. The experiments were conducted in October, 1988 and January, 1989 by the Institute of Marine Research, Bergen using a Norwegian sampling trawl. As before, the cv's within an experiment are fairly constant. Though the estimated cv's are rather imprecise for each experiment, taken together there is little evidence that any precision was gained by increasing tow duration.

In order to determine the most efficient sampling unit size for marine surveys, it is necessary to develop an empirical formula relating the sampling variance and mean. The sampling distribution converges to the Poisson for sufficiently small unit sizes (Bliss, 1971). If for larger units the cv is roughly constant, then the relationship between the sampling variance and the mean for varying unit sizes would be adequately described by

$$\sigma_t^2 = \mu_t + b\mu_t^2 \ , \qquad (1)$$

where b is a constant, and hence,

$$cv_t = (1/\mu_t + b)^{1/2} \ . \qquad (2)$$

Given the large sampling variance of marine surveys, the predicted increase in cv_t as sampling unit size decreases [equation (2)] would be difficult to detect except at low population densities or for relatively small unit sizes. (Fig 1) Figure (1) is a plot of the estimated cv_t's for ocean pout (Macrozoarces americanus), which had a relatively low density (data are from the Georges Bank trawl experiment), and shows an apparent increase in cv_t as tow length decreases.

The negative binomial distribution frequently provides a good description of the sampling distribution of many marine populations (Taylor, 1953; Bliss and Owens, 1958; Roessler, 1965; Lenarz and Adams, 1980). If a negative binomial distribution is generated by a heterogeneous Poisson process, equation (1) will hold (with constant b).

In some situations another variance function may be more appropriate. An empirical formula, which often describes the relationship between the variance and the mean as population varies and unit size is fixed, is Taylor's power law (Taylor, 1961, Seber, 1986),

$$\sigma^2 = a\mu^b \ . \tag{3}$$

Seber (1986) observes that b is usually between 1 and 2 and is fairly constant over time and space for many species. Now if for a particular fixed population, formula (3) also provides a model of the relationship between the mean and the variance as unit size varies, then as unit size increases the cv will decrease to zero if $b < 2$, if $b > 2$ it goes to infinity, and is constant if $b = 2$.

2. Determining the optimum length of tow

For fish or shellfish surveys, it is convenient to measure cost in terms of ship time. Since the cost per day of running a ship does not, for the most part, change with cruise duration, the total cost of a survey is directly proportional to the number of days at sea. The catch normally is processed while steaming between stations, and therefore, a survey's shoreside expenses are relatively minor. Other overhead expenses, such as for equipment and travel to and from the survey region, are fixed and will not be considered here as part of the total cost. Having selected the appropriate sampling gear, the statistical problem is to determine an appropriate length of tow (i.e., what is the size of the sampling unit) at each station (sample site).

At each station it takes on average a certain time, c_1, to set and haul in the sampler. Thus for a tow of length t, the total time at a station is c_1+t. If n stations are selected randomly in a region, and a cruise track of approximately minimum length is followed, then the total travel time between stations will be approximately proportional to $\sqrt{n}$ (see, e.g., Cochran, 1977, p.96 and p.244). Thus for a random survey, the total time, C, to conduct the survey is given by the cost function

$$C = (c_1+t)n + c_2 \sqrt{n} \ , \qquad (4)$$

where the value of the constant, c_2, depends on the area of the survey region and the cruising speed of the ship. If the survey design is a grid of equally spaced stations, then total travel time will also be approximately proportional to $\sqrt{n}$ (Hansen, Hurwitz, and Madow, 1953, p.273), and formula (4) will hold. For surveys for which total travel time is fixed, such as sampling along a transect, $c_2 = 0$, and C becomes the time available for sampling, while n is the number of stations which can be taken along the transect.

Let the total amount of resources, C, be fixed. Assuming that the mean and sampling variance are related as in equation (1), denoting $cv/\sqrt{n}$ by k and setting $\mu = mt$, then

$$k = [\ 1/mt + b\ ]^{1/2}/\sqrt{n}\ . \qquad (5)$$

Minimizing equation (5) subject to the constraint (4) results in the optimum value of t being the solution of the equation

$$(c_1+t)/t(1+mbt) + [\ 1 + 4(c_1+t)C/c_2^2\ ]^{-1/2} = 1\ . \qquad (6)$$

The sample size is from equation (4) given by

$$n_t = \{[(\ c_2^2 + 4(c_1+t)C\ )^{1/2} - c_2\ ]/2(c_1+t)\}^2\ . \qquad (7)$$

The solution of equation (6) can be found numerically, or iteratively solved, since

$$t = [\ (c_1 + c_2/2\sqrt{n_t})/mb\ ]^{1/2}\ , \qquad (8)$$

along with equation (7) also defines the optimum value of t. An initial value for the iterative procedure is $t = [\ c_1/mb\ ]^{1/2}$, the optimum length of tow if $c_2 = 0$, which is substituted into equation (7) and the resulting value of n_t into equation (8), and so on, till convergence.

Conversely, to obtain a desired level of precision, k, at minimum cost the optimum length of tow is the solution of the equation

$$c_2 k = 2[\ bmt^2 - c_1\ ][\ 1/mt + b\ ]^{1/2} \quad . \qquad (9)$$

The sample size is from equation (5) given by

$$n_t = [\ 1/mt + b\ ]/k^2 \ , \qquad (10)$$

and the total cost is given by equation (4). Again equation (9) may be solved iteratively using equations (10) and (8).

It is apparent from equation (8) that the optimum towing time decreases if: the preparation time, c_1, at a station decreases; the travel time parameter, c_2, decreases as a result, e.g., of a reduced habitat area; sample size increases due to, e.g., an increase in resources, C; the density coefficient, m, increases; or, if heterogeneity, as measured by b, increases.

A similar analysis can be made if another variance function is more appropriate. For example, if Taylor's relationship [equation (3)] is more suitable, then for a fixed cost, the iterative form of the solution for the optimum value of t is given by

$$t = (2-b)/(b-1)[\ c_1 + c_2/2\sqrt{n}\ ] \quad .$$

and equation (7), for 1 < b < 2. It may be noted that for this case the optimum length of tow is independent of the species density.

3. Examples

Surveys are routinely conducted on Georges Bank by the National Marine Fisheries Service to assess the abundance of various fish and scallop stocks. For groundfish surveys, a standard otter trawl is towed at 3.5 knots for 30 minutes at each selected station. It takes, on average, 30 minutes to set and retrieve the net, i.e., $c_1 = 30$. The travel time necessary to sample various numbers of stations was measured using data from previous surveys. Based on this data, the form of the travel component of the cost function was verified and values of c_2 were determined for areas of interest. If there are insufficient data, or the survey design changes, then map studies can be combined with experience to estimate c_2.

Currently, 5 days are spent surveying a particular region on Georges Bank, or C = 7200 min. For this area, $c_2 = 465$ min., and the variance function is assumed to be given by equation (1). Then for ocean pout, $\hat{b} = 1.60$ [the average value of $(s^2-\bar{x})/\bar{x}^2$ for the experimental data, (figure 1)] and $\hat{m} = .208$. For the gear presently in use, the optimum tow length is, from equation (6), 13.1 min. Figure 2 is a plot of the precision, k, of the density estimate versus tow length [equation (5)]. If it is desired to reduce k to .1 at minimum cost, the optimum length of tow would be 11.8 minutes [equation (6)] and the total time necessary to conduct the survey would be 14,979 min. When a species abundance is high, the optimum length of tow is much shorter. For example, for haddock, $\hat{b} = 2.43$, $\hat{m} = 10.769$ (Table 2), and

(Fig 2)

the optimum tow duration is 1.44 min..

In practice, groundfish surveys monitor a number of species which have wide fluctuations in abundance. For example, the survey catch per tow of haddock has declined to levels similar to the experimental catches of ocean pout. One way to select the tow duration for multispecies surveys with varying populations over time is to note that the optimum length of tow is monotonic in b and m. Historically, the important commercial stocks on Georges Ba had values of $b > 1.50$ and of $m > .17$. For $b = 1.50$ and $m = .17$, the optimum tow duration is approximately 15 min.. That is, for the Georges Bank example, a 15 min. tow duration would be more efficient than any greater tow length for these species and abundance levels.

For a typical survey of sea scallops on the north-east part of Georges Bank, $c_1 = 5$ and $c_2 = 614$. At each station a standard commercial sea scallop dredge is towed 15 min. at 3.5 knots and samples 3954 m^2. If $\mu = 5.79/100\ m^2$ and $b = 1.82$, then for a survey of 5 days, the optimum sampling unit is 291 m^2. By decreasing the unit size from 3954 m^2 to the optimum, the number of stations occupied increases by 37%, and k decreases from .15 to .13. Using the optimum unit, the total area actually dredged is $3.27 \times 10^4\ m^2$ as compared with $3.24 \times 10^5\ m^2$ for the 3954 m^2 unit. Conversely, keeping the level of precision fixed at $k = .15$, then the length of the survey can be reduced by 15% if the optimum unit size, which for this case is 311 m^2, is used instead of the 3954 m^2 unit.

4. Discussion

Decreasing tow duration, if appropriate, not only saves survey time, but also reduces operating expenses. For example, 70 tows of fifteen minutes each produce a density estimate for ocean pout which is as precise as 60 tows of thirty minutes but with considerably less total towing time. Gear and equipment wear is a function of tow length, and more fuel is consumed while dragging a trawl. Furthermore, with shorter tows there is less of a chance that an obstruction will cause a tow to be aborted or damage the gear.

An additional benefit from reducing tow duration is the resultant smaller catches which require less sorting time and allow more time for taking other biological measurements. Gear saturation, the filling of the sampler with animals or debris before the tow is completed, is also less likely to be a problem if tow length is reduced.

Other considerations may also influence the choice of unit size. In addition to density estimates, a certain number of animals is sometimes needed for other biological studies such as determining the age structure and growth rate of a population. Total survey catch is a function of the actual area sampled, and if total survey costs are fixed, this area decreases with unit size. Thus the size of the smallest practical unit may depend on the expected total catch for a given unit size. However, it is not necessarily true that more animals from fewer sampling locations will provide a better estimate of a population parameter than will fewer animals from more stations. Whatever the case, such factors should be taken into account along with cost when deciding on the size of the sampling unit.

Acknowledgements

We would like to thank Marvin Grosslein (NMFS, Woods Hole) for providing us with the Georges Bank tow duration data. J.H.V.'s work was supported by grant from the Norwegian Fisheries Research Council. The hospitality and cooperation of the staff at the Population Dynamics Branch at NEFC are also greatly appreciated.

References

Bliss, C.I. (1971). The aggregation of species within spatial units. In Spatial Patterns and Statistical Distributions, Statistical Ecology Series, Vol. 1, G.P. Patil, E.C. Pielou and W.E. Waters (eds), 311-335. University Park, Pennsylvania: The Pennsylvania State University Press.

Bliss, C. I. and Owens, A. R. G. (1958). Negative binomial distributions with common k. Biometrika 45, 37-58.

Cochran, W. G. (1977). Sampling Techniques, 3rd edition, New York: Wiley.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953) Sample Survey Methods and Theory. New York: Wiley.

Lenarz, W. H. and Adams, P. B. (1980). Some statistical considerations of the design of trawl surveys for rockfish (Scorpaenidae). Fisheries Bulletin 78, 659-674.

Roessler, M. (1965). An analysis of the variability of fish populations taken by otter trawl in Biscayne Bay, Florida. Transactions of the American Fisheries Society 94, 311- 318.

Seber, G. A. F. (1986). A review of estimating animal abundance. Biometrics 42, 267-292.

Serchuck, F. M. and Wigley, S. E. (1986). Evaluation of USA and Canadian research vessel surveys for sea scallops (Placopecten magellanicus) on Georges Bank. Journal of the Northwest Atlantic Fisheries Science 7, 1-13.

Taylor, C. C. (1953). Nature of variability in trawl catches. U. S. Fish and Wildlife Service, Fisheries Bulletin 54, 145-166.

Taylor, L. R. (1961). Aggregation, variance and the mean. Nature 189, 732-735.

Taylor, L. R., Woiwood, I. P. and Perry, J. N. (1979). The negative binomial as a dynamic ecological model for aggregation, and the density dependence of k. Journal of Animal Ecology 48, 289-304.

Table 1

Estimated coefficient of variation ($cv = s/\bar{x}$) with approximate standard errors (in brackets) for various sampling unit sizes for some sea scallop populations on Georges Bank.

Area	Country Conducting Survey	Average Catch/100m^2	Average cv	Sampling unit size(m^2)	Number of samples
North-east part,	USA	5.79	1.41 (.15)	3954	235
1982-84	CANADA		1.34 (.08)	3013	589
South Channel,	USA	14.42	1.64 (.52)	3954	32
1983	USA		1.55 (.47)	1318	32
All areas,	USA	3.42	1.63 (.16)	4943	343
1975,1977,1978					

Table 2.

Estimated coefficient of variation for haddock (*Melanogrammus aeglefinus*) from a tow duration experiment on Georges Bank in Jan., 1965. Each estimate is based on 16 tows. The trawl swept 1145 m^2 of bottom per minute.

Length of tow (min.)	15	30	60	120
Avg. catch per tow	190	394	627	1341
$\widehat{cv}$	1.16	1.90	1.56	1.53

Table 3.

Estimated coefficient of variation for haddock (Melanogrammus aeglefinus) from two tow duration experiments in the Barents Sea. Each estimate in the first experiment is based on 20 tows, and in the second experiment on 8 tows. The trawl swept 1574 m^2 of bottom per minute.

Length of tow (min)	5	15	30	60
Avg. catch per tow[1]		184	319	438
$\hat{cv}$		.94	1.17	.80
Avg. catch per tow[2]	33		126	
$\hat{cv}$	.72		.68	

[1] Oct., 1988; [2] Jan., 1989.

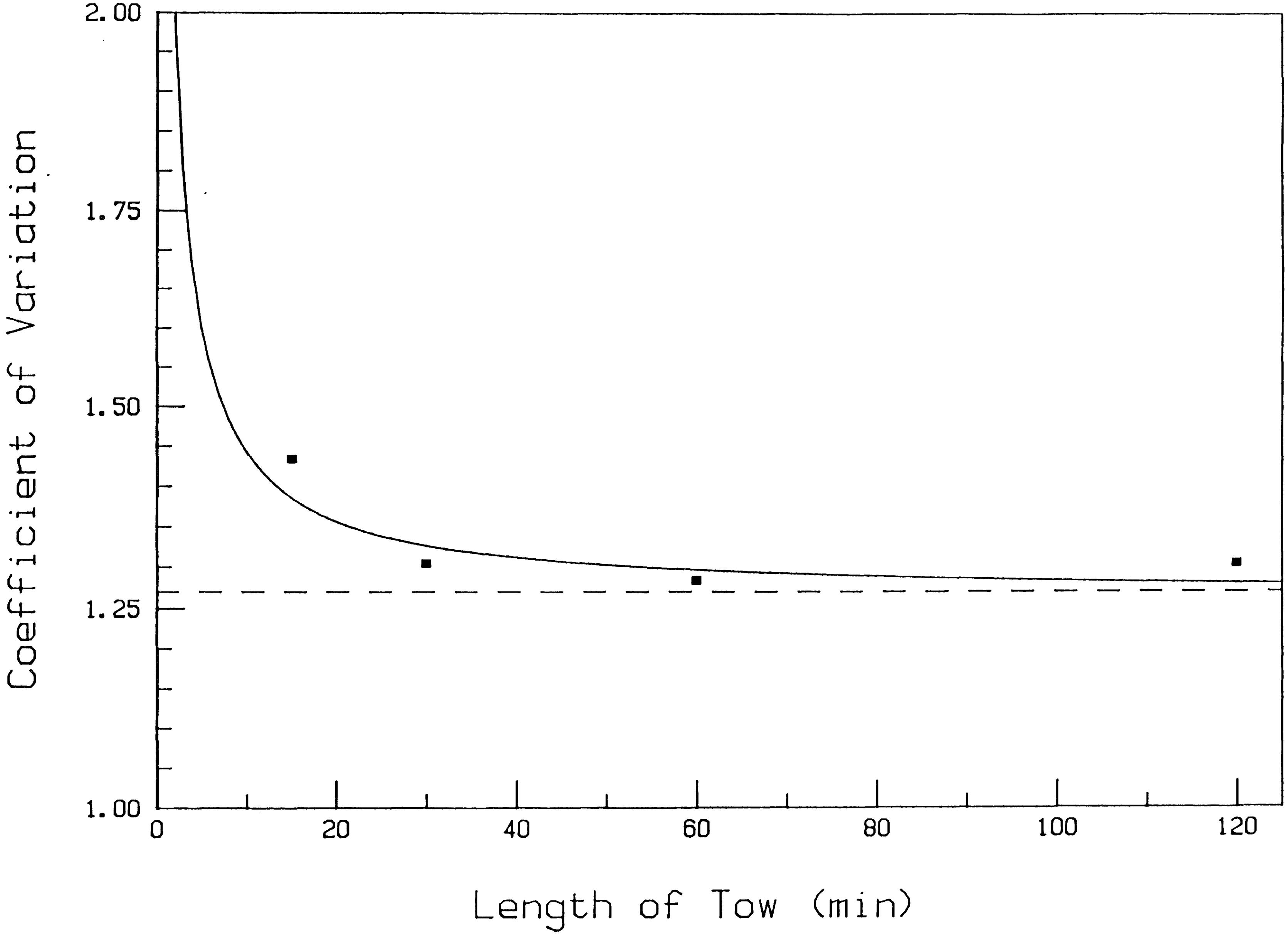

Figure 1. The coefficient of variation as a function of tow length for ocean pout. The graph is given by equation (2), and the points are estimates, each of which are based on 16 tows. The average numbers of fish caught per tow was 2.3, 7.5, 13.1, and 24.9 for the 15, 30, 60, and 120 min tow respectively.

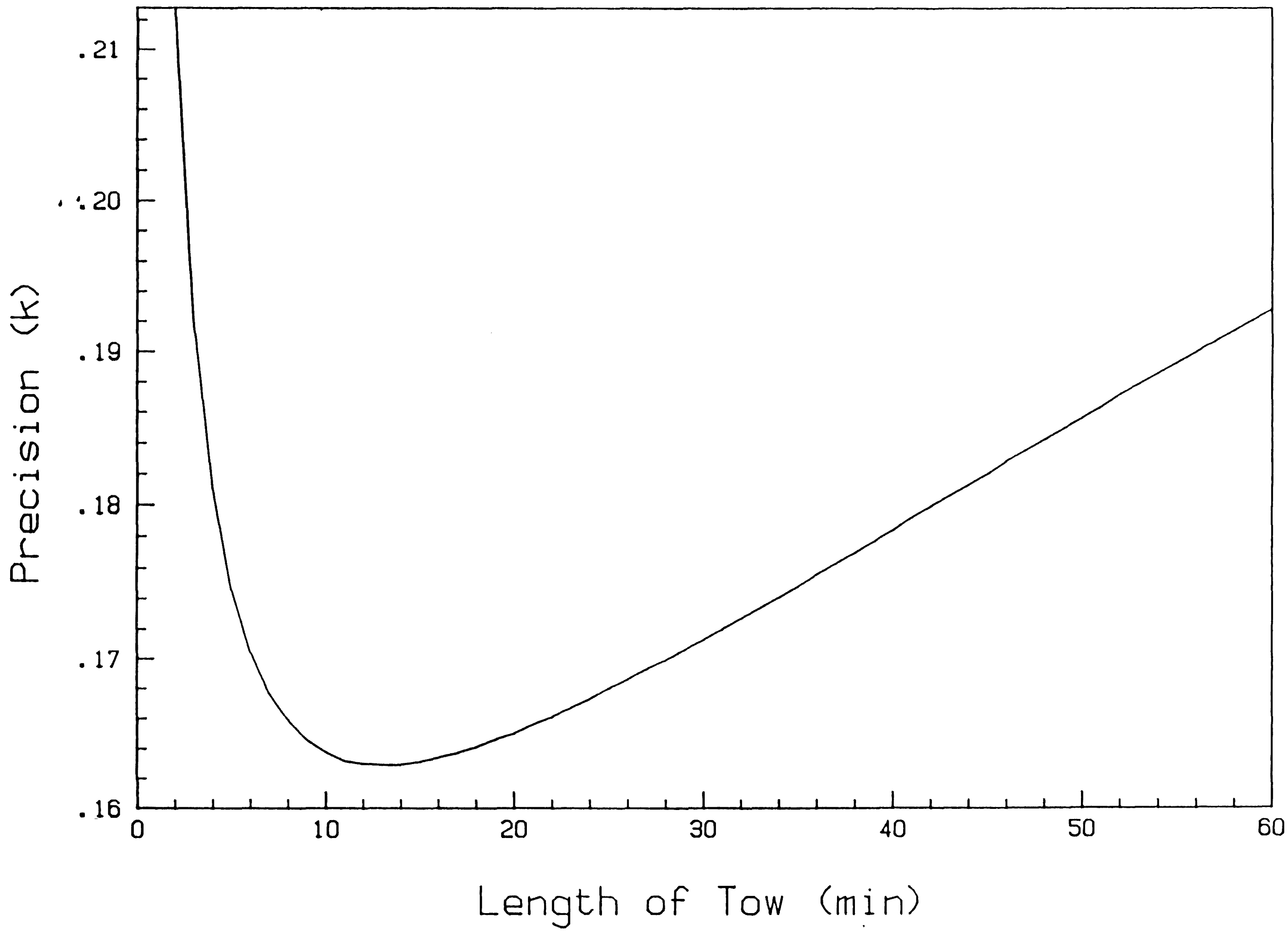

Figure 2. The precision, $k=cv/\sqrt{n_t}$, versus tow length for a survey of ocean pout with fixed total cost.