SOUTHWEST FISHERIES CENTER

NATIONAL MARINE FISHERIES SERVICE SOUTHWEST FISHERIES CENTER TIBURON LABORATORY 3150 PARADISE DRIVE TIBURON, CALIFORNIA 94920

JUNE 1990

A TEST OF USING WIND VELOCITY DATA WITH A PRODUCT ESTIMATOR TO IMPROVE THE EFFICIENCY OF SPORT SALMON LANDINGS ESTIMATOR

by

Norman J. Abramson

ADMINISTRTIVE REPORT T-90-04

Southwest Fisheries Center Administrative Report T-90-04

A Test of Using Wind Velocity Data with a Product Estimator to Improve the Efficiency of Sport Salmon Landings Estimation

Norman J. Abramson

Tiburon Laboratory
Southwest Fisheries Center
National Marine Fisheries Service, NOAA
3150 Paradise Drive
Tiburon, California 94920

June 1990

Introduction

A procedure which increases the efficiency of estimating fish landings by yielding more precise estimates for given sampling effort or, conversely, returns estimates of given precision with less sampling effort and lower cost would be quite valuable. Among sample survey methods, ratio and product estimators utilize auxiliary data which are available at low cost to improve the precision of estimates if there is correlation between the variable of interest and the auxiliary data; the ratio estimate is appropriate when the correlation is positive and the product estimate would be the choice when the variables are negatively correlated (Cochran, 1977). From general observations that sportfishing boats do not operate when the seas are rough, intuitively it seemed quite likely that a large negative correlation would exist between sportfish landings and some function of wind velocity and that a product estimator involving wind would improve sportfish catch estimates.

Area of Study

To test the hypothesis that wind velocity information would be useful in improving sportfish estimates, I chose to evaluate the utility of applying a product estimator to salmon landings of private boats (skiffs) at the ports of Crescent City and Eureka. While the California Department of Fish and Game (CDFG) samples salmon landings at most northern California ports, these sites were selected because of the relative simplicity of the sampling situation and the availability of nearby wind velocity

information. Wind velocity data were available from several weather buoys located off the California coastline. Buoys which I believed would be associated with sea conditions at Crescent City and Eureka, respectively, were St. George Reef (40.8°N, 124.5°W) and Eel River (41.8°N, 124.4°W). The landings sampling data used here were furnished by Joe Lesh, CDFG, Eureka, CA, and the wind data were provided by David Husby, NMFS, PFEG, Monterey, CA.

Description of Data

The weather data from the two buoys were in different formats and contained somewhat disparate information but both did contain the items of interest to me: hourly observations on wind velocity in knots or in m/sec. for each day. Programs were written to produce files with a common format, to convert wind velocity from m/sec to knots, and to calculate a mean velocity for each date.

Private boat landings data were collected with a two-stage cross-stratified random sampling scheme with the strata being port complexes and weekday or weekend days within semi-monthly periods (Lesh, 1977). Within strata, the first-stage sampling units are days and second-stage units are subports. Items of interest in the fishery data file are numbers of boats (trips), anglers, silver salmon , and king salmon. The fishery observations used herein were collected beginning May 28, 1988. Seventy-eight dates were observed at Eureka and 82 at Crescent City.

Estimation procedures

CDFG estimates total catches and other characteristics of the landings using mean-per-unit estimates and the ultimate cluster technique (Hansen, Hurwitz, and Madow, 1953). With this method, observations on the second-stage units (subports) are expanded, based on the subsampling fraction, to estimated totals for each first-stage unit (day) in the sample. In the computations these estimated totals are treated as if they were observations on the entire contents of the first-stage units. Resulting estimates and their estimated variances are unbiased so long as unbiased procedures are followed at each stage; however, variance contributions from separate stages cannot be estimated. For each port they estimate total landings of each species and total number of angler days using standard sample survey methods for stratified random samples(Lesh, 1977).

A product estimator is of the general form

$$\bar{Y}^* = \bar{y}\bar{x}/\bar{X}, \qquad (1)$$

where $\bar{Y}^*$ is the product estimate of the population mean, $\bar{X}$ is the population mean of the auxiliary variable, and $\bar{y}$ and $\bar{x}$ are the corresponding sample means; $N\bar{Y}^*$ is then the estimated population total when N is total number of sampling units in the population. In this case the y's correspond to the quantities being estimated and the x's to the known wind measurements.

With a stratified sample, one might estimate the mean and total for each stratum and sum the totals to estimate the population total. But because $\bar{Y}^*$ is a biased estimator and the

stratum-sample sizes are small, relative bias would build up during the summation. In this case a combined product estimator would be preferable if the correlation between the x's and y's can be assumed constant over the strata (Kaur, 1984). This combined estimator is of the form

$$\overline{Y}^*_c = (\Sigma p_i \overline{y}_i)(\Sigma p_i \overline{x}_i)/\overline{X} \qquad (2)$$

where $p_i = N_i/N$ and $\overline{y}_i$ and $\overline{x}_i$ are the sample means from the ith stratum. The larger combined sample size substantially reduces relative bias compared to summing stratum product estimated totals (Sukhatme, Sukhatme, Sukhatme, and Asok,1984).

Before actually computing the product estimates, it is possible to determine if conditions are present such that product estimation will be more efficient than the standard procedure. This testing process involves considerably less computation than calculating the estimates so I proceeded to carry out the test. Sukhatme, et al. (1984) show that the product estimator is more efficient than the mean-per-unit estimator if

$$r < -\tfrac{1}{2}\, C_x/C_y \ , \qquad (3)$$

where r is the correlation coefficient and C_x and C_y are coefficients of variation for x and y, respectively.

For each port, daily samples were expanded to estimated port-day totals of numbers of king salmon, silver salmon, angler days, and boat days. Because the cube of wind velocity is generally assumed proportional to sea height (D. Husby, Pacific Fisheries Environmental Group, P.O. Box 831, Monterey, CA 93942, pers. commun., 1989), this quantity was calculated for each date

in the sample and both mean wind velocity and its cube were tried as auxiliary variables in conjunction with the daily estimated totals mentioned above. For the purpose of testing the efficiencies of the product estimators, all of the observations for each port were pooled into a single sample; i.e., the strata were collapsed.

Results

The values of r, the sample correlation coefficient, and the statistic cvr = $-\frac{1}{2}\ C_x/C_y$ were computed for each of the sixteen pairs of daily estimated totals and auxiliary wind variables described previously (Table 1). The results of these calculations were surprising to me; in only one of the sixteen cases was $r < cvr$.

Examining Table 1, we see that the correlation coefficient is negative in all of the categories for Eureka and in half of them for Crescent City. But, except for the Crescent City silver salmon-wind pairing, the cvr statistic is smaller than r. From equation (3) we can see that the efficiency of the product estimator relative to the mean per unit estimator is dependent not only on a negative r but also on the ratio of C_x to C_y. The coefficients of variation are shown in Table 2; examining these in conjunction with the contents of Table 1, it is apparent that the lack of efficiency of the product estimators was caused by factors ranging from a positive correlation between wind and king salmon catches at Crescent City to too large values for C_x relative to C_y.

In summary it appears that the use of a product estimator to improve the efficiency of estimating recreational salmon catches, or other parameters associated with the fishery, is not a viable option. The outcome of this work now may join many, many others in the huge, undocumented library of negative results.

Table 1.--Sample correlation coefficients (r) and coefficient of variation ratio statistics (cvr) for Crescent City and Eureka.

CRESCENT CITY

	Wind		$Wind^3$	
	r	cvr	r	cvr
King salmon	0.101	-0.188	0.113	-0.476
Silver salmon	-0.182	-0.130	-0.174	-0.329
Boat days	-0.098	-0.318	-0.063	-0.809
Angler days	0.022	-0.309	0.038	-0.784

EUREKA

	Wind		$Wind^3$	
	r	cvr	r	cvr
King salmon	-0.201	-0.228	-0.215	-0.661
Silver salmon	-0.270	-0.311	-0.275	-0.899
Boat days	-0.393	-0.470	-0.353	-1.360
Angler days	-0.359	-0.437	-0.324	-1.263

Table 2.--Coefficients of variation of variables and auxiliary variables for the ports of Crescent City and Eureka.

	CRESCENT CITY	EUREKA
King salmon	1.007	1.540
Silver salmon	1.456	1.134
Boat days	0.593	0.750
Angler days	0.612	0.808
Wind	0.378	0.705
$Wind^3$	0.959	2.040

Citations

Cochran, W.G.
1977 Sampling techniques, 3d ed. Wiley, New York, 428 p.

Hansen, M.H., W.N. Hurwitz, and G.W. Madow.
1953 Sample survey methods and theory, vol 1. Wiley, New York, 638 p.

Kaur, P.
1984 A combined product estimator in sample surveys. Biom. J. 26:749-753.

Lesh, E.W.
1977 A stratified random sampling program used to estimate landings in the northern California ocean skiff recreational fishery. Calif. Dept. Fish and Game, Anad. Fish. Br. Admin. Rep. 77-2, 9p.

Sukhatme, P.V., B.V. Sukhatme, S. Sukhatme, and C. Asok.
1984 Sampling theory of surveys with applications, 3d ed. Iowa State Univ. Press, Ames, Iowa, 526 p.