J. mar. biol. Ass. India, **46** (2) : 204 - 208, July - Dec., 2004

Bootstrap evaluation of the sampling scheme to estimate the marine fish landings in Madras Fisheries Harbour

Somy Kuriakose and M. Srinath

Central Marine Fisheries Research Institute, Cochin-682 018, India

Abstract

Bootstrap methodology was applied for the evaluation of the sampling design with reference to the coefficient of variation of the estimate. The data on trawl landings at the Madras Fisheries Harbour during 2002 was considered for the study. The bootstrap sampling was done for the first stage units (days) and the Coefficient of Variation was estimated for 1000 bootstrap replication. The optimum sample size (number of days) for the estimation of marine fish landings at a desired level of precision has been estimated.

Key words: Bootstraping, optimum sample size, marine fish landings

Introduction

Exploited fish stocks are assessed with the help of micro analytic and macro models (Alagaraja, 1990). Catch in numbers (age specific or length specific) or in weight and the corresponding fishing effort expended are the main inputs to the fish population models. For a proper evaluation of the stock, statistics of catch and effort along with the relevant biological characteristics are essential over time and space. The quality of this input data governs the performance (predictive or interpolate) of the models and determines the relevance of management options derived from the stock assessment.

Catches usually are estimated from sampling of commercial landings. These sampling schemes are often complex and multistage in nature. In India marine fish catch statistics are collected by the Central Marine Fisheries Research Institute, Cochin through a sampling system based on the theory of sampling (Banerji, 1971). Most of the catches are from the inshore regions and landed at about 1400 landing centres spread all along the coast line in the various maritime states of India. Keeping pace with the changing pattern of the fishery the mode of collection has also undergone change periodically without any significant alterations in the basic structure of the sampling design.

The sampling design followed by the Central Marine Fisheries Research Institute during 1970s and before was explained by Kutty (1973). With the spurt in the implementation of mechanization in the fishing industry the quantity and quality of data to be collected increased tremendously. Taking this into account the concept of Single Centre Zone was introduced meaning a particular centre at which there was intense mechanized activity. The mode of collection during the late 1970s and early part of 1980s was described by Jacob (1983). Later, the mode of collection underwent slight change with respect to selection of crafts and the modified sampling scheme was given by Alagaraja (1990).

It is known that if the first stage units are selected with replacement and the second stage units selected systematically then the estimate of the variance reduces to that among the first stage units (Sukhatme and Sukhatme, 1970). However, in this case only the landing centres are selected with replacement and the landing center days are selected without replacement. Thus, in this case estimation of variance poses a problem. Another important aspect is the sample size. It is important to know the optimum sample size for a desired level of precision. Are the currently observed number of days and the boats selected on the selected day adequate enough for estimating the total catch for a specified level of precision? This question can be answered if an estimate of the variance is available with us (here the total cost of the survey is not considered). This paper presents an approach in an attempt to answer the above question.

Material and methods

The Monte Carlo Bootstrap methodology was applied to evaluate the sampling scheme in terms of estimates of the coefficient of variation and determining the number of days for observation. Bootstrap evaluation involves computer-intensive methods of statistical analysis that use simulation to calculate standard errors, confidence intervals and significance tests (Efron, 1979,1991). Kimura and Balsiger (1985) pointed out that one could spend considerable time and effort fitting these data into classical sampling theory. Alternatively, the bootstrap method uses the well-defined structure of the survey to define an empirical process. This sample is processed repeatedly using Monte Carlo methods and the resulting variability analyzed. According to Efron (1982) the important theme of resampling methods such as Bootstrap is the substitution of computational power for theoretical analysis. The bootstrap can routinely answer questions which are too complicated for traditional statistical analysis. Bootstrap experiments were carried out with 1000 bootstraps. The software used for this study was developed in C language.

Madras Fisheries Harbour one of the most important landing centers in Tamil Nadu, where large number of mechanized boats operate is chosen. The data collected during January to December 2002 formed the material for anaysis. At this center, catches from trawlers and gillnetters are landed. Of these, the landings from the trawlers formed the major component and thus only trawl catches were considered for estimation of variance and also for determining the optimum sample size. A month was divided into 3 groups of ten days each. From the initial five days of the first group, a day was selected at random. Starting from this day, 3 clusters of two days each were formed. From the remaining two 10 day-groups , the clusters were selected with an interval of 10 days. For example, from the first five days if the day selected was 3, then the three clusters in the first ten day group were (3,4),(5,6) and (7,8). Then from the next two 10 day-groups, clusters would be (13,14), (15,16), (17,18), (23,24), (25,26) and (27,28). Thus we have 9 clusters of 2 days each resulting in 18 days of observation in a month which could be considered as a simple random sample without replacement. On each selected day, the landings from a certain number of boats were observed depending upon the number of boats landed (Alagaraja,1990). Here also it was assumed that the boats were selected without replacement though in practice they were usually selected systematically.

Results and discussion

The monthwise number of fishing days (No. of days) and the observed number of days at the Madras Fisheries Harbour during 2002 are given Table 1. Although 18 days per month were selected, observations could not be made on some days due to various reasons and only the effective number of days observed were considered for the study. Thus the scheme of collection of catch statistics for the purpose of this study can be assumed to be that of a classical two-stage design with the days forming the first stage units and the boats landing their catches being the second stage units.

Table 1. *Number of fishing days and number of days observed at Madras Fisheries Harbour during 2002*

Month	NF	NDS
Jan.	31	14
Feb.	28	15
Mar.	31	14
Apr.	15	6
May	0	0
Jun.	30	12
Jul.	31	14
Aug.	31	14
Sep.	30	16
Oct.	31	16
Nov.	30	16
Dec.	31	15

NF: Number of fishing days in a month
NDS: Number of days observed

Ideally, the bootstrap evaluation in this case should be carried out in two stages, one for the days and the other on the number of boats on the selected day. However, the bootstrap sampling was done only among the first stage units because on analysis it was found that the percentage contribution of the variance due to the second stage units to the total variance was not large enough to be considered and the major contribution to the total variance was from among the first stage units only. The bootsrap variance was estimated using bootstrap samples of size 1000. The monthly coefficients of variation for different sample sizes were estimated. Only those months where the number of observations is more than 8 days were considered for analysis. The monthly coefficient of variations ranged from 4% to 30% for 18 days to 3 days observation. In most of the months the coefficient of variation ranged between 10% to 15% for 10 or more days of observation per month (Fig.1). If a precision level of 10% to 15 % is assumed to be satisfactory for estimating the total landings from a centre , it could be concluded that 10-12 days observation would be sufficient to estimate the catch statistics. The conclusions about the optimum number of observations for a desired level of precision cannot obviously, be generalized to all the single centre zones in the country. Besides, these results are applicable only to the trawl fishery of the selected centre and may not be valid for covering all other types of

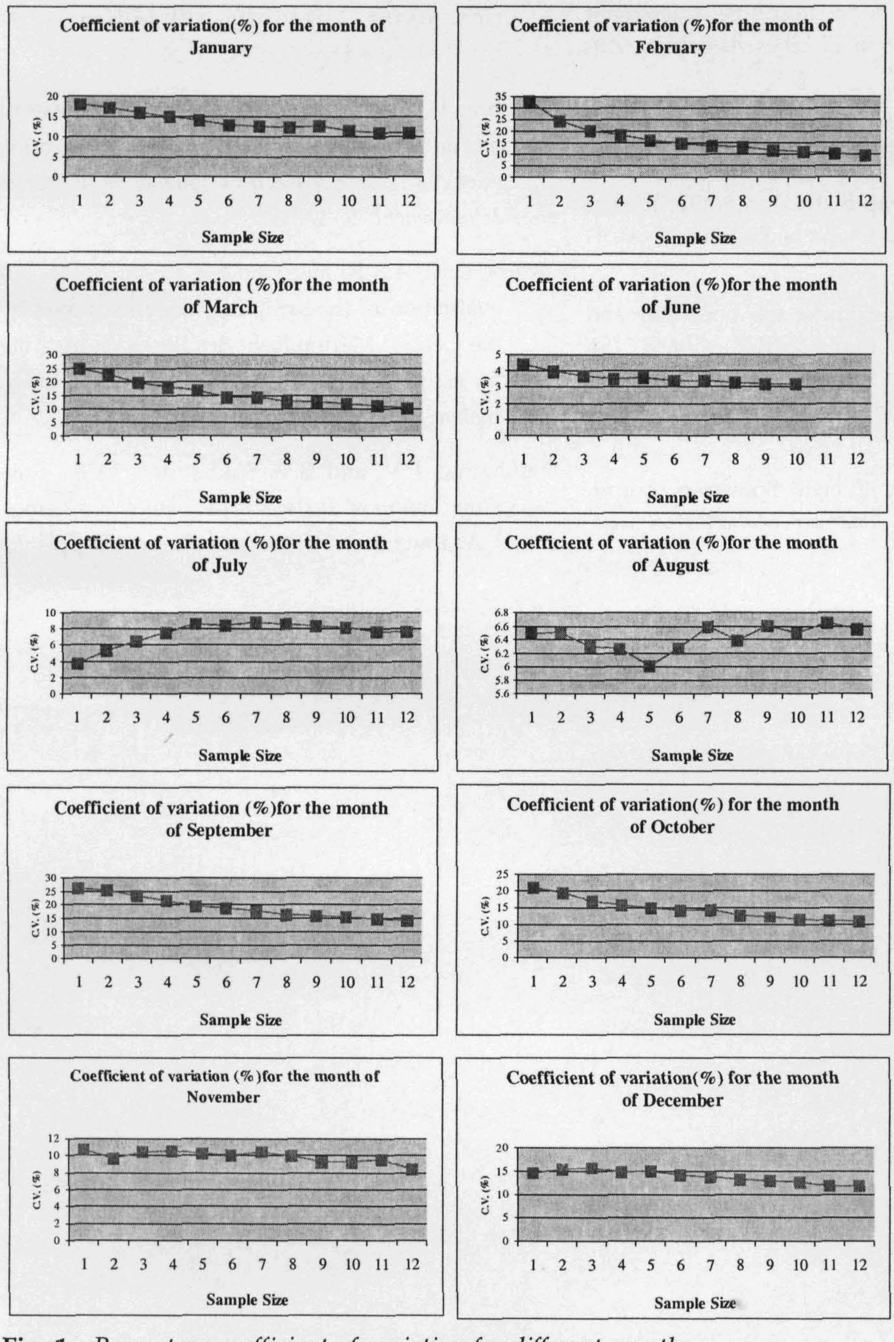

Fig. 1. *Percentage coefficient of variation for different months*

fishery such as gillnet, ring-seine and purse-seine.

References

Alagaraja, K. 1990. Fisheries Resource Assessment - A Glance. *Industrial Fisheries Association Annual.* **7**: 5-11.

Banerji, S.K. 1971. Fishery Statistics. *FAO Publ.* No. FC/DEV/71/5: 15 pp.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. *The Annuals of Statistics.* **7** (1): 1-26.

________1982. The jackknife, the bootstrap and other resampling plans. *SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics,* No.38: Philadelphia, Pennsylvania, USA.92pp.

_________1991. More efficient Bootstrap computations. *Journal of American Statistical Association,* **85**: 78-89.

Jacob, T., K. Alagaraja and K.N. Kurup. 1983. Marine fisheries statistics in India - Present status. *In* : Proceedings of the Workshop on Acquisition and Dissemination of data on marine living resources of Indian seas. *Mar.Fish.Infor.Serv.T&E.Ser.,* No. **46**: 6-11.

Kimura, D.K. and J.W. Balsiger. 1985. Bootstrap methods for evaluating Sablefish pot index surveys. *North American Journal of Fisheries Management,* **5**: 47-56.

Kutty, D.K., A.K.K. Nair and S.Z. Qasim. 1973. An evaluation of the sampling design adopted by the Central Marine Fisheries Research Institute for estimating marine fish production in India. *Indian J. Fish.,* **22**: 16-34.

Sukhatme, P.V. and B.V. Sukhatme. 1970. *Sampling Theory of Surveys with Applications,* Ames. I A: Iowa State University Press. 452 pp.