1. Introduction

The approaches described in the paper address the question of sampling accuracy when sample-based data collection is performed under operational constraints, a frequent concern of fishery administrations with limited budget and human resources. Such a question is usually directly related to the frequency and extent of field operations for data collection and a number of classical methods are available for determining appropriate sample size on the basis of population parameters that have been derived at an earlier stage from the same or similar populations (see Cochran, 1977; Thompson, 1992; for discussion). This study focuses on *a priori* determination of safe sample size using the same methods and adjusting them to respond to specific target populations.

During the design phase of large-scale catch/effort assessment surveys for artisanal fisheries the question often arises as to what should be the appropriate sample size guaranteeing an acceptable level of reliability for the estimated population parameters. In some circumstances, and particularly at the early stages of implementing a fishery statistical monitoring programme, very little is known about the distribution and variability of the target population. Consequently, statistical developers tend to initially operate on a large-sample basis, with the intention of scaling down data collection as soon as some guiding statistical indicators, used for improving the cost-effectiveness of the sampling schemes, become available. Usually such indicators can only be formulated and verified after a complete operational cycle of a fishery statistical programme, which means that for long periods data collection is performed at high operational capacity. Generally, lack of any *a priori* guidance on sample size tends to increase the size and complexity of field operations and this, in turn, has a direct impact on the logistical aspects of data collection and data management procedures.

Most of the discussions in this publication concern a relative index of proximity between a sample mean and the population mean that

derives from a maximum allowable difference between a true population value and its estimate. This index, referred to as *accuracy* A throughout the paper, has several statistical and geometrical properties that are only a function of the population size or the knowledge that the population under study is large or infinite. Using these properties it is possible to formulate accuracy boundaries that can be used for predicting a lower limit for accuracy at any sample size. Construction of these boundaries is fairly simple and can be quickly achieved through the use of standard computing tools (such as spreadsheets) that are available on most personal computers.

Formulation of *a priori* boundaries for sampling accuracy consists of:

a) Guessing the general shape of the distribution of the target population in catch/effort assessment surveys (or accepting that no guessing is possible); and

b) Setting-up global accuracy boundaries that are only a function of the population size.

Before introducing the underlying concepts of *a priori* accuracy indicators, this introductory section will deal with some general aspects of sample-based fishery surveys with emphasis on:

(a) Basic fishery data.
(b) Justification for regular collection of basic fishery data.
(c) Scope and utility of basic fishery data.
(d) Need for fishery surveys to be cost-effective and sustainable.
(e) The key role of survey planners and statistical supervisors in the monitoring and evaluation of the performance of a data collection system.

1.1 Utility of basic fishery data

The definition of basic fishery data is rather empirical and based on the traditional method of setting-up general-purpose datasets that are subsequently used by a variety of application-specific systems.

In this handbook basic fishery data refer to catch, fishing effort, catch by species, first-sale prices (i.e. prices at landing), values, and fish size (in weight units).

A fishery statistical programme collecting basic fishery data should not be an end in itself. People involved in such programmes are sometimes challenged to provide a justification for regularly conducted (and thus costly) fishery surveys. From a long list of potential uses of basic fishery data, the following applications may be considered as representative:

1.1.1 Food Security

Food security is an over-riding concern for policy-makers, planners and administrators of natural living resources. In many communities, particularly in developing countries, fish is the major source of animal protein and people are dependent on fish as a food source.

Food balance sheets constitute a principal source of information for studies concerned with food security. Estimated total production of fish, combined with data on catch disposition, imports and exports, constitutes the basis for calculating *per capita* consumption of fish, which is subsequently used in the formulation of food balance sheets.

Basic data involved: Estimated total catch, estimated catch by species. Estimations are usually based on sampling approaches.

1.1.2 Fishing mortality

Fishing effort is one of the variables used to estimate fishing mortality. Fishing mortality is a fundamental variable in stock assessment, representing the proportion of stock that is removed due to fishing. Effort is used in setting most fishing controls. Changes in total fishing effort may be an indication of stock status or fishing profitability.

Basic data involved: Estimated fishing effort. Estimation is usually based on sampling approaches.

1.1.3 Fishing operations

Fishing operations indicators describe the composition of fishing fleets and fishing patterns and are the basis of most management decisions. They are important for monitoring compliance and in analyses involving fishing effort.

Basic data involved: Thematic maps of homeports and landing sites, numbers of fishing units by gear category, estimated fishing effort by boat/gear category. Effort estimates are usually based on sampling approaches.

1.1.4 Gear selectivity

It is often useful to obtain data indicative of the species that are targeted by different boat/gear categories and/or fishing methods, together with other information relating to the size of the fish being caught. These datasets are used for a wide variety of in-time and in-space comparisons of gear selectivity indicators.

Basic data involved: Species composition, average weight of fish. Such indicators are usually based on sampling approaches.

1.1.5 Abundance and exploitation

Catch-Per-Unit-Effort (CPUE) or catch rate is frequently the single most useful index for long-term monitoring of a fishery. It is often used as an index of stock abundance, where some relationship is assumed between the index and the stock size. It can also be used in monitoring economic efficiency. Catch rates by boat and gear categories, often combined with data on size at capture, permit a large number of analyses relating to gear selectivity and indices of exploitation.

Basic data involved: Sample Catch Per Unit Effort (CPUE).

1.1.6 Importance to national economy

For national and local policy-making and planning it is essential to describe the contribution of fisheries to the economy. Assessment of the economic contribution of fisheries needs to take into account several important variables and indicators, among which product prices and gross value of production.

Basic data involved: Estimated total catch, estimated catch by species, sample prices, estimated values. *Estimations are usually based on sampling approaches*.

1.1.7 Fleet performance and profitability

Boat profitability is a vital micro-economic indicator of fishery performance since it provides a measure of economic sustainability of artisanal fleets. Prices at landing, combined with data on investment and operational costs can provide indices of fleet performance.

Basic data involved: Overall Catch-Per-Unit-Effort (CPUE), unit value (average price) of catch. Such indicators are usually based on sampling approaches.

1.1.8 Socio-economic studies

Time series of fishing effort, catch, Catch-Per-Unit-Effort (CPUE) and prices are often used in socio-economic studies. Such data are indicative of declining or increasing trends of fisheries in districts and regions.

Basic data involved: Monthly data on catch, effort, CPUEs, prices and values. Estimations are usually based on sampling approaches.

In the description of the basic data involved in each of the listed applications, sample-based approaches are used for the estimation of key variables and the formulation of fishery indicators. Thus it becomes evident that the utility of the applications and their impact on fishery planning and management depends directly on the reliability of estimates resulting from sample-based data collection operations. This reliability and its relation to size and frequency of samples, together with its measurement and control constitute the main focus of this paper.

1.2 Cost-effective fishery surveys

Regularly conducted fishery surveys are costly since they involve salaries and wages of field and office personnel, recurring field operations costs and other overhead and maintenance costs relating to office infrastructure and operations. In many developing countries these costs constitute a major constraint and it is thus important for fishery statistical programmes to be as effective as possible at lowest cost. Sample-based fishery surveys are cost-effective when:

(a) They are economical in data collection effort and yet capable of producing reliable estimates.
(b) They make good use of existing human and financial resources involved in data collection and processing.
(c) They respond to users' needs in a timely and reliable manner.

The above three criteria indicate the need for realistic survey planning, well-defined sampling programmes and regular monitoring of survey results by means of meaningful and simple statistical indicators. As it will be seen in the coming sections of this document, most of these indicators are directly related to sampling accuracy.

SUMMARY

In this introductory section readers have been introduced to:

(a) The need for *a priori* statistical indicators that will guide the use of sufficient and appropriate samples and guarantee a desired level of accuracy in the estimates.
(b) Importance and utility of basic fishery data such as catch, effort, prices and values and a list of commonly used applications that make use of such data.
(c) The need for sample-based fishery surveys to be cost-effective and sustainable and some criteria for evaluating them from these two standpoints.