

Why Counting Every Fish Isn't Necessary: The Role of Statistical Sampling in Aquaculture

*Jayasri Ravi¹ and L. Surulive^{1,2}

¹PG Scholar, Department of Aquaculture, Dr. M.G.R. Fisheries College and Research Institute, TNJFU, Ponneri, Tamil Nadu, India

²Assistant Professor, Department of Fisheries Extension, Economics & Statistics, Dr. M.G.R. Fisheries College and Research Institute, TNJFU Ponneri, Tamil Nadu, India

*Corresponding Author's email: jayasriravi340@gmail.com

It is rarely possible to completely enumerate an aquatic population by counting and measuring each individual in a pond, cage, or fishery, and when it can, it is rarely desired. For all but the smallest or most valuable populations, a complete census is unfeasible due to stress, mortality risk, labour costs, and time. A scientifically validated alternative is statistical sampling, which allows researchers and farm managers to estimate population parameters like mean weight, length-frequency distribution, disease prevalence, and abundance with a known and controllable margin of error by selecting a representative subset of a population and measuring it under a predetermined design. The main sampling strategies utilised in fisheries and aquaculture science, the justification for sampling over census in aquaculture, and the fundamentals of sample size are all covered in this article. In order to demonstrate that sampling is the methodologically proper way to understand a population well enough to manage it, rather than a compromise imposed by convenience, the discussion draws on standard reference books in fisheries statistics as well as recent primary research and stock-assessment procedures.

Keywords: Sampling Methods, Statistics, Fisheries Science, Sampling Methods, Sampling Design, Complete census.

Introduction

Aquaculture and fisheries management are, at their core, exercises in making decisions about populations that cannot be fully observed. A grow-out pond may hold tens of thousands of fish; a coastal fishery, millions spread across hundreds of kilometres. Decisions on feeding rate, stocking density, harvest timing, disease surveillance, and stock status all depend on knowing something about the average condition and variability of that population. The intuitive solution such as counting and measuring every individual is termed complete enumeration, and while it eliminates sampling error entirely, it is seldom practical. Complete enumeration is expensive in labour and time, and in the case of live aquatic animals, it is also stressful and potentially harmful to the very population being assessed (Cadima et al., 2005). Fisheries science therefore developed, over the course of the twentieth century, a formal body of sampling theory adapted specifically to aquatic populations, where individuals are mobile, often invisible below the water surface, and costly to physically recover. This article examines why sampling, rather than counting, is the accepted standard in both capture fisheries and aquaculture, and outlines the key statistical concepts such as sampling designs, sample size determination, and effort optimisation.

The Case Against Complete Enumeration

A census requires that every unit of the population be located, captured, and measured. For farmed aquatic populations this generally means fully draining or netting a pond, cage, or tank. Beyond the practical difficulty of doing this at scale, the process itself is destructive: it causes acute physiological stress, injury, and, in disease-compromised populations, elevated mortality (Cadima et al., 2005). For wild fisheries, complete enumeration of a stock is not physically possible at all, a stock's true abundance can only ever be estimated, never directly counted. Cadima et al. (2005) frame this trade-off explicitly in FAO's standard sampling manual: complete enumeration is reserved for situations where full information is essential and resources permit it, whereas sample-based estimation is the default wherever a reliable estimate can be produced at substantially lower cost and disturbance. This is a statistical judgement, not merely a logistical shortcut. A well-designed sample carries a "known" and "quantifiable" margin of error, which a haphazard partial count does not.

What Statistical Sampling Actually Involves

Sampling is the deliberate selection of a representative subset of individuals from a defined population, measured so that the results can be generalised back to that population with a stated level of confidence. Cadima et al. (2005) describe this using the concept of the "three worlds" of sampling theory: the real population, the particular sample drawn from it, and the theoretical distribution of all possible samples of that size that could have been drawn. Provided selection is random and the sample size is adequate, the sample mean is an unbiased estimator of the population mean, and the associated sampling variance allows a confidence interval to be calculated around that estimate. This distinguishes rigorous sampling from casual observation. A statement such as "mean body weight was 45 g, 95% CI $\pm$ 3 g, n = 60" is a testable, reproducible scientific claim; an estimate based on glancing at a few fish scooped from the feeding area is not, because it carries an unknown and likely biased error.

Sampling Designs Used in Fisheries and Aquaculture

Different designs suit different population structures and logistical constraints (Cadima et al., 2005)

- **Simple random sampling** gives every individual an equal chance of selection.
- **Stratified random sampling** divides the population into meaningful strata, pond zones, cage depths, or size classes and samples each separately, improving precision when the population is not uniform.
- **Cluster sampling** selects whole groups (entire cages, tanks, or landing sites) rather than individuals, which is often more logistically convenient though potentially less statistically efficient per animal sampled.
- **Multistage sampling** combines these approaches, for instance, first selecting a subset of ponds or landing sites, then sampling individuals within each and it is common in large-scale fisheries data collection programmes.

The choice of design is a deliberate trade-off between statistical precision, field practicality, cost, and animal welfare, not an arbitrary matter of convenience.

Determining an Adequate Sample Size

A central question in practice is how large a sample must be to yield a trustworthy estimate. Larger samples reduce sampling error but increase handling stress, labour, and cost, so sample size must be deliberately calculated rather than assumed. In aquatic animal health surveillance, conventional sample sizes of 30, 60, and 150 individuals are widely used to detect infection at assumed minimum prevalence levels of 10%, 5%, and 2% respectively, based on the historical convention of treating aquatic populations as effectively infinite (de Blas et al., 2020). Using Monte Carlo simulation, de Blas et al. (2020) tested the validity of these conventional formulas against varying population sizes and spatial distributions of infected animals, and found that while the conventional sample sizes perform well under simple random sampling regardless of prevalence or clustering, their reliability changes once

the sampling method or the spatial distribution of the trait of interest departs from these assumptions. The practical implication for aquaculture is that sample size formulas cannot be applied mechanically; they must be matched to the specific population size, the expected distribution of the characteristic being measured, and the sampling method actually used in the field.

Optimising Sampling Effort Without Losing Precision

Beyond deciding how many individuals to sample, a growing body of fisheries statistics research addresses how to extract maximum information from a given sample, directly relevant to reducing cost and animal handling in both wild-stock and farmed contexts. Length-frequency data are a standard sampling output because length is quick and inexpensive to record compared with age or reproductive status, and length-frequency distributions underpin growth-parameter estimation and stock assessment more broadly (Wischnewski et al., 2020). Wischnewski et al. (2020) proposed a distribution-free framework, the “Admissible Dissimilarity Value” that identifies the smallest length subsample still capable of preserving an accurate length-frequency distribution, enabling national and regional sampling programmes to optimise effort without compromising assessment quality. The same reasoning applies at farm scale: a smaller but carefully designed and analysed sample can be more informative than a larger, haphazardly collected one.

From Wild Stock Assessment to the Farm Pond

The statistical framework described above originated in capture-fisheries stock assessment, where fishery-independent surveys sample only a small fraction of a stock's geographic range using standardised gear at a limited number of stations, yet still produce abundance indices precise enough to set catch limits for an entire fishery. If a few hundred trawl stations can responsibly inform management decisions across a stock spanning thousands of square kilometres, the same logic scales down naturally: a properly designed sample of a few dozen individuals can responsibly inform feeding, health, or harvest decisions for a single pond or cage. The underlying principles of random or stratified selection, an adequate and justified sample size, and a stated margin of error remain identical; only the spatial and operational scale changes.

Conclusion

Sampling in aquaculture is not a compromise adopted because counting every fish is inconvenient, it is the methodologically correct approach to characterising a population that cannot, and should not, be fully enumerated. A well-designed sample, drawn using an appropriate design (simple random, stratified, cluster, or multistage), sized according to the population and objective at hand, and analysed with attention to precision and effort optimisation, yields estimates whose reliability can be stated explicitly rather than assumed. From routine feed-rate adjustment and growth monitoring to disease surveillance and formal stock assessment, the statistical logic is the same at every scale, from a single grow-out pond to an entire capture fishery. Understanding this logic, not merely the mechanics of taking a sample is what separates trained fisheries and aquaculture practice from guesswork.

References

1. Cadima, E. L. (2005). Sampling methods applied to fisheries science: a manual (No. 434). Food & Agriculture Org..
2. De Blas, I., Muniesa, A., Vallejo, A., & Ruiz-Zarzuela, I. (2020). Assessment of sample size calculations used in aquaculture by simulation techniques. *Frontiers in veterinary science*, 7, 253.
3. Wischnewski, J., Bernreuther, M., & Kempf, A. (2020). Admissible dissimilarity value (ADV) as a measure of subsampling reliability: case study North Sea cod (*Gadus morhua*). *Environmental monitoring and assessment*, 192(12), 756.