



FishStat

FAO FISHERY AND AQUACULTURE STATISTICS

2. Data collection and sampling

2.3. How to collect data (data collection strategies and methods)

IOTC species identification and sampling workshop 2025

Kochi, India, September 29th to October 3rd, 2025

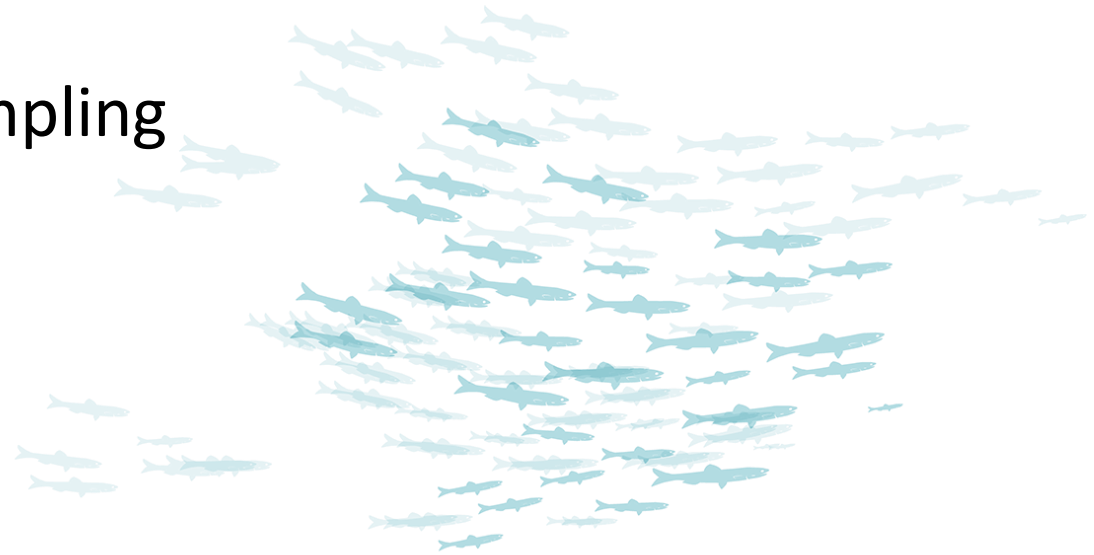
Yann Laurent

Yann.laurent@fao.org



OUTLINE

1. Objectives of the module
2. Introduction
3. Census vs sampling
4. Sample-based approach and sampling
5. Stratification
6. Data collection methods



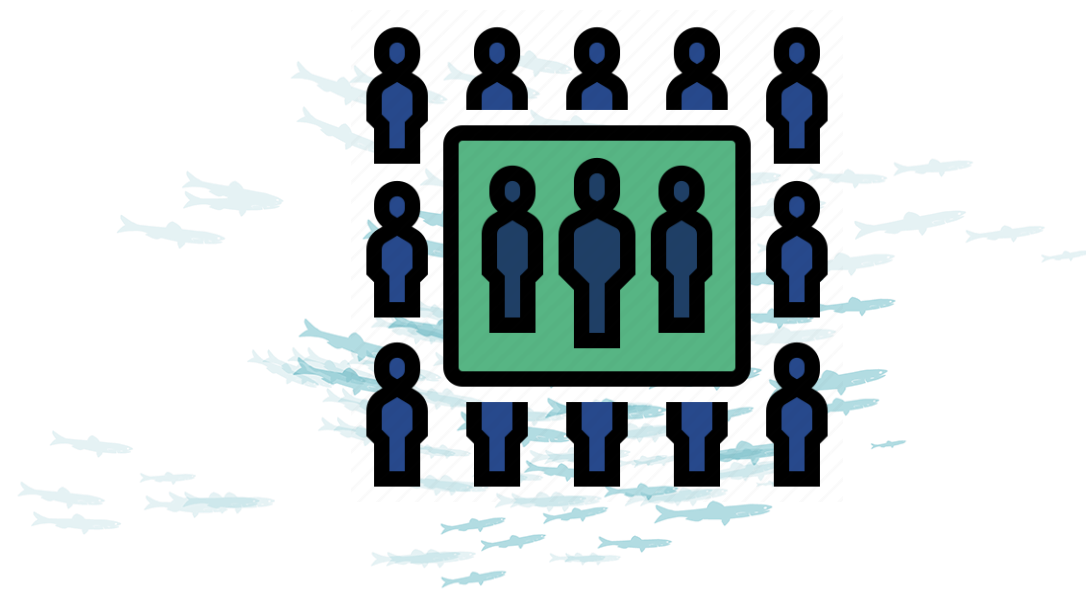
1. Objectives of the module

- Understand basic statistical methods and concepts
- Know when to apply full census approach or a sample-based approach
- Understand stratification concept and use
- What can go wrong at various stages of data collection process?



2. Introduction

- Strategies for design of data collection programmes varies among fisheries
- There are two strategies:
 - complete enumeration = census
 - sampling



Introduction

Data collection strategy - steps

Evaluate

- Evaluate existing data sets in relation to the objectives of the programme, including accessibility of the data;

Describe

- Describe the operating characteristics of the sector or subsector (e.g. fishery, market, fleet, community, institutional environment), aka the census or frame survey;

Decide

- Decide on the approach to be taken: complete enumeration or sampling, including cost/benefit and cost effectiveness analysis and an evaluation of operational considerations;

Design

- Design methods according to the approach adopted, including the form of stratification to be used in sampling;

Implement

- Implement a test phase to validate the method, including participation by other stakeholders;

Feedback

- Establish a continuing feedback mechanism between data sources and data users to ensure that data types, quantity, quality and origin are consistent with the requirements for determination of the performance indicator in question.

3. Census vs Sampling



- Data collection is **the recording of one or more data variables** (catch, value, fishing duration, etc.) from members of a population of “data-units” (the population of fishing vessels, fishers, etc.).
- Two basic data collection approaches are possible:
 - by **complete enumeration**, where all members of the whole population are measured; e.g. CPUE, Prices, Vessel registers.
 - by **sampling**, where only a proportion of members of the whole population are measured e.g. species comp.

Basic Statistical Methods and Concepts Practice

❖ Lake Bosumtwi?

- Area 100 ha, 20 boats, 1 landing site, well trained staff and unlimited funds.
- How do you monitor this small lake to get most accurate results?

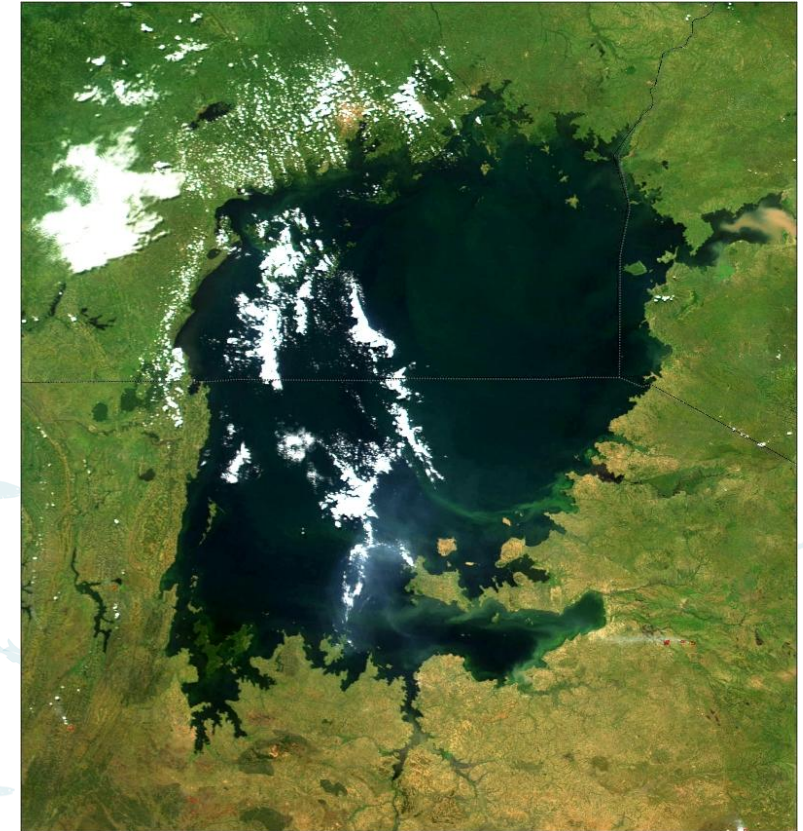
Plenary discussion



Basic Statistical Methods and Concepts Practice

❖ Lake Victoria?

- Area 35,000 km², 30,208 boats, 634 landing sites, limited staff and limited funds
- How would you monitor this big lake to get most accurate result?



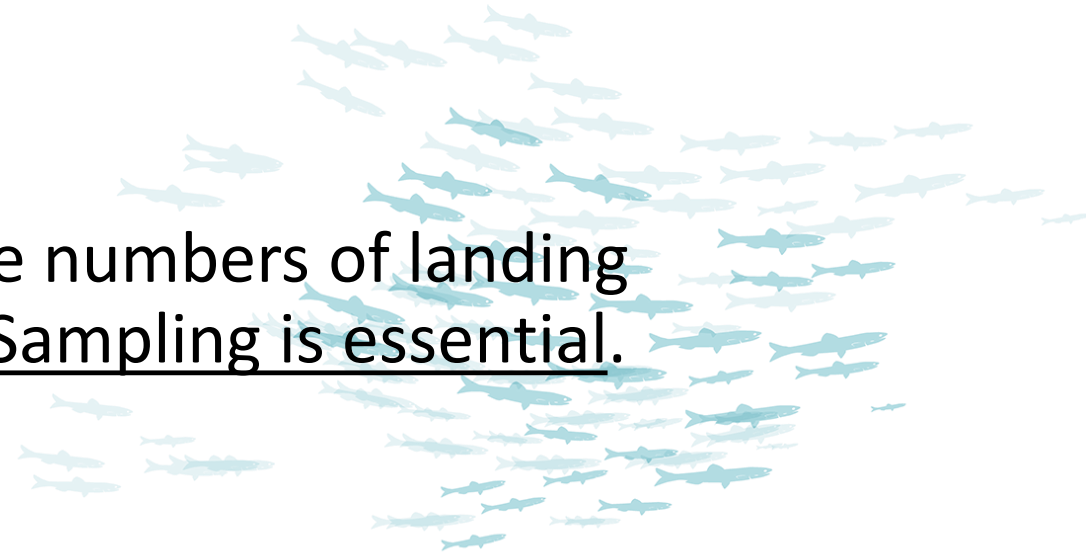
Basic Statistical Methods and Concepts Practice

Conclusion

- Lake Bosumtwi: Small area, limited number of boats, you can cover completely:

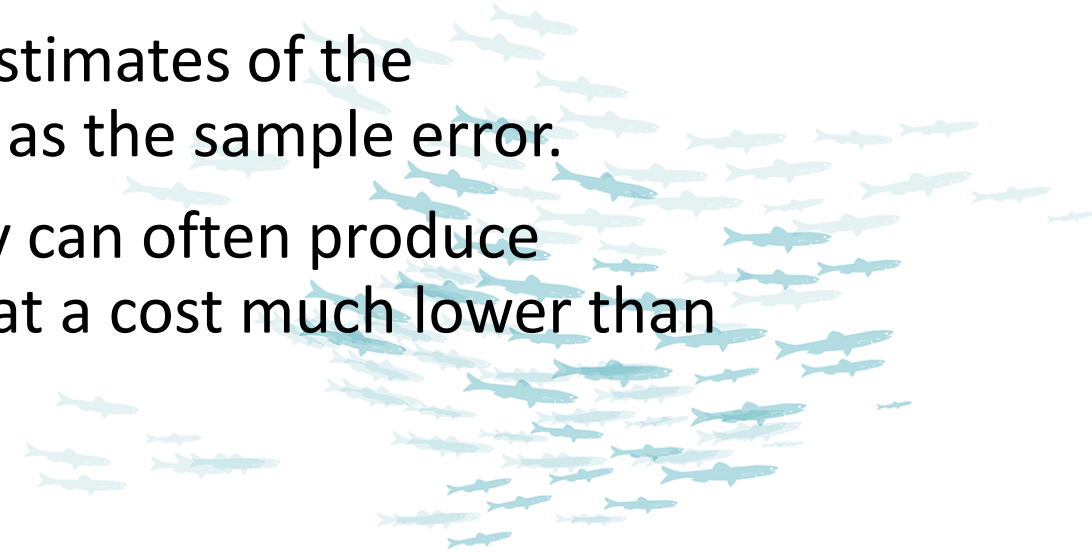
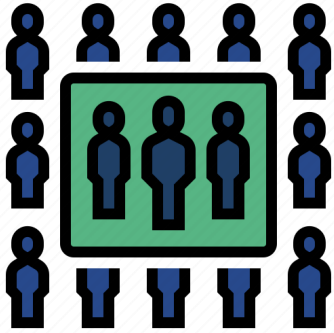
Full Enumeration or Census

- Lake Victoria: Large area, large numbers of landing sites, large number of boats: Sampling is essential.



4. Sample-based approaches

- We sample if it is impossible, difficult or expensive to observe all the elements of a target population.
- Sample surveys can operate on selected subsets of the **target population** and, using a number of assumptions regarding the distribution of the population.
- Sample based surveys provide estimates of the parameters under study, as well as the sample error.
- A well-designed sampling survey can often produce accurate and reliable estimates at a cost much lower than that of complete enumeration.



Objectives of sampling

- To obtain data from stocks and their exploitation (e.g. to find out what fish exist in the body of water being studied, roughly how many these are, of what species?)
- To analyse the characteristics of resources (e.g. what are the more striking characteristics (size, feeding habits, etc.) of each species?).
- To analyse the effects of exploitation on the abundance of resources, and
- To determine appropriate fishing levels to obtain best possible catches at present and during future years.



Challenges in sampling

❖ Major issues in sampling: **the representativity of sample**

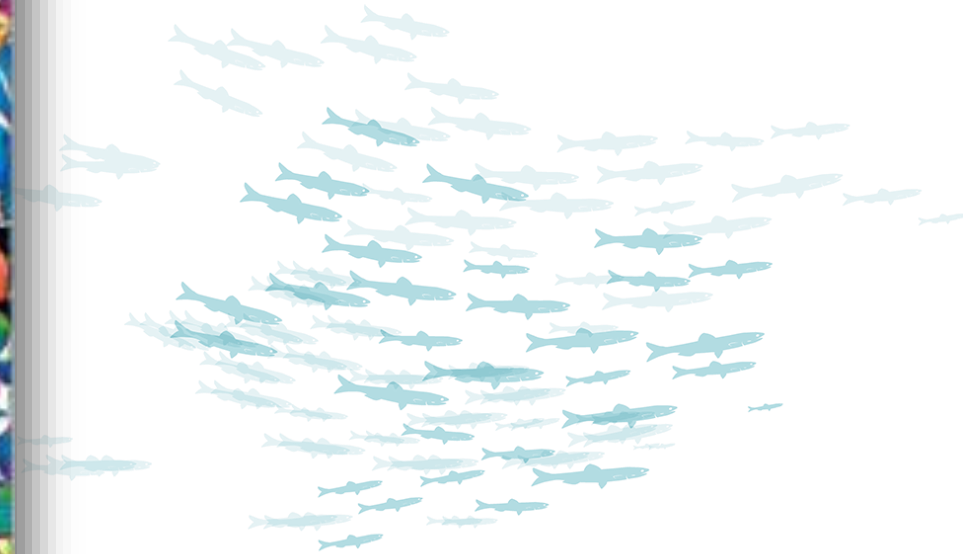
❖ Questions before sampling:

- What is sampling?
- Objective?
- Sample size
- Sampling frequency
- Sample methods
- Design/select sample sites/stratification



What is sampling?

- Sampling is a 'short-cut' method for investigating a whole group.



What is sampling?

Let's say it costs 1\$ to investigate one square of the puzzle

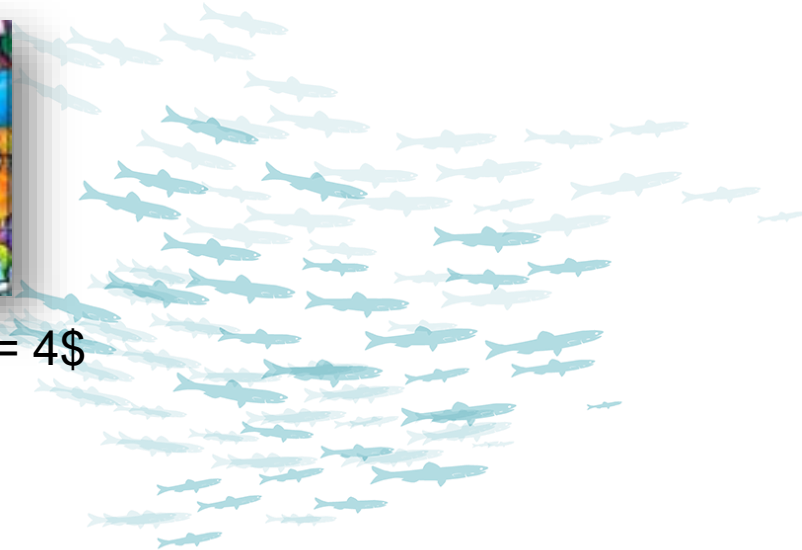


Total cost = 16\$



Cost of sample = 4\$

Good
estimate



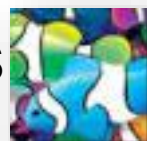
What is sampling?

Can we reduce the cost and still get an acceptable estimate?



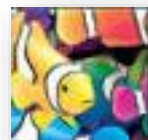
Total cost = 16\$

1\$



Biased
estimates

1\$



2\$

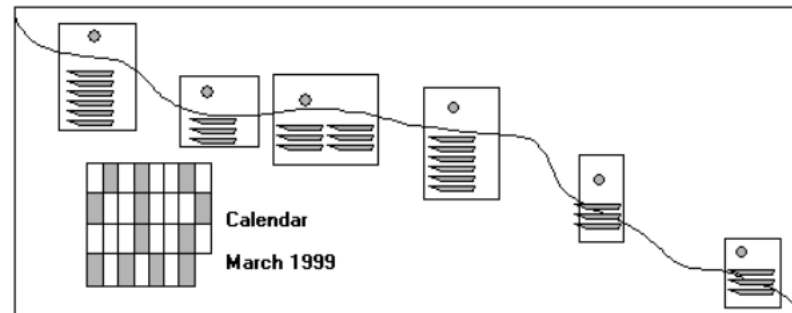


Combined
sample

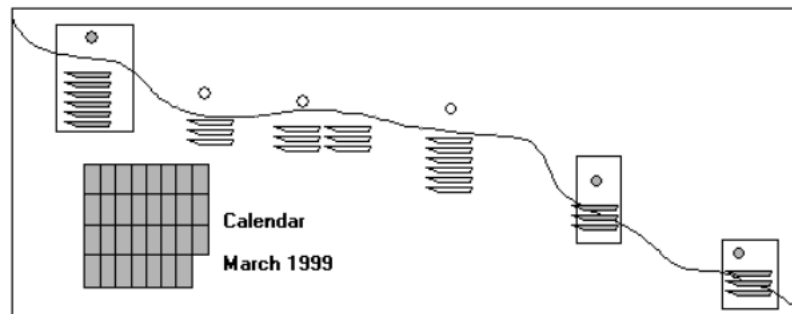
A good
enough
estimate?

Sampling options

- Sampling in time, census in space

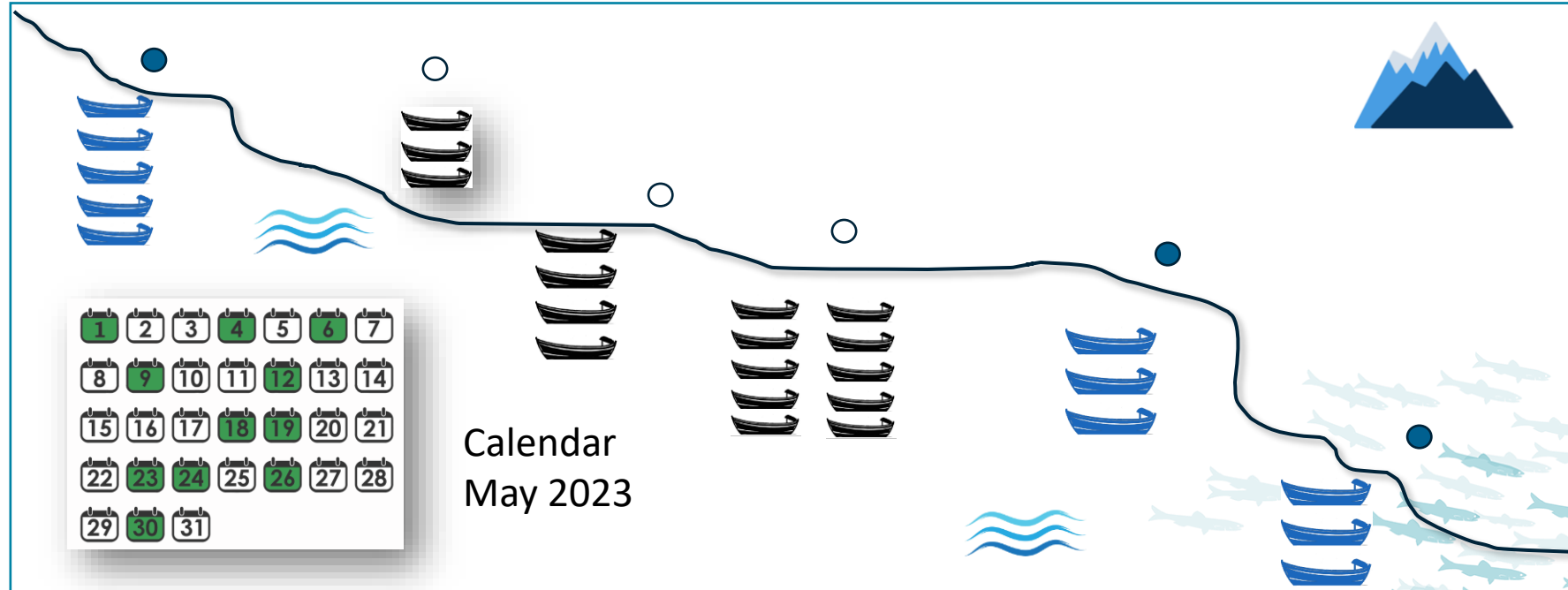


- Sampling in space, census in time



Sampling options

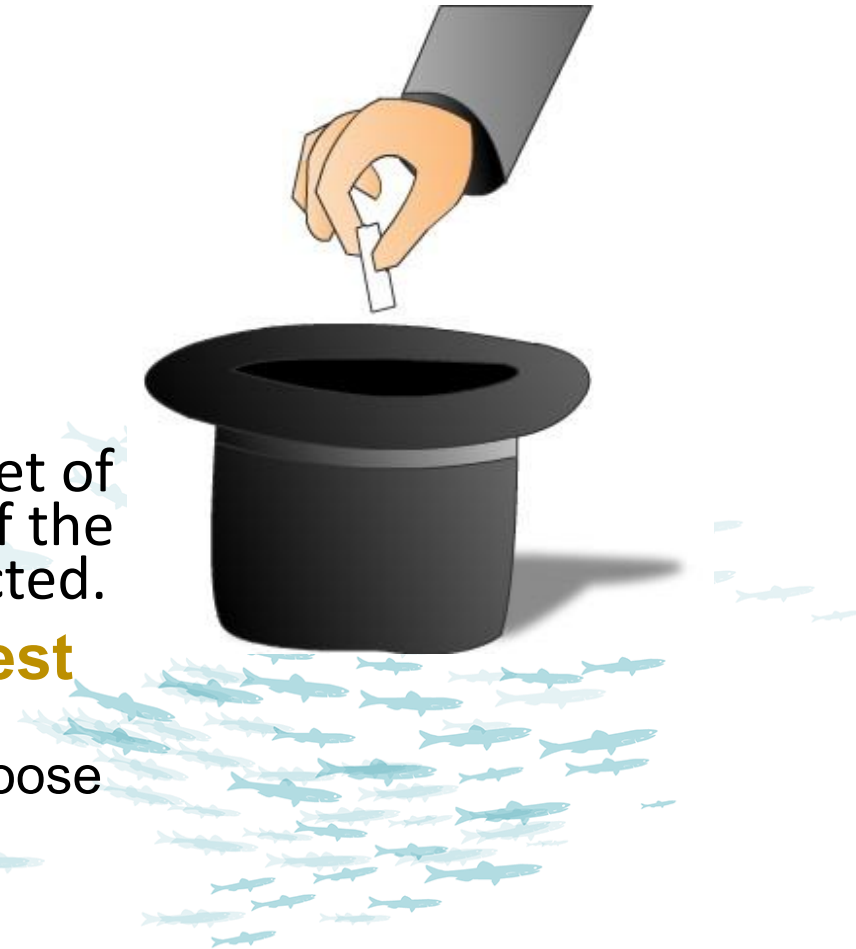
Sampling in Time and Space



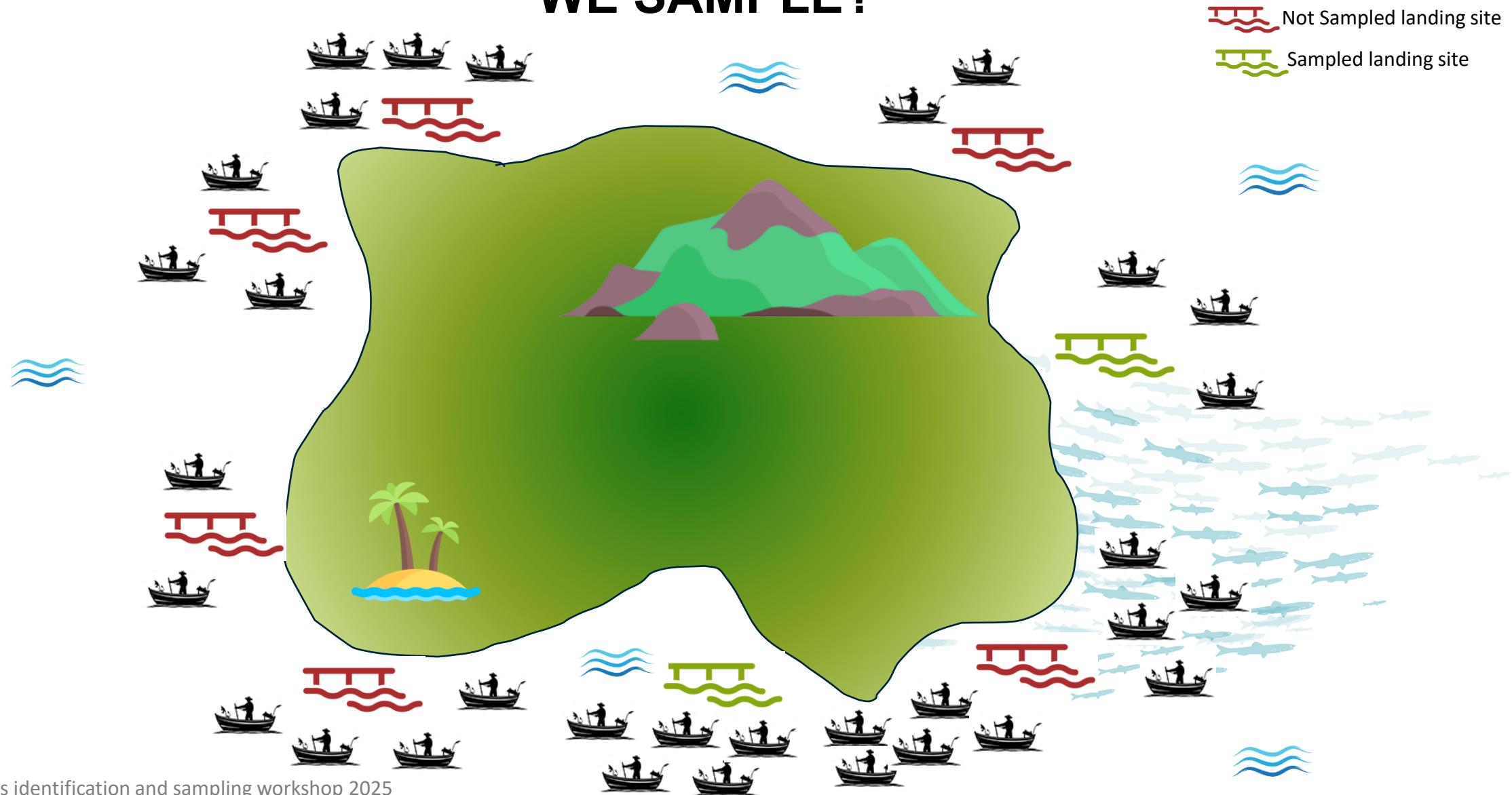
- Most commonly used
- At some sites, boats are sampled (**Space**)
- Sampling is done at certain days only (**Time**)

Why do we often speak about random sampling ?

- Simple Random Sampling
or
- Stratified Random Sampling
- Because **It matters greatly how the sampling is done**
- A **simple random sample** is a randomly selected subset of population. In this sampling method, each member of the population has an exactly equal chance of being selected.
- **Random sampling (i.e. not choosing) is the best scientific way to sample.**
 - It prevents investigator from using his judgement to choose sites, time, etc.
 - It allows to use most of the common statistical test



8 LANDING SITES, WHERE AND WHAT DO WE SAMPLE?



Random sampling

In this case, if we select at Random:

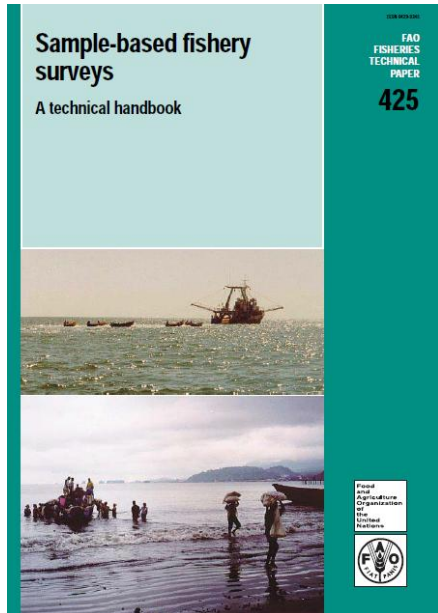
- Random picking of two landing sites -> high probability of only picking red sites
- It can work for the boats but absolutely not for landing sites

➔ So we use systematic site selection through stratification

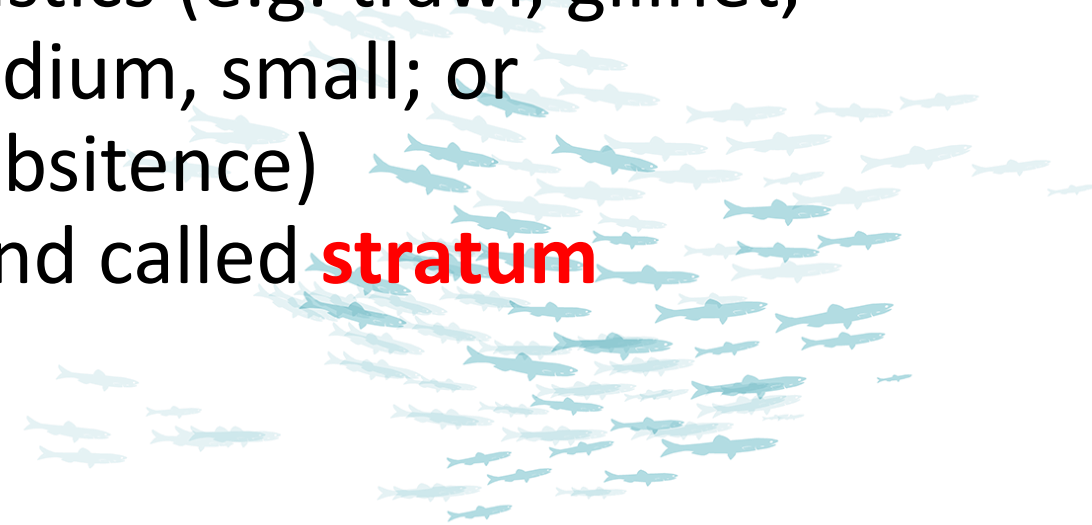


5. Stratification

What is stratified sampling?

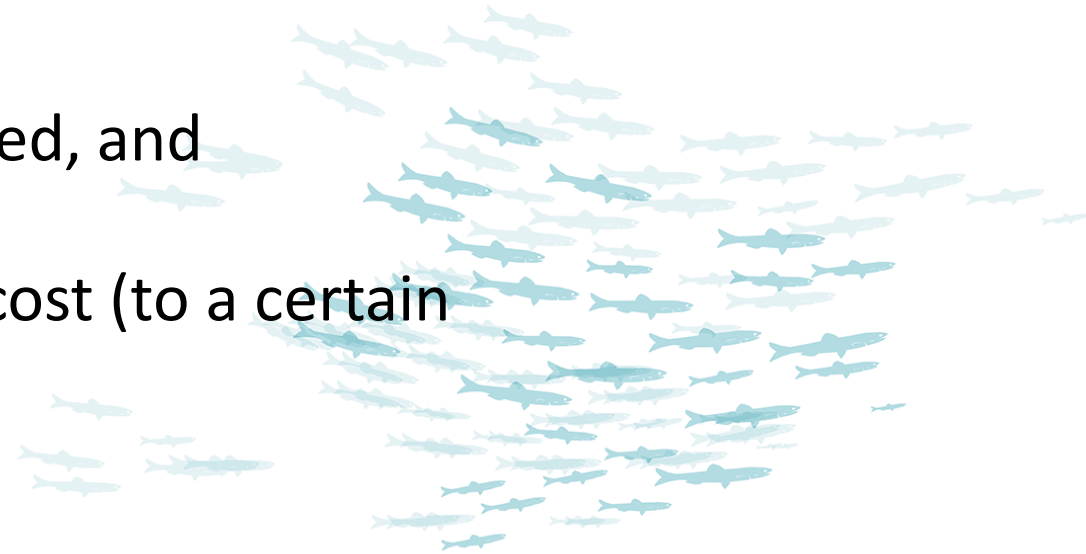


Stratification is the process of **partitioning** a target data population (e.g. all fishing vessels) into a number of more **homogeneous sub-sets** based on their characteristics (e.g. trawl, gillnet, purse-seine; or large, medium, small; or commercial, artisanal, subsistence) (Stamatopoulos, 2002) and called **stratum**

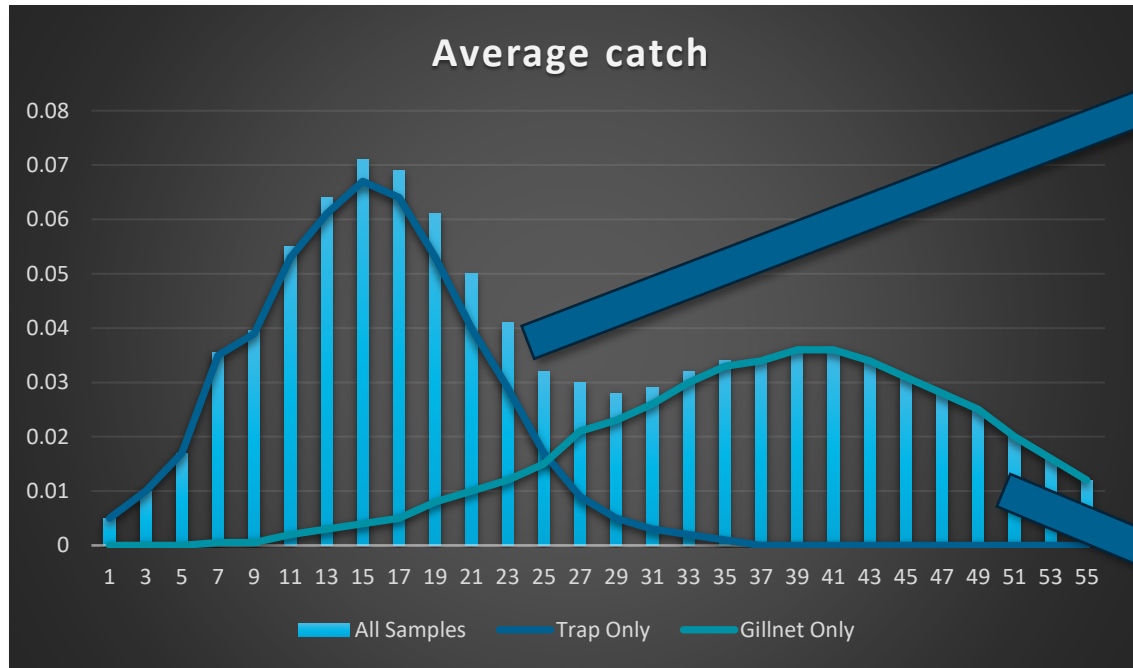


Why using stratified sampling?

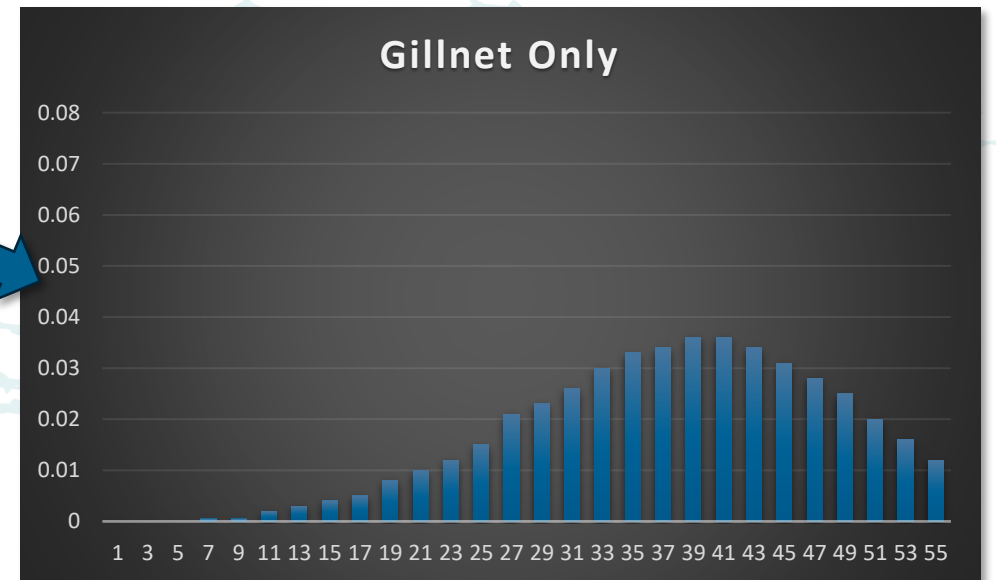
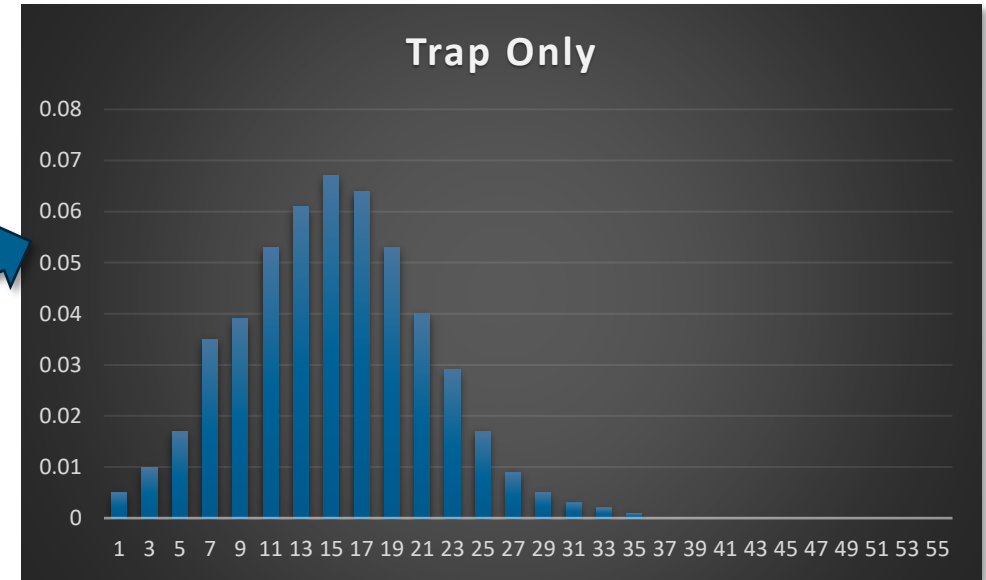
- Stratification aims to reduce the variability in each. Strata of the sampled data, which will:
 - Improve the reliability of the collected data: CPUE, total catch, species composition
 - Reduce the number of samples needed, and
 - Consequently, reduce the sampling cost (to a certain point)
 - Save money
 - Save resources and personnel



Why using stratified sampling?

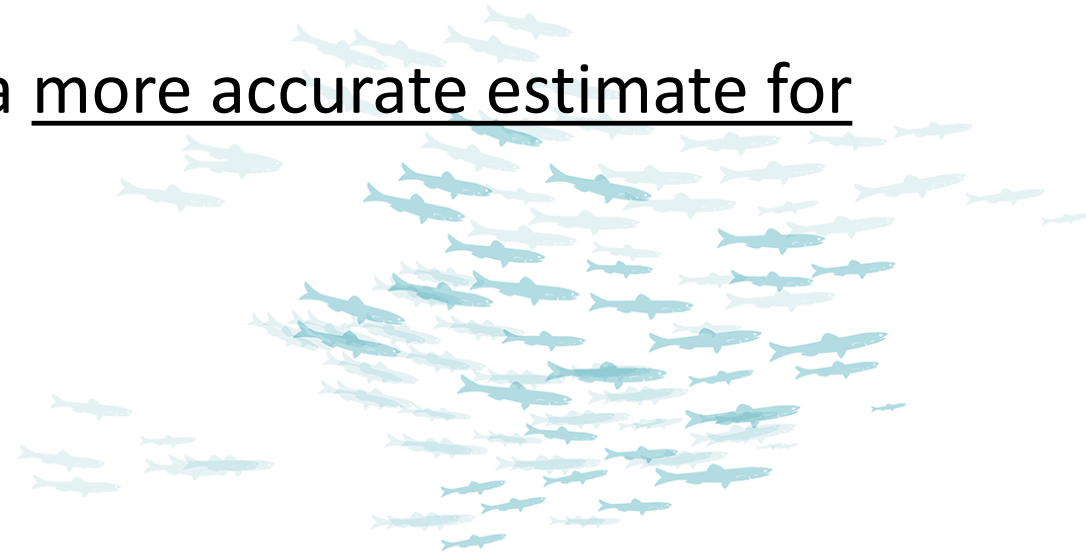


Combined sampling of traps and gill nets
= Large variation !



Why this stratification?

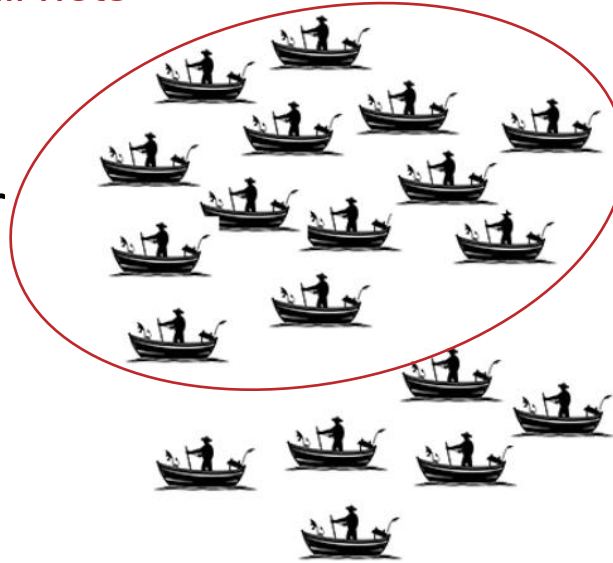
- Even if you take a lot of samples from Traps and Gillnets combined you still wouldn't get accurate estimates
- By separating Traps and Gillnets you get a more accurate estimate for less costs (smaller samples needed)



Structure of sampling

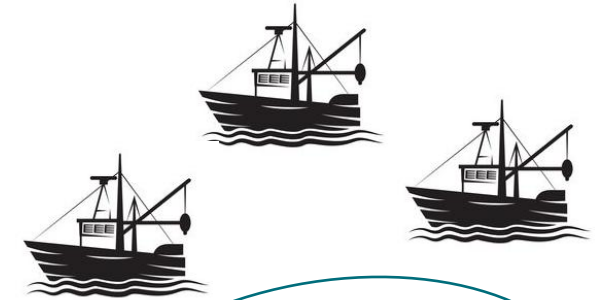
Gill nets

- 1. Similar boats and similar gear are already grouped and counted separately in Frame survey



- 2. **Samples** are taken from all gear and boat groups, so that all fish species and all fish sizes have a chance to be included

Gill nets



Trawl

Stratification for fisheries

Stratification

- **Major strata** – are done for Administrative classification, e.g. Region, District
- **Minor strata** – to improve sampling for higher accuracy and to reduce costs
 - Ecological/Geographical/seasonal strata
- **Fishing units** – minimal stratification level

=> Knowledge of the fisheries helps to build a good stratification

Fishing unit definition

- Fishing unit definition is a key point of sampling design
- That the minimum (and often the only) stratification level
- Group of vessel that have the same fishing activity
 - Boat type: size, number , characteristics, fishing trip characteristics
 - Gear type, number
 - Targeted species
- **Knowledge of the fisheries helps to build a good stratification**
 - Number and type of ports, landing sites, docks
 - Number of vessels per size category, gear type, fishing strategy

} Frame survey or
reliable registries

“The vessel register is the backbone of fisheries statistics. The creation and maintenance of a vessel register has first priority among the different fisheries data”, (de Graaf, G.J. et al., 2015)

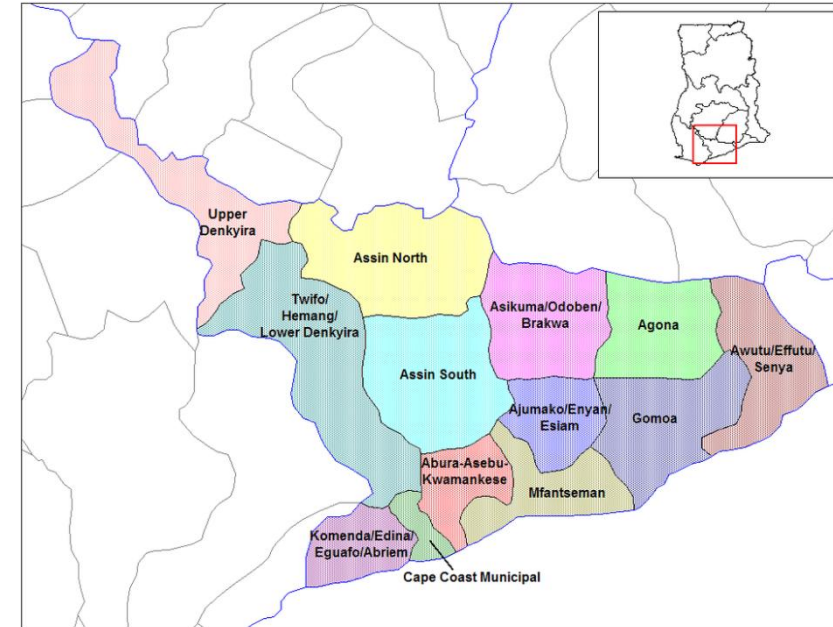
Stratification - Example

Example

Major stratum: Gomoa District

Minor strata: River Tana

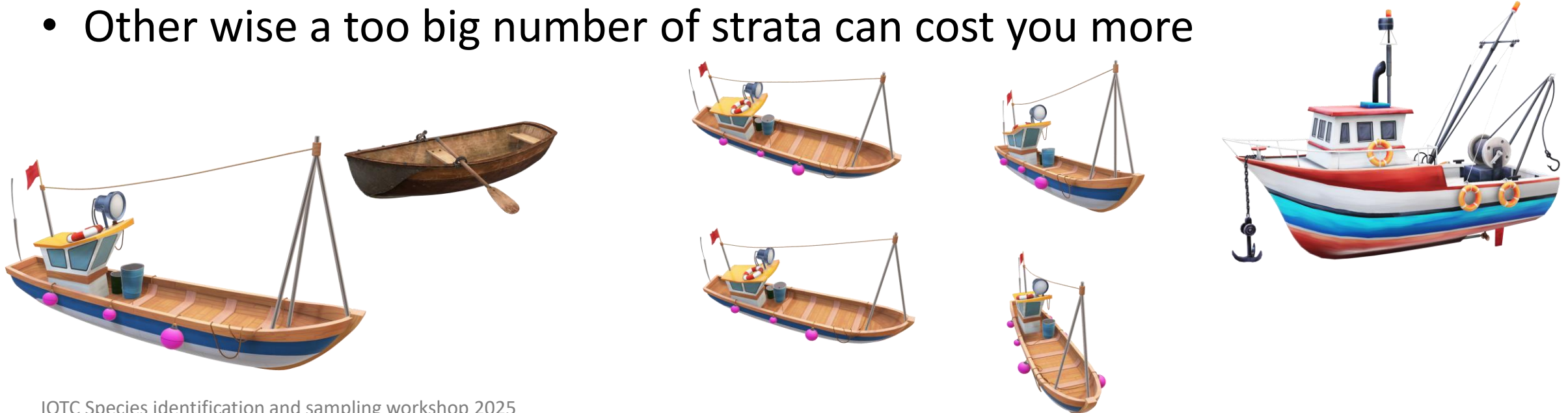
Fishing units: Boats with traps
Boats with gillnets



- When **sampling** -> raw data must be **raised/extrapolated** to produce estimates at the country level
 - Unbiased sample are crucial
 - Raw data quality must be verified
- Extrapolation are done **by stratum**

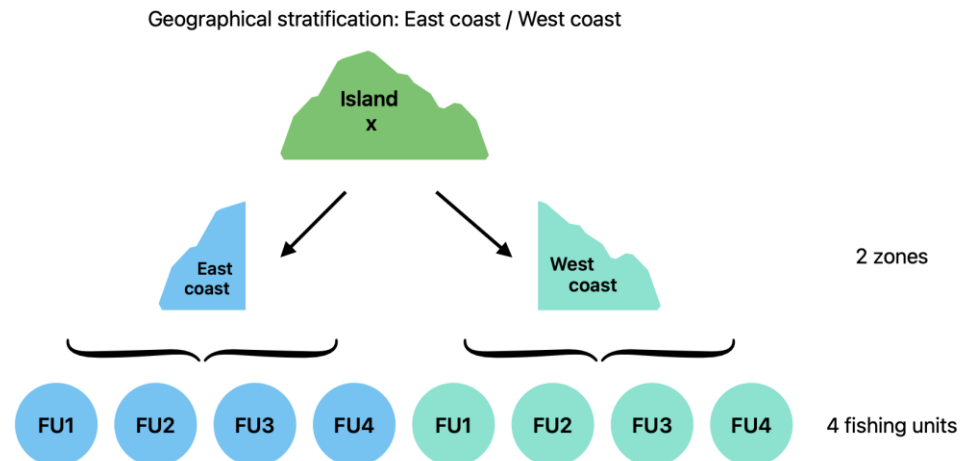
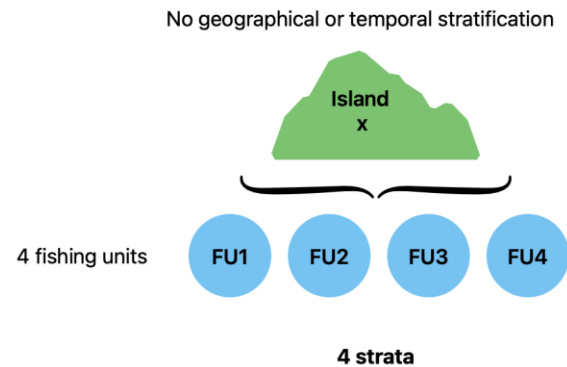
Warning

- Stratification sometimes may be complicated by the need to reconcile two conflicting objectives
 1. To select strata with the maximum degree of homogeneity
 2. To minimise the number of strata (usually in view of operational constraints)
- Other wise a too big number of strata can cost you more



Warning

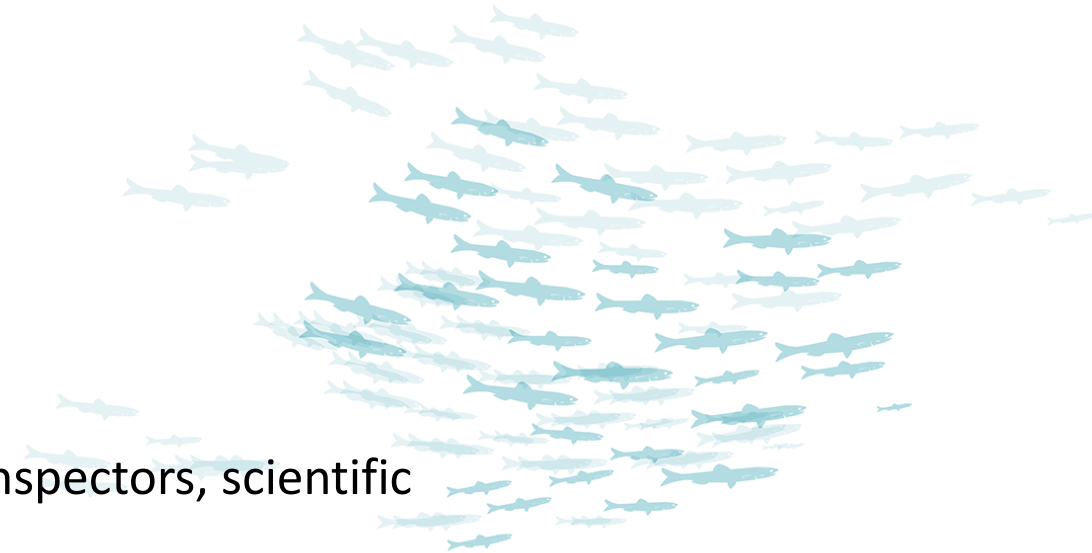
- Why? Because:
 - The estimates are produced by stratum
 - Each stratum must be sampled with a minimal sample size
- So, each time you add a stratification level, you increase the number of strata to sample



=> Find the good
compromise
between
accuracy and cost

6. Data Collection Methods

- Choice of collection method influenced by:
 - data collection strategy
 - type of variable (collection in routine, *ad hoc* survey)
 - accuracy required
 - collection point
 - type of fisheries activity
 - skill of the enumerator
- Main data collection methods are:
 - Registration
 - Questionnaires
 - Interviews – open ended, structured
 - Direct observation – enumerators, observers, inspectors, scientific research, data logging
 - Reporting – harvest, post harvest, logbook, self report





Practice

- Have you already designed a stratification plan in your country?
 - What are your stratification criteria
- Do you find it operational?
- Have you ever tested its statistical robustness?



Food and Agriculture
Organization of the
United Nations

Thank you ▪ Merci
Благодарю ▪ ¡Muchas gracias!
謝謝 ▪ شكرا

Yann Laurent

✉ yann.laurent@fao.org

