9th National Convention on Statistics (NCS)
EDSA Shangri-la Hotel
1 Garden Way, Ortigas Center, Mandaluyong City
October 4-5, 2004

Towards the Development of an Aquaculture Master Sample

by

Reynaldo Q. Vallesteros, Jr.

For additional information, please contact:

Author's Name : Reynaldo Q. Vallesteros, Jr.
Designation : Chief, Survey Planning and Programming Section
Affiliation : Bureau of Agricultural Statistics
Address : 4/F Ben-Lor Bldg. 1184 Quezon Avenue, Quezon City, Philippines
Telefax : (632) 3712075
Cell : 0918 5248382
Net : smrdrqv@yahoo.com

Towards the Development of an Aquaculture Master Sample

by
Reynaldo Q. Vallesteros, Jr.

ABSTRACT

The importance of master sampling frame in generating accurate and timely production survey estimates amidst resource constraints has been recognized. Initially, the list of aquaculture operators from the quick count results of the 2002 Census of Fisheries was compared to the aquaculture farms inventory list to ascertain whether the latter is suitable as a frame. A prototype sampling frame, master sample, and production survey designs for aquaculture is developed. The paper illustrates the feasibility of employing the master sample design to aquaculture production surveys that can produce quality fishery statistics in Central Luzon.

I. Introduction

Since 1990, the Department of Agriculture-Bureau of Fisheries and Aquatic Resources (BFAR) has been extending project fund to sustain various fisheries activities of the BAS. Indeed, the BFAR Regional Field Unit III and BAS entered into a Memorandum of Agreement on June 11, 2002 to implement the project "Strengthening the Fisheries Database through Aquaculture Farm Inventory and Profiling in Central Luzon". The project aims primarily to strengthen the fisheries database to effectively support the development and management of the fisheries sector in Central Luzon.

This study investigates the development of an aquaculture master sample (AMS) using the results of 2002 Aquaculture Farms Inventory in Central Luzon. Specifically, the study aims to (1) design and construct a prototype master sampling frame for aquaculture surveys; (2) analyze the characteristics of fresh water farms, brackish water farms, marine water aquaculture farms, and mussel and oyster farms; (3) identify auxiliary variables needed in the stratification and estimation; and (4) to simulate the prototype AMS design.

The aquaculture master sample is perceived to gain the following significant benefits:(1) costs savings in frame updating and maintenance; (2) samples for individual aquaculture surveys can be selected more quickly and economically; (3) improvement of the reliability of aquaculture survey results from different rounds of the program; and (4) possibilities for integration and substantive linkages between surveys and survey rounds.

The study covered six (6) provinces of Central Luzon (Region 3) namely, Bataan, Bulacan, Nueva Ecija, Pampanga, Tarlac, and Zambales.

In estimating the regional as well as the provincial level, the BAS tried to develop an aquaculture master sampling design that employs regional domain

and generates provincial estimates though of lesser precision by implicit stratification.

II. Sources of Data and Data Format

The main data source is the results of 2002 Aquaculture Farms Inventory in Region 3. The data compiled in MS Excel format is converted into STATA v.8 file for data management and statistical analysis.

For every aquaculture operator, the CDCs gathered information such as follows: Demographic characteristics, Type of aquaculture farm, Ownership of Aquaculture farm, Area of Aquaculture farm, and Location of Aquaculture farm (inside or outside the barangay)

Another data source is the Philippine National Statistics Office that provides the Enumeration Areas Reference File (EARF) and results of Quick Count for Region 3. The data files for Region 3 contained the names and the number of fishing and aquaculture operators per listed sample barangay in the 2002 Census of Fisheries. Since the focus of the study is to develop a master sample for aquaculture, the 2002 CF database is used as auxiliary information to validate if the prototype master sampling frame is complete, accurate, and current.

III. Conceptual Framework

An **Aquaculture Master Sampling Frame (MSF)** is a single unified sampling frame of all aquafarm operators engaged in domestic production in the freshwater, brackishwater, and marinewater farms including mussels and oysters farms. For this study, MSF is the consolidation of the four aquafarm frames into only one master frame which is generated from the 2002 Aquaculture Farms Inventory.

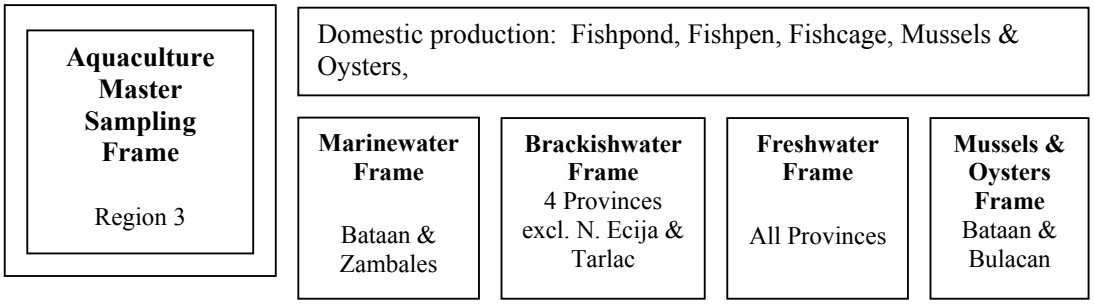

Figure 1. Composition of an Aquaculture MSF

In relation to the survey target population and the constructed **aquaculture MSF,** the **aquaculture master sample (MS)** is defined as a large sample of aquafarm from which sub-samples can be selected to serve the needs of ad-hoc Cost and Returns Surveys for Fishery and of the three aquaculture

major surveys – the Survey of Mariculture Farms, Survey of Freshwater Aquaculture Farms, and Survey of Fishponds.

An **aquaculture MS design** encompasses the generation of master sampling frame, the construction of a master sample, as well as, the selection of sub-samples to serve the needs of more than one survey or survey rounds either in a continuous regular interval or in ad-hoc manner. For this purpose, the full master sample is intended to be used for production survey of freshwater fishpond, freshwater fishcage, brackishwater fishpond, brackishwater fishpen, brackishwater fishcage, marine fishpen, marine fishcage and mussel and oyster farms. The sub-samples can be generated to serve smaller but detailed specialized surveys such as Cost and Returns Survey for tilapia, milkfish, etc.

IV. Results and Discussion

A. The design of an Aquaculture master sampling frame

An aquaculture master sampling frame (MSF) is conceptualized and defined as a single unified sampling frame of all aquafarm operators engaged in domestic production in the freshwater, brackishwater, and marinewater farms including mussels and oysters farms. For this study, MSF is the consolidation of the four aquafarm frames into only one master frame which is generated from the 2002 Aquaculture Farm Inventory.

The Universe. Of the total fish production, aquaculture accounted for 40 percent while commercial and municipal fisheries contributed shares of 31 percent and 29 percent, respectively. The production from aquaculture registered a growth rate of 9.65 percent in 2002 as compared in 2001. The major contributors in the hefty increase in aquaculture came from freshwater and brackish water fishpond.

A total of 16,672 aquafarm operators were listed in Central Luzon. The largest number at 5,140 was recorded in Pampanga while the smallest at 1,027 was listed in Zambales. About 71 percent of the aquafarm operator was engaged in freshwater farming.

Of the total number of aquafarm operators, about 87 percent of which is engaged in fishpond, 6% in small farm reservoir, and 4% in ricefish culture. Others remaining were engaged in fishcage, fishpen, mussels and oysters, and hatcheries. Mussel farmers were found only in Bataan with 631 operators while Bulacan was the sole province with oyster farm operation with 214 producers listed.

About 64 percent of the aquafarm operators was engaged to freshwater fishpond operation; 24 percent in brackishwater fishpen.

The ratio of total aquafarm operators with the total number of aqua farms is almost equal to 1:1 simply because only few aquafarm operators have two or more types of aquafarm. Though minimal in number, this frame characteristic should be taken care of. Region 3 has 16,895 aquafarms operated by 16,672

operators, so the average number of aquafarm per operator is only 1.013376. It would be appropriate to consider aquafarms as the USUs rather than the operators. Aquafarms are more stable overtime than the operators. In this study, operators are used as the USUs because of the limitations on the file structure of the aquaculture databases in which the aquafarms id were not readily sorted and linked to operator's id.

B. Construction of an Aquaculture master sampling frame.

The Aquaculture MSF. In this research, the survey target population is limited to aquafarm operators who are engaged in food production and at the same time those that are closely monitored by the BAS in generating "official" statistics for fishery sector. Government-owned aquafarm, freshwater and brackishwater hatcheries, rice fish, and small farm reservoir were excluded from the master sampling frame.

The results of 2002 Aquaculture Farm Inventory show that the average number of operator per barangay is 9.5 with a minimum and maximum value of 1 and 110 operators, respectively (Table 1).

Table 1. Summary statistics of the number of operators, aquafarms, barangays, municipalities, by province: Region 3, July 2002

		Total				Number of operator per barangay		
PC	Province	Mun	Brgy	Aquafarm	Operator	Mean	Min	Max
08	Bataan	12	111	1,501	1,487	13.4	1	101
14	Bulacan	23	245	2,511	2,426	9.9	1	76
49	Nueva Ecija	32	410	2,557	2,557	6.2	1	35
54	Pampanga	22	370	5,107	5,045	13.6	1	110
69	Tarlac	18	348	2,989	2,985	8.6	1	85
71	Zambales	13	153	1,044	1,010	6.6	1	39
	Region 3	120	1,637	15,709	15,510	9.5	1	110

Because of time constraints, limited field resources, and unavailability of mapping software, the researcher could not able to pursue the formation of PSUs (merging and splitting) to meet the minimum size requirement. In reference to the survey of related literature, the national surveys minimum size requirement of PSUs (barangays) intended for aquaculture surveys might be set to at least 100 farm operators to accommodate rotation of samples. Result shows that this is not attainable; for fresh water alone, the average operator per barangay is 7.81 (Table 2).

Perhaps, municipality might be used as the PSU. The average number of operators per municipality is 129, which is close to the appropriate sample size of the PSU (100 operators). According to NSO (2003), though there are 1,513 municipalities in the country and 122 for Region 3, still it is too few as PSU for a national survey.

Table 2. MS Frame. Summary statistics of the number of operators, aquafarms, barangays, municipalities, by environment: Region 3, March 2004

Type of Environment	Total					Number of operator per barangay			
	Prov	Mun	Bgy	Farm	Operator	Mean	Std	Min	Max
Marine water	2	4	20	41	41	2.05	2.60	1	12
Mussels & Oysters	2	12	53	846	845	15.94	20.45	1	85
Brackish water	4	35	822	4,118	3,976	12.34	16.99	1	110
Freshwater	6	114	1,364	10,704	10,648	7.81	10.07	1	99

Table 3. Characteristics of the PSUs: Municipality

Province	Number of Municipality	Number of operator per municipality		
		Mean	Min	Max
Bataan	12	123.9	8	303
Bulacan	23	105.5	7	555
Nueva Ecija	32	79.9	5	262
Pampanga	22	229.3	32	805
Tarlac	18	165.8	28	510
Zambales	13	77.69	5	168
TOTAL	120	129.3	5	805

C. The design of an Aquaculture master sample

The prototype aquaculture MS. The Aquaculture MSF, which was discussed in section 3.1, provided the frame for the selection of the Aquaculture MS. The design is quite different from that of 2003 MS simply because it employs four modules using four mutually exclusive frames – Marinewater, Brackishwater, Freshwater, and Mussel & Oyster Farms. The basic frame unit in the MSF could in some cases be used as the PSU for the master sample.

The main steps in the design of an Aquaculture master sample are as follows:

Sample Size. First step is to decide what would be the sampling domain. For this purpose, the region is the domain. Result, as shown in Appendix A, shows that the mean of the PSUs between provinces, type of aqua farm, environment, status of ownership are significantly different except for the following cases: (1) total physical area per municipality and per barangay for type of aquafarm and ownership status, and (2) number of operator per municipality for status of ownership. To address this issue on heterogeneity of the PSUs, systematic sampling (implicit stratification) has been employed.

Sampling Design. Second step is to decide on the most appropriate sampling schemes for the four different sampling frames. The criterion used was the sampling scheme that yields the lowest CV and optimum sample size is the best alternative.

For marinewater. (Fishpen and Fishcage). The marinewater fishpen and fishcage was considered as rare commodity. Since there were only 1 marine fishpen and 40 marine fishcage listed in the provinces of Bataan (1 operator) and Zambales (40), complete enumeration or all units of the Marinewater farms were taken with certainty as one stratum.

For brackishwater farms and mussels & oyster farms. The sampling scheme used is systematic (equal weights) sampling of the PSUs. We did simulation of pre-determined sample sizes, that is, 5%, 10%. 15%, 20%, and 25% of the total sampling domain.

Result shows that at 10% samples (398) of brackishwater operators could yield a CV of 8.89% at the regional level (Table 4). It is also probable to yield acceptable estimates even at the provincial level, though of lesser precision (less than 20% CV).

Table 4. Brackishwater Coefficient of Variation

Sampling (N = 3976)	Sample Size	Mean Squared Error (MSE)	Coeff. Of Variation (CV) in %
5%	199	13,468,620	15.07
10%	398	4,686,126	8.89
15%	596	1,895,998	5.65
20%	795	3,658,213	7.85
25%	994	1,506,043	5.04

The survey on mussels and oysters at 15% sampling scheme could yield acceptable estimates (10.80% CV) at the regional level and about 20% CV at the provincial level (Table 5).

Table 5. Mussels & Oysters Coefficient of Variation

Sampling (N = 845)	Sample Size	Mean Squared Error (MSE)	Coef. Of Variation (CV) in %
5%	42	376,751	16.30
10%	85	189,725	11.56
15%	127	165,468	10.80
20%	169	174,594	11.09
25%	211	34,862	4.96

For freshwater farms. The sampling scheme used is a two-stage systematic (equal weights) sampling of the PSUs with implicit stratification. The simulated sample sizes of PSU are 5%, 10%. 15%, 20% and 25% of the regional total number of aquafarm barangays. Within a sample PSUs, the three approaches planned to be simulated for selection of SSUs are as follows: (1) complete enumeration, (2) 50% SRS, and (3) 25% SRS. However, due to time constraint, approaches 2 and 3 were not pursued. These will be part of the Future Directions in section IV.

Table 6 shows that the freshwater production survey with a 20% samples (273 barangays, and 100% operators in the sample barangay) yield a 10.89% CV at the regional level. Generally, provincial estimates could be generated with a lesser precision of about 20% CV.

Table 6. Freshwater: Coefficient of Variation
SYSTEMATIC (equal weight)

Sampling (N = 1364)	Sample Size	Mean Squared Error (MSE)	Coeff. of Variation (CV) in %
5%	68	48,733,178	22.85
10%	136	25,960,017	16.68
15%	205	14,888,283	12.63
20%	273	11,057,898	10.89
25%	341	10,295,514	10.50

D. The precision of an aquaculture master sample design

Table 7 and 8 show the changes in the number of sample sizes of operators, barangays and municipalities. These figures show a significant decrease in mean and maximum values. In this case, better sampling schemes can be looked at such as: probability proportional to size (pps), multistage stratified systematic sampling, and the one recommended by Pettersson, i.e., to order the PSUs within strata according to a stratification criterion and to select the sample systematically (implicit stratification). The best sampling scheme that may yield most precise estimates could be determined if further simulation will be pursued.

Table 7. MS Frame. Summary statistics of the number of operators, aquafarms, barangays, municipalities, by environment: Region 3

Type of Environment	Total					Number of operator per barangay			
	Prov	Mun	Bgy	Farm	Operator	Mean	Std	Min	Max
Marine water	2	4	20	41	41	2.05	2.60	1	12
Mussels& Oysters	2	12	53	846	845	15.94	20.45	1	85
Brackish water	4	35	822	4118	3,976	12.34	16.99	1	110
Freshwater	6	114	1364	10704	10,648	7.81	10.07	1	99
TOTAL			2259		15,510				

The number of operators for mussels and oysters farms in the MS (mean: 3.53 vs 15.94 and maximum value: 14 vs 85) were absolutely different with that of the MS frame. Even if the number of master sample barangays is about 22.84% of the survey target population (MSF) (516/2,259), other statistics such as deff and actual CV will be determined as part of the future directions. The quality of estimated statistics is of utmost priority while cost-savings gained from small sample sizes is nonetheless of second priority.

Table 8. Master Sample. Summary statistics of the number of operators, aquafarms, Barangays, municipalities, by environment: Region 3

Type of Environment	Total					Number of operator per barangay			
	Prov	Mun	Bgy	Farm	Operator	Mean	Std	Min	Max
Marine water	2	4	20	41	41	2.05	2.60	1	12
Mussels& Oysters	2	11	36		127	3.52	3.51	1	14
Brackish water	4	32	187		398	2.13	2.02	1	13
Freshwater	6	105	273		1,658	6.07	6.42	1	40
TOTAL			516		2,224				

There will be a tremendous savings in taking a surveys based on master sample approach rather than conducting an annual inventory if the master sample is properly designed and maintained. However, master sample can be cost-effective and efficient if the sampled geographic areas will remain stable for a long period of time. It is hypothesized that characteristics of fishery barangays (topography and total physical area devoted to aquaculture farming) in the country as a whole or even within a region did not abruptly change as compared to characteristics of aquaculture operators who may changed occupational status instantly.

V. Future Directions

The important topics that must be addressed by concerned BAS statisticians are the following:

1. Studies on the choice of sampling units at various levels (farm vs. operator, barangay vs. municipalities, etc);
2. Studies on other sampling schemes;
3. Splitting and/or merging of basic sampling units;
4. Sample Size and Sample Allocation needs further studies.
5. Rotation of samples - Explore practical methods on rotation of samples
6. Frequency of collection - It should be determined based on historical data; the seasonality of production should be studied to schedule the data collection: monthly, quarterly, biannual or annual.
7. Precision of estimates - Studies on weighting and estimation procedures; coefficient of variation, and design effect.
8. The use of auxiliary information or other available information to improve reliability of estimator.

VI. Summary, Conclusions and Recommendations

The survey designs for household in the areas of demographic, social and economic statistics are presumed applicable also in fishery statistics. The operational definitions, description and properties of master sampling frame and master sample were presented. The design and maintenance of a master sample scheme were tackled and the main steps were summarized as follows:

- Choice of PSUs (and possibly for SSUs) for the master sample
- Combining/Splitting area to reduce variations in PSU sizes
- Stratification of PSUs
- Allocation of the master sample to strata
- Sampling of PSUs, sample sizes
- Rotation of samples

In summary, the findings of this study are:

About the master sampling frame:

1. The list of operators and aquafarms in the 2002 Aquaculture Farm Inventory is a good source of the aquaculture master sampling frame.
2. The barangay serves as the primary sampling unit (PSU) when two-stage sampling scheme was employed in the survey of freshwater aquafarms.
3. Operators of marine fish pen and fish cage were taken as certainty PSU.
4. On the issue on the choice of PSUs, municipality was considered in lieu of barangay.
5. On the issue on the formation of PSUs, whether municipalities or barangays serve as the PSUs, study shows that the average number of aquafarm operators is significantly different among and between provinces, municipalities, and barangays. To address this issue, splitting and/or merging of PSUs is necessary to improve the precision of the survey estimates.

About the master sample:

1. The region is used as sampling domain. By implicit stratification, there is no need to determine the boundaries in order to generate provincial level estimates, though of lesser precision. Municipal and barangay level estimates could be approximated if weights are properly designed and properly maintained overtime.
2. For an integrated aquaculture surveys, a total of 516 sample barangays can be utilized for a full master sample. This master sample can represent the survey target population of about 2,259 barangays in the 6 provinces of Central Luzon.

In view of the foregoing, **it can be concluded that** an integrated aquaculture survey using master sample frame is feasible within a regional domain. It is also probable to yield acceptable estimates even at the provincial level for selected characteristics. It is important however, to pursue studies on the choice of sampling units at various levels, that is, aquafarm vs. operator, barangay vs. municipality, etc. Furthermore, after activities listed in Section 5.0 (Future Directions) have been accomplished, determination of best alternative sampling scheme should be pursued prior to conduct of pilot testing of an aquaculture production survey based on master sample approach.

Field resources, such as computer hardware and software including mapping application programs are significant tools for the efficient and effective use of an Aquaculture MS.

For the full implementation of a nationwide survey, institutional support particularly from PNSO and BFAR is the prime requirement for the development of a master sample for aquaculture farm households.

There are many ways of developing a master sample depending on the field resources (including logistics and manpower), technical expertise of the lead personnel involved, time frame, and most of all the right attitudes of the intended users toward how MS should be utilized. The following are some recommendations:

1. Good estimates rely on the goodness of the sampling frame. Thus, the master sample must be kept complete, accurate, and current at all times.

2. Computer hardware and software including proper training of the competent GIS oriented personnel are deemed necessary to facilitate the development of an aquaculture MS.

3. For a national survey, the 2002 Census of Fisheries databases should be properly evaluated and reviewed for use as the sampling frames for all the fishery surveys.

References

Barrios, Erniel B. 2003. *Statistical Methods Plan: Research and Development Agenda of Bureau of Agricultural Statistics*.

[No Author]. 1984. *Handbook of Household Surveys. Studies in Methods.* Series F No. 31 Department of International Economic and Social Affairs, Statistical Office, United Nations, New York.

[No Author]. 2002. *Enumerator's Manual: 2002 Census of Fisheries.* Philippine National Statistics Office, Manila, Philippines

[No Author]. 2003. *2003 Master Sample, ADBTA 3656-PHI: IMPROVING poverty Monitoring Surveys.* Philippine National Statistics Office, Manila, Philippines

[No Author]. 2002. *Manual of Operations: Aquaculture Farms Inventory.* A Joint Project of the Bureau of Fisheries and Aquatic Resources and the Bureau of Agricultural Statistics, Department of Agriculture.

Petterssen, Hans. 2003. *Chapter 5: The design of master sampling frames and master samples for household surveys in developing countries.* Household Surveys in Developing and Transition Countries: Design, Implementation and Analysis. United Nations Statistics Department.

Appendix A

ONE-WAY ANALYSIS OF VARIANCE

Hypothesis Testing:
To determine if the means of PSUs data set are equal within group

1. The means of total physical area per municipality are equal among provinces, type of aquafarm environment, and status of ownership.
2. The means of total physical area per barangay are equal among provinces, type of aquafarm environment, and status of ownership.
3. The means of total physical area per operator are equal among provinces, type of aquafarm environment, and status of ownership
4. The means of number of operator per municipality are equal among provinces, type of aquafarm environment, and status of ownership.
5. The means of number of operator per barangay are equal among provinces, type of aquafarm environment, and status of ownership.

Table 1. Testing of Means of Data Set within Group

	GROUP	PSU	F	P-value	Remark
1.	PROVINCE	AREA PER MUNICIPALITY	3.49	0.0056	**
		AREA PER BARANGAY	18.32	0.0000	**
		AREA PER OPERATOR	165.87	0.0000	**
		NO. OF OPERATOR / MUN.	5.14	0.0000	**
		NO. OF OPERATOR / BGY.	18.11	0.0000	**
2.	TYPE OF AQUAFARM	AREA PER MUNICIPALITY	1.33	0.2381	ns
		AREA PER BARANGAY	1.57	0.1380	ns
		NO. OF OPERATOR / MUN.	5.51	0.0000	**
		NO. OF OPERATOR / BGY.	7.46	0.0000	**
3.	ENVIRONMENT	AREA PER MUNICIPALITY	5.06	0.0023	**
		AREA PER BARANGAY	47.43	0.0000	**
		NO. OF OPERATOR / MUN.	1.42	0.2388	ns
		NO. OF OPERATOR / BGY.	17.51	0.0000	**
4.	OWNERSHIP STATUS	AREA PER MUNICIPALITY	2.11	0.0999	ns
		AREA PER BARANGAY	1.07	0.3592	ns
		NO. OF OPERATOR / MUN.	19.05	0.0000	**
		NO. OF OPERATOR / BGY.	7.69	0.0000	**

SOURCE: BAS, 2002 AQUACULTURE FARMS INVENTORY

	GROUP	DATA SET	F	P-value	Remark
1.	PROVINCE	NO. OF OPERATOR / MUN.	6.78	0.0000	**
		NO. OF OPERATOR / BGY.	4.43	0.0006	**
		NO. OF OPERATOR / EA	5.53	0.0001	**
2.	TYPE OF AQUAFARM	NO. OF OPERATOR / MUN.	4.60	0.0000	**
		NO. OF OPERATOR / BGY.	2.47	0.0119	**
		NO. OF OPERATOR / EA	2.01	0.0422	*

SOURCE: NSO, 2002 CENSUS OF FISHERIES, Form 32A Quick Count Results

Note: ** - highly significantly different; * - significantly different;
ns - not significant