Fishery Statistical Metodology: Onboard Fishing[1]

Guilherme Guimarães Moreira (Guilherme Moreira)

Brazilian Institute of Geography and Statistics, Department of Methods and Quality, guilherme.moreira@ibge.gov.br

Marcos Paulo Soares de Freitas (Marcos Freitas)

Brazilian Institute of Geography and Statistics, Department of Methods and Quality, marcos.freitas@ibge.gov.br

Antonio José Ribeiro Dias (Antônio Dias)

Brazilian Institute of Geography and Statistics, Department of Methods and Quality, vermelho@ibge.gov.br

Aristides Pereira Lima Green (Aristides Green)

Brazilian Institute of Geography and Statistics, Department of Agriculture, aristides.lima@ibge.gov.br

ABSTRACT

In 2008, the Special Secretariat of Aquaculture and Fisheries (SEAP), nowadays Ministry of Fisheries and Aquaculture (MPA), created a working group to discuss the fishery statistics conducted by ESTATPESCA (IBAMA, 1995 and ARAGÃO, 2006), a statistical monitoring program implemented by the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA), and by other state agencies.

After 4 years of discussions, simulations and pilot surveys, this group produced a technical report (LIMA-GREEN and MOREIRA, 2012) on a methodology of a probabilistic sample survey to measure the variability of estimates of production and effort of fisheries.

Since May 2011, this methodology has been applied in Espírito Santo, a Brazilian Federation Unit (UF), bringing very encouraging results.

In this paper, we describe the sampling plan proposed and its advantages and differences in relation to the previous methodology (ESTATPESCA methodology). Furthermore, we present the

[1] The IBGE is exempt from any responsability for the opinions, information, data and concepts presented in this report which are exclusive responsability of the authors.

results of point and variability estimates of fishery production and effort of the Federation Unit of Espírito Santo.
Keywords: Fishery Monitoring; Sampling; Production, Fishery Effort

Background

In 1962, when the Superintendency for Fishery Development (SUDEPE) was created, fishing was recognized as a staple industry for financial support by official credit lines. From then to the mid-1980s a number of fishery programs were developed and applied especially in order to develop the fishing activity in Brazil.

With the recognition of fishing as a staple industry, the sector's statistics now became more relevant from the economic-financial viewpoint. They had been formerly calculated and published by the Production Statistics Service of the Ministry of Agriculture and by IBGE (PEIXOTO FILHO and DIAS, 1988), and were now attributed to SUDEPE in 1967, which undertook and completed this task until 1979. In 1980, this activity was then again allocated to IBGE in conjunction with SUDEPE, which adopted the methodology created by the latter, and which used three questionnaires: one for the fishing industry, another for colonized fishing and a third for the non-colonized fishing. This administrative-methodological model prevailed until 1989, in which IBGE and SUDEPE produced and published the fishery statistics.

That same year, Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) was created in the sphere of the Ministry of the Environment, and this new institute absorbed SUDEPE duties and personnel. As a result of this merger process the production of statistics was interrupted for some years. It was only in 1991 that fishery landing data collection was resumed by IBAMA, when it created and developed a fishery methodology of its own called the Fishery Statistics Program, or simply EstatPesca. At the start, this program was put in place in the Northeastern Brazilian States, and was later extended to all Brazilian coastal States.

In 1995, the then President of the Republic Fernando Henrique Cardoso signed Decree no. 1694, creating the National Information System for Fisheries and Aquaculture (SINPESQ). Once again IBGE was given the task to coordinate production, systematization and publication of fishery statistics in Brazil. The purpose of this system is to collect, aggregate, process, analyze, exchange and disseminate information about the national fishery sector.

SINPESQ, under the decree, should basically contain data and information produced by the Brazilian Institute of Geography and Statistics (IBGE) and the Ministries of Agriculture, Supply and Land Reform, the Environment, and Tourism, for example, as well as those available from the other federal, state and local agencies, learning and research institutions and agencies involved in the fishing sector. Moreover, it should prepare a working plan defining the duties and those responsible for actions arising from its adoption. Most of the activities and information required refer to the statistical monitoring of fishery and aquaculture, namely, the ongoing systematic collection of information relating to the fishing activity throughout Brazil.

Although IBGE is officially given this task, it no longer collects the information nor produces the sector's statistics. IBAMA and other regional institutions carried on with these activities and only in some States does IBGE participate only in closing the annual results. In other words, the 1995 decree remained with no practical consequences, a situation that lasted until the mid-2000s, when the Special Board of Fishery and Aquaculture (SEAP) was created under the Provisional Measure 103 dated January 1, 2003, and then later on June 29, 2009, adopted ministry status to become the Ministry of Fishery and Aquaculture (MPA).

As soon as SEAP was created, it felt the need for more detailed data on the fishing activity, which culminated in 2007 in a seminar, attended by various fishery-related agencies, as well as by representatives of the civil and academic societies. At this seminar some problems were raised regarding the methodology used by IBAMA for estimating fishery statistics: EstatPesca (IBAMA,

1995). As a result of these comments, SEAP sought the help of IBGE to be able to solve the problems with EstatPesca.

Some studies were carried out and a new methodology was suggested to correct the earlier problems. A pilot survey was then undertaken between May and July 2009 in the States of Ceará and Pernambuco (DIAS and MOREIRA, 2010). Slight adjustments were made and the methodology concluded in 2010 and published by IBGE in 2012 (LIMA-GREEN and MOREIRA, 2012).

In January 2011 the then Ministry of Fishery and Aquaculture, in an agreement with the Federal University of Espírito Santo (UFES) resolved to implement the Fishery Study to be carried out under the new methodology developed by IBGE but not yet published, along the coast of Espírito Santo State. There, the Fishery Statistical Methodology: Onboard Fishing, or simply MEPE, was fully put in place, the team was trained by IBGE and MPA specialists and the data collected from April of that year. The results point to an estimated total fish catch of 8,473 t for the 12-month period - April 2011–March 2012 - with a variation coefficient of 2.4%.

That same year, 2011, another agreement, this time with the Federal University of Rio Grande (FURG), was reached to estimate the volume of sea fish caught and/or landed in Patos Lagoon. In 2013, after the successful implementation of MEPE in Espírito Santo, it was now fully adopted by FURG. IBGE specialists visited Rio Grande, in Rio Grande do Sul State, to train the coordination, supervision and collection teams.

In 2014, it is expected to put MEPE in place all along the Brazilian coast.

After this brief background of the fishery statistics in Brazil, the next sections will address the problems that triggered the request for a new collection and estimation methodology for production and fishery information by SEAP, now MPA, and some of the main results from the experiment in Espírito Santo will be presented.

Problems with the EstatPesca Sample Design

The sample design adopted in EstatPesca is the “Partial control of the fleet in some locations” type, with stratification of the locations by the size measured by the total number of registered boats. The ports with a large number of vessels formed the Management Stratum, in which all locations were selected. Other ports with fewer vessels formed Sampled Strata, in which some of the locations were selected.

It was found that:

- There was effective control in some of the landing locations but the stratification criteria by size and their selection were not only unexplained but apparently were only adopted in some of the landing locations. For example, large locations - even larger than others in the correct stratum – are not controlled.
- At each selected location some of the vessels were selected (in some cases census-based), depending on the size of the location and type of fishing practiced there. The probability selection criteria are not explained and it is not certain whether it was done at random. From the interviews with the collectors working in the visited ports, there are reasons to believe that this did not strictly occur or, at best, only partly. Some of the controlled fleet was apparently selected at random, taking into consideration a subjective criterion of the collector.
- Although recommendable from the methodological viewpoint, there is no sign that a rotation of part of the selected fleet was provided for control purposes.

The sample design adopted in EstatPesca foresees that catch estimates result in the product of another two estimates:

- The first is estimated based on the data collected in the monitoring according to the sampe design that has just been summarized above.
- The second is the result of registration update that must be done monthly, according to which the design provides this in all fishery landing locations in the State and not only where monitoring is undertaken.

This is undoubtedly one of the weakest characteristics of every sample design applied in EstatPesca. A monthly survey of the number of active vessels operating in each port of the State is provided. Not to mention that this information, even if perfect, would still be insufficient since as proxy for the fishery effort, which is the role it plays in this case, is not taking into account the key characteristics of the production unit, namely the number of fishermen, number of fishery equipment items, fishery strategy, etc. So one of the estimation portions is a result of a very fragile procedure, which undermines the overall estimates.

New methodology and its benefits for EstatPesca

The new fishery statistical research methodology for onboard fishing, MEPE, solved the problems of the old methodology with regard to the sample design and data expansion.

The MEPE, in statistical terms, is a production statistical survey that has a complex sampling design, which uses phase sampling. This design has three phases as follows:

- Phase 1: Ports and/or landing locations in each Brazilian State are grouped in strata according to their size in relation to the volume of fish landed there. There may be a management stratum, in which the larger ports and those that desire special treatment (e.g. Santo Antonio, the only fishery landing location on the archipelago of Fernando de Noronha). In the other strata a sample of ports is selected.
- Phase 2: In the port and/or landing location strata other than the management stratum, a simple random sample is taken of ports and/or landing locations within each stratum.
- Phase 3: In all ports and/or landing locations selected to be part of the sample the landings will be stratified in two strata as follows: those from local large vessels, where a census will be taken; and those from local small vessels where a systematic sampling of the vessels according to their order of arrival at the location.

This design solves the three aforementioned problems, as follows:

- Selection criteria of the sample of landing locations in this case are clear. The ports in the management stratum are self-representative, namely, they represent themselves and their selection probability is therefore 1. Now those belonging to the sampled strata of location will be selected according to a simple random sample, and there selection probability will be given by the ratio of the number of locations allocated to this stratum by the total number of places contained therein.
- Since the sample involves actual landings rather than active production units, this could raise a major problem regarding registration. This problem was solved using in the last selection stage a systematic sampling of landings occurring at the location of question. The benefit of this type of sampling is that it is the only type of probability sampling where there is no need to have a list of units beforehand, built up during the sample section process. This selection strategy has also solved the third problem of the rotation of the sample of the selected units, since the permutations of landings occurring in one day are different from those on the other days and a totally different sample has to be taken daily from the day before.

- ❖ With regard to calculating the estimated catch production, whether total or by fishery type, since there is a probability sample design, such problems were solved since the selection probabilities of ports and landings are fully explicit and anyone can calculate them.

Calculating the estimates of the total and their variability

The major benefit of using a probability sample design, is that we can calculate the size of a required sample in order to have a sampling error of the size we want. If we want a small sample error, then we will have to have a relatively large sample, since if the sample error could be larger then the sample will be smaller.

In the case of the State of Espírito Santo, we did not know the level of the variability of the main variable of interest: all fish unloaded during landing. To attempt to solve this problem, a high number of locations, known to have a large volume of fish landed there, were placed in the management stratum. Moreover, in 23 of the 24 ports selected as part of the port sample of the State a census sampling was taken of local small vessels, because at these locations there were very few landings of this category, on average less than 40 landings a day. Therefore, the variation coefficient of all landed fish in Espírito Santo was approximately 2.4%, with an estimated total volume of 8,473t over the 12 months between April 2011 and March 2012.

The estimated total fish caught in the State is calculated by:

$$\hat{Y} = \sum_{h=1}^{H} \frac{M_h}{m_h} \sum_{p=1}^{m_h} \left(\sum_{i=1}^{N_{hp}^{(G)}} y_{hpi}^{(G)} + k_{hp}^{(P)} \sum_{i=1}^{n_{hp}^{(P)}} y_{hpi}^{(P)} \right) \quad (1)$$

where,

H is the total number of location strata, $h = 1, \ldots, H$;

M_h is the total number of landing places of the *h-th* location stratum;

$m_h = 1, \ldots, M_h$ is the number of places in the sample belonging to the *h-th* location stratum;

$N_{hp}^{(G)}$ is the total number of landings from locally considered large vessels of the *p-th* port of the *h-th* location stratum;

$n_{hp}^{(P)}$ is the number of sampled landings from locally considered small vessels of the é *p-th* port of *the h-th* location stratum;

$k_{hp}^{(P)}$ is the leap in systematic sampling associated with the *p-th* port of the *h-th* location stratum;

y_{hpi} is the total fish caught and landed by the *i-th* unloading at the *p-th* port of the *h-th* location stratum.

The estimated variance of the total is calculated by the sum of the variations within each stratum, which is given by the following expression

$$\hat{V}\left(\hat{Y}\right) = \sum_{h=1}^{H} \hat{V}\left(\hat{Y}_h\right) \quad (2)$$

where, $\hat{V}\left(\hat{Y}_h\right)$ is the estimate of the variance of the estimate of all fish unloaded in the *h-th* location stratum.

The estimate of the variance of the estimation of the total for the management strata, each port belonging to this stratum being considered as only one stratum, is given by the following expression:

$$\hat{V}\left(\hat{Y}_h\right)=\sum_{p=1}^{1}\left[\left(k_{hp}^{(P)}\cdot n_{hp}^{(P)}\right)^2\left(1-f_{hp}^{(P)}\right)\frac{\sum_{i=1}^{n_{hp}^{(P)}}\left(y_{hpi}^{(P)}-\bar{y}_{hp}^{(P)}\right)^2}{n_{hp}^{(P)}-1}\right] \quad (3)$$

where,

$\bar{y}_{hp}^{(P)}$ is the simple sampling average of the fish caught from considered small landings of the *p-th* port in the *h-th* location stratum;

$f_{hp}^{(P)}=\frac{n_{hp}^{(P)}}{N_{hp}^{(P)}}=\frac{1}{k_{hp}^{(P)}}$ is the correction factor of a finite population for the *p-th* port in the *h-th* location stratum;

$N_{hp}^{(P)}$ is the total number of landings from locally considered small vessels of the *p-th* port in the *h-th* location stratum,

On the other hand, for the samples location strata, the variance of the estimated total is estimated by the following expression:

$$\hat{V}\left(\hat{Y}_h\right)=\frac{M_h}{m_h}\sum_{p=1}^{1}\left[\left(k_{hp}^{(P)}\cdot n_{hp}^{(P)}\right)^2\frac{\hat{V}\left(\hat{Y}_{hp}\right)}{n_{hp}}\right]+\frac{M_h^2}{m_h}\left(1-f_h\right)\frac{\sum_{p=1}^{m_h}\left(\hat{Y}_{hp}-\bar{Y}_h\right)^2}{M_h-1} \quad (4)$$

where,

$\bar{Y}_h=\frac{\sum_{p=1}^{m_h}\hat{Y}_{hp}}{m_h}$ is the estimate of the average total of fish caught in the *h-th* location stratum;

f_h is the correction factor of a finite population for the *h-th* location stratum, and the estimate of the variance of the estimated total of fishery catch in the *p-th* port of the *h-th* location stratum is estimated by the following expression:

$$\hat{V}\left(\hat{Y}_{hp}\right)=\left(k_{hp}^{(P)}\cdot n_{hp}^{(P)}\right)^2\left(1-f_{hp}^{(P)}\right)\frac{\sum_{i=1}^{n_{hp}^{(P)}}\left(y_{hpi}^{(P)}-\bar{y}_{hp}^{(P)}\right)^2}{n_{hp}^{(P)}-1} \quad (5)$$

Actual estimates calculated using the MEPE

The MEPE was fully used for the first time to estimate the fishery production in the State of Espírito Santo. There the information was first collected in April 2011 and proceeded until late December 2012. The MPA provided the microdata for the first collection year: April 2011- March 2012. For this one-year period the fishery production of a Brazilian State was estimated using a probability sample for the first time.

This allowed us to calculate for each estimate, statistics to indicate the quality of information estimated therein. This quantity is the variation coefficient of the estimate. In this way we estimated

that the landed fish all along the Espírito Santo coast was 8,473t, and its estimated variation coefficient was 2.4%. This means that the estimate error for all landed fish along the Espírito Santo coast is at most more or less 4.7%.

The table below shows the estimates of total fishery of the 12 main species landed there, for which the total estimated landing exceeds the 100 t mark and their variation coefficients.

Espécie		Estimativa de Captura	
Nome Popular	Nome Científico	Peso (t)	CV (%)
Total		8.473	2,4
Camarão Sete Barbas	*Xiphopenaus kroyeri*	2.046	3,8
Dourado	*Coryphaena spp.*	1.750	0,2
Albacora Lage	*Thunnus albacares*	650	0,1
Bonito	*Sarda sarda*	397	0,7
Pargo	*Pagrus pagrus*	339	15,4
Cação	*Selachimorpha spp.*	290	3,2
Pescadinha	*Isopisthus parvipinnis*	267	13,8
Corvina	*Micropogonias furnieri*	257	5,3
Baiacu	*Lagocephalus laevigatus*	210	9,1
Peroá	*Balistidae spp.*	142	3,7
Espadarte	*Xiphias gladius*	124	0,3
Sarda	*Scombrinae spp.*	103	2,9

Table 1 – *Estimates of the total fishery catch in Espírito Santo State for the main species*

As Table 1 shows, the estimates for the total fishery per species are very good, since they present a variation coefficient of less than 10% for almost all species, except sea bream and whiting. The variation coefficients for the estimated variation of the total sea bream and whiting landed on the Espírito Santo coast are higher than the others due to the fact that these species of fish are mostly caught by smaller vessels, and because quite widespread fishing of these species are found off the State coast.

Since this kind of vessel forms the bulk of the fleet of small and medium-size ports and/or locations, and since they are allocated to the sampled strata, we have, therefore, an explanation for the high variability associated with fishing sea bream and whiting: their landings are associated with the vessel size strata and locations associated with the variability of the process, since unit sampling is taken in these strata – whether landings or ports.

Conclusions and next steps

Having completed the pilot tests and put the MEPE in place in the States of Espírito Santo and Rio Grande do Sul, where the collection is being currently done using this methodology, we consider the results very satisfactory, especially since we can use a technique that permits the calculation of the estimates and their sampling errors, and consequently, a measure of quality for each calculated estimate.

The variation coefficients are small, which indicates that the sample size is suitable for estimating these quantities.

MPA is working to put the MEPE in place all along the Brazilian coast, with the start of the collection expected in the first half of 2014.

REFERENCES

Aragão, J. A. N. (2006) Sistema de geração de dados estatísticos da pesca – EstatPesca: Manual do usuário. *Tecnichal Report.* 40p.

Dias, A. J. R. and Moreira G.G. (2009). Relatório sobre o projeto piloto: Estimação da Pesca Marinha no Ceará e em Pernambuco. *Tecnichal Report*.

IBAMA (1995). Manual do EstatPesca. *Tecnichal Report.* IBAMA: Tamandaré/PE. 57p.

Peixoto Filho, C. V. and Dias, A. C. B.(1988) *Retrospectiva sobre a estatística da pesca,* IBGE: Rio de Janeiro.

Lima Green, A. P. and Moreira, G. G. (2012) Metodologia Estatística da Pesca: Pesca Embrcada, *Textos para Discussão.* Tecnichal Report of IBGE, Department of Agriculture and Department of Methods and Quality: Rio de Janeiro. 52pp.