

Estimation of fleet and stock related variables in the Greek fisheries under Regulation (EC) No 2017/1004



**Hellenic Centre for Marine Research
Institute of Marine Biological Resources & Inland Waters**



**Fisheries Research Institute
Hellenic Agricultural Organization 'Demeter'
Hellenic Ministry of Rural Development and Food**



Athens – Kavala, 2021

Contributors:

HCMR: Stefanos KAVADAS, Danai MANTOPOULOU-PALOUKA, George TSERPES, Dimitrios DAMALAS

FRI: Konstantinos TOULOUMIS, Angeliki ADAMIDOU, Manos KOUTRAKIS

Table of Contents

| | |
|--|----|
| Table of Contents..... | 2 |
| Summary | 3 |
| Preface..... | 4 |
| 1. Effort-Landings | 6 |
| 1a. Fisheries Research Institute, Estimation of fleet and stock related variables applied by Fisheries Research Institute in the central-north part of the Aegean Sea (GSA22) | 6 |
| Calculation of fishing effort for the multi-gear small scale Greek fisheries | 6 |
| Sampling plan | 6 |
| Fishing Activity..... | 7 |
| Fishing Effort calculation | 8 |
| Example | 8 |
| Assessment of the applied method..... | 11 |
| Total landings per species calculation..... | 17 |
| 1.b Hellenic Centre for Marine Research,..... | 18 |
| Estimation of fleet and stock related variables applied by the Institute for Marine Biological Resources of and Inland waters of HCMR in the south Aegean Sea (GSA22), Crete (GSA23) and the eastern Ionian Sea (GSA20)..... | 18 |
| Effort | 18 |
| Effort by rectangle (FDI table I) | 19 |
| Evaluation of method accuracy | 20 |
| Landings..... | 21 |
| Landings by rectangle (FDI table H)..... | 22 |
| 2. Discards | 23 |
| 3. Biological Data | 24 |
| Age-Length Key (ALK)..... | 24 |
| 4. Demographics | 25 |
| Length frequency distribution | 25 |
| Age frequency distribution | 28 |
| Conclusion | 29 |
| References | 31 |
| Annex to estimation of fleet and stock related variables applied in the Institute for Marine Biological Resources and Inland waters..... | 33 |

Summary

During the meeting on the implementation of DCF in Greece (18th of November 2019), DG MARE noted that, apart from data transmission failures linked to data gaps, there were also data quality issues reported from end users deterring proper assessment of Greek stocks.

In collaboration with the local authorities (DG of Fisheries - Ministry of Rural Development and Food) a "Working Group on quality assurance" has been set up for this purpose involving scientists from all institutes implementing Greek DCF (HCMR and FRI).

The Hellenic Centre for Marine Research (HCMR – www.hcmr.gr) is responsible for monitoring: (i) the southern part of Aegean Sea (GSA22), (ii) Crete (GSA23) and (iii) the eastern Ionian Sea (GSA20), while the Fisheries Research Institute (FRI – www.inale.gr) is monitoring the north-central part of the Aegean Sea (GSA22). The aforementioned regions encompass specific characteristics that led to the adoption of somehow diverse estimation procedures. More specifically, the stratification of Northern Aegean Sea (FRI approach) is based on the registered Primary Gear, while for the Southern Aegean and Ionian Seas (HCMR approach) it is based on the activity level.

The goal of this WG was to compile a technical document describing the sampling scheme and statistical estimation procedures. This document is a result of the efforts of the specific WG.

Herein the approaches adopted by the Institutes responsible for realizing DCF in Greece to estimate fleet & stock related variables to service the Mediterranean Data Calls are provided in detail. Accuracy and precision of the aforementioned implementations were assessed by employing a series of simulations.

Preface

This document is the outcome of the ‘Working Group on quality assurance’ which was established in the framework of the ‘Priority Actions on Data Collection in Greece’.

The aim of this document is to describe the approaches implemented to derive estimates for the various Greek fish stocks monitored under the National Data Collection Program, Council Regulation 2017 (EC, 2017) and Commission Delegated Decision 910/2019 (EC, 2019) in line with the requirements of the EU Common Fisheries Policy (EC, 2013).

The document expands on an existing document describing the Greek sampling scheme and data quality assurance framework (Anonymous, 2019) and relies heavily on the data collection and sampling protocols provided therein.

In general, the assessment of the state of exploitation of fish stocks and the impact of fishing on their reproductive capacity and the ecosystem requires a set of information regarding the activities of fleets, catches, abundances and demography of fish stocks. Council Regulation 2017 (EC, 2017) and Commission Decision 910/2019 (EC, 2019), establish a community framework for the collection, management and use of data for the purpose of creating a solid basis for carrying out scientific analysis on fisheries and enabling the formulation of sound scientific advice for the implementation of the common fisheries policy (CFP).

The ultimate goal is to provide administrations, national and European, necessary scientific support to ensure proper decision making (e.g. undertaking interventions, planning and launching management measures). Such interventions and measures are necessary to guarantee a rational use of fishery resources, both biologically and at the same time economically sustainable.

In the Mediterranean, such competent scientific groups in charge of assessing stock status and informing decision makers include:

- the Scientific Advisory Committee (SAC) on Fisheries of GFCM
- the Expert Working Groups (EWG) on Mediterranean Stock Assessment of STECF
- ICCAT Stock Assessment Groups

The input required by these groups is usually estimations of effort, landings, discards, size/age distribution and maturity, at a fleet/area level, for the stocks that data collection is carried out. To this end GFCM, STECF and ICCAT issue periodic data calls to which member states (or contracting parties) have an obligation to respond by submitting aggregated data (estimations).

Note: Source of data for all the following algorithms is the data collected under Council Regulation (EC) No 2017/1004 and (EC) No 199/2008. These data are stored in the Fisheries Database of HCMR (Kavadas et al., 2013) and FRI Database (a version of STOCKMAN database modified for the Greek data).

Structure of the data in the DB follows the aggregation level defined in Commission Delegated Decision (EU) 2019/910:

6 levels of Fishing activity-metiers

Fleet segmentation

Geographical stratification

Species/stocks defined in the multi-annual Union Programme (see Anonymous, 2019)

For the Greek fisheries the classification is as follows:

Metier: list of metiers monitored in the Greek DCF (see Anonymous, 2019)

Geographical stratification: GSA (GFCM FAO subareas): 20, 22, 23

Subareas (Greek DCF): N-ION, C-ION, S-ION, THERM, THR-LIM, VOL-SPOR, EVIA, ARGSAR, CHI-MIT, CYCL, DODEC, CRETE

Fleet segmentation: 0-6m, 6-12m, 12-24m, 24-40m, >40m

Quarter (year): 1, 2, 3, 4

Species: list of species monitored in the Greek DCF (see Anonymous, 2019)

The following pages provide a detailed description of the methodologies employed within the Greek DCF. Different approaches for the estimation of effort-landings for the SSF were deemed necessary, due to differences in the characteristics of the SSF fleets between the North and South Aegean. The corresponding methodologies are described in the effort-landings section.

1. Effort-Landings

1a. Fisheries Research Institute, Estimation of fleet and stock related variables applied by Fisheries Research Institute in the central-north part of the Aegean Sea (GSA22)

Contributors: *Touloumis Konstantinos, Adamidou Angeliki, Koutrakis Manos*

Calculation of fishing effort for the multi-gear small scale Greek fisheries

Declarative data, such as Logbook and VMS data are available only for a fraction of the existing Small-Scale Fisheries (SSF) fleet [vessels >10m and vessels with a special fishing license (i.e. vessels targeting large pelagic species and Boat seiners). Since Greece has the largest fleet in the EU, in terms of number of vessels, and the vast majority of these are SSF vessels (~91.5% of the total, see figure 1), for this part of the fleet an Effort-Landings assessment survey is conducted, based on a spatially and technically stratified sampling scheme estimate effort and landings per species.

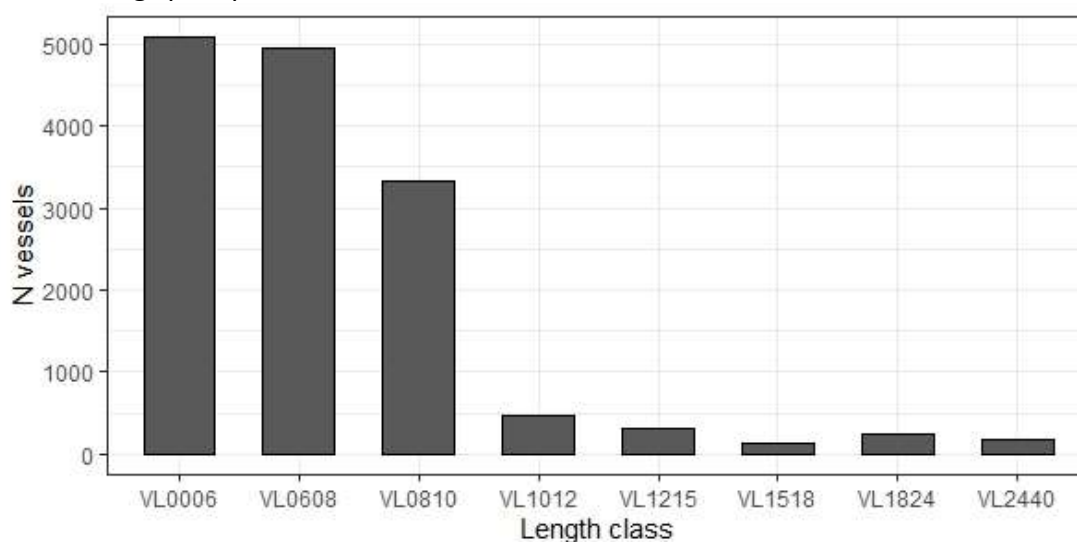


Figure 1. Vessels present in the Greek fishing fleet on 15/11/2020 per vessel length category, Date Source: https://webgate.ec.europa.eu/fleet-europa/search_en Situation as at November 2020.

Sampling plan

The Effort-Landings assessment survey for the Greek SSF is based on a spatially and technically stratified random sampling scheme. The spatial stratification is the same with the one applied on biological data sampling scheme (see Anonymous 2019). The technical stratification includes all SSF fishing gears, as well as the classification of the fleet based on total vessel length (for SSF: 0-6 m and 6-12 m) (Anonymous 2019).

The Sampling Frame of the survey is the National Fleet Register (NFR). This data frame registers all the fishing vessels of the Greek fleet and their main characteristics (total length, gross tonnage, engine power, registration port, main and the secondary fishing gear etc.). Fishing vessels are classified to different strata based on their registration port in NFR (area

(a)), on the main gear declared in NFR (main gear (**mg**)) and on their overall length (vessel length category (**vl**)).

Within each of the aforementioned strata, vessels are selected randomly for sampling. The number of samples within each stratum is, in general terms, proportional to the total number of vessels within the stratum. In order to cope with possible refusals on behalf of the fishermen, a list of randomly vessels is formulated within each stratum. In a case of a refusal, the vessel is replaced with the following in the list.

For the estimation of effort, Primary Sampling Unit is the days at sea per month/ fishing gear/vessel. Apart from days at sea, additional data such as net length, number of hooks, number of traps etc. are collected. The Effort-Landings assessment survey for the Greek SSF is conducted through structured questionnaires - personal interviews with the fishers.

To summarize, based on the above, the stratification hierarchy of the applied sampling scheme is:

GSA → subArea(a) → Main Gear(**mg**) → Vessel length (**vl**) → Month (**m**) → Fishing Gear (**fg**) and the PSU sampled is days at sea per fishing Gear (**d_{fg}**).

The goal of the Effort estimation survey is the calculation of d_{fg} per $a/vl/m$.

It has to be noted, that the main gear declared in NFR (**mg**) is often different than the fishing gear (**fg**). The Greek SSF is a multi-gear fishing fleet, and fishers are not restricted to use only the registered gears. As a result, they often use three or even four different fishing gears throughout the year and, frequently, these are none of the registered ones. However, the outcomes of the yearly landings-effort survey indicate that vessels registered under the same main gear and area exhibit similar patterns of fishing activity.

Fishing Activity

It is well documented that in some EU member states, including Greece, for various reasons, the number of active fishing vessels is different (less) than the official number registered in the National Fleet Register (STECF, 2019). As a result, an Effort-Landings assessment survey using the National Fleet Register without any treatment to calculate the total effort and landings per species for SSF for a specific year, would lead to overestimations.

In order to improve the accuracy of the Sampling Frame and calculate the actual active fleet operating, an extensive field work is conducted throughout **each year** to identify the active vessels, by frequent visits on the main registration ports. These data are cross-referenced with official documents involving special license to operate with specific fishing gears (such as Drifting longlines (LLD) as well as with data coming from local fishery supervising authorities. The above information is assessed in order to assign each vessel in the Active or Inactive category. **A vessel is considered active when it is adequately justified that it has been operating and merchandising even for a single day within the year.**

Based on the field work survey, the activity ratio per stratum is calculated:

$$A_j = \frac{AC_j}{AC_j + IN_j} \quad (1)$$

where the activity coefficient A_j for the stratum j is equal to the ratio of the active vessels AC_j to the sum of active AC_j and IN_j , inactive vessels. As a result, the total fishing vessels per stratum N_j , is equal to:

$$N_j = A_j \times NFR_j \quad (2)$$

where NFR_j is the sum of fishing vessels registered in the Sampling frame as part of the stratum j .

Fishing Effort calculation

Raising fishing effort from sample follows the sampling scheme stratification (a/vl/mg/m/fg), the PSU raised are the days per fishing gear (d_{fg}) and the goal of the raising is to calculate the total days per fishing gear per a/vl/m/fg.

For each fishing gear **fg** and within each stratum **j** (**a/vl/mg/m**), the mean fishing days \bar{d}_{fgj} is equal to:

$$\bar{d}_{fgj} = \frac{1}{n_j} \times \sum_{i=1}^{N_j} d_{fgi} \quad (3)$$

where n_j is total number of vessels sampled in stratum **j** and d_{fgi} is the fishing days for fishing gear **fg** in vessel **i**.

Sample variance $Var(Y_{gj})$ is equal to:

$$Var(\bar{d}_{fgj}) = \frac{1 - \frac{n_j}{N_j}}{n_j} \times s_{fgj}^2 \quad (4)$$

where n_j is total number of vessels sampled in stratum **j**, N_j is the total active fleet of the stratum **j**, and s_{gj} is standard deviation of the sample for gear **g**.

Standard error $se(\bar{d}_{fgj})$ of the sample mean is equal to:

$$se(\bar{d}_{fgj}) = \frac{\sqrt{Var(\bar{d}_{fgj})}}{\sqrt{n}} \quad (5)$$

And coefficient of variation $CV(\bar{d}_{fgj})$ is equal to

$$CV(\bar{d}_{fgj}) = \frac{se(\bar{d}_{fgj})}{\bar{d}_{fgj}} \quad (6)$$

Within each stratum, total days effort per fishing gear and standard error respectively, are equal to:

$$D_{fgj} = N_j \times \bar{d}_{fgj} \quad (7)$$

$$se(D_{fgj}) = \frac{N_j \times \sqrt{Var(\bar{d}_{fgj})}}{\sqrt{n_j}} \times \sqrt{1 - (n_j/N_j)} \quad (8)$$

CV is calculated as in equation (6).

Finally, if **L** is the total number of main gear categories within each stratum **j**, the total days per a/vl/mare equal to:

$$TotalD_{fg} = \sum_L D_{fgj} \quad (9)$$

The Variance and standard error are calculated as in equation (8) and (9) respectively. All the above equations are from Cochran (1977).

Example

Hypothetically, the total effort per fishing gear is required to be calculated for the following stratum:

| GSA | subArea | Vessel length | Main Gear | Month |
|-----|---------|---------------|-----------|-------|
| 22 | THR-LIM | VL0612 | GTR | 2 |

Based on the National Fleet Registration 362 vessels are registered in this stratum. Field work for the assessment of the activity of this stratum identified 219 vessels (no positive or negative information were available for the remaining 143 vessels), from which 171 were active. As a result, activity coefficient is set to 0.781 (143/219) and the active fleet $N_j=282.66$ ($A \times Fleet_N$).

| Sub_Area | Gear.Main.Code | VL | Fleet_N | ACT | IN | A | N |
|----------|----------------|--------|---------|-----|----|-------|---------------|
| THR-LIM | GTR | VL0612 | 362 | 171 | 48 | 0.781 | 282.66 |

The raw sample data for this stratum are shown in table 1. In this case, the sample consist of 16 vessels, one of which (Vess_16) didn't operate within this month. From the remaining, nine used only one gear within the month and another five used two different gears.

Table 1. Raw data

| Reg_Code | Lan_Gear | Lan_Days |
|---------------|----------|----------|
| Ves_01 | GNS | 12 |
| Ves_01 | GTR | 24 |
| Ves_02 | FPO | 1 |
| Ves_02 | GTR | 7 |
| Ves_03 | GNS | 20 |
| Ves_03 | GTR | 20 |
| Ves_04 | GNS | 15 |
| Ves_04 | GTR | 20 |
| Ves_05 | GNS | 22 |
| Ves_06 | GTR | 9 |
| Ves_07 | GTR | 14 |
| Ves_08 | GNS | 12 |
| Ves_09 | GTR | 15 |
| Ves_10 | FPO | 7 |
| Ves_10 | GTR | 20 |
| Ves_11 | LLS | 18 |
| Ves_12 | GTR | 12 |
| Ves_13 | GTR | 10 |
| Ves_14 | GNS | 5 |
| Ves_15 | GNS | 12 |
| Ves_15 | LLS | 12 |
| Ves_16 | - | 0 |

In table two, the information is reformed in a way that the distribution of fishing days between gears for each vessel is better depicted (zero for a gear that has not been used). By using the equations 3-5 we calculate Mean fishing days, std error and CV within the stratum.

Table 2. Calculation of Mean fishing days, std error and CV.

| Reg_Code | GTR | GNS | FPO | LLS |
|----------|-----|-----|-----|-----|
| Ves_01 | 24 | 12 | 0 | 0 |
| Ves_02 | 7 | 0 | 1 | 0 |
| Ves_03 | 20 | 20 | 0 | 0 |
| Ves_04 | 20 | 15 | 0 | 0 |
| Ves_05 | 0 | 22 | 0 | 0 |
| Ves_06 | 9 | 0 | 0 | 0 |
| Ves_07 | 14 | 0 | 0 | 0 |
| Ves_08 | 0 | 12 | 0 | 0 |
| Ves_09 | 15 | 0 | 0 | 0 |
| Ves_10 | 20 | 0 | 7 | 0 |
| Ves_11 | 0 | 0 | 0 | 18 |
| Ves_12 | 12 | 0 | 0 | 0 |

| | | | | |
|---------------------|-------------|-------------|-------------|-------------|
| Ves_13 | 10 | 0 | 0 | 0 |
| Ves_14 | 0 | 5 | 0 | 0 |
| Ves_15 | 0 | 12 | 0 | 12 |
| Ves_16 | 0 | 0 | 0 | 0 |
| \bar{d}_{fgj} | 9.44 | 6.13 | 0.50 | 1.88 |
| $se(\bar{d}_{fgj})$ | 0.53 | 0.49 | 0.11 | 0.32 |
| $CV(\bar{d}_{fgj})$ | 0.06 | 0.08 | 0.21 | 0.17 |

In Table 3, by using the equations 7-8 we calculate total days effort per gear and standard error within the stratum. Following equation 6, we calculate CV.

Table 3. Calculation of Total fishing days and std error within the stratum

| MAIN_Gear==GTR | GTR | GNS | FPO | LLS |
|-----------------------|----------------|----------------|---------------|---------------|
| D_{fgj} | 2667.59 | 1731.28 | 141.33 | 529.98 |
| $se(D_{fgj})$ | 145.60 | 133.80 | 29.12 | 87.36 |
| $CV(D_{fgj})$ | 0.05 | 0.08 | 0.21 | 0.16 |

Applying the same approach in the FPO part of the fleet in the stratum:

| GSA | subArea | Main Gear | Vessel length | Month |
|------------|----------------|------------------|----------------------|--------------|
| 22 | THR-LIM | FPO | VL0612 | 2 |

(only Main Gear is different than before)

leads to the results shown in table 4(the steps we follow for the calculation is the same and are not show here. Furthermore, for simplicity we omit the results for the remaining gears – (GNS, LSS etc.).

Table 4. Calculation of Total fishing days and std error within the stratum FPO.

| MAIN_Gear==FPO | GTR | GNS | FPO | LLS |
|-----------------------|--------------|--------------|--------------|------------|
| D_{fgj} | 45.73 | 82.32 | 731.7 | 0 |

Moving upwards the stratification scale by “removing” main gear:

| GSA | subArea | Vessel length | Month |
|------------|----------------|----------------------|--------------|
| 22 | THR-LIM | VL0612 | 2 |

And by using the equation 9, we get the results of Table 5.

Table 5. Total days per fishing gear for within the stratum. For simplicity we assume that only GTR and FPO vessels operate in the area.

| Stratum days | GTR | GNS | FPO | LLS |
|---------------------|------------|------------|------------|------------|
| $TotalD_{fg}$ | 2713.32 | 1813.6 | 873.03 | 529.98 |

Provided that additional Main Gear categories operate in the area, the process is repeated for all Main Gear categories within the stratum to calculate total days per fishing gear.

Assessment of the applied method

Description of the assessment method

One possible caveat of the aforementioned raising method, as well as of the applied sampling scheme, is the assumption that vessels under the same area, vessel length and main registration gear exhibit similar patterns of fishing activity, in terms of the interchangeability of the different fishing gears within a year. Additionally, apart from this method, others have been proposed and applied. In order to be able to estimate the efficiency of each applied method, in terms of the precision of the related calculations, an unbiased method of assessment should be applied, properly constructed to be able to assess objectively the efficiency of every applied method.

However, the effort calculation sampling methods are usually applied to calculate variables that are otherwise unknown (like total effort of for segment of a fleet with no declarative data). Thus, their efficiency could not be assessed through the comparison of their outcomes with the actual, real values. To overcome this obstacle, a bootstrap-simulation method of efficiency assessment is proposed and applied. This method relies on the formulation of an artificial dataset with substantial effort data upon which a random bootstrap technique simulating the applied sampling scheme could be applied. In this way, the total sum of effort of the vessels included in the artificial dataset could be regarded as the “real” numbers to be calculated and each tested method could be assessed in through the comparison of their outcomes with these numbers.

To that end, we pooled together effort assessment samples from six years (2013 to 2016, 2018 and 2019) and four subareas (THR-LIM, THERM, CHIO-MIT and VOL-SPOR) to create a single, artificial dataset of effort data. In this dataset, we included only SSF vessels and gears, and only vessels included in the 0-6 m, 6-12 m and 12-18 m categories. Since some vessels were sampled more than once throughout the years, we set as unique vessel ID the combination of year and vessel CFR (vesselID=year_CFR). This resulted in creating a dataset containing information from 881 different vessel IDs. Spatial stratification as well as the years of sampling were not taken into account; as a result, in this dataset the stratification hierarchy that is available is

Main Gear → Vessel length → Month → Fishing Gear

and the PSU is still days at sea per fishing Gear (dfg). Furthermore, we constructed a pseudo-NFR by collating the relative information for the vessels included in our dataset, from each years NFR.

In the Table 6, the datasets available vesselIDs per vessel length category and main fishing gear are presented. The “real” effort of this fleet, i.e. the days at sea per fishing gear of the whole artificial dataset are presented in Table 7. These numbers are resulted from adding up the days at sea per fishing gear.

Table 6. VesselIDs per vessel length category and main gear in the dataset.

| Gears | VL0006 | VL0612 | VL1218 | Total |
|--------------|---------------|---------------|---------------|--------------|
| <i>DRB</i> | 0 | 1 | 0 | 1 |
| <i>FPO</i> | 12 | 72 | 11 | 95 |
| <i>GNS</i> | 39 | 133 | 12 | 184 |
| <i>GTN</i> | 4 | 18 | 5 | 27 |
| <i>GTR</i> | 85 | 229 | 40 | 354 |

| | | | | |
|--------------|------------|------------|------------|------------|
| <i>LHP</i> | 0 | 4 | 0 | 4 |
| <i>LLD</i> | 3 | 3 | 6 | 12 |
| <i>LLS</i> | 38 | 116 | 50 | 204 |
| Total | 181 | 576 | 124 | 881 |

Table 7. Days at sea for the fishing gears in the data set.

| <i>Lan_Gear</i> | Days at sea |
|-----------------|--------------------|
| <i>GTR</i> | 48510 |
| <i>GNS</i> | 34076 |
| <i>LLS</i> | 20094 |
| <i>FPO</i> | 14337 |
| <i>LLD</i> | 1035 |
| <i>LHP</i> | 767 |
| <i>DRB</i> | 678 |
| <i>GNC</i> | 297 |
| <i>GTN</i> | 281 |

Finally, the days at sea per fishing gear and month are depicted in figure 2.

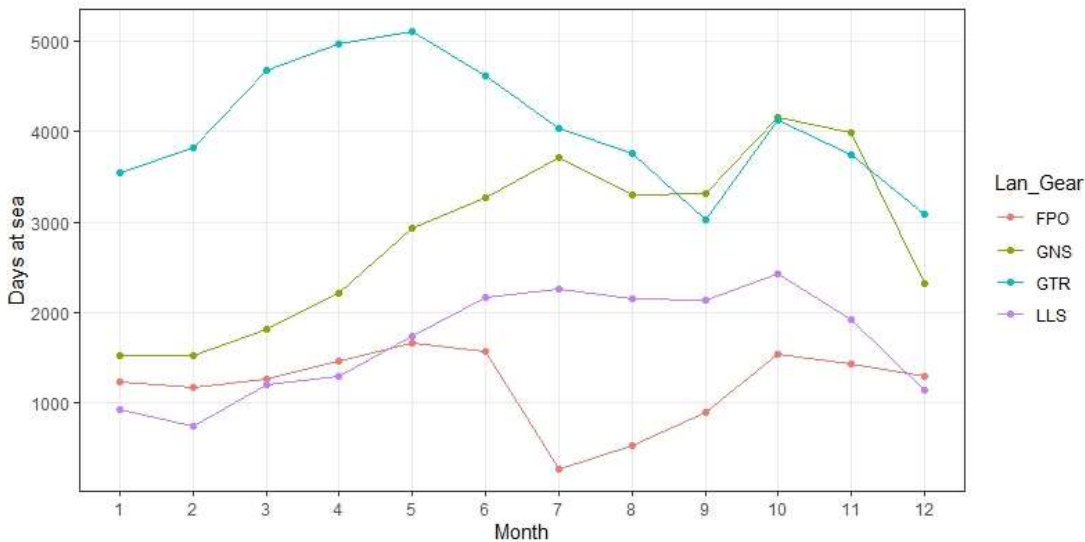


Figure 2. Days at sea per month and fishing gear for the fleet included in the artificial dataset.

To assess the applied raising method described in the previous section, we sample randomly from the dataset, following the sampling strategy described above. More specifically, we randomly select a proportion of the available vesselIDs per vessel length and main gears. To simplify the process, the sampling proportion was set equal to a constant 0.05 (in the actual sampling plan, although sampling effort per gear and vessel length varies, a total of 861 vessels are sampled, which correspond to ~0.05 of the total population of vessels). On this sample, the raising method described above is applied, in order to calculate Total days at sea per fishing gear, vessel length category and month. An iterative bootstrap simulation loop is applied, and the process is repeated n times (n=1000 times). The effectiveness of the method is assessed on the basis of comparing the calculated effort with the “real” ones, as described in the Table 4 and Figure 2. All the above analyses were conducted within R programming environment. An extension of the above described method, to apply the same

technique on landings per species as well, is under construction and will be available soon. The relevant R scripts as well as technical support, are available upon request.

Results

Applying a 0.05 sampling proportion on the datasets results in a sampling effort distribution as the one shown in Table 8.

Table 8. Total vessels in the dataset and sampling effort for the simulations per main fishing gear and vessel length. When total vessels*sampling proportion <1, Sample is fixed to 1.

| <i>Main gear</i> | <i>VL</i> | <i>Total vessels</i> | <i>Sample</i> |
|------------------|-----------|----------------------|---------------|
| <i>DRB</i> | VL0612 | 1 | 1 |
| <i>FPO</i> | VL0006 | 12 | 1 |
| <i>FPO</i> | VL0612 | 72 | 4 |
| <i>FPO</i> | VL1218 | 11 | 1 |
| <i>GNS</i> | VL0006 | 39 | 2 |
| <i>GNS</i> | VL0612 | 133 | 7 |
| <i>GNS</i> | VL1218 | 12 | 1 |
| <i>GTN</i> | VL0006 | 4 | 1 |
| <i>GTN</i> | VL0612 | 18 | 1 |
| <i>GTN</i> | VL1218 | 5 | 1 |
| <i>GTR</i> | VL0006 | 85 | 4 |
| <i>GTR</i> | VL0612 | 229 | 11 |
| <i>GTR</i> | VL1218 | 40 | 2 |
| <i>LHP</i> | VL0612 | 4 | 1 |
| <i>LLD</i> | VL0006 | 3 | 1 |
| <i>LLD</i> | VL0612 | 3 | 1 |
| <i>LLD</i> | VL1218 | 6 | 1 |
| <i>LLS</i> | VL0006 | 38 | 2 |
| <i>LLS</i> | VL0612 | 116 | 6 |
| <i>LLS</i> | VL1218 | 50 | 2 |

In figure 3, the outcomes of the bootstrap-simulation (n=1000) are depicted. More specifically, for each iteration we summed up the total days per fishing gear (throughout the whole year and vessel length category). In figure 3, the density distributions of the total days at sea per each fishing gear are depicted. With dashed lines, the “real” values (as in Table 7) per fishing gear are shown. In general terms, the values are normally distributed around the “real” values per fishing gear. To quantify the spread of the simulated values around the mean, we calculated Coefficient of Variation (as standard deviation/mean value) for each fishing gear. CV values, as well as the mean and median values per fishing gear, are presented in table 9. As far as the four main fishing gears of the Greek SSF are concerned (GTR, GNS, LLS, FPO) CV value is lower on GTR (CV=0.16) and it gradually increases as the mean days per year increase (GNS CV=0.19, LLS CV=0.25 and FPO CV=0.32). Finally, in figure 4, the fluctuations of the calculated days per month and fishing gear (boxplots) are contrasted to the “real” values. It is evident that, the calculated values follow adequately the periodical fluctuation of the “real” values.

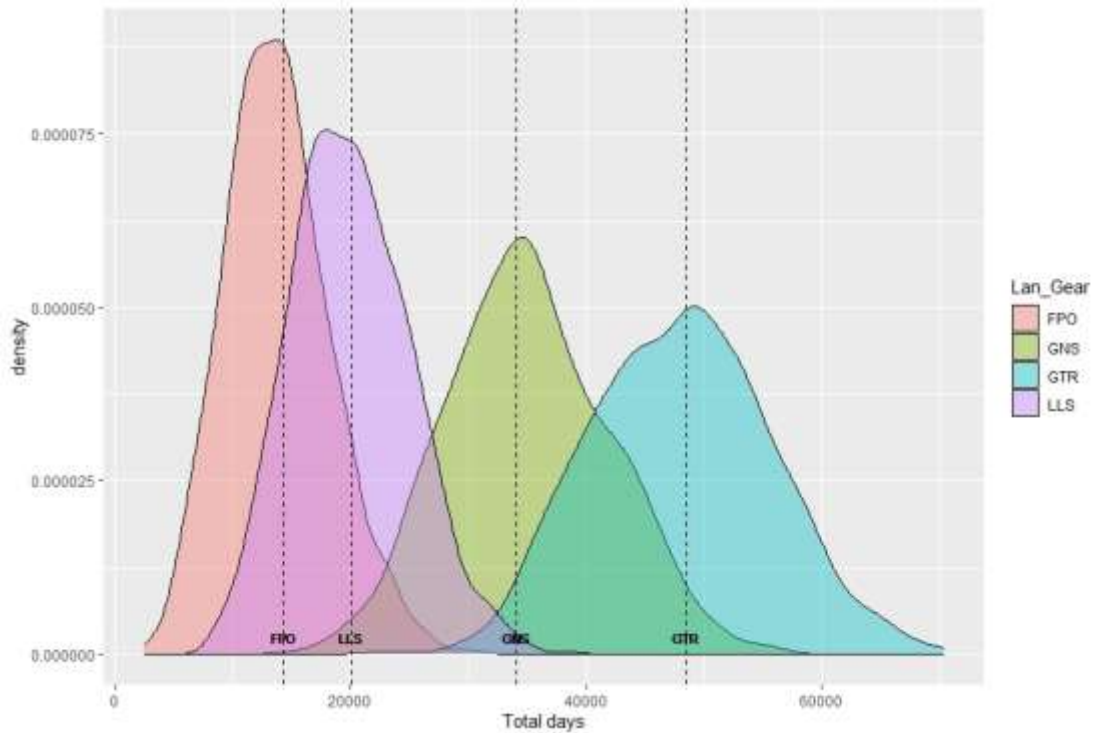


Figure 3. Density distribution of the calculated effort in 1000 iterations, for the four main gears of the Greek SSF. In dashed lines the “real” numbers are depicted.

Table 9. Mean and median days, standard deviation and CV for n=1000 iterations. The “real” days are also included, as a comparison set point.

| FishingGear | Meandays | Mediandays | sd | CV | "real" days |
|-------------|----------|------------|---------|------|-------------|
| <i>GTR</i> | 48061.21 | 48191.26 | 7586.39 | 0.16 | 48510 |
| <i>GNS</i> | 35087.19 | 34788.29 | 6841.47 | 0.19 | 34076 |
| <i>LLS</i> | 20031.35 | 19768.32 | 4971.16 | 0.25 | 20094 |
| <i>FPO</i> | 13800.04 | 13577.76 | 4420.35 | 0.32 | 14337 |
| <i>LLD</i> | 1308.81 | 1047.38 | 1003.34 | 0.77 | 1035 |
| <i>LHP</i> | 1026.89 | 810.58 | 795.28 | 0.77 | 767 |
| <i>DRB</i> | 878.96 | 731.50 | 655.79 | 0.75 | 678 |
| <i>GNC</i> | 660.50 | 496.17 | 450.82 | 0.68 | 297 |
| <i>GTN</i> | 422.07 | 208.18 | 427.63 | 1.01 | 281 |

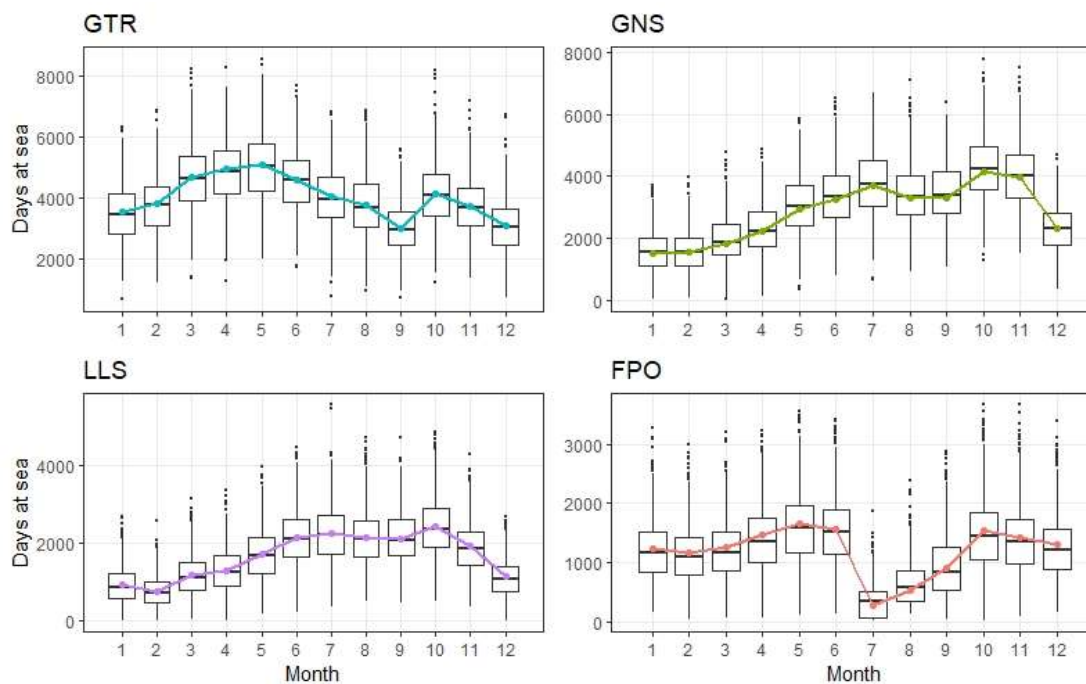


Figure 4. The calculated days per month and fishing gear (boxpots) are contrasted to the “real” values, for the four main gears of the Greek SSF.

Based on the above, the application of the bootstrap simulation technique revealed that the applied effort calculation method should be considered accurate and unbiased. Mean and median values of the simulated calculations per fishing gear are only marginally different to the “real” days of the artificial dataset. The seasonal fluctuations of the applied effort per fishing gear is calculated accurately. However, the moderate to high CV values for some fishing gears indicate that the applied sampling plan could be improved in order to increase the accuracy of the above calculations (see next section).

Sampling plan and effort assessment accuracy

Naturally, an increase in the sampling effort it is expected to consequently increase the accuracy of the effort assessment. To indicate that, we applied the bootstrap-simulation technique on the artificial dataset by setting the sampling proportion to 0.1 (double than the one applied before). In table 10, the mean, median, standard deviation and CV of this simulation are presented. As expected, CV values decrease (by 0.05 in GTR, 0.06 in GNS, 0.07 in LLS and 0.1 in FPO), meaning that the accuracy of the effort assessment is substantially improved. However, since Greek SSF fleet consist of ~10000 vessels, doubling the sampling effort would not be feasible under the available resources (in terms of the corresponding cost, effort, available personnel and time).

Table 10. Results of the simulation with sampling proportion =0.1

| FishingGear | Meandays | Mediandays | sd | CV | "real" days |
|-------------|----------|------------|---------|------|-------------|
| GTR | 47706.82 | 47556.01 | 5168.93 | 0.11 | 48510 |
| GNS | 34081.12 | 33896.64 | 4458.82 | 0.13 | 34076 |
| LLS | 20243.81 | 20151.52 | 3544.94 | 0.18 | 20094 |
| FPO | 14214.72 | 14112.94 | 3130.14 | 0.22 | 14337 |
| LLD | 1103.61 | 1000.62 | 707.28 | 0.64 | 1035 |

| | | | | | |
|-----|--------|--------|--------|------|-----|
| LHP | 826.97 | 704.54 | 555.99 | 0.67 | 767 |
| DRB | 702.80 | 638.40 | 435.45 | 0.62 | 678 |
| GNC | 486.70 | 410.66 | 311.86 | 0.64 | 297 |
| GTN | 370.15 | 255.00 | 277.43 | 0.75 | 281 |

Nevertheless, the accuracy of the effort assessment for SSF could be improved through some adjustments on the current applied sampling scheme. More specifically, currently random selection of the vessels to be monitored for their fishing effort, is conducted once per year, meaning that for every month of the year the same part of the fleet is monitored. This strategy, although applied due to cost, effort, time and personnel constrains, it violates the independency of the month-to-month samples, thus decreasing the accuracy of the effort assessment method. In any case, however, the statistically sound strategy, i.e. to select a different random sample each month, would be unfeasible under the current constrains described above.

A possible compromise between the intended accuracy and the availability of resources, would be the selection of a different, independent sample each quarter. To explore the possible benefits of this approach, we performed a bootstrap-simulation run in which a **different random sample of vessels (per vessel length category and main gear) is selected each quarter**. The sampling proportion remained equal to 0.05. The resulted density distributions per fishing gear are depicted in the figure 5 and the mean, median, standard deviation and CV of this simulation is presented in table 11. From table x we observe that CV values decrease substantially, (by 0.06 in GTR, 0.06 in GNS, 0.09 in LLS and 0.12 in FPO). By comparing CV values from tables x and y, we see that values in y are lower. **In other words, keeping the same sampling effort but increasing the independency of the samples would be more beneficial (and with less cost) than doubling the sampling effort (but keeping low samples independency)**. In any case, a cost-benefit analysis should be conducted in order to estimate whether it would be possible to increase the independency of our samples.

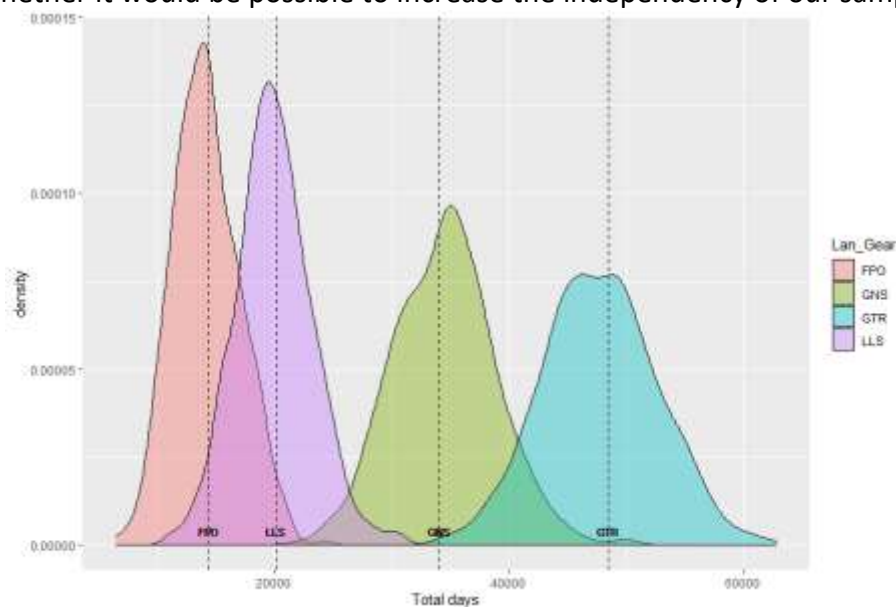


Figure 5. Density distribution of the calculated effort in 1000 iterations and by selecting different sample each quarter (sampling proportion =0.05), for the four main gears of the Greek SSF. In dashed lines the “real” numbers are depicted.

Table 11. Results of the simulation with sampling proportion =0.05 and by selecting different sample each quarter

| FishingGear | Meandays | Mediandays | sd | CV | "real" days |
|-------------|----------|------------|---------|------|-------------|
| TR | 47770.64 | 47707.14 | 4858.60 | 0.10 | 48510 |
| GNS | 34537.22 | 34586.48 | 4327.79 | 0.13 | 34076 |
| LLS | 19966.02 | 19832.12 | 3222.48 | 0.16 | 20094 |
| FPO | 14176.55 | 14070.10 | 2830.43 | 0.20 | 14337 |
| LLD | 1117.46 | 956.88 | 752.53 | 0.67 | 1035 |
| LHP | 871.32 | 743.39 | 568.11 | 0.65 | 767 |
| DRB | 772.79 | 676.67 | 519.45 | 0.67 | 678 |
| GNC | 431.88 | 310.00 | 387.22 | 0.90 | 297 |
| GTN | 385.43 | 246.00 | 320.68 | 0.83 | 281 |

Total landings per species calculation

Apart from the calculation of fishing effort, total landings per species for the SSF, since no declarative data are available. For landings estimation, PSU is total landings weight, per species/month/gear/vessel. The raw data consist of the number of total landings per month, per species per fishing gear and fishing days per month and per fishing gear for each vessel included in the sample. These data are collected through personal interview with the fishers. Here the stratification hierarchy is:

GSA → subArea → Main Gear → Vessel length → Month → Fishing Gear → Species

For this estimation, we use the approach of ratio estimators (from Cochran, 1977). First, within each stratum we calculate the ratio:

$$R_{gjs} = \frac{Lan_{fgjs}}{\bar{d}_{fgj}} \quad (11)$$

where Lan_{fgjs} are total landings per species s per fishing gear fg in stratum j and \bar{d}_{fgj} is mean fishing days per fishing gear fg and per stratum j .

The variance for R_{gjs} is equal to:

$$VAR(s_{fgj}^2) = \frac{1}{n_j \times \bar{d}_{fgj}^2} \times (s_{Lan_{fgjs}}^2 - 2R_{fgjs}Cov_{Lan\bar{d}_{fgj}} + R_{fgjs}^2 s_{\bar{d}_{fgj}}^2) \quad (14)$$

where n_j are total number of sample vessels in stratum j , \bar{d}_{fgj} are mean fishing days per fishing gear fg and per stratum j , $s_{Lan_{fgjs}}$ is standard deviation for Lan_{fgjs} , $Cov_{Lan\bar{d}_{fgj}}$ is covariance between Landings and mean days (see equation 15) and $s_{\bar{d}_{fgj}}$ is standard deviation for \bar{d}_{fgj} .

Covariance $Cov_{Lan\bar{d}_{fgj}}$ is calculated through:

$$Cov_{Lan\bar{d}_{fgj}} = \frac{1}{n_j - 1} \times \sum_{i=1}^{n_j} (Lan_{fgji} - \bar{Lan}_{fgj})(d_{fgji} - \bar{d}_{fgj}) \quad (15)$$

Finally, the total Landings per stratum is equal to:

$$TotL_{fgj} = R_{fgjs} \times TotalD_{fg} \quad (16)$$

Total D_{fg} is calculated with equation (9).

**1.b Hellenic Centre for Marine Research,
Estimation of fleet and stock related variables applied by the Institute for
Marine Biological Resources of and Inland waters of HCMR in the south
Aegean Sea (GSA22), Crete (GSA23) and the eastern Ionian Sea (GSA20)**

Contributors: *Stefanos KAVADAS, Danai MANTOPOULOU-PALOUKA,
George TSERPES, Dimitrios DAMALAS*

Effort

Estimation of effort exerted by the fleets is fundamental in the whole statistical estimation scheme of HCMR, since all other estimates (Landings, Discards, Biological parameters) depend largely on it.

Effort estimates are derived by applying the Fishing Effort Statistical Survey (FESS) and Fishing Zeal Statistical Survey (FZSS) methods described in detail in the manuscript of Bazigos & Kavadas (2007); they were also provided in the early versions of the Greek National Plans (Anonymous 2010).

Estimations of Effort in days at sea for small scale fisheries and for boats <12m are based on two tables:

- Fleet table
- Effort samples from a representative number of vessels up to vessel length category and activity level, randomly sampled through gears.

The final estimations of Effort are in terms of days at sea by month, area, vessel length category and gear.

Fleet table contains information about number of vessels (**Nvess**) by month (**m**), area (**a**), vessel length category (**cl**) and activity level (**ac**), (aggregation level **s**)

Effort sampling raw data in days at sea (**d**) by month (**m**), area (**a**), gear (**g**), vessel length category (**cl**) and activity level(**ac**)

Assign for each record (i) in Effort table an activity level as follows:

$$ac_i = \begin{cases} FT_i, & \text{if } \max(d) \geq 20 \\ PT_i, & \text{if } 20 > \max(d) > 10 \\ OC_i, & \text{if } \max(d) \leq 10 \end{cases}$$

note: the thresholds of FT, PT and OC are not fixed.

where 10, 20: thresholds for days at sea; FT : full time; PT : part time; OC: occasional. Although NT: null time i.e inactive vessels, exists as an activity level in Fleet table, there is no sampling for this category.

Note: Thresholds FT, PT and OC are assigned on an individual vessel level, which means that independent of the gear or gears being used, **a vessel is categorized based on the days at sea operating with the gear used more often.**

Step 1: Create aggregated table with average days at sea from Effort table as:

$$\bar{d} = \sum_j \frac{d_{pj}}{n_p}$$

where :

p is the aggregation level i.e. $p = (m, a, g, cl, ac)$, j each vessel in the aggregation level p and

n_p the number of vessels in aggregation level p .

Step 2: calculate the estimated effort by month, area, gear and vessel length category (aggregation level I):

$$E_l = \sum_{ac} \bar{d}_{p,ac} \times N_{vess_{s,ac}} \quad ac \in (FT, PT, OC)$$

Formulas for calculation of errors around the point estimates are provided in the Annex based on the classical textbook of Cochran (1977).

Note: For the trawlers (OTB), purse seines (PS), boat seines (SB) and other artisanal vessels having specific license to fish (e.g. large pelagic fishing), the fishing effort is estimated using ERS, as well as VMS data according to the method proposed by Kavadas et al., 2014, Maina et al., 2016)

Effort by rectangle (FDI table I)

Fishing effort distribution by: GFCM squares (at 0.5*0.5 degrees resolution), quarter, vessel length, fishing technique, gear type, sub region (GSA). Currently, only the spatial information for trawlers (OTB), purse seines (PS), boat seines (SB) and other artisanal vessels having specific license to fish (e.g. large pelagic fishing) can be provided because only those vessels are equipped with a positioning system. The next year, a method combining the estimated fishing effort and the fishing footprint index estimated by a Multi Criteria Decision Analysis (Kavadas et al., 2015) from small scale fisheries, will be used to distribute the spatial effort.

Required tables (main)

- fishing effort estimates (based on methods proposed by Kavadas et al., 2014, Maina et al., 2016),
- fishing_rectangle_2x2km,
- ers_rectangle (GFCM squares)

Algorithm

Step 1 - calculate the **sum(days_at_sea)** by GFCM squares, quarter, vessel length, fishing technique, gear type, GSA

Step 2 - join fishing effort estimates, fishing_rectangle_2x2km & ERS_rectangle

Evaluation of method accuracy

To assess the abovementioned applied method, we created a virtual fleet file and sampled randomly, following the sampling strategy described above. In this exercise, sampling intensity was set to 3% in order to calculate Days at Sea (DAS) per fishing gear for a specific selection of area and vessel length category over a period of 12 months.

A simulation loop applied the process 1000 times and the estimated effort was compared against the 'true' effort. Results are shown in the figure below in a form of density plots per gear.

It can be deduced that the method performs adequately, exhibiting low CVs and relative errors, at sample sizes $\geq 3\%$.

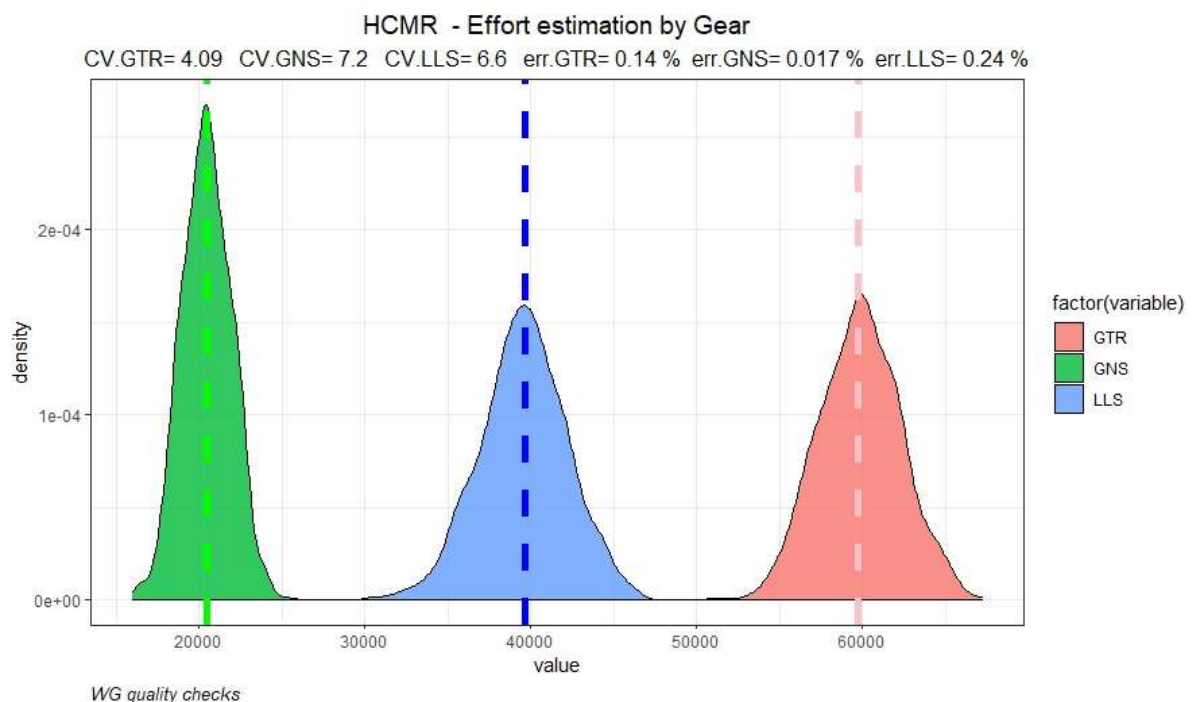


Figure 1. Density distribution of estimated effort in 1000 iterations, for three main gears sampled. 'True' values are overlaid as dashed vertical lines.
(GTR-trammel nets, GNS-gillnets, LLS-static longlines)

Landings

Estimation of landings (total) is derived from:

- landings samples from a representative number of vessels
- estimated fishing effort (*see 1. Effort*)

The goal is to calculate average weight landed per vessel (for a specific species, area, gear, vessel length class, quarter), which will be multiplied by the relevant effort.

The algorithm applies to all vessels <12m of length; (for vessels >=12m landings are obtained directly through [ERS](#))

Reference table in the DB:

raw data of landed weight (**weight**) and fishing days (**land_days**) by area **a**, gear **g**, vessel **v**, vessel length category **cl**, quarter **q**, species **s**

Step 1: Create aggregated table with aggregation:

r = (area, gear, vessel_length_category, quarter)

Step 2: calculate weight landed per day per vessel category:

$$w_r = \sum_{i \in r} \frac{weight_i}{land_{days}_i}$$

Step 3: Calculate number of all distinct vessels in **r**

$$Nves_r = \sum_{i \in r \forall vessel_i} 1$$

Step 4: Calculate total weight landed per day in **r**

$$t_r = \sum_{i \in r} w_i$$

Step 5: Calculate total weight landed per day per vessel in **r**

$$t_r = \frac{t_r}{Nves_r}$$

Step 6: Calculate total weight landed in **r**

$$Tland_r = t_r \times effort_r$$

Formulas for calculation of errors around the point estimates are provided in the Annex based on the classical textbook of Cochran and William (1977).

Note: to estimate landings at species level, the abovementioned aggregation **r** should include **species** column from the DB.

Landings by rectangle (FDI table H)

Landings distribution by: species, GFCM squares (at 0.5*0.5 degrees resolution), quarter, vessel length, fishing technique, gear type, sub region (GSA). Currently, only the spatial information from trawlers (OTB), purse seines (PS), boat seines (SB) and other artisanal vessels having specific license to fish (e.g. large pelagic fishing) can be provided because only those vessels are obligated to keep ERS.

Required tables

- ERS logbook data,
- ers_rectangle (GFCM squares),
- prices (by Kg)

Algorithm

Step 1: - create table spatial_landings as: calculate the **sum(land_weight)** by species, GFCM squares, quarter, vessel length, fishing technique, gear type, GSA from ERS logbook joining ERS logbook data & ers_rectangle

Step 2: - update spatial_landings set value:=calculated land_weight * price
join spatial_landings& prices

2. Discards

“The portion of the total organic material of animal origin in the catch which is taken out of the water, thrown away or dumped in the water for whatever reason”, is considered discard (Pérez Roda, 2019).

Based on the above definition, monitoring of discards requires the placement of on-board observers, as discards are not usually declared in the official forms handed to the authorities by the fishers (e.g. ERS reporting). The available data within the premises of Data Collection Framework are collected based the biological sampling scheme.

The algorithm (given below in pseudocode and mathematical notations) estimates the discard ratio (DR) as the fraction of discarded weight over the total weight of the catch (by species, area, fleet segment, vessel length category, year and quarter) and the relative error around the estimation.

Estimates of absolute values of discards in weight is derived at a later stage by taking into account the effort exerted and landings.

Reference table in the DB:

Raw data of catch (**Catch**) and discards (**Discards**) by species **s**, haul **h**, cruise **c**, vessel **v**, area **a**, quarter **q**, year **y**

Calculate for each record **i**:

$$x_i^2 = \text{catch}_i^2$$

$$x_i y_i = \text{catch}_i * \text{discards}_i$$

$$y_i^2 = \text{discards}_i^2$$

Calculate discard ratios for the preferred aggregation level (Official Data Call aggregation is by Species, GSA, Vessel length category, Gear, quarter):

Create aggregated table with aggregation:

p = (Species, GSA, Vessel Length Category, Gear, Quarter):

Discards Ratio:

$$DR_p = \frac{\text{discards}_p}{\text{catch}_p}$$

The following are needed for the estimation of errors¹ around the DR estimate (see Annex)

$$Nb_p = \sum_{i \in p} i$$

$$\text{catch}_p = \sum_{i \in p} \text{catch}_i$$

$$\text{discards}_p = \sum_{i \in p} \text{discards}_i$$

$$x_p^2 = \sum_{i \in p} x_i^2$$

¹ Formulas for calculation of errors around the point estimates are provided in the Annex based on the classical textbook of Cochran (1977).

$$y_p^2 = \sum_{i \in p} y_i^2$$

$$x_p y_p = \sum_{i \in p} x_i y_i$$

Often, it is required to report a single per year and fishing gear Discard Ratio per species. Taken into account that, as stated above, Discards ratio is calculated at lower aggregation level, DR_y , discard ratio per species/fishing gear/year is calculated as the weighted mean of DR_p :

$$DR_y = \sum DR_p \times \frac{Lan_p}{Lan_y}$$

Where Lan_p are total landings per stratum p , while Lan_y are total landings per species, fishing gear and year.

3. Biological Data

Age-Length Key (ALK)

Needs estimated length frequency distribution (*see 5. Demographics*).

The goal is to allocate age readings to length classes by constructing a table of relevant proportions. ALK is later on used to derive age frequency distribution of the catch.

Reference table in the DB:

raw data of age readings (**age**) by
area **a**, vessel **v**, vessel length category **cl**, quarter **q**, species **s**, sex **x**

Create aggregated table with aggregation:

- $p = (\text{area, vessel_length_category, quarter, species, age, length, sex})$
- $r = (\text{area, vessel_length_category, quarter, species, sex})$

Step 1: Calculate the number of records in p (Nb of individuals with age i by length bin j):

$$N_{i,j} = \sum_{i \neq \emptyset} Nb_{i,j}$$

Step 2: Calculate the number of records in r (Nb of individuals with age i over all ages in each length bin j):

$$n_j = \sum_i n_{i,j}$$

Step 3: Calculate the proportion of each age to each length bin:

$$p_{alk} = p_{i,j} = \frac{n_{i,j}}{n_j}$$

where in the following table we have:

$$\sum_{i=1}^N p_{i,j} = 1$$

We have now reached a table of the following form:

| (i, j) | Age _i | Age _{i+1} | Age _{i+2} | ... | Age _N | Totals |
|-----------------------|------------------|--------------------|--------------------|-----|------------------|--------|
| Length _j | $p_{i,j}$ | | ... | ... | $p_{i,N}$ | |
| Length _{j+1} | | $p_{i+1,j+1}$ | ... | ... | ... | |
| ... | | | ... | ... | ... | |
| Length _k | | | ... | ... | $p_{k,N}$ | |
| Totals | | | | | | |

Which is to be used in the construction of the Age frequency distribution (see 5. *Demographics*)

4. Demographics

Involves statistical study of population characteristics in order to reason on differences/changes occurring over time.

Length frequency distribution

Estimation of length frequency distribution is derived from:

- measurements taken both on board and at landing locations
- estimated landings (see 2. *Landings*)
- estimated discard ratio (see 3. *Discards*)
- length-weight relationships

The goal is to allocate measured individual specimen to appropriate length classes (for a specific species, area, gear, vessel length class, quarter) and construct a size frequency table raised to the total catch.

Reference table in the DB:

raw data of length measurements (**length**) by
 area **a**, gear **g**, haul **h**, date **d**, vessel **v**, vessel length category **cl**, quarter **q**,
 *market_category **m**, species **s** (*: market_category = 'D' for discarded, 'L' for landed)

Create aggregated table with aggregation:

- p =(area, gear, vessel_length_category, quarter, market_category, species)
- r =(area, gear, vessel_length_category, quarter, market_category, species)

Step 1: Set an appropriate length class l_{cl} (depending on the species and data call requirements e.g. 2 mm) and calculate number of length classes $N_{classes}$ and all median length classes $med-l_{cl_i}$

$$N_{classes} = \frac{l_{cl} = constant}{\frac{max(length) - min(length)}{l_{cl}}}$$

Step 2: Calculate the median length classes i :

$$med_{i \in N_{classes}} = min(length) + (l_{cl} * (i - 1))$$

Step 3: Calculate the number of individuals measured in each length class $N_{bl_{cl_i}}$

$$N_{bl_{cl_i}} = \sum_{j \in N_{bl_{cl_i}}} j$$

Step 4: Calculate weight at given length class i

(where a, b are the coefficients of the length weight relationship $W = a * L^b$)

$$ind_{weight_i} = a * med_i^b$$

Step 5: Calculate total weight of each length class i

$$w_{cl_i} = ind_{weight_i} * N_{bl_{cl_i}}$$

Step 6: Sum weights over all length classes in aggregation p

$$w_{cl} = \sum_{i \in p} w_{cl_i}$$

Step 7: Calculate the percentage contribution of each length class weight in summed total weights

$$perc_{cl_i} = \frac{w_{cl_i}}{w_{cl}}$$

Step 8: Calculate the biomass at each length class i (multiply **percentage contribution** of each length class to total by **total landings weight** for each length class - from landings table)

$$biomass_{cl_i} = perc_{cl_i} * tot$$

Step 9: Calculate number of individuals **landed** per length class i

$$N_i = \frac{biomass_{cl_i}}{ind_{weight_i}}$$

Step 10: Generate the final output by assigning the number of individuals landed (N_i) to each corresponding length class (med_i)

| Length classes | Nb of individuals |
|----------------|-------------------|
| med_i | N_i |
| med_{i+1} | N_{i+1} |
| ... | ... |
| med_n | N_n |

For the calculation of the **discarded** length distribution we need the discard ratio (**DR**) calculated in 3. *Discards*.

Step 1: Sum weights over all length classes in aggregation p

$$lnd_{biomass\ p} = \sum_{i \in p} biomass_{cl\ i}$$

Step 2: Calculate the total discarded biomass

$$disc = \frac{lnd_{biomass\ p}}{1 - DR} - lnd_{biomass\ p}$$

Step 3: Calculate the discarded biomass at each length class i (multiply **percentage contribution** of each length class to total by **total discarded weight** for each length class). The % contribution of each length class discarded weight is calculate as previously for the landings

$$discards_{cl\ i} = perc\ _i * disc_{biomass\ p}$$

Step 4: Calculate number of individuals **discarded** per length class i

$$N_i = \frac{discards_{cl\ i}}{ind_{weight\ i}}$$

Step 5: Generate the final output by assigning the number of individuals discarded (N_i) to each corresponding length class (med_i)

Age frequency distribution

Estimation of age frequency distribution is derived from:

- Age Length Key (see 4. *Biological data*)
- Length frequency distribution (see 5. *Demographics*)

The goal is to allocate measured individual specimen to appropriate age classes and construct a size frequency table raised to the total catch.

Reference tables in the DB (join):

- Length frequency distribution table (LFD) by area **a**, year **y**, species **s**
- ALK table

Create aggregated table with aggregation:

- **p =(area, year, species, age, length_class)**

Calculate numbers@age class using ALK and LFD.

let N_j be the number of individuals in length class j in LFD table, then we calculate N_i (number of individuals in age class (i)) as:

$$N_i = \sum_j p_{i,j} N_j$$

where j is overall length classes.

Conclusion

This document provided in detail the approaches employed by the Institutes involved in the realization of DCF in Greece to estimate fleet and stock related variables to service the Mediterranean Data Calls. Furthermore, it assessed the accuracy and precision of the aforementioned implementations by employing a series of simulations.

The distinct characteristics of the small-scale fleets in the marine regions in which each Institute is responsible, led to somehow different approaches for the estimation of *Effort* and *Landings*. Estimation of *Discards*, *Biological* data and *Demographics* follow identical procedures.

Since estimation of *Effort* is key to all other estimates, we provide below a synoptic comparative table of the two approaches.

Table A. Comparison table summarizing the two approaches employed by each Institute to estimate *Effort*.

| | FRI | | HCMR | |
|-----------------------------------|--|---|--|---|
| Stratification | <ul style="list-style-type: none"> Area (a) Main gear (mg) Vessel length (l) Month (m) Fishing gear (fg) | | <ul style="list-style-type: none"> Area (a) Vessel length (cl) Month (m) Activity level (ac) | |
| Sampling | <p>Population = Nat. Fleet Register (NFR)</p> <p><i>Random sampling</i> samples within each stratum are proportional to the total number of vessels within the stratum</p> $k = (a, v, m, mg)$ | | <p>Population = Nat. Fleet Register (NFR)</p> <p><i>Random sampling</i> samples within each stratum are proportional to the total number of vessels within the stratum</p> $t = (a, cl, m)$ | The gears appearing in the sample are considered representative of the population |
| Active vs Inactive vessels | <p>Variable activity</p> <p>Activity ratio per stratum based on field survey</p> | | <p>Four different activity levels (based on days per month)</p> | $ac = \begin{cases} FT, \text{ if } \max(d) \geq 20 \\ FT, \text{ if } 20 > \max(d) > 10 \\ OC, \text{ if } \max(d) \leq 10 \\ NT \text{ (inactive vessels)} \end{cases}$ |
| Main gear vs fishing gear | <p>assumption that vessels under the same area, vessel length and main registration gear exhibit similar patterns of fishing activity</p> | | <p>the assumption is that the primary (declared) fishing gear only accidentally coincides with the actual fishing gears that are being used</p> | |
| Steps | <ol style="list-style-type: none"> 1. Calc Activity Ratio 2. Calc active vessels 3. Calc mean days at sea per stratum j 4. Calc total effort per fishing gear in j 5. Calc total effort of all gears in j | $A_j = \frac{AC_j}{AC_j + IN_j}$ $N_j = A_j \times NFR_j$ $\bar{d}_{fBj} = \frac{1}{n_j} \times \sum_{i=1}^{N_j} d_{fBj}$ $D_{fBj} = N_j \times \bar{d}_{fBj}$ $TotalD_{fB} = \sum_i D_{fBj}$ | <ol style="list-style-type: none"> 1. Calc mean DAS per vessel j at level p 2. Calc DAS at level l | $\bar{d} = \sum_i \frac{d_{p_j}}{n_p}$ $E_l = \sum_{ac} \bar{d}_{p,ac} \times N_{vess_{z,ac}}$ |

The areas in which each Institute is responsible for, encompass specific characteristics that led to the adoption of diverse approaches for the estimation of *Effort* and *Landings*: the stratification of northern-central Aegean Sea (FRI approach) is based on the registered Primary Gear, while for the southern Aegean and the Ionian Seas (HCMR approach) it is based on the activity level.

Nevertheless, outcomes of the simulation tests suggest that both methods:

- (i) perform quite well for adequate sample sizes
- (ii) converge to similar results

However, since the different methods are implemented based on the specific characteristics of the corresponding fisheries, it is advised that North Aegean data should be processed based on FRI approach, while South Aegean and Ionian Sea data should be processed based on HCMR approach, in order to have the most accurate outcomes.

References

- Anonymous, 2010. GREECE National Fisheries Data Collection Program 2011-2013. (In application of EC Decision 93/2010). Ministry of Rural Development and Food Directorate General for Fisheries. 63 pp.
- Anonymous, 2019. GREECE - Sampling scheme and Data Quality Assurance Framework. National Data Collection Programme 2019. 13 pp. Available at: https://inale.gr/wp-content/uploads/2019/10/Sampling_scheme_data_quality.pdf
- Bazigos, G. & Kavadas, S., 2007. Optimal sampling designs for large-scale fishery sample surveys in Greece. *Mediterranean Marine Science*, 8(2), 65-82.
- Cochran, W.G., 1977. *Sampling Techniques*. New York: John Wiley & Sons.
- EC, 2013. Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC
- EC, 2017. Council Regulation (EC) No 2017/1004 of 17 May 2017: on the establishment of a Union framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the common fisheries policy and repealing Council Regulation (EC) No 199/2008 (Recast).
- EC, 2019. Commission Delegated Decision (EU) 2019/910 of 13 March 2019: establishing the multiannual Union programme for the collection and management of biological, environmental, technical and socioeconomic data in the fisheries and aquaculture sectors (applies from 1 January 2020)
- Kavadas, S., Damalas, D., Tserpes, G., Georgakarakos, E., Papaconstantinou, C., Maravelias, C., 2013. IMAS-Fish - Integrated Management System to support the sustainability of Greek fisheries resources. A multidisciplinary web-based database management system: implementation, capabilities, utilization and future prospects for fisheries stakeholders. *Mediterranean Marine Science* 14(1): 109-118.
- Kavadas, S., Barberá C., Belardinelli A., Carpi P., Cataudella S., Croci C., D'Andrea L. et al. 2014. Common methodological procedures for analysis of VMS data, including web-based GIS applications related to the spatial extent and intensity of fishing effort. PERSEUS Project Report. 40 pp. +annexes; ISBN: 978-960-9798-14-3
- Kavadas, S.; Maina, I.; Damalas, D.; Dokos, I.; Pantazi, M.; Vassilopoulou, V. Multi-Criteria Decision Analysis as a tool to extract fishing footprints and estimate fishing pressure: Application to small scale coastal fisheries and implications for management in the context of the Maritime Spatial Planning Directive. *Mediterr. Mar. Sci.* 2015, 16, 294–304.
- Maina, I., Kavadas S., Katsanevakis S., Somarakis S., Tserpes G., Georgakarakos S. 2016. A methodological approach to identify fishing grounds: a case study on Greek trawlers. *Fisheries Research*, 183: 326–339.
- Pérez Roda, M.A. (ed.), Gilman, E., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M. and Medley, P. 2019. A third assessment of global marine fisheries discards. *FAO Fisheries and Aquaculture Technical Paper No. 633*. Rome, FAO. 78 pp.
- Scientific, Technical and Economic Committee for Fisheries (STECF) – Assessment of balance indicators for key fleet segments and review of national reports on Member States efforts to achieve balance between

fleet capacity and fishing opportunities (STECF-19-13), Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-11286-0, doi:10.2760/300448, JRC119006

Annex to estimation of fleet and stock related variables applied in the Institute for Marine Biological Resources and Inland waters

Confidence Intervals for algorithms output estimations

1 Regression coefficients (a & b)

$$Y = a + b \cdot X$$

$$CI_a = a \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) \cdot s_a$$

$$CI_b = b \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) \cdot s_b$$

where:

$$s_b^2 =$$

$$s_a^2 = s_b^2 \left[\frac{n-1}{n} s_x^2 + \bar{x}^2 \right]$$

$$s_x = \sqrt{\frac{1}{n-1} \sum_i (x_i - \bar{x})^2} \quad s_y = \sqrt{\frac{1}{n-1} \sum_i (y_i - \bar{y})^2}$$

$n =$

Number of observations

$\bar{x} =$

average of x_i

$t_{\alpha, \nu}$ or $Z_{\alpha}:$

values from Table, depending on significance level α and degrees of freedom $\nu = n-1$

2 Average or mean \bar{x}

$$\bar{x} = \frac{\sum x_i}{n}$$

$$CI = \bar{x} \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) \cdot s_x$$

s_x as defined in 1. above

3 Ratios (R)

Note:

R is an estimate derived from the ratio:

$$R = \frac{\bar{y}}{\bar{x}} = \frac{\sum_i y_i}{\sum_i x_i}$$

$$CI = R \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) \cdot s(R)$$

where:

$$s(R) = \frac{1}{\bar{x}\sqrt{n}} \sqrt{\frac{\sum_i (y_i^2) - 2R \sum_i (y_i x_i) + R^2 \sum_i (x_i^2)}{n-1}}$$

\bar{x} , nas defined in 1. aboveo 1.

4 Proportion (ρ)

$$p = \frac{x}{n}$$

$$CI = \rho \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) * s_{\rho}$$

$$s_{\rho} = \sqrt{\frac{p(1-p)}{n-1}}$$

If weneed an estimate of x then:

$$CI = x \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) * s_{\rho} * n$$

5 Total (τ)

$$\tau = N\bar{x} = \frac{N \sum x_i}{n}$$

$$CI = \tau \pm (t_{\alpha, \nu} \text{ or } Z_{\alpha}) * s_{\tau}$$

where: $s_{\tau} = \sqrt{N^2 \frac{s^2}{n}}$

Table I. Critical Values of the Student's-t Distribution for $v = n-1$ degrees of freedom at significance level (α).

| | v | α | | |
|-------------------------|----------------------------|----------|--------|--------|
| | | 0.10 | 0.05 | 0.01 |
| t - Distribution | 1 | 6.314 | 12.706 | 63.656 |
| | 2 | 2.920 | 4.303 | 9.925 |
| | 3 | 2.353 | 3.182 | 5.841 |
| | 4 | 2.132 | 2.776 | 4.604 |
| | 5 | 2.015 | 2.571 | 4.032 |
| | 6 | 1.943 | 2.447 | 3.707 |
| | 7 | 1.895 | 2.365 | 3.499 |
| | 8 | 1.860 | 2.306 | 3.355 |
| | 9 | 1.833 | 2.262 | 3.250 |
| | 10 | 1.812 | 2.228 | 3.169 |
| | 11 | 1.796 | 2.201 | 3.106 |
| | 12 | 1.782 | 2.179 | 3.055 |
| | 13 | 1.771 | 2.160 | 3.012 |
| | 14 | 1.761 | 2.145 | 2.977 |
| | 15 | 1.753 | 2.131 | 2.947 |
| | 16 | 1.746 | 2.120 | 2.921 |
| | 17 | 1.740 | 2.110 | 2.898 |
| | 18 | 1.734 | 2.101 | 2.878 |
| | 19 | 1.729 | 2.093 | 2.861 |
| | 20 | 1.725 | 2.086 | 2.845 |
| | 21 | 1.721 | 2.080 | 2.831 |
| | 22 | 1.717 | 2.074 | 2.819 |
| | 23 | 1.714 | 2.069 | 2.807 |
| | 24 | 1.711 | 2.064 | 2.797 |
| | 25 | 1.708 | 2.060 | 2.787 |
| | 26 | 1.706 | 2.056 | 2.779 |
| | 27 | 1.703 | 2.052 | 2.771 |
| | 28 | 1.701 | 2.048 | 2.763 |
| | 29 | 1.699 | 2.045 | 2.756 |
| | 30 | 1.697 | 2.042 | 2.750 |
| | Normal Distribution | >30 | 1.645 | 1.960 |