Northwest Atlantic Fisheries Organization

Serial No. N7145 **NAFO SCR Doc. 20/069**

SCIENTIFIC COUNCIL MEETING – NOVEMBER 2020

Improvements in the methodology to study the bottom fishing footprint in the NRA using VMS and logbook data

By

Sacau, M.[1], Durán-Muñoz[1], P., Garrido, I[2] and Baldó, F[3].

[1] Instituto Español de Oceanografía (IEO). C. O. Vigo. Subida a Radio Faro, 50. 36390 Vigo. Spain
[2] Organización de Productores de buques congeladores de merlúcidos, cefalópodos y especies varias (OPPC-3), Vigo, Spain
[3] Instituto Español de Oceanografía (IEO). C. O. Cádiz. Puerto Pesquero, Muelle de Levante, s/n. 11006 Cádiz, Spain

Abstract

To understand the extent of fishing activities within NAFO Regulatory Area (NRA), a characterization of distribution of fishing effort for a four-year period (2016 to 2019) was carried out. This characterization was done on the basis of two data sources: Haul by haul logbook information and Vessel Monitoring System (VMS) data. The results indicate that logbook data and VMS are complementary and the coupling of both datasets is a powerful methodology for describing the spatial distribution of fishing activity.

The objective of this document is to describe some problems detected on the original "*coupling VMS with Logbook data"* methodology (NAFO, 2017) and suggest the improvements developed in order to tackle them by the implementation of a new "*coupling VMS with Logbook data"* methodology. Problems detected in the quality of VMS and Logbook data are also raised by this study.

Keywords: Fishing effort, logbook data, NAFO, Vessel Monitoring System (VMS), Vulnerable Marine Ecosystems (VMEs).

1. Introduction

During the 10th NAFO Working Group on Ecosystem Science and Assessment (WGESA) meeting the "*coupling VMS with Logbook data"* original methodology (NAFO, 2017) was described and presented to make a characterization of distribution and intensity of fishing effort from 2016 onwards. This methodology was possible to put in place since in 2016, a new logbook data format was implemented as an improvement over 2015, by including fishing timestamps, geographic coordinates for gear deployment and retrieval, as well as the catch and discard weight for each species caught.
In 2020 when working with the original methodology (NAFO, 2017) in additional studies on fishing effort (Garrido *et al*., 2020), some technical problems were detected. The present document describes the original methodology and its benefits and presents some improvements made on it to tackle several issues detected in the original methodology.

2. Material and methods: Characterization of fishing effort in the NAFO Regulatory Area

The distribution and intensity of fishing effort during the period 2016-2019 in the NAFO Regulatory Area were estimated based of two data sources: Vessel Monitoring System (VMS) and logbook information data.

2.1 Vessel Monitoring System (VMS)

The NAFO *Vessel Monitoring System (VMS)* is a satellite-based monitoring system that provides data on the location, heading and speed of licensed fishing vessels. The transmission of such data occurs approximately every hour, called a "ping", providing high resolution positions recorded at higher frequencies when compared to logbook reporting.

VMS data used in this work were obtained from NAFO Secretariat who has responsibility for collecting and maintaining VMS data from fishing vessels in the NAFO Regulatory Area. In addition to be an integral part of NAFO´s Monitoring, Control and Surveillance (MCS) scheme, the VMS data is also used for scientific purposes, e.g. for the assessment of Significant Adverse Impacts (SAIs) on Vulnerable Marine Ecosystems (VMEs) and fish stock assessments.

VMS data includes the following information: NAFO Vessel Identification; Flag State; Radio (vessel call sign); UTC date and Time of the vessel position; vessel position by latitude and longitude; speed and heading.

2.2 Haul-by-haul catch data (logbook data)

Haul-by-haul catch data is logbook data collected during fishing vessel activities. Specifically, timestamps and geographic coordinates for gear deployment and retrieval are recorded, as well as the catch and discard weight for each species caught. This data format was implemented in 2016, and is an improvement over 2015 where the data was recorded only for the top three species by weight and did not include fishing time-stamps.

Haul-by-haul logbook data used for this work was also provided by NAFO Secretariat. It provides details for each vessel on catch and discard characteristics, date, type of gear used, and geographic position collected during fishing vessel activities. The collection of these data is the responsibility of the skipper of each vessel.

2.3 Methodologies: Simple speed filter vs Coupling VMS and Logbook data

While applying a “simple speed filter” is a very common method for extracting VMS points associated with fishing, there will inevitably be some points that are misclassified at a rate that is difficult to quantify. In previous years (NAFO, 2015), a simple speed filter of 1 – 5 knots (rounded to the nearest integer) was used to filter VMS points and assign them as fishing activities, but this presented challenges in terms of threshold speeds across entire fleets.

Logbook data and VMS are complementary and the coupling of both datasets has already proven powerful for describing the spatial distribution of fishing activity at a much finer resolution (NAFO, 2017). Figure 1 illustrates the flowchart with the main steps involved in the procedure of linking VMS with logbook data. The entire framework is a modular structure where each step has been developed in open-source statistical computing environment R (R Code Team, 2020).

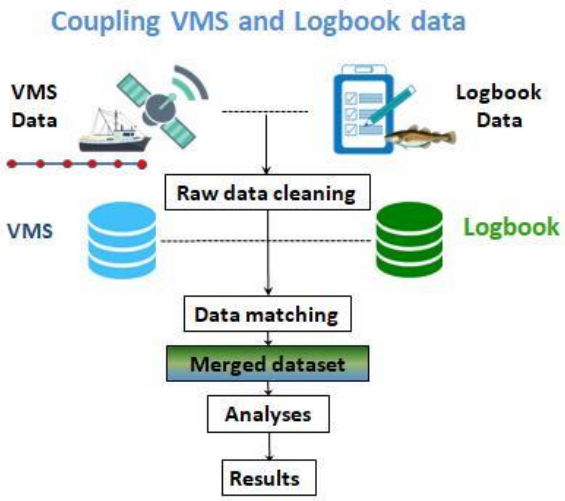

Figure 1. Flowchart with the main steps involved on the procedure of coupling VMS and logbook data

The first important step is "Raw Data Cleaning". In many instances, both VMS and logbook data contain erroneous entries namely: points with incomplete timestamps; wrong vessel positions; duplicated records; headings outside a compass range, etc. These errors should be removed or flagged.

Once the cleaning has been performed both datasets are ready for the "Data Matching" by using the NAFO Vessel ID and the Date as common fields between both databases. This step is particularly important as all subsequent analyses depend on the success of the linking. From the "Merged dataset" we can start to do the "Analyses" and get the final "Results".

Use of the haul-by-haul data permits VMS pings to be assigned as "fishing" or "non-fishing" based on whether or not they fall within fishing time intervals reported in the haul-by-haul catch data (match in time window, see Figure 2). That is, start and end of fishing timestamps from the logbooks are used to extract relevant VMS points which are then mapped in space to represent fishing effort. Because these VMS points are directly within the reported fishing times interval, they are considered to be associated with fishing activity.

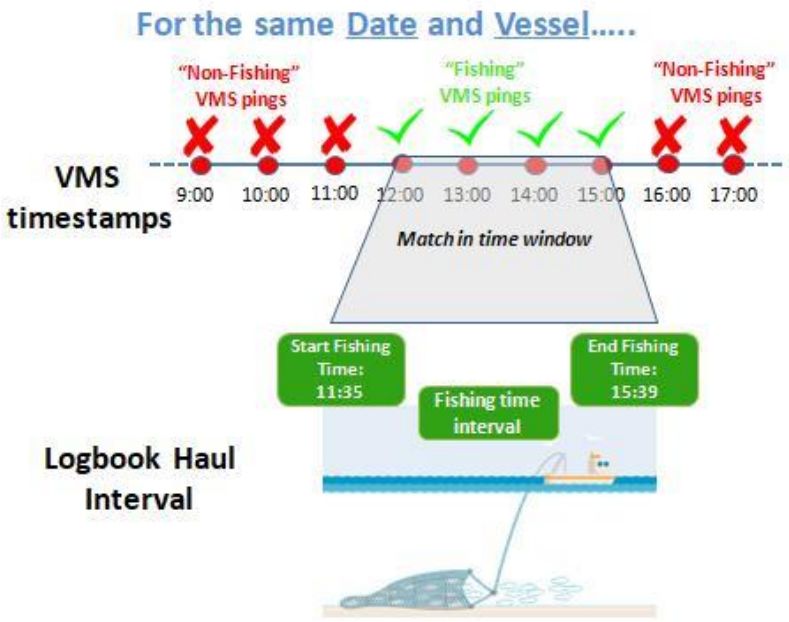

Figure 2. Match in time window procedure

Through this analysis, fishing footprint layers were created for fisheries-specific and cumulative fishing effort using VMS data and new haul-by-haul catch data (logbook).

To create fishery-specific effort maps, VMS points were assigned to a fishery based on the species with the highest retained catch weight in the logbook during the corresponding logbook fishing time interval. This definition of fishery is based solely on the main species in the catch and in some cases, the main species may differ from the main species sought.

Filtered VMS points were assigned a "ping-time" interval to represent the duration of fishing. This value was calculated as the forward difference in time between VMS points. Typically, ping intervals were approximately one hour, so if the interval exceeded 2 hours, it was assigned to be 2 hours to avoid inflating effort within a cell. The last VMS point in a vessel's series was assigned the mean ping-time interval for that vessel. The VMS points were aggregated over a 0.05 x 0.05 degree grid and the ping-time intervals were summed to represent the hours fished in each cell. However, Garrido *et al*., 2020 showed that around 3% of the sets have under or overestimation pings problems and 25% of the received pings have frequencies different to one hour.

2.4 Original "Coupling VMS and Logbook data" methodology: Benefits

Outputs from "*simple speed filter*" and "*coupling VMS with Logbook data*" methodologies were compared side by side and visually examined for congruence.

Overall, the areas represented by the "*coupling VMS with Logbook*" original method and the "*simple speed filter*" old method showed fishing activities in the same general areas with similar patterns of intensity (NAFO, 2017). However, the footprint from the "*coupling VMS with Logbook*" original method was considered an improvement because it tended to have fewer spurious points outside of the main footprint area. With the original "*coupling VMS with Logbook"* method, there were also fewer cells displaying fishing effort within the vulnerable marine ecosystem (VME) closures, and if we assume the closures are being respected, this would indicate that the *"simple speed filter"* old method over

represents fishing effort in some cells, particularly where effort appears to be low. In the "*coupling VMS with Logbook*" original filtered maps there were still some points outside of the NAFO fishing footprint, in deep waters, likely due to VMS points associated with "steaming". This probably occurred because of an incorrect start/end time.

2.5 New methodology: Issues found in the "coupling VMS with Logbook" original method

Despite the fact that the "*coupling VMS with Logbook*" original methodology has been shown to improve the description of the spatial distribution of fishing activity at a much finer resolution, some problems were detected in some steps of the original methodology. The following improvements were implemented with the aim to tackle such problems.

2.5.1 Calculation of bottom longline footprint

Bottom trawl and bottom longline fishing gears can produce negative impacts on VMEs, but technical and operative characteristics of both gears are very different. Consequently, the parameters needed to describe their footprints and associated impacts are very different too. For this reason, trawl and longline cumulative fishing effort need to be calculated separately.

Despite these issues, the original "coupling VMS with Logbook" methodology (NAFO, 2017) calculates the *"cumulative fishing effort"* taking into account available data from bottom trawlers and bottom longliners. It also calculates the *"longline specific fisheries footprint"* (e.g. Atlantic halibut and cod) taking into account available data from bottom longliners only. Moreover, information from longline activity currently available in the logbooks is insufficient to describe appropriately the longline footprint[1] and the associated impacts.

According to the discussions of the SC 2020 Annual Meeting, "... *in the case of longline fisheries, collection and compilation of additional information would be crucial to start the process of defining a more precise fishing bottom longline footprint..... since with the information that is currently available, it is not possible to obtain the real footprint for this fishery"*. As a consequence, currently is not possible to analyze adequately such footprint in the NRA until the necessary information be available. Therefore, *"cumulative fishing effort"* calculated should include only pings from trawl fisheries, and *"longline specific fisheries footprint"* cannot be calculated adequately. In line with this, the new proposed methodology removes VMS pings from longliners for the calculation of the *"cumulative fishing effort"* analysis.

At present time, taking into account the available data and the improvements of the new methodology, only the *"bottom trawl cumulative fishing effort"* and *"bottom trawl specific fisheries footprint"* can be calculated adequately.

2.5.2 Missing "Fishing VMS pings"

Garrido *et al*., 2020 found that with "*coupling VMS with Logbook*" original methodology (NAFO, 2017) there were many missing pings. This is due to the existence of many hauls with start fishing time one day and finish fishing time next day (e.g. Haul 1; Start day: 2 January; Start time 23:45; End day: 3 January; End Time 04:00). In these cases, VMS pings from the second day of the same haul were not taken into account and therefore missed from the analysis. This is not unusual so, the new proposed methodology takes into account those missing VMS pings considering all "fishing pings" comprised

[1] Meanwhile in trawl fisheries is enough to know the start and the end of the haul, in longline fisheries is necessary to know the start and end of the line set as well as the start and end of the line haul. Currently, in the logbook only information on the start and end of the haul is recorded, being not possible to know the exact location where the longline was deployed over the seabed.

between start and end of the haul, including when the dates for the same haul and vessel are not the same day. Solving this issue is really important as all subsequent analyses depend on the success of the linking with the selection of all the "fishing pings".

2.5.3 Bottom trawlers fishing speeds

It has been noted that with original methodology there were many vessels with very high speeds. New proposed methodology applies a speed filter in order to remove those vessels with speeds equal or higher than 6 knots, as it is considered that those are non trawler fishing speeds. New proposed methodology considers that bottom trawlers are classified as "fishing" at speed intervals lower than 6 knots. Other speeds should be classified as "steaming".

3. Results and Discussion

Maps below (Figure 3), represent on a yearly basis (from 2016 to 2019), the distribution of "fishing VMS pings" obtained with the original "*coupling VMS with Logbook*" methodology (left map; red dots) and the new "*coupling VMS with Logbook*" (right map; green dots).

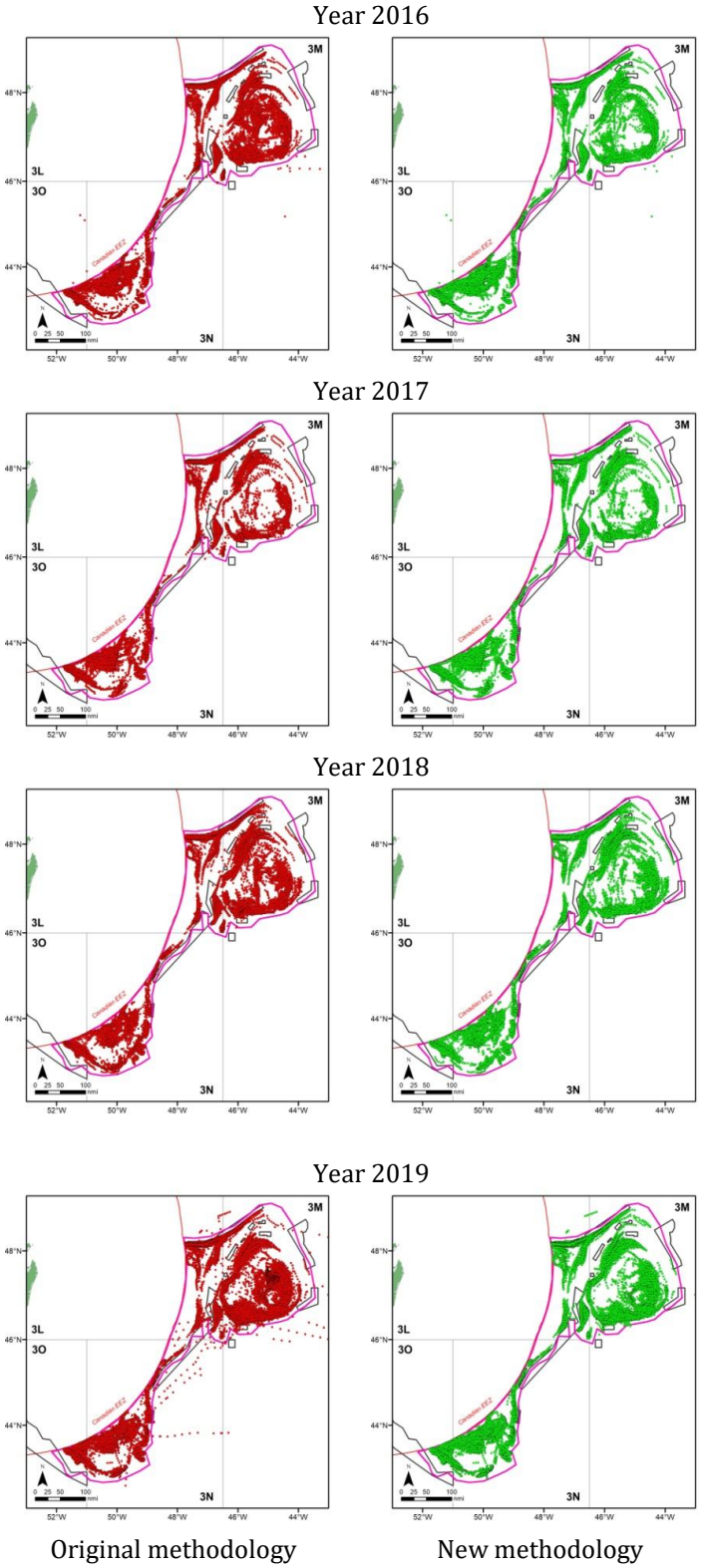

Figure 3. 2016 to 2019 Fishing VMS pings (Left: original "*coupling VMS with Logbook*" methodology Right: new "*coupling VMS with Logbook*" methodology). Red line: EEZ; Purple line: Footprint polygon; Black line: High Sponge and Coral Concentration Area Closures

Overall, both methods showed "Fishing VMS pings" in the same general areas. However, new "*coupling VMS with Logbook*" methodology separates "fishing VMS pings" corresponding to longliners , mainly located in Division 3M (Atlantic Cod fishery) and to a lesser extent in the Tail of the Bank, Division 3N (Atlantic Halibut fishery). This new methodology also eliminates fishing VMS pings with speeds equal or higher than 6 knots as they are considered as non fishing pings.

The following table presents the comparison of ping numbers calculated by both "*coupling VMS with Logbook*" methodologies (original *versus* new), missing pings (not taken into account by the original methodology), pings for longliners (removed) and pings with speeds higher or equal to 6 knots (removed).

New methodology **(B)** tackles the three aspects mentioned earlier: 1) separates fishing VMS pings corresponding to longliners **(D)**; 2) retrieve missing VMS pings **(C)** and; 3) applies a filter that removes speeds equal or higher than 6 knots **(E)**. According to table 1: **B=A+C-D-E**

Table 1. Number of "VMS pings" for both methodologies (A and B), where B=A+C-D-E

	2016	2017	2018	2019
A. Total pings original methodology	34 565	26 845	45 150	58 697
B. Total pings new methodology	36 002	29 649	48 397	57 413
C. (+) Missing pings	5104	4709	8013	6686
D. (-) Pings for Longliners	2821	995	3561	5563
E. (-) Pings with speeds ≥ 6 knots	846	910	1205	2407

Except for year 2019, Table 1 illustrates that even after applying the filter for removing longliner VMS fishing activity and implement a filtering for speeds higher than 6 knots, number of "Fishing VMS pings" is considerably higher when using the new "*coupling VMS with Logbook*" methodology. This means that the original "*coupling VMS with Logbook*" methodology was underestimating the number of "Fishing VMS pings" in the studied area.

Table 2. Number of pings per year that were obtained in the original "coupling VMS with Logbook" methodology, number of missing pings and percentage of pings that were underestimated with the original methodology.

	Total pings original methodology	Missing pings	Perc. Diff.
2016	34 565	5104	14.7% underestimation
2017	26 845	4709	17.5% underestimation
2018	45 150	8013	17.7% underestimation
2019	58 697	6686	11.4% underestimation
TOTAL	**165 257**	**24 512**	**14.8% underestimation**

Table 2 shows that, for the 2016 to 2019 period, there is a ping underestimation of 14.8%, very similar to the 14.4% that was calculated by Garrido *et al.*, 2020. Moreover, the original "*coupling VMS with Logbook*" methodology was taking into account longliner VMS pings when calculating the cumulative fishing effort maps. In the new "*coupling VMS with Logbook*" methodology "Fishing VMS pings" from

trawlers and longliners were considered separately. These issues are particularly important because they affect all the subsequent analyses like the cumulative fishing intensity calculations that are summing pings and applied to a 0.05 x 0.05 degree grid. In this sense, it is important to bear in mind that the impact of the bottom longline is related with the location where the longline was deployed over the seabed and current available data is insufficient to describe this appropriately.

Moreover, many issues must be improved in terms of quality of data. The analysis made by Garrido *et al*., 2020 for testing the coverage of the merging VMS and logbook method based on the effort compiled by the Spanish Scientific Observers shows that around 20-30% of VMS pings are misreported. This significant amount of missing data and high level of misreporting may occur from unknown failure of the VMS device. Furthermore, with regards to logbook data, it was found a very large number of records with errors in date and time (e.g. start time was later than end time) as well as misreported hauls or fishing trips (e.g. last hauls of a fishing trip are missing or several hauls are grouped in a single haul). According to Garrido *et al*., 2020 this error occurs in approximately 3% of the total hauls for the 2016-2019 period.

4. Conclusions

The new "*coupling VMS with Logbook*" methodology has demonstrated to improve the identification of "fishing VMS pings" by taking into consideration missing pings that were not taken into account by the original methodology. This new methodology also considers the fact that the parameters needed to describe the footprints and associated impacts of trawlers and longliners are different and therefore, their corresponding "fishing VMS pings" must be considered separately when calculating the cumulative fishing effort.

Nevertheless, information from longline activity that is currently available in the logbooks is insufficient to describe appropriately the longline footprint and the associated impacts. At present time, taking into account the available data and the improvements of the new methodology, only the *"bottom trawl cumulative fishing effort"* and *"bottom trawl specific fisheries footprint"* can be calculated adequately. In the case of longline fisheries further studies are needed to implement a specific methodology. Moreover, collection and compilation of logbook additional information (i.e. start and end of the line set as well as the start and end of the line haul) would be crucial to start the process of defining a more precise fishing bottom longline footprint.

Even though the new methodology has been found to improve the original, refining the spatial distribution of bottom fishing activity, many issues were raised in terms of quality of data, compromising a better quality of the results. Therefore, misreporting and errors found in VMS and Logbook data should be further analysed (e.g. through a previous quality control check process).

All this improvements will help to increase the quality of data (VMS and Logbook) that is being used, among other analysis, to better understand if and how fishing effort is changing over the years in the NAFO Regulatory Area.

Acknowledgements

This work was conducted as part of the NEREIDA project funded by the European Commission under Grant Agreements SI2.770786; SI2.793318 and SI2.827558. We would like to thank Corinna Favaro (DFO, St. John´s) who kindly shared the original R code used in the present study. We would also like to thank to NAFO Secretariat that provided access to VMS and logbook data.

References:

- Garrido, I., González-Costas, F and González-Troncoso, D (2020). Analysis of the NAFO VMS and logbook data. Serial No. N7144 NAFO SCR Doc. 20/068.

- NAFO 2015. Report of the 8th Meeting of the NAFO Scientific Council (SC) Working Group on Ecosystem Science and Assessment (WG-ESA) [Formerly SC WGEAFM]. NAFO SCS Doc. 15/19, Serial No 6549, 176 pp.

- NAFO 2017. Report of the 10th Meeting of the NAFO Scientific Council (SC) Working Group on Ecosystem Science and Assessment (WGESA). NAFO SCS Doc. 17/21, Serial No N6774.

- R Core Team (2020). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.