

PRELIMINARY ANALYSES OF THE ICCAT VMS DATA 2010-2011 TO IDENTIFY FISHING TRIP BEHAVIOR AND ESTIMATE FISHING EFFORT.

Mauricio Ortiz¹, Ana Justel-Rubio¹, Alberto Parrilla¹

SUMMARY

In 2003, the Commission recommended the ICCAT Vessel Monitoring System (VMS). The ICCAT VMS system has been fully operational since 2008. Besides its monitoring and compliance use, the VMS data offer detailed information on the fishing activity that can be useful for scientific purposes. This manuscript reviews the current ICCAT VMS data and provides preliminary analyses on 2010 and 2011 to identify fishing behavior and estimate potential fishing effort for the main fleet monitored. Results indicated that it is feasible to estimate a probability of active fishing operations given an identification of the main gear-fleet vessel.

RÉSUMÉ

En 2003, la Commission a recommandé le système de surveillance des navires (VMS). Le système VMS de l'ICCAT est entièrement opérationnel depuis 2008. Outre son emploi pour le suivi et l'application, les données de VMS fournissent des informations détaillées sur les activités de pêche qui peuvent s'avérer utiles à des fins scientifiques. Le présent document passe en revue les données actuelles du VMS de l'ICCAT et fournit des analyses préliminaires au titre de 2010 et 2011 afin d'identifier le comportement de pêche et estimer l'effort de pêche potentiel de la principale flottille objet de suivi. Les résultats indiquaient qu'il était faisable d'estimer la probabilité d'opérations de pêche actives une fois identifié le type d'engin/de flottille principal.

RESUMEN

En 2003, la Comisión recomendó el establecimiento del Sistema de Seguimiento de Buques (VMS). El VMS de ICCAT ha estado plenamente operativo desde 2008. Además de su uso para el seguimiento y el cumplimiento, los datos del VMS ofrecen información detallada sobre las actividades pesqueras, que puede ser útil para fines científicos. Este documento examina los datos actuales del VMS de ICCAT y proporciona análisis preliminares de 2010 y 2011 para identificar el comportamiento pesquero y estimar el posible esfuerzo pesquero de la principal flota objeto de seguimiento. Los resultados indicaron que es viable estimar la probabilidad de operaciones pesqueras activas siempre que se identifique el tipo de arte-flota principal.

KEYWORDS

VMS, fishing effort

¹ ICCAT Secretariat, c/Corazón de María, 8. 28002 Madrid, Spain.

1. Introduction

Following the recommendations from the ICCAT Commission [Rec 03-14], the Secretariat established a VMS data center in April 2008. This VMS data center received information from vessels involved in fishing Eastern Bluefin tuna primarily, as required in Rec 06-05 and Rec 07-08. The VMS information was initially classified as confidential data primarily used for compliance purposes exclusively. Then, after three years, the Commission authorized the SCRS to use the VMS data for scientific purposes. Finally, with the adoption of the Data Confidentiality Policy by the Commission in 2010, it authorized the SCRS the availability of VMS data for scientific purposes.

Vessel Monitoring Systems (VMS) have quickly developed and expanded in terms of implementation, with the primary purposes of monitoring, control and surveillance (Kourti *et al.* 2005). The VMS data itself reports only the geographical position of the boat, the time of the signal emission and a vessel identifier to the reception station(s) or satellite. There is not strictly associated information such as vessel catch, or gear deployment, that can classify the type of fishing activity and or estimate fishing effort, or catch associated. From a scientific point, the detailed operation of fishing activity requires to discriminate between active fishing (e.g. searching, gear deployment, active fishing gear, gear retrieval and catch associated) and non-fishing time (e.g. traveling, port deployment, etc.).

The objective of this document is to review the available ICCAT VMS data, to carry out preliminary analyses trying to identify and separate fishing trips, fishing behavior and to estimate, if possible, effective fishing effort. Because of the size of the DB, analyses were restricted to last two years 2010 and 2011.

2. Data

The ICCAT VMS receives the data transmitted by CPCs to a database center, in a specified format NAF. The ICCAT VMS DB contains over 2.25 and 1.64 million VMS records for 2010 and 2011, respectively (**Figure 1**). They represent messages from 22 CPCs, and over 17 different gear/vessel type categories (**Figure 2**). By month, most VMS are received from March through July, with a peak in May-June (**Figure 3**), coinciding with the peak in bluefin fishery season. The main type of gear/vessel category from which VMS messages are received are from the longline fleet, purse-seiners, trawlers, and pole & line (baitboat). However, there are also VMS messages from tugboats, freezers, assisting vessels and “other” vessels category that account for at least 35% of the data. In 2010, there were registered 998 different vessels based on the ICCAT registration number, and 698 different vessels in 2011. However, there are VMS messages received that don't have an ICCAT Registration Number associated. The main ID in the VMS transmission is actually the Radio Call Sign (RCS) code, indeed in 2010 there were 1330 different RCS registered, and 799 in 2011. Therefore, the ICCAT VMS receive messages from vessels not registered in the ICCAT list of vessels; although the proportion of non-registered VMS is relatively low. In 2011, for example 20,184 VMS messages (< 1.25 %) were from vessels with no ICCAT registration number and only RCS. The list of ICCAT registered vessels objective is to keep an updated list of vessels in the different ICCAT fisheries. However, the lack of a truly unique identifier for each vessel that can keep track of changes of flag/gear or fishing activity makes it difficult to track historically all the changes associated with a particular vessel. Therefore, it is possible that the same vessel may appear with two or more ICCAT registration numbers, if for example changes of vessel name and or flag are not reported to the Secretariat.

Further analyses were restricted to ICCAT VMS records for which the ICCAT Vessel registration number was available. **Table 1** shows the actual input variable available in the ICCAT VMS DB (variables 1 to 13). Most of this information is received from the vessel's VMS unit and then retransmitted via satellite or a land station to a center which re-routes the messages to the ICCAT Secretariat. There are, however, CPCs that do send the VMS information via electronic files directly to ICCAT in a NAF format. The input data include a unique message identifier (MDGID), vessel code(s) (Vessel-ID, ICCAT registration Number, RCS), vessel name, date of VMS origination (date and hour, in GMT), date of VMS reception, and latitude and longitude position of each message. At the VMS DB, each record is linked to the vessel registration to assign the Gear/vessel type, size of vessel (LOA), and using the latitude-longitude position to match to a list of know landing ports. For matching lat-lon to port, it uses a radio of approximately 2 km within the port position.

3. Methods

Analysis of VMS data are based primarily in the estimated trajectory (distance, bearing), time and speed of the vessels activity. This information is not actually transmitted from the boat (although it may be requested directly in the NAF format), therefore it was required to estimate it. VMS records were first sorted by the ICCAT registration number (ICCATRegN), and then by the date/time of the VMS signal transmission. At this stage, records were checked for duplication and duplicated messages were deleted. A VMS record was considered duplicate only if it had the same ICCATRegN and the exact same date/time of VMS transmission.

Subsequently, the time between messages in hours was estimated between two consecutive records from the same vessel ICCATRegN. Then, the distance (km) between two consecutive records was calculated using the following formulation prescribed by the FCC (47 CFR 73.208), recommended for distances not exceeding 475 km.

$$D = \sqrt{(K_1\Delta\phi)^2 + (K_2\Delta\lambda)^2};$$

Where D = distance in kilometers; $\Delta\phi$ and $\Delta\lambda$ are the differences of longitude and latitude (degrees), respectively, and

$$K_1 = 111.13209 - 0.56605 \cos(2\phi_m) + 0.00120 \cos(4\phi_m);$$

$$K_2 = 111.41513 \cos(\phi_m) - 0.09455 \cos(3\phi_m) + 0.00012 \cos(5\phi_m).$$

With ϕ_m is expressed in units compatible with the method used for determining $\cos(\phi_m)$; with the K_1 , and K_2 parameters adjusted to the *meridional* and its perpendicular (*normal*) radii of curvature of the Earth, using the Clarke 1866 reference ellipsoid (“Geographical distance”, 2013).

Also, the bearing (or initial heading direction) was estimated between two consecutive records as:

$$\theta = \text{atan}(\sin(\Delta\text{lon}) * \cos(\text{lat}2), \cos(\text{lat}1) * \sin(\text{lat}2) - \sin(\text{lat}1) * \cos(\text{lat}2) * \cos(\Delta\text{lon}))$$

where lat1, lon1 refer to the point of origin. Bearing angles were expressed in reference to 0-360° degrees. Finally, having distance and time, speed in nautical miles per hour was also estimated. Because the list of ports available in the ICCAT VMS is limited, a second list of known harbors from the EU-Fisheries Center was also used to match lat-lon positions (Hintzen *et al.*, 2012). Similarly, a radio of approximately 2 km was used as overlapping between the VMS position and the harbor lat-lon center position. For records that have been identified as located in harbor or port, no estimates of distance, speed or bearing were calculated.

The frequency of VMS messages was initially requested at every six (6) hours during fishing season. However, not all vessels send full time, or continuously. Furthermore, with the current data format it is not identified if given sequences of messages are from a single fishing trip, partial or several trips. Thus, the next step was to identify for each vessel the number of trips and separate the consecutive records into individual vessel-trip(s). For this, the following two assumption rules were followed:

- a. Once a vessel entered or was in port it was considered the end of a fishing trip. Then, the first following VMS record of this vessel out of port was considered a new trip.
- b. If a vessel was not in port, a record was considered a new trip if the time between the current record and the immediately previous VMS record was over 48 hours of difference.

Once identified the 1st record for each vessel-trip, estimates of time between signals, distance, speed and bearing were set to null for this 1st record, with values only for the second to last VMS of each trip.

Based on the geographical locations, VMS signals at sea were then evaluated for “clustering” distribution assuming that during active fishing (e.g. searching or gear deployment) with lower speed operations and in general shifting directions, the VMS signals will cluster. In contrast, during cruising the points will tend to follow a pre-determined path. Nonparametric bivariate density estimation was used to identify the ‘clusters’ (Faraway 2006). The kernel density function was used to estimate the bivariate probability density function of the latitude longitude VMS positions for each vessel-trip at sea. Then the density function was calculated using the inverse Fourier transformation.

4. Results

There is a wide range in terms of VMS messages received at the ICCAT DB by vessel. In 2010, for example, regarding the 934 different vessels (ICCATRegNum), the frequency of VMS messages varied from 1 to 19,000; with a median of 1721 and a clearly skewed distribution (**Figure 4**). If restricted to the main gear-fleets, longline (154 vessels), purse seine (116) and Baitboat (pole & line, 62) in 2011; the number of VMS messages ranged from 30 to 15,000 with some median around 2000 (**Figure 5**). For 2011 main gear-fleets, the VMS messages represented over 1.6 million hours of monitoring. The distribution of these messages between at sea and in port was roughly equal for the purse-seiners and the pole line fleets, while it showed 61% at sea and 39% in port for the longline fleet (**Figure 6**). There were at least 244 different harbors reported with VMS messages, of which Kali, Sete, Marsala, Hondarribia, Salerno, Marsaxlokk, Guetaria, Licata and Svetikajo are the top ten ports in terms of VMS messages in 2011 (**Figure 7**). The mean time between VMS messages also varied by main gear-fleet, with purse-seiners having on average VMS signals each 1.3 hours while longliners send VMS signals each 2.6 hours (**Figure 8**). In general, the frequency of signals is below the 6 hours original requirement.

Figure 9 shows the frequency distribution of vessel speed (NM/h) of VMS signals at sea by main gear-fleets in 2011. Although there are some obvious outlier points, the average speed for longliners is 4 NM/h, with a bimodal distribution and with peaks at low (< 3 NM/h) and high (8 NM/h) speeds, which show a substantial overlapping. A similar shape distribution was observed for the pole & line fleet with a more clear definition of the bimodal feature. Instead, for the purse seine there was a high proportion of lower speeds (< 1 NM/h) and a less defined second peak at 8 to 10 NM/h range. This bimodal distribution has been previously used as an indicator for the fishing activity at low speed ranges, and the ‘cruising’ or non-fishing operation at higher speed. It is important to note, however, the different features for each main vessel type. **Figure 10** shows a scatterplot of the total hours at sea (2011) versus the distance covered (km) at sea (e.g. excluding all in port VMS) grouped by main gear type. The longliners show overall a wider range of hours and distances covered. By comparison, the pole & line fleet has much more restricted ranges of hours at sea and distance, but the linear relationship is similar to the longliners. In contrast, the purse seiners cover on average shorter distances than longliners or pole line vessels, and have different slope in the linear relationship with total hours at sea. Another indicator commonly used is the changes in bearing; it is expected that during cruising the vessel does change little in their heading direction, while more erratic direction is observed during fishing searching or fishing deployments. **Figure 11** shows a time series (Apr –Jun 2011) for some longline vessels and the change in heading directions. The change in directions (DtaBea) was estimated as the \sin of $0.5 * 2$ -lag difference between consecutive VMS messages from a vessel. Thus, a value of 0 in the y-axis indicates no change in heading, while values closer to 1 indicate opposite heading.

The estimated number of trips per vessel changed in function of the fishing gear type (**Figure 12**). For example, in 2011, the number of trips for longliners averaged 30 but the distribution is skewed towards less than 40 trips per year. In contrast, the bait boat fleet has on average 90 trips with a 80% quantile between 66 and 120 trips. Finally, much fewer trips were estimated for the purse seine fleet, averaging 20 trips per year with also a predominant lower (<30 trips) overall. Plots in **Appendix 1** show general examples of selected vessels and their estimated annual trips distribution. Overall, in the ICCAT VMS data there are indicators that can be used to characterize fishing behavior such as speed and changes of bearing (heading). It is important, however, to note that these indicators did vary with respect to the main gear type of the vessel and should be estimated separately. Usually classification of time at sea in fishing or cruising has used these indicators of speed and bearing, as well as auxiliary information (Russo et al 2011, Vernard et al 2010). Some approaches use Bayesian with priors by fleet type (Vernard et al 2010). A more complex approach uses artificial neural networks (Russo et al 2011) incorporating a relatively large set of auxiliary information such as license type, sea depth of operation and vessel class. In our case we have no auxiliary information, thus alternative methods were explored to classify vessel time at sea. **Figure 13** presents an example of a nonparametric bivariate density approach, in this case using the lat-lon positions directly; the non-parametric kernel estimates a density bivariate function. The assumptions under this approach are: a) that geographical positions in proximity are positively correlated (clustered) with active fishing while when on cruising the positions are non-clustered, and b) the sampling of the positions are equally distributed in time, independent of the location. The second assumption holds for most of the ICCAT VMS data at the vessel-trip level. Preliminary results of the bivariate density approach show results consistent with the indicators of speed and changes of heading for the vessel-trips evaluated (**Figure 14**). The results from the bivariate nonparametric density are a classification of each position in a cluster level with high cluster for high density aggregations (**Figure 13**).

After the analyses of speed, change of heading and clustering of positions, we normalized each of these variables to a scale 0 to 1; with 0 representing low speeds, high variation in heading direction and high cluster density.

Figure 15 shows in a ternary plot the distribution of VMS signals from a single longline vessel and the corresponding lat-lon map of the same information. Given equal weighting to all three variables, it is then possible to estimate a probability for each VMS position summing all three variables. **Figure 16** shows the frequency distribution of this probability for our prior example and the cumulative density function. In this example, 75% of the 2765 VMS records have ≥ 0.41 probability of being in an active fishing behavior base on speed, heading and clustering.

In summary, the preliminary analyses and review of the ICCAT VMS data indicated the following:

- High percent of the VMS data is associated with vessel ID and are registered in the ICCAT Vessel DB.
- Given the main gear-fleet declaration of each vessel, there are differences in the derived variables of speed, time at sea and overall fishing behavior between the longline, pole and line and purse seine fleets. There is however a considerable portion of vessels with VMS records for which main gear is not declared. It is recommended to update this information and maintain it up to date with changes in the ICCAT fleet registration. Or alternatively classify these vessels within main gear types using indicators such as the total distance covered versus their total hours at sea.
- The proposed algorithms for trip identification proved consistent and robust. However, validation should be done with auxiliary information such logbooks or reports from observers at board.
- The combination of indicators such as speed, changes of heading and clustering of positions appear to give consistent results when evaluated for each main gear-fleet. The probability of fishing effort should be also validated with auxiliary information from logbooks or observer programs.
- Independently of the fishing effort estimates, there is not at present a direct link with catch for the ICCAT VMS database. In the case of bluefin tuna, possible links with BCD files can provide the catch component and catch size distribution information.

Literature cited

- Faraway, J. J. 2006. Extending the linear model with R. Generalized linear mixed effects and nonparametric regression models. Chapman & Hall/CRC Text in Statistical Science.
- Geographical distance. (2013, March 7). In Wikipedia, The Free Encyclopedia. Retrieved 09:41, April 9, 2013, from http://en.wikipedia.org/w/index.php?title=Geographical_distance&oldid=542697392.
- Kourti, N. I. Shepherd, H. Greidanus, H. Greidanus, M. Alvarez, E. Aresu, T. Bauna, J. Chesworth, G. Lemoine and G. Schwartz. 2005. Integrating remote sensing in fisheries control. Fisheries Management and Ecology 12:295-307.
- Murawski, S.A., S.E. Wigley, M.J. Fogarty, P.J. Rago and D.G. Mountain. 2005. Effort distribution and catch patterns adjacent to temperate MPAs. ICES journal of Marine Science 62:1150-1167.
- Niels Hintzen, Francois Bastardie and Doug Beare (2012). vmstools: For analysing fisheries VMS (Vessel Monitoring System) data. Package version 0.59. <http://CRAN.R-project.org/package=vmstools>
- Russo, T., A. Parisi, M. Prorgi, F. Boccoli, I. Cignini, M. Tordoni and S. Cataudella. 2011. When behavior reveals activity: Assigning fishing effort to métiers based on VMS data using artificial neural networks. Fish. Res. 111:53-64.
- Vernard, Y., E. Rivot, S. Mahevas, P. Marchal and D. Gascuel. 2010. Identifying fishing trip behavior and estimating fishing effort from VMS data using Bayesian Hidden Markov Models. Eco. Modelling. 221:1757:1769.

Table 1. Variables recorded in the ICCAT VMS database.

1. MSGID	Unique identifier for each VMS message received by the ICCAT VMS center
2. FlagCPC	Corresponding flag associated with the vessel ID as available in the ICCAT Vessel records
3. VesselName	ICCAT vessel name recorded
4. VesselID	CPC vessel identification code
5. ICCATRegNumber	ICCAT Vessel identification code
6. RCS	Radio Call Sign code
7. Gear	Main gear type reported for registered vessels
8. DateRec	Date and time of the VMS reception at ICCAT
9. DateVMS	Date and time of the VMS transmission from the vessel unit to the main reception center
10. LOA	Length overall registered in ICCAT vessel DB
11. LatY	Latitude position of VMS
12. LonX	Longitude position of VMS
13. Puerto	An estimated variable base on Lat-Lon that indicate if a signal is within an approximate 2 km radio from a know port (lat-lon) position, The value is the name of the port
14. DATE	Date of the VMS transmission (dd-mm-year) format
15. TIME	Time of the VMS transmission (hh-mm-sec) format
16. TBM_Hrs	Estimated time (in hours) between two consecutive VMS from the same vessel
17. DupPing	Variable to define if a VMS signals is duplicated. Duplicate = same Date & Time for same vessel ID (ICCATRegNumber). If duplicate = TRUE, these records were deleted
18. Dkm	Estimated distance in km between 2 consecutive VMS signals from the same Vessel ID (ICCATRegNumber) and same trip.
19. NMhr	Estimated speed (Nautical Miles per hour, or knots) of the vessel between two consecutive pings from the same Vessel ID and trip
20. Bearing	Estimated bearing direction (0-360 degrees) between two consecutive pings from same VesselID and trip
21. Harbour	A second estimation of a position of a vessel within an approximate 2 km radio from a known port's (lat-lon) position. In this case, the EU list of ports was used, a more comprehensive database than the ICCAT VMS

22. Atport	Identify if a vessel is within a port (Puerto or Harbour)
23. trip	A trip for a given vessel was estimated as: If time transmission between consecutive VMS signals is over 48 h, it was considered a new trip/ If a vessel returns to port, next signal not in port is considered a new trip
24. Bytrip	ID code for unique Vessel ID and Trip number
25. EEZLibya	Overlap with the EEZ shape file of Lybia
26. Month	Month of the Date variable

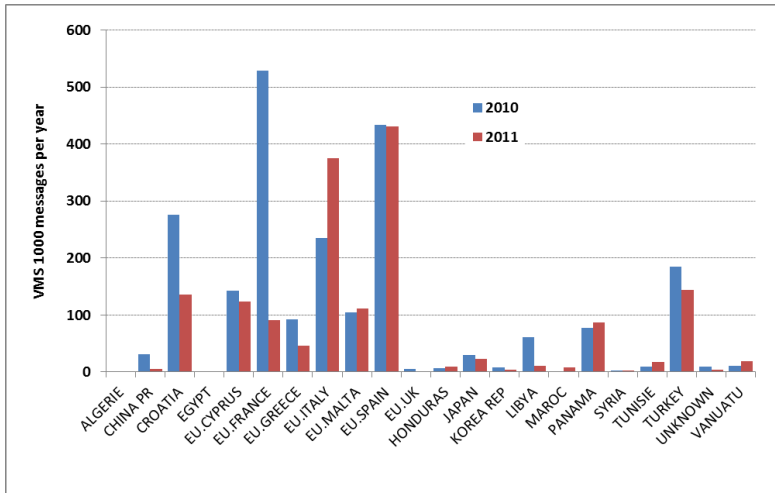


Figure 1. Number of VMS messages received at the ICCAT VMS center by year and CPC.

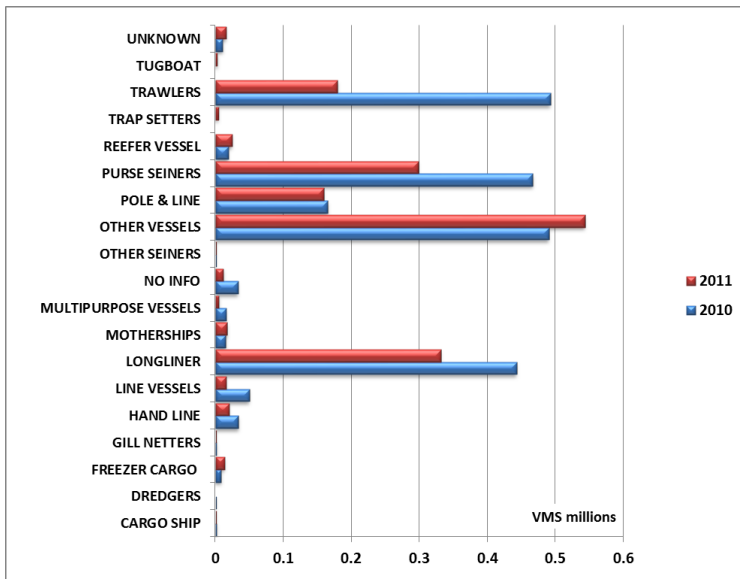


Figure 2. Number of VMS messages by gear vessel category and year.

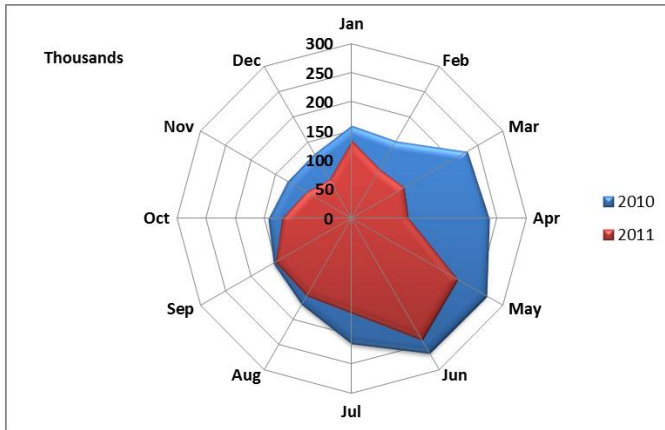


Figure 3. Monthly distribution of VMS messages received in 2010 and 2011 at ICCAT VMS.

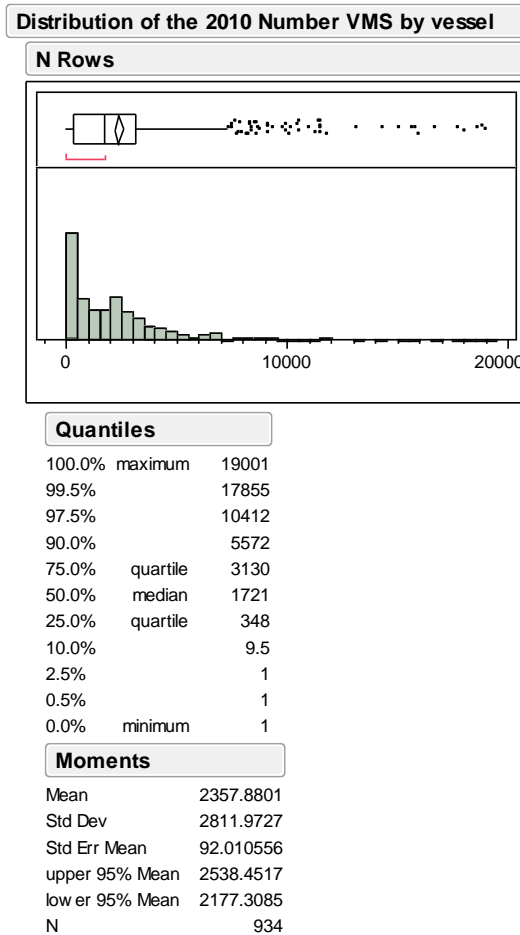


Figure 4. Frequency distribution of number of VMS messages by vessel sent to ICCAT VMS during 2010.

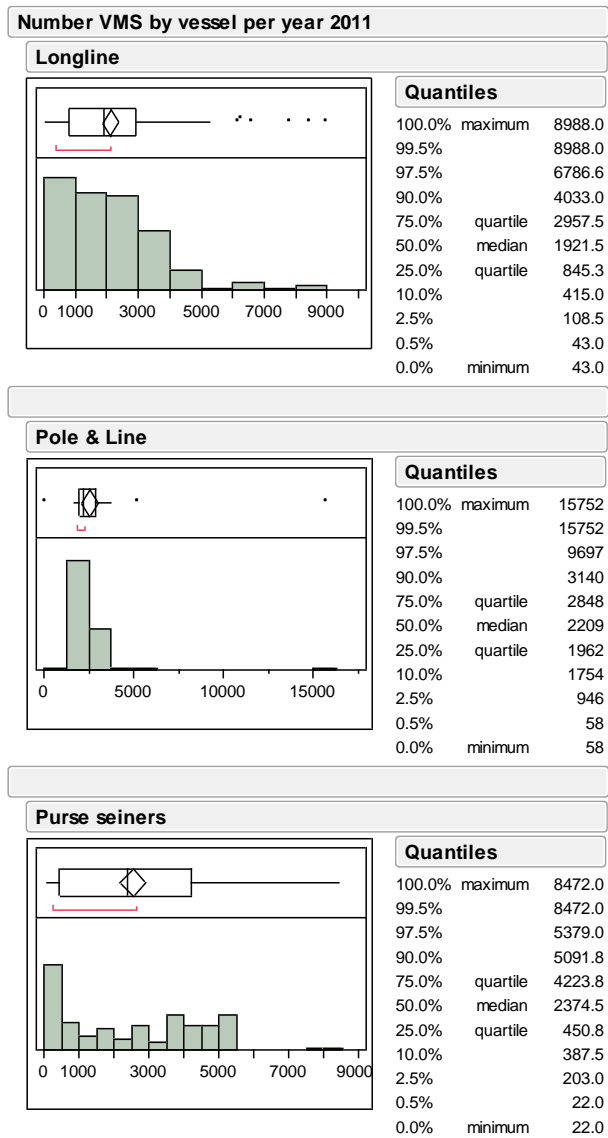


Figure 5. Frequency distribution of number of VMS messages by vessel for the main gear-fleet categories in 2011.

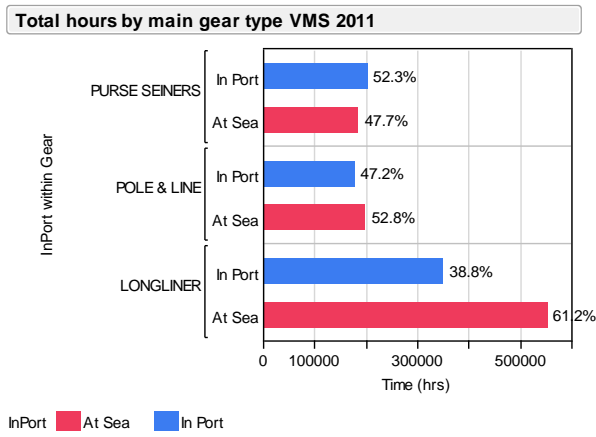


Figure 6. Total hours VMS monitoring by main gear type and whether in port or at sea for 2011.



Figure 7. Identified ports of operation from the 2011 ICCAT VMS DB.

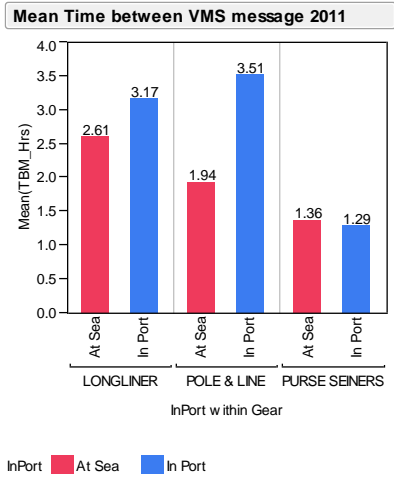


Figure 8. Mean time between VMS messages received by ICCAT VMS in 2011 by main gear fleet and location.



Figure 9. Frequency distribution of vessel speed (NM/hr) of VMS at sea positions by main gear-fleet 2011.

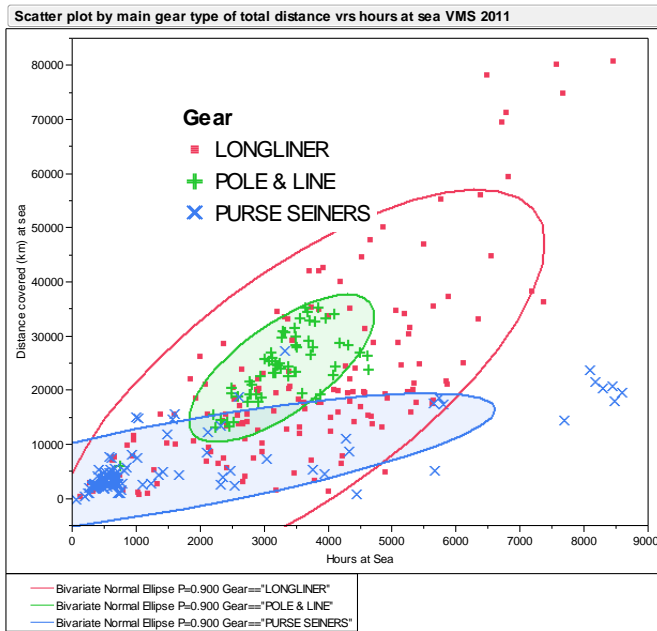


Figure 10. Scatter plot of total hours at sea (2011) versus the distance covered (km) at sea by main gear-fleets. The shaded areas represent the 90% bivariate normal ellipses.

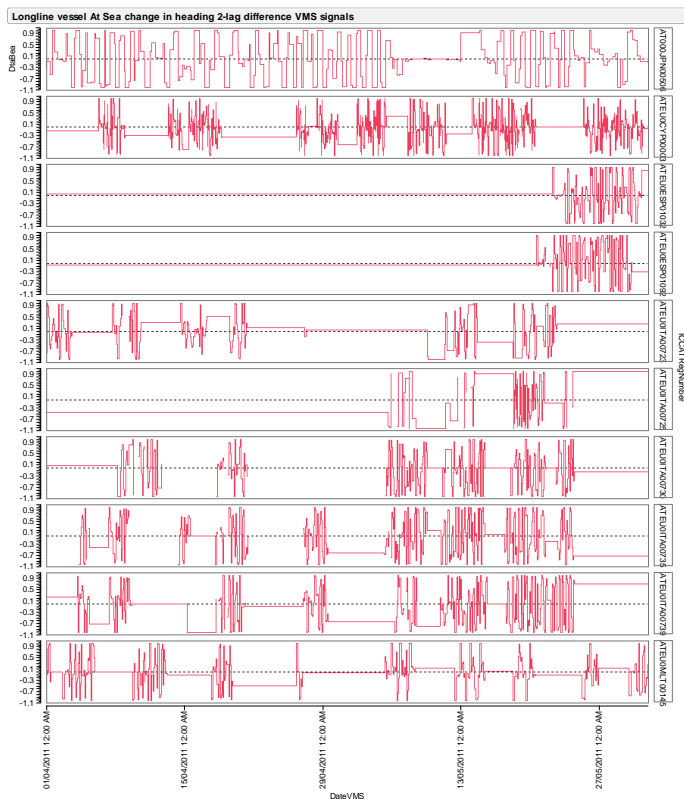


Figure 11. Time series April through June 2011 for selected longline vessels and their change in heading direction. In the y-axis, a value of zero corresponds with no change in heading for 2-lag consecutive VMS signals from the same vessel; values closer to 1 or -1 represent maximum heading change for 2-lag VMS signals.

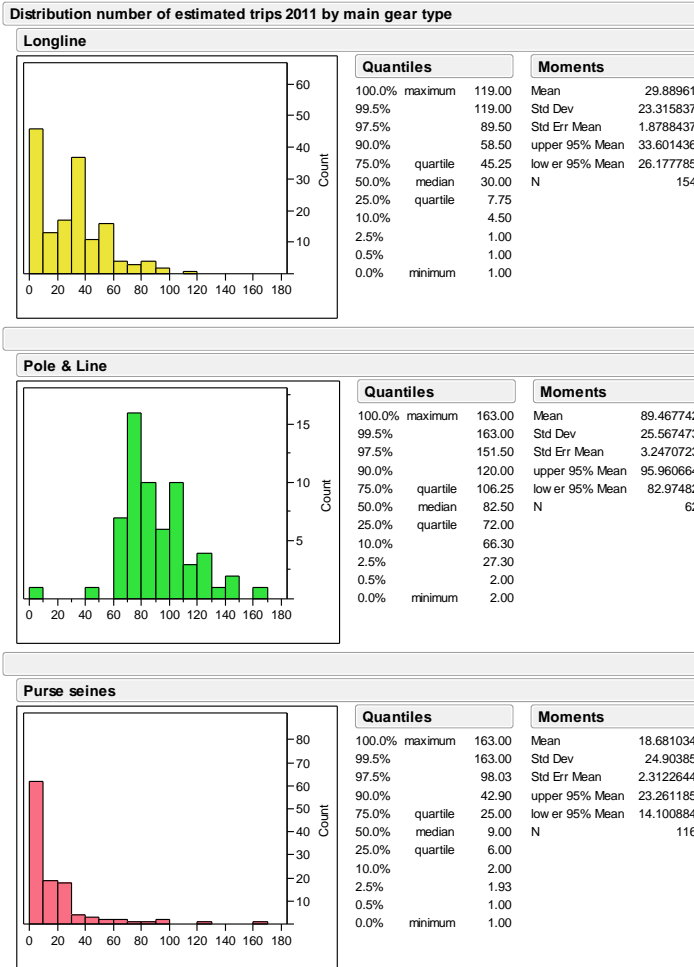


Figure 12. Frequency distribution of estimated number of trips per vessel in 2011 grouped by main gear-fleet type.

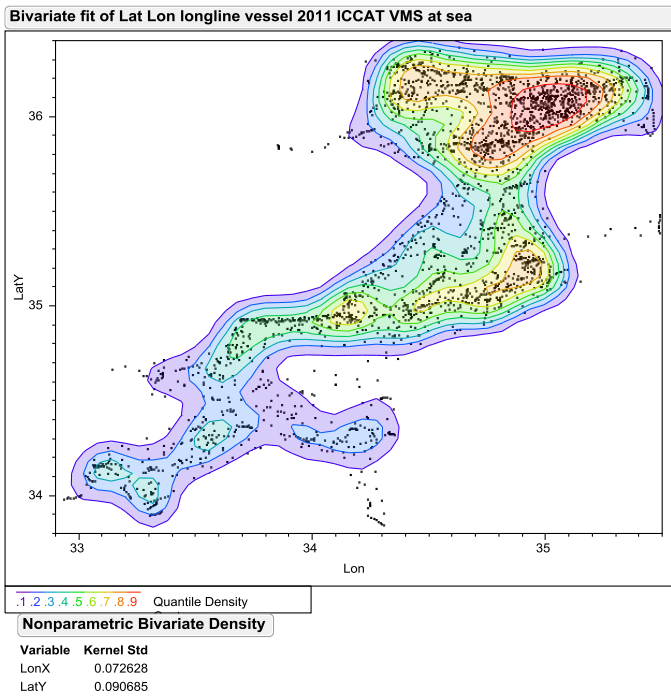


Figure 13. Nonparametric bivariate density fit to VMS lat-lon positions at sea from a longline vessel 2011. Shaded areas correspond to the estimated quantile density “clusters”.

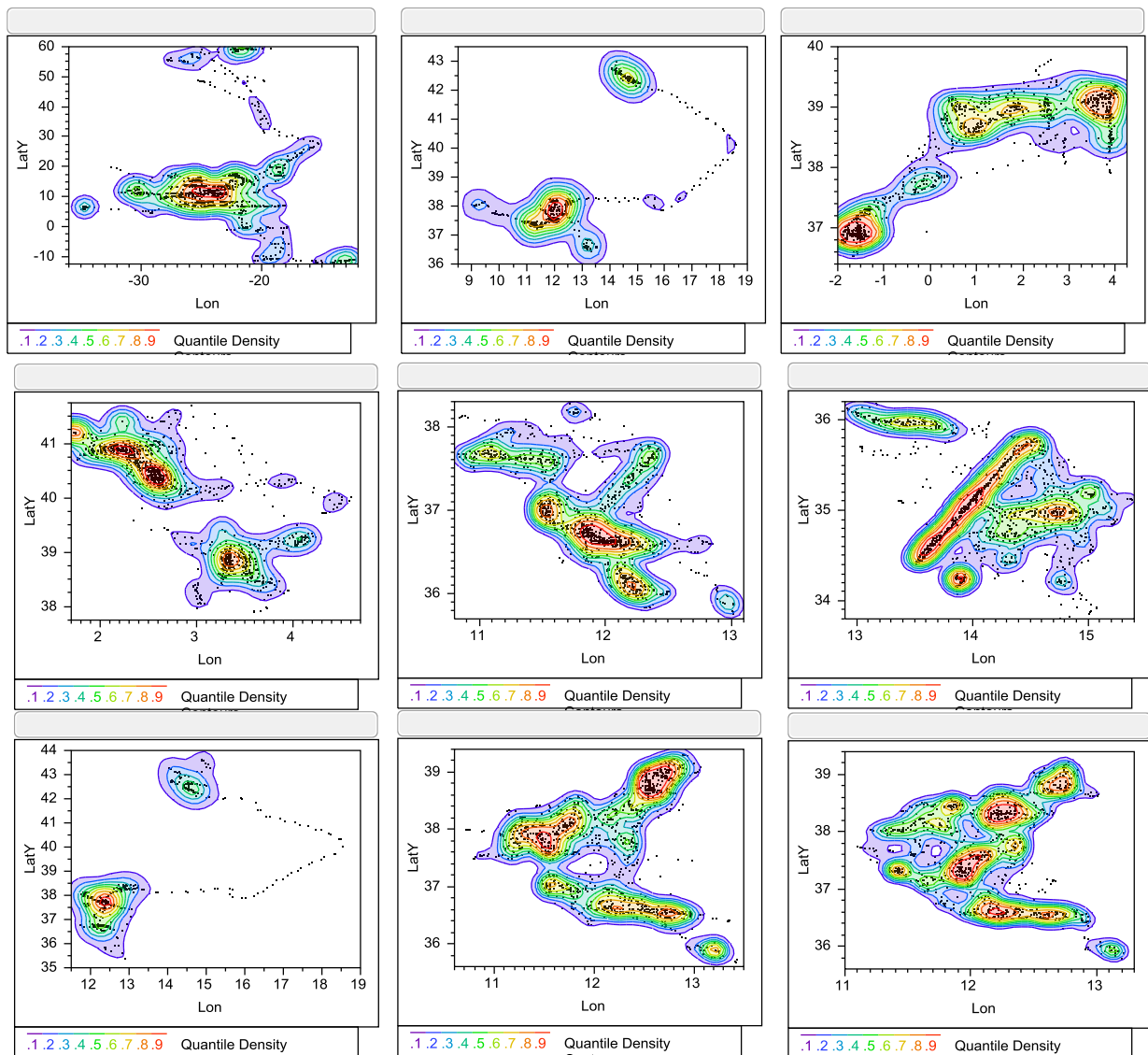


Figure 14. Examples of nonparametric bivariate density fit for lat-lon positions of at sea VMS signals 2011. Contour represent the estimated cluster level according to the estimated density function, points are the actual VMS locations.

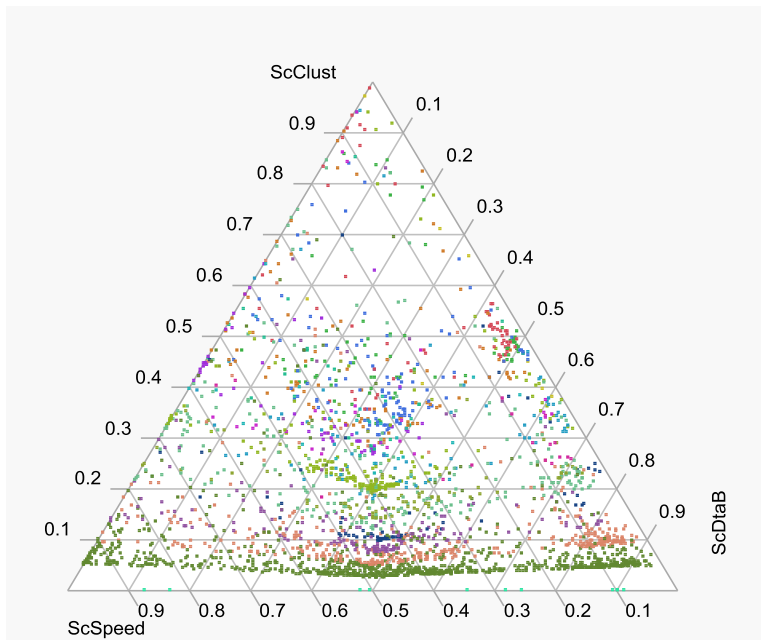


Figure 15. Ternary plot of the distribution of VMS signals from a single longline vessel at sea in reference to vessel speed (ScSpeed), changes in heading (ScDtaB) and nonparametric cluster bivariate fit (ScClust).

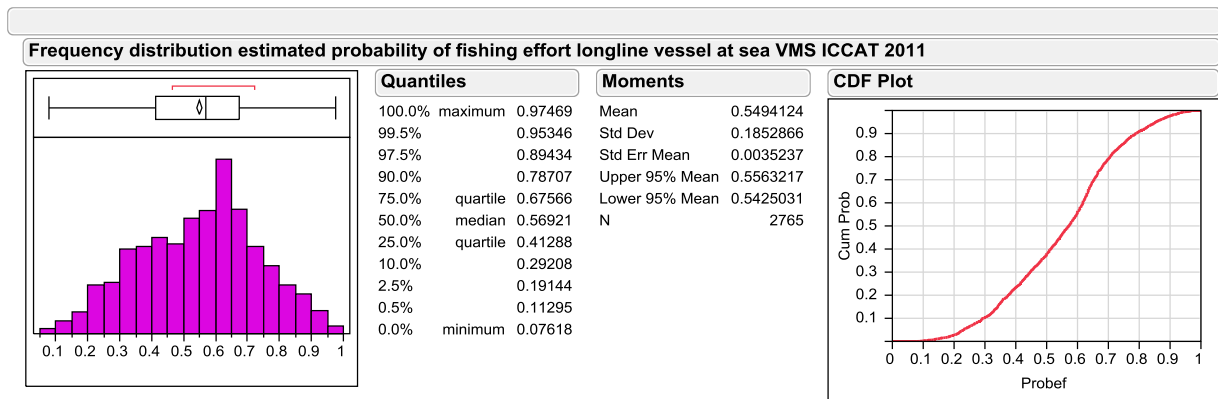
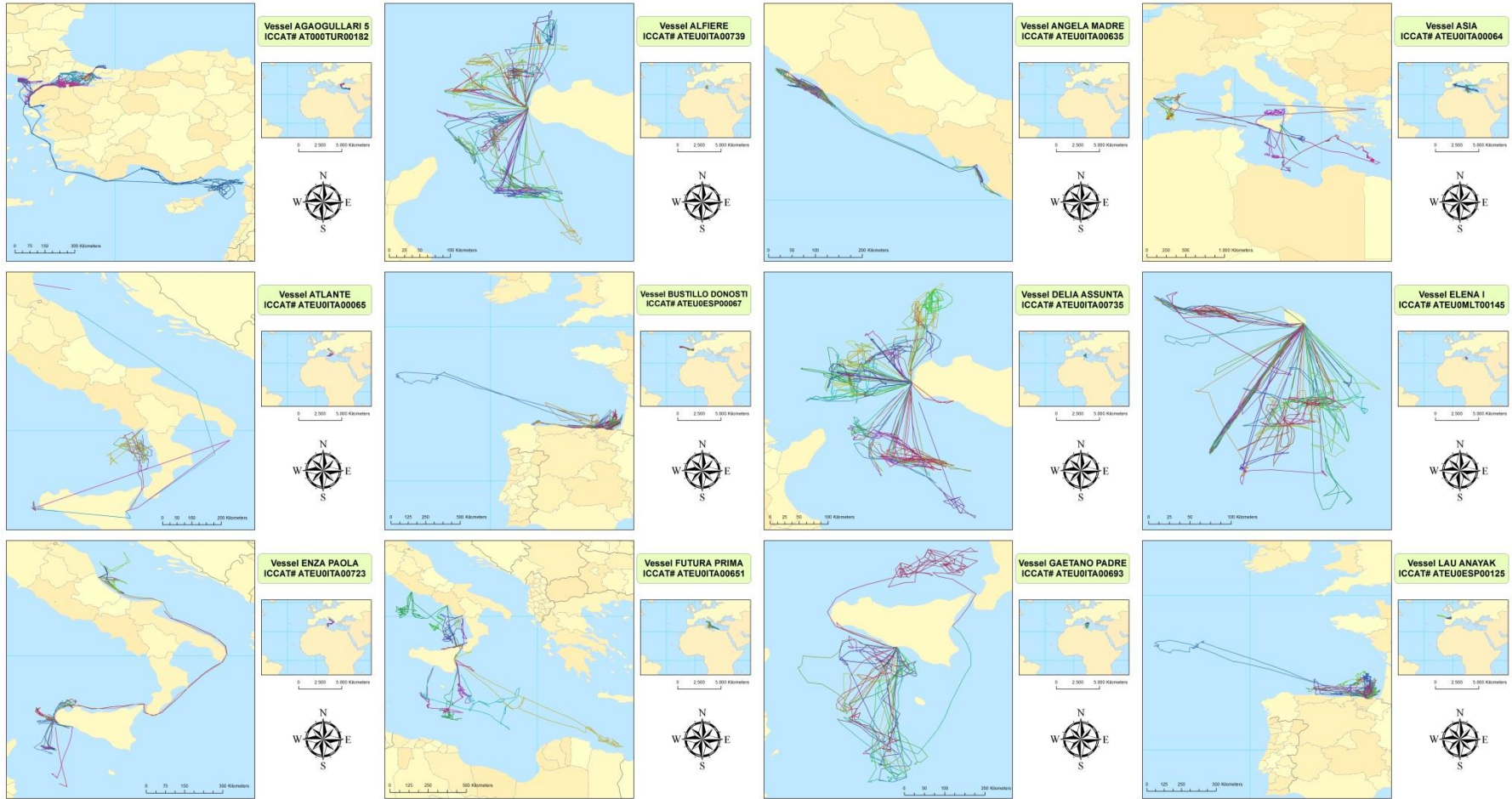
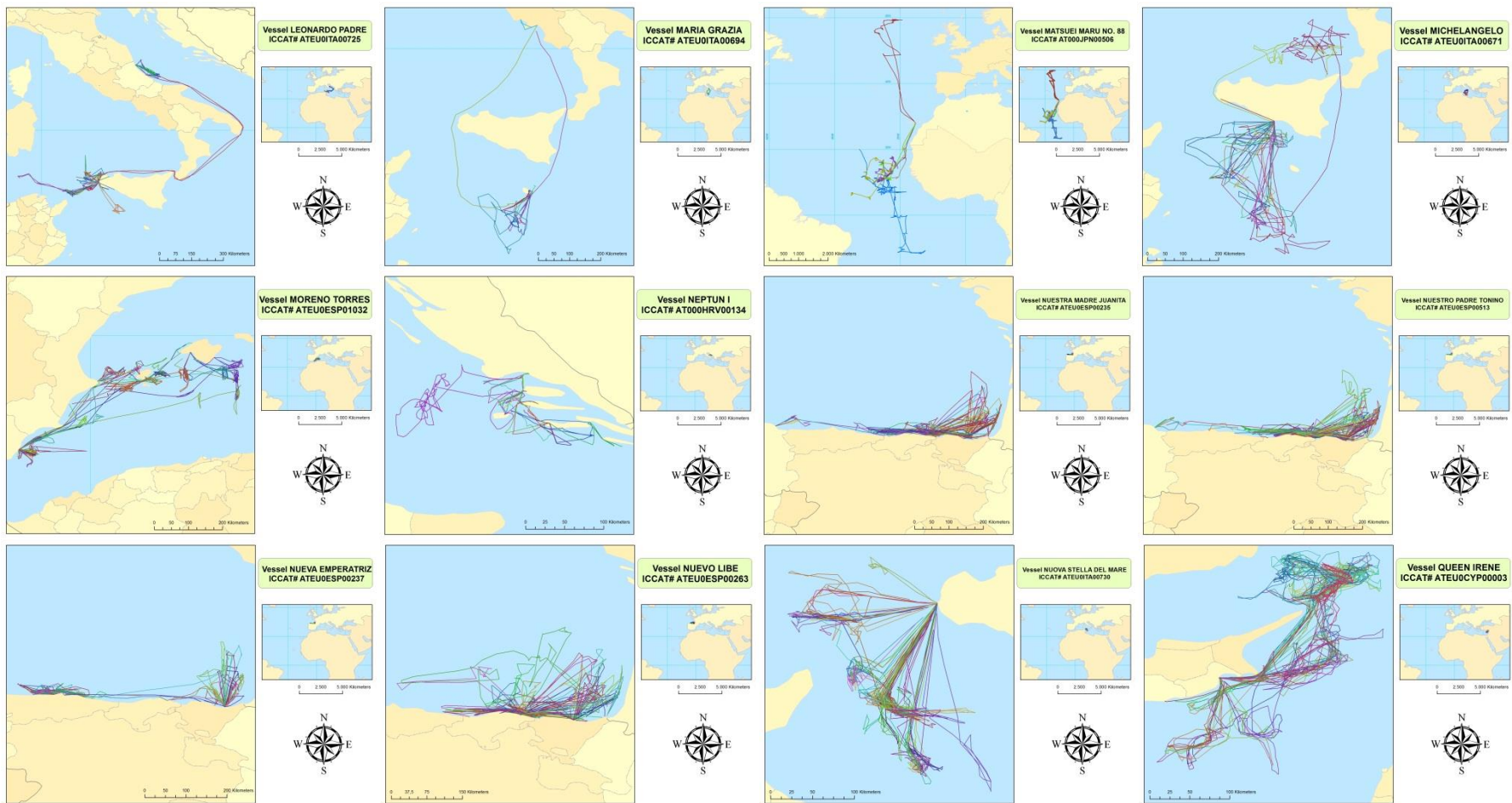
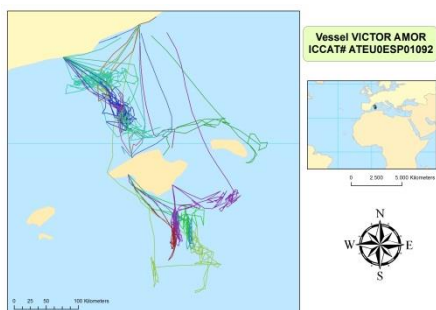


Figure 16. Frequency distribution of the estimated fishing effort probability for a longline vessel at sea from ICCAT VMS 2011.







Appendix 1. Examples of 2011 VMS trajectories for ICCAT registered vessels and the estimated trip distribution as shown by different color track.