Fisheries Research 188 (2017) 1–5

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Short communication

# Mismatch between VMS data temporal resolution and fishing activity time scales

I. Katara[a,b,*], A. Silva[a]

[a] IPMA-Instituto Português do Mar e da Atmosfera, I.P. Av. Brasília, 6, 1449-006, Lisboa, Portugal
[b] Centre de Recherches Halieutiques Méditerranéenne et Tropicale Sète (IRD, CRH, BP 171) Avenue Jean Monnet BP 171, Sète, 34203, France

ARTICLE INFO

*Article history:*
Received 16 March 2016
Received in revised form
26 November 2016
Accepted 28 November 2016
Handled by Prof. George A. Rose
Available online 14 December 2016

*Keywords:*
Vessel monitoring system
Purse seine fishery
Fishing effort
Fishing grounds
Bias

ABSTRACT

The use of Vessel Monitoring System (VMS) data to map fishing activity is challenged by the mismatch between the temporal resolution of position records (typically 2 h) and the time scale of fishing activity in fisheries with short trips and short fishing operations such as purse seining for small pelagics. We analysed the first five years of VMS and logbook data for the Portuguese purse seine fleet, when 10 min resolution VMS data were available, to evaluate bias and errors in fishing trips and fishing sets' identification related to the mismatch. We adapted the standardised VMS analysis workflow for EU fleets to the characteristics of the fishery and developed a framework to quantify bias for different VMS-based products, by resampling the 10 min VMS dataset at 20, 30, 60 min and 2 h intervals. For the Portuguese purse seine fishery, a 2 h time interval resulted to 42% missed fishing trips compared to the 10 min time interval data and a bias towards longer fishing trips. For trips that were correctly identified in the 2 h resampled dataset, 7% of the fishing sets were missing and fishing locations were identified with an error of approximately 2.36 km. The general spatial patterns of fishing operations – i.e. fishing grounds – were not significantly altered by the decrease of the data temporal resolution. Our framework is applicable to other fisheries and could become a useful tool for managers using VMS data.



## 1. Introduction

The emergence of global positioning surveillance in fisheries presented an exceptional opportunity to map fishing activity at a high spatial resolution (Bastardie et al., 2010; Bertrand et al., 2008; Fonseca et al., 2008; Gerritsen and Lordan, 2011; Gerritsen et al., 2012; Jennings and Lee, 2012). In the search for a standardised and comprehensive framework to analyse VMS and logbook data, fisheries scientists developed algorithms for data processing, metier identification, fishing set identification, coupling VMS with logbook data, and mapping fishing effort (Hintzen et al., 2012). But when it comes to fishing activity identification, a one-fits-all approach across fisheries or fleets is lacking and methods need to adapt to the specific practices and behaviours of each fleet or metier.

In Europe, fishing vessels typically transmit VMS data at 2 h intervals (EC, 2009). To achieve reliable results, a matching between the temporal resolution of VMS tracking data and the timescales of fishing activities is required (Postlethwaite and Dennis, 2013). Vermard et al. (2010) illustrated the need for observations to be collected at a frequency that ensures synchrony with the timing of behavioural switches (e.g. from steaming to fishing). The analysis of VMS data has been challenging (Bastardie et al., 2010; Vermard et al., 2010), especially for fisheries such as purse seining that fishing trips and fishing operations are typically short and take place close to harbours (e.g. Stratoudakis and Marçalo, 2002; Tsitsika and Maravelias, 2008). Discrepancy between the temporal scale of fishing activity and typical VMS data resolution may result in unrecorded or misidentified fishing trips and fishing sets and possibly bias towards longer lasting trips (Postlethwaite and Dennis, 2013; Vermard et al., 2010). Despite the wealth of studies developed in the past ten years using both VMS and logbook data to support fisheries and ecosystem policies (e.g. Campbell et al., 2013; ICES, 2011; Martín et al., 2014; Witt and Godley, 2007) limited attention has been paid to the quantification of such errors.

The Portuguese purse seine fishery is a characteristic example of a single-gear, single-species fishery, targeting European sardine, *Sardina pilchardus* (Walbaum)(Stratoudakis and Marçalo, 2002). The fleet consists of approximately 180 vessels, the majority with an overall length between 18 and 24 m, with an average gross tonnage of 50 and an average engine power of 200 kW (Silva et al., 2015). Fishing trips take on average 8 h and fishing sets (shooting,

* Corresponding author at: IPMA-Instituto Português do Mar e da Atmosfera, I.P. Av. Brasília, 6, 1449-006 Lisboa, Portugal.
*E-mail address:* isidora10@yahoo.com (I. Katara).

http://dx.doi.org/10.1016/j.fishres.2016.11.023
0165-7836/

closing and hauling the net and fish transfer on-board) last approximately 1 h but some trips can last less than 4 h and a fishing set can conclude in 30 min (Feijó, 2012; Stratoudakis and Marçalo, 2002). In this study we develop a framework – hybrid of spatial analysis and statistics – to evaluate bias and errors in fishing trips and fishing sets' identification, related to the mismatch between VMS data temporal resolution and fishing activity time scales, and apply it to the Portuguese purse seine fishery.

## 2. Materials and methods

VMS position, date/time and speed, and logbook trip date/time data by vessel (anonymous identification) for Portuguese purse seiners in the period 2000–2005 were provided by the Portuguese Fisheries Directorate (DGRM). For the vessels equipped with VMS during this period, common records for logbooks were available for 2003–2004.The number of vessels equipped with VMS increased from 26 to 85 during the period. The first five years of VMS data have a time interval between consecutive records of 10 min; since 2005, the interval becomes 2 h, the maximum complying with EU regulations (EC, 2009).

### 2.1. Fishing trip and fishing set identification

Standardised processes for the analysis of VMS and logbook data were adapted to the specificities of the Portuguese purse seine fishery. Two algorithms were applied to increase the chance of identifying fishing trips. The first algorithm identifies trips based on trip starting and ending dates from logbook data (*logbook* algorithm, Bastardie et al., 2010). Since misreporting in logbooks can result in missing trips a second algorithm (developed for trawlers by Hintzen et al., 2012) based on locations flagged as inside or outside of harbours (here termed *harbour* algorithm) was applied; a trip is defined as consecutive "outside of harbour" locations, confined by two "inside harbour" locations. To avoid missing short trips when fishing occurs very close to the harbour, as may happen in purse seine fishing, the threshold distance to harbour (a distance used to identify points "inside/outside harbour") was decided by testing distances from 1 to 2 km and evaluating the results by cross-validating randomly selected identified trips with logbooks. Points in harbours were defined as those less than 1.5 km from harbour locations. Fishing trips were identified by first applying the *logbook* algorithm and adding to its results any missing trips identified by the *harbour* algorithm.

The identification of fishing sets was based on speed patterns (Hintzen et al., 2012). The beginning of the set is signalled by a rapid drop in velocity from 7 to 9 knots to lower than 3 knots; low speed is maintained for 30–60 minutes, while the net is set and hauled, and the end of the set is signalled by a rapid increase in velocity (Diana Feijó, personal communication). Consecutive drops and increases in speed (difference between consecutive points >5 knots) were flagged as the beginning and end of fishing sets. This complex speed pattern was sought because absolute speed values (e.g. speed <3 knots) could depict other activities, such as repairs.

### 2.2. Bias and error quantification

The 10 min VMS data were re-sampled to create datasets of 20 min, 30 min, 60 min and 2 h, referred to as *resampled* datasets (Supplementary Material). The fishing trip and fishing set identification algorithms (Section 2.1) were applied on each dataset. The following statistical approaches were employed to compare the 10 min dataset with each of the *resampled* datasets. To quantify missing trips, we calculated the percentage of trips not identified in the *resampled* dataset and trip duration for each dataset. To quantify trip duration bias, the Wilcoxon rank sum test was applied to assess if trip duration differs between the 10 min and each of the *resampled* datasets. To quantify the error in fishing set identification the 10 min dataset was compared to each of the *resampled* datasets using the McNemar's test for paired observations (McNemar, 1947). The McNemar's test is a nonparametric test used to compare the effect of a factor on related or paired samples. Common points (geographic positions) between the 10 min and the *resampled* datasets were treated as paired samples. The factor, i.e. the time interval between consecutive points, is hypothesized to affect the identification of a point as fishing set (or not). The test is based on a standardised normal test statistic: $Z = (r_{12}-r_{21})/\mathrm{sqrt}(r_{12}-r_{21})$, where $r_{12}$ is the number of points identified as non-fishing points in the 10 min dataset but as fishing points in the *resampled* dataset (false positives) and $r_{21}$ is the number of points flagged as non-fishing points in the *resampled* dataset but as fishing points in the 10 min dataset (false negatives).

To quantify the error of the identified location of the fishing sets, for each trip, the distance between the fishing sets in the 10 min and each of the *resampled* datasets was calculated. The derived distance vectors were interpreted as systematic errors of the fishing set location estimates, due to lower data resolution (without accounting for error of trip identification).

To test if the fishing set identification algorithm gives similar spatial (point) patterns at different temporal resolutions of the VMS data, we applied an approach that combines the average Nearest Neighbour Distance statistic (NND, distance between each point and its closest neighbouring point) and a randomization process (Monte Carlo; Manly, 2006). If the patterns are the same, we can hypothesize that fleet-level average NND between points labelled as "fishing" in the 10min and the *resampled* datasets is zero. The probability distribution for NND was retrieved applying Monte Carlo permutations iteratively at the vessel level and averaging NDD values for the whole fleet (Supplementary Material).

All workflows were developed in R 2.15.3 (R Core Team, 2013) and ArcGIS (ESRI, 2011).

**Table 1**
Percentage of trips identified in *resampled* datasets, and corresponding median trip duration, assuming the 10 min data allow for the identification of all trips-100%.

| Time interval (min) | Identified trips (%) | Median Trip Duration (h) |
|---|---|---|
| 10 | – | 4 |
| 20 | 84 | 10 |
| 30 | 77 | 11 |
| 60 | 67 | 13 |
| 120 | 58 | 18 |

## 3. Results

Using the combination of the *logbook* and the *harbour* algorithms increased the number of identified fishing trips in comparison to the use of the *logbook* algorithm alone. For the 2 h dataset, the number of trips identified increased by 19%, when we added the trips identified with the *harbour* algorithm to those identified by the *logbook* algorithm (the analysis was restricted to 2003–2004, when logbooks are available for VMS equipped vessels).

The comparison of VMS data at different temporal resolutions showed that the number of identified trips decreased and median trip duration increased significantly ($p < 10^{-15}$ for all Wilcoxon tests) as the interval between consecutive records increases from 10 min to 2 h (Table 1). 58% of the trips with a four times longer median duration were identified in the 2 h *resampled* dataset compared to the 10 min dataset.

The McNemar's test showed that as the temporal resolution decreases so does the percentage of points "correctly" identified as fishing operations ($p < 0.001$; Table 2). For example, 7% of the fishing sets are missed in the 2 h *resampled* dataset in

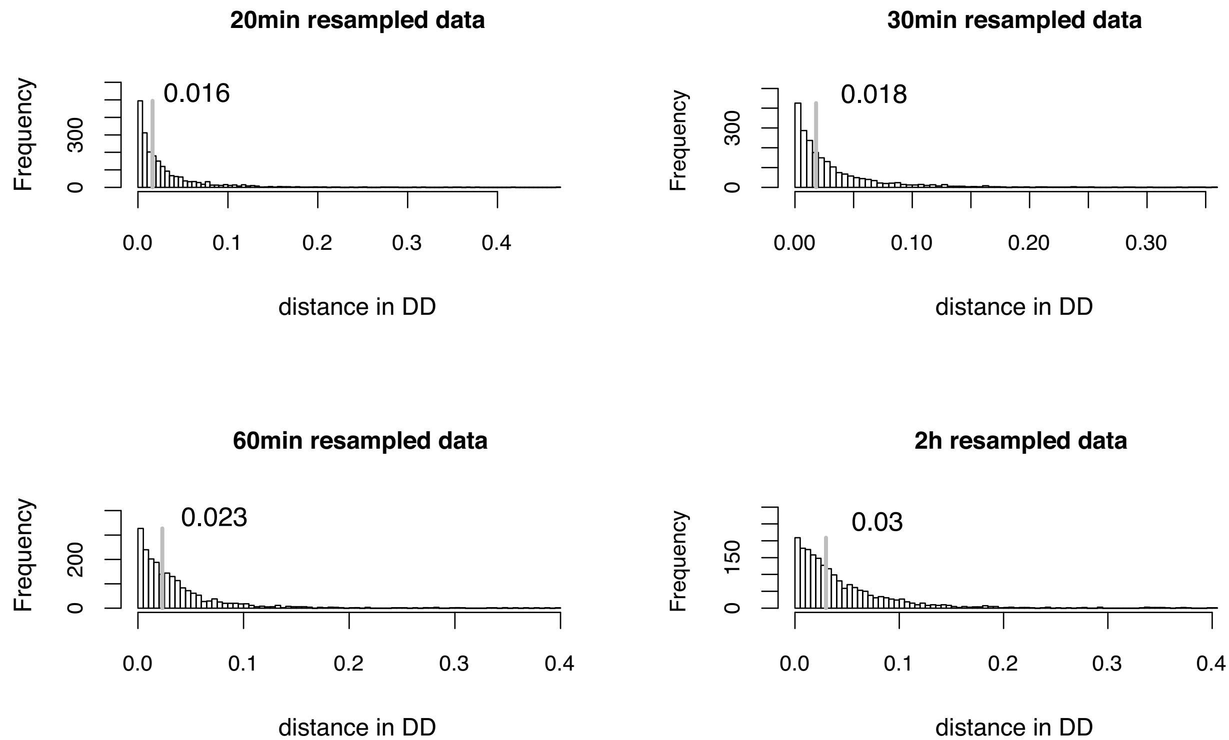


**Fig. 1.** Distribution of distance (in Decimal Degrees) between fishing operations identified in 10 min data and each of the *resampled* datasets (per fishing trip). The median values are indicated with grey lines.

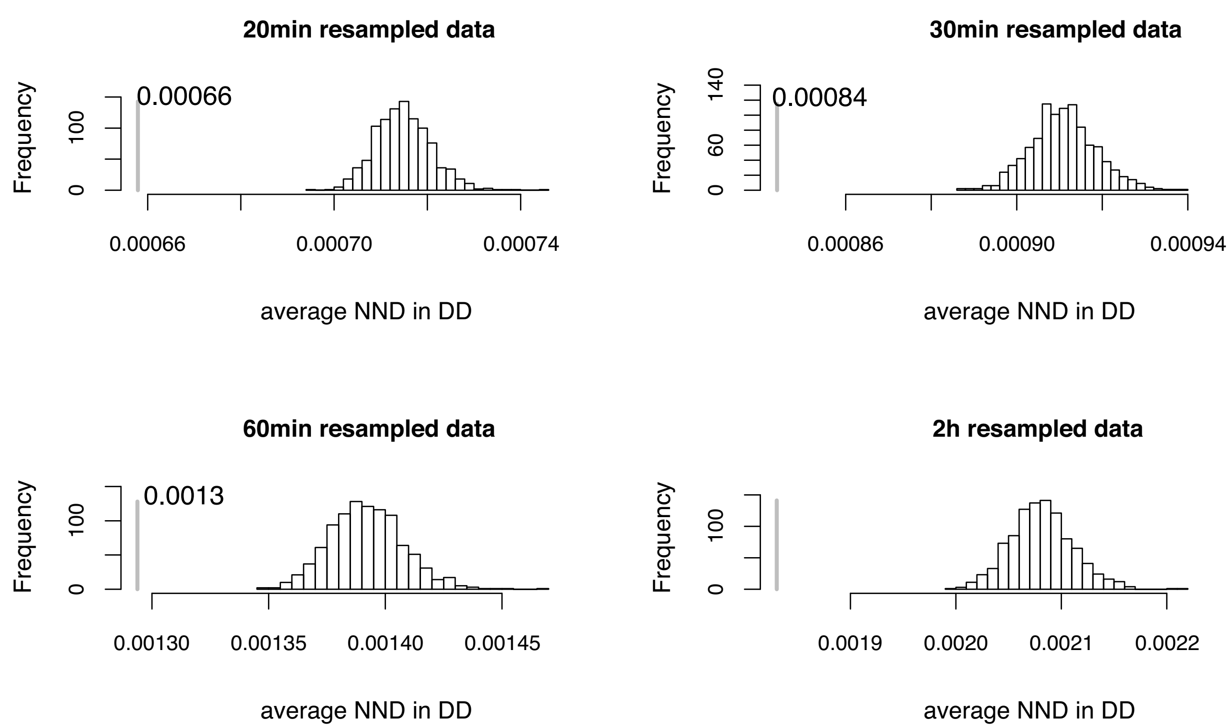


**Fig. 2.** Distribution of Nearest Neighbour Distance values between 10 min data and each of the *resampled* datasets after random re-labelling derived from the Monte Carlo randomisation approach (see Supplementary Material). The values of NND statistics are indicated with grey lines.

comparison with fishing sets identified in 10 min dataset. The median value of the distance between identified fishing operations increased from 0.016 decimal degrees between 10 min and 20 min *resampled* datasets (approximately 1.26 km) to 0.030 decimal degrees between 10 min and 2 h *resampled* datasets (approximately 2.36 km) (Fig. 1). The error follows a lognormal distribution, typical of cases when sources of variation accumulate multiplicatively. Finally, for all *resampled* datasets, the statistic NND was significantly lower than the simulations average ($p = 0.001$; Fig. 2), indicating that, even with a time interval of 2 h, the algorithm is

**Table 2**
McNemar's test: Percentage of points flagged as "non-fishing" in *resampled* datasets but as "fishing" in the 10 min data ($r_{21}$, false negatives) and as "fishing" in the *resampled* datasets but as "non-fishing" in the 10 min data ($r_{12}$, false positives).

| Time interval (min) | False negatives (%) | False positives (%) |
|---|---|---|
| 20 | 0.4 | 1.2 |
| 30 | 2.7 | 1.7 |
| 60 | 3.7 | 2.8 |
| 120 | 7 | 3 |

producing better results than a random assignment of fishing sets: the general spatial pattern derived by the fishing set identification algorithms is informative regardless of the time interval between consecutive locations.

## 4. Discussion

Our results showed the combination of logbook information on trip dates (Bastardie et al., 2010), with the criterion of proximity to a port derived from VMS data(Hintzen et al., 2012) increased the probability of identifying fishing trips by a non-negligible factor for a fleet that typically carries out short trips in the neighbourhood of harbours.

We developed an approach to quantify bias induced by the mismatch between data temporal resolution and trip or fishing operation timescales. The analysis confirmed the hypothesis that lower temporal resolution would lead to misidentifications of trips and fishing operations. For the Portuguese purse seine fishery, an increase of the time interval between consecutive locations from 10 min to 2 h resulted to a loss of 42% of the trips that take place and a bias towards longer trips. This loss should be taken into account when sample sizes are calculated, since even if the whole fleet is equipped with VMS, only 58% of the trips are recorded. Added to this error, 7% of the fishing sets are not identified in 2 h VMS data.

Different uses of VMS data require different temporal resolution of the recorded data. VMS data can reveal violations of area closures (Davies et al., 2007), monitor illegal and unreported fishing (Chang et al., 2010; Detsis et al., 2012; Gribble and Robertson, 1998; Sousa et al., 2006), and promote maritime safety (Shih et al., 2010). In these cases, inability to capture trips and fishing sets by more than 40%, as in the case of the Portuguese purse seiners, is problematic.

As a sampler of fishing activity, VMS provides a far better coverage than alternative methods such as on-board observations (0.25% coverage of fishing trips; Feijó, 2012). For the Portuguese purse seine fishery, the 2 h VMS data proved sufficient to identify fishing grounds since the general spatial patterns of fishing effort were not significantly affected by the temporal resolution of the data.

Given the similarity of purse seine fleets and practices for small pelagics in European waters, the present findings are likely to apply to several other fisheries both in the Atlantic and in the Mediterranean waters (Silva et al., 2015; Tsitsika and Maravelias, 2008). Our approach can be followed to evaluate the levels of uncertainty and bias of fishing set identification for those fisheries characterised by short fishing trips and sets, and according to the end goal of the VMS product. The framework could be used by researchers and managers to identify the optimal data resolution required for each fishery, develop a suitable data collection scheme with regard to the time scales of fishing activities, and ensure the use of VMS and logbook data to their full extent and at a minimum cost.

## Acknowledgements

This study was supported by the Life +MarPro project (Ref. LIFE 09/NAT/PT/000038). VMS and logbook data were provided by *Direcção-Geral de Recursos Naturais Segurança e Serviços Marítimos (DGRM, Portugal)*. The authors thank Diana Feijó and Ana Marçalo for providing information on the operational behaviours of fishers, and Laura Wise for useful suggestions on an early version of the manuscript.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fishres.2016.11.023.

## References

Bastardie, F., Nielsen, J.R., Ulrich, C., Egekvist, J., Degel, H., 2010. Detailed mapping of fishing effort and landings by coupling fishing logbooks with satellite-recorded vessel geo-location. Fish. Res. 106, 41–53, http://dx.doi.org/10.1016/j.fishres.2010.06.016.

Bertrand, S., Díaz, E., Lengaigne, M., 2008. Patterns in the spatial distribution of Peruvian anchovy (Engraulis ringens) revealed by spatially explicit fishing data. Prog. Oceanogr. 79, 379–389, http://dx.doi.org/10.1016/j.pocean.2008.10.009.

Campbell, M.S., Stehfest, K.M., Votier, S.C., Hall-Spencer, J.M., 2013. Mapping fisheries for marine spatial planning: gear-specific vessel monitoring system (VMS), marine conservation and offshore renewable energy. Mar. Policy, http://dx.doi.org/10.1016/j.marpol.2013.09.015.

Chang, S.-K., Liu, K.-Y., Song, Y.-H., 2010. Distant water fisheries development and vessel monitoring system implementation in Taiwan—history and driving forces. Mar. Policy 34, 541–548, http://dx.doi.org/10.1016/j.marpol.2009.11.001.

Davies, A.J., Roberts, J.M., Hall-Spencer, J., 2007. Preserving deep-sea natural heritage: emerging issues in offshore conservation and management. Biol. Conserv. 138, 299–312, http://dx.doi.org/10.1016/j.biocon.2007.05.011.

Detsis, E., Brodsky, Y., Knudtson, P., Cuba, M., Fuqua, H., Szalai, B., 2012. Project Catch: a space based solution to combat illegal, unreported and unregulated fishing: part I: vessel monitoring system. Acta Astronaut. 80, 114–123, http://dx.doi.org/10.1016/j.actaastro.2012.06.009.

EC, 2009. Council Regulation (EC) No. 1224/ 2009 of 20 November 2009 establishing a Community control system for ensuring compliance with the rules of the Common Fisheries Policy, amending Regulations (EC) No. 847/96, (EC) No. 2371/2002, (EC) No. 811/2004, (EC) No. 768/2005, (EC) No.2115/2005, (EC) No.2166/2005, (EC) No. 388/2006, (EC) No.509/2007, (EC) No.676/2007, (EC) No. 1098/2007, (EC) No. 1300/2008, (EC) No. 1342/2008 and repealing Regulations (EEC) No. 2847/93, (EC) No. 1627/94 and (EC) No. 1966/2006.

ESRI, 2011 ArcGIS Desktop: Release 10.

Feijó, D., 2012. Caraterização Da Pesca Do Cerco Na Costa Portuguesa (Mestrado). Faculdade de Ciências da Universidade do Porto, Porto.

Fonseca, T., Campos, A., Afonso-Dias, M., Fonseca, P., Pereira, J., 2008. Trawling for cephalopods off the Portuguese coast—Fleet dynamics and landings composition. Fish. Res. 92, 180–188, http://dx.doi.org/10.1016/j.fishres.2008.01.015.

Gerritsen, H., Lordan, C., 2011. Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. ICES J. Mar. Sci. 68, 245–252, http://dx.doi.org/10.1093/icesjms/fsq137.

Gerritsen, H.D., Lordan, C., Minto, C., Kraak, S.B.M., 2012. Spatial patterns in the retained catch composition of Irish demersal otter trawlers: high-resolution fisheries data as a management tool. Fish. Res. 129–130, 127–136, http://dx.doi.org/10.1016/j.fishres.2012.06.019.

Gribble, N.A., Robertson, J.W.A., 1998. Fishing effort in the far northern section cross shelf closure area of the Great Barrier Reef Marine Park: the effectiveness of area-closures. J. Environ. Manage. 52, 53–67, http://dx.doi.org/10.1006/jema.1997.0160.

Hintzen, N.T., Bastardie, F., Beare, D., Piet, G.J., Ulrich, C., Deporte, N., Egekvist, J., Degel, H., 2012. VMStools: open-source software for the processing, analysis and visualisation of fisheries logbook and VMS data. Fish. Res. 115–116, 31–43, http://dx.doi.org/10.1016/j.fishres.2011.11.007.

ICES, 2011. Report of the Workshop on the utility of commercial CPUE and VMS data in assessment (WKCPUEFFORT), 5–7 April 2011. ICES, Copenhagen, Denmark, ICES CM 2011/ACOM:49. 71 pp. http://www.ices.dk/communit/Documents/PGCCDBS/WKCPUEFFORT%202011.pdf.

Jennings, S., Lee, J., 2012. Defining fishing grounds with vessel monitoring system data. ICES J. Mar. Sci. 69, 51–63, http://dx.doi.org/10.1093/icesjms/fsr173.

Manly, B.F., 2006. Randomization, bootstrap and Monte Carlo methods in biology. CRC Press.

Martín, P., Muntadas, A., de Juan, S., Sánchez, P., Demestre, M., 2014. Performance of a northwestern Mediterranean bottom trawl fleet: how the integration of landings and {VMS} data can contribute to the implementation of ecosystem-based fisheries management. Mar. Policy 43, 112–121, http://dx.doi.org/10.1016/j.marpol.2013.05.009.

McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12, 153–157.

Postlethwaite, C.M., Dennis, T.E., 2013. Effects of temporal resolution on an inferential model of animal movement. PLoS One 8, e57640.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Shih, Y.-C., Chou, C.L., Chiau, W.-Y., 2010. Maritime safety for fishing boat operations and avoidable hijacking in Taiwan. Mar. Policy 34, 349–351, http://dx.doi.org/10.1016/j.marpol.2009.07.007.
Silva, A., Moreno, A., Riveiro, I., Santos, B., Pita, C., Rodrigues, G.C., Villasante, S., Pawlowski, L., Duhamel, E., 2015. Resource Assessment and Social and Economic Situation Sardine Fisheries (No. IP/B/PECH/IC/2015_133).
Sousa, L.P., Brito, A., Abdula, S., Caputi, N., 2006. Research assessment for the management of the industrial shallow-water multi-species shrimp fishery in Sofala Bank in Mozambique. Fish. Res. 77, 207–219, http://dx.doi.org/10.1016/j.fishres.2005.10.009.
Stratoudakis, Y., Marçalo, A., 2002. Sardine slipping during purse-seining off northern Portugal. ICES J. Mar. Sci. J. Cons. 59, 1256–1262, http://dx.doi.org/10.1006/jmsc.2002.1314.
Tsitsika, E.V., Maravelias, C.D., 2008. Fishing strategy choices of purse seines in the Mediterranean: implications for management. Fish. Sci. 74, 19–27, http://dx.doi.org/10.1111/j.1444-2906.2007.01492.x.
Vermard, Y., Rivot, E., Mahévas, S., Marchal, P., Gascuel, D., 2010. Identifying fishing trip behaviour and estimating fishing effort from VMS data using Bayesian Hidden Markov Models. Ecol. Modell. 221, 1757–1769.
Witt, M.J., Godley, B.J., 2007. A step towards seascape scale conservation: using vessel monitoring systems (VMS) to map fishing activity. PLoS One 2, e1111.