7.2 Artículos

Artículo 1

Título: Using mobile device's sensors to identify fishing activity

Autores: **M. M. Galotto-Tebar**, A. Pomares-Padilla, I. A. Czerwinski y
J. C. Gutiérrez-Estrada.

Revista: Journal of Marine Science and Technology

Volumen, número, páginas: 25(3), 978-989

Año: 2019

DOI: 10.1007/s00773-019-00694-5

Indice de impacto JCR (2019): 1.446

Recibido el 8 de febrero de 2019; aceptado el 12 de noviembre de 2019; publicado el 20 de noviembre de 2019; fecha de edición septiembre 2020.

Categoría de JCR	Clasificación en la categoría	Cuartil en la categoría
Engineering, Marine	**7 de 14**	**Q2**

Artículo 2

Título: Is the vessel fishing? Discrimination of fishing activity with low-cost intelligent mobile devices through traditional and heuristic approaches

Autores: **M. M. Galotto-Tebar**, A. Pomares-Padilla, I. A. Czerwinski y
J. C. Gutiérrez-Estrada.

Revista: Expert Systems With Applications

Indice de impacto JCR (2020): 6.954

Enviado el 12 de noviembre de 2020; primera revisión enviada el 26 de marzo de 2021; segunda revisión enviada el 20 de septiembre de 2021.

Categoría de JCR	Clasificación en la categoría	Cuartil en la categoría
Engineering, Electrical and Electronic	**24 de 273**	**Q1-D1**

Using mobile device's sensors to identify fishing activity

M. M. Galotto-Tebar, A. Pomares-Padilla, I. A. Czerwinski & J. C. Gutiérrez-Estrada

Journal of

Marine Science and Technology

Volu

ONLINE FIRST

Springer

Springer

ORIGINAL ARTICLE

Using mobile device's sensors to identify fishing activity

M. M. Galotto-Tebar[1] · **A. Pomares-Padilla[1]** · **I. A. Czerwinski[2]** · **J. C. Gutiérrez-Estrada[3]**

Abstract

Fisheries management is generally based on the regulation of the fishing effort, by limiting fishing capacity and activity. The fishing capacity can be quantified objectively, however, the calculation of the fishing activity requires knowing the effective fishing time, for which it is essential to monitor the vessels activity. The European Vessel Monitoring System (VMS) only describes the geographical position, course and speed of the vessel at 2-h intervals, it is an expensive system used only in vessels over 12 m in length and there are no common criteria to infer the fishing activity from the VMS data. To evaluate more precisely the fishing activity, we propose to incorporate new sensors in the vessels that provide additional information. The sensors of mobile devices offer an economic solution that would allow their implementation throughout the fishing fleet. The objective of this work is to evaluate whether the most common sensors integrated in current mobile devices: GPS, accelerometer, gyroscope, and magnetic field, offer relevant information to identify the different phases of bottom trawling fishing activity. The results obtained indicate that these sensors detect, with very high precision, foreseeable changes in the movement of the vessel during the towing manoeuvre.

Keywords Vessel monitoring system · Mobile device · Sensor · Trawl fishing · Ship's behaviour

1 Introduction

The sustainable exploitation of fishery resources is based on the use of common indicators to measure fishing activity [1]. The European Union, under the Common Fisheries Policy, continuously analyses how to improve the fishing capacity and effort indicators of its fishing fleet [2]. The fishing capacity can be objectively quantified from the vessels' technical characteristics and the fishing gears; however, the calculation of the fishing activity requires knowing the effective fishing time [3]. Therefore, an accurate estimate of the fishing effort requires knowing the time when the fishing gear is working, for which it is essential to monitor the vessels' activity.

With the aim of improving the effectiveness and efficiency of monitoring, control and surveillance of fishing activity, the European Union regulated in 2009 the use of vessel monitoring system (VMS) [4]. Thus, from January 1, 2012, every vessel with a total length of more than 12 m has a VMS transmitter installed, which provides each State Fishing Monitoring Centre with real-time information on the geographical position, course and speed of the vessel.

VMS data have become an indispensable tool to research on the control and planning of fishing activity [5, 6]. For example, they have been used to complete and verify the information provided by the electronic recording and reporting system (ERS) and to establish spatial and temporal maps of fishing activity, to study patterns of fishermen's behaviour [7] as well as to accurately assess the vessels' trajectories [8–14]. Several authors such as Lee [15] and Szostek [16] have developed and validated fishing effort estimation

✉ M. M. Galotto-Tebar
m.galotto@umh.es

A. Pomares-Padilla
pomares@umh.es

I. A. Czerwinski
ivone.czerwinski@uca.es

J. C. Gutiérrez-Estrada
juanc@uhu.es

1 Dpto. de Ingeniería de Computadores, Universidad Miguel Hernández, Campus de Elche, Avenida del ferrocarril s/n, 03202 Elche, Alicante, Spain

2 Dpto. de Biología, Facultad de CC. del Mar y Ambientales, Universidad de Cádiz, Campus de Puerto Real, 11510 Puerto Real, Cádiz, Spain

3 Dpto. de Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Campus de El Carmen, Universidad de Huelva, 21007 Huelva, Spain

methods based on the recorded VMS data in United Kingdom vessels. On the other hand, Marzuki [17] used the VMS data for the monitoring and identification of unauthorised fishing activities. In this sense, Gerritsen [5] analysed the effects on CPUE (catch per unit of effort) of the seasonal closure of cod (*Gadus morhua*) from the VMS data. Currently the information provided by the VMS is available on platforms such as those developed in the Global Fishing Watch projects whose purpose is to provide a tool that allows governments, fishery management organisations, scientists, private industry, and non-governmental organisations to implement rules and regulations that will ensure a sustainable and abundant ocean (www.globalfishingwatch.org).

Despite the undoubted advantages of the availability and use of VMS data, there are certain aspects that hinder to interpret the information they provide. In the first place, as indicated in [4], the installation of transponders is only mandatory for vessels with a length exceeding 12 m, so that the information coming from a large part of the artisanal fleet is not available. On the other hand, the information transmitted is relatively scarce since the devices send data with a very low frequency (intervals of at least 2 h) limiting the reliability and precision of the obtained results. Likewise, the use of the VMS data for the determination of trajectories, spatial and temporal fishing maps, and estimation of the fishing effort requires a validation with the auxiliary information provided by the logbooks and/or the on board observers reports [11]. Another VMS deficiency is that it does not provide information on the state of the sea to assess its influence on the trawl gear performance [18] and improve the fishing effort estimation. In addition, the VMS only provides the vessel's speed over the ground and does not report the vessel's true speed over the surface, obtaining erroneous results in areas with high sea currents. Therefore, we think that to evaluate fishing effort more accurately, it is necessary to develop new devices that provide better quality information and whose cost favours their implementation throughout the fishing fleet.

Considering that the presence of an active trawl gear affects the dynamic behaviour of the vessel, as it is affirmed and mathematically modelled by Sun [19], a device equipped with sensors and sufficient processing capacity could record the movement of the vessel and identify the different phases of fishing activity. This task could be carried out by a mobile device (tablet or Smartphone), thanks to the sensors it usually incorporates, its high processing capacity and facility for the development of specific applications (APP). At present, there are many fishing APPs, most of them for sport fishing, which take advantage of the sensors potential to georeference the fishing activity. For example, Fishing Status (https://play.google.com/store/apps/details?id=com.FishingStatus) is an application that provides resources and information (local fishing reports and hotspots with GPS

coordinates) around 100 miles of the current device location. For its part, Fishing Log (https://play.google.com/store/apps/details?id=com.WebAndPrint.FishDiary) is a logbook application that uses the GPS to georeference the information. Therefore, from the technical point of view, we consider it interesting to explore the use of mobile devices on the control and planning of fishing activity.

The objective of this work is to evaluate if the most common sensors integrated in current mobile devices: GPS, accelerometer, gyroscope and magnetic field are capable of offering relevant information that allows the identification of the different phases of fishing activity.

2 Materials and methods

2.1 Mobile device, sensors, variables and data

Nowadays, most mobile phones and tablets incorporate GPS, accelerometer, gyroscope and magnetic field sensors. In this work, we used a Samsung SM-P600 tablet with a 6-axis inertial sensor (Bosch Sensortec BMI055) consisting of a digital tri-axial 12 bit acceleration sensor and a digital tri-axial 16 bit gyroscope, a 3-axis electronic 16 bit compass sensor (Asahi Kasei Microdevices AK8963C) and a GPS receiver. From these sensors, a total of 11 variables can be obtained with a sampling period ranging from 50 ms to 1 s (Table 1).

The accelerometer measures the acceleration applied to the device ($\overrightarrow{\text{Ace}}$). Conceptually, it does so by measuring forces (F) applied to the mass of the sensor. In particular, the force of gravity ($\vec{G}$) is always influencing the measured acceleration (Eq. 1). By immobilising the device on the vessel, it moves in solidarity with the vessel, so the mass of the equation corresponds to the sum of the mass of the device plus that of the vessel and this term can be disregarded and consider that the accelerometer will record the components of the gravity vector ($\vec{G}$) at each instant:

$$\overrightarrow{\text{Ace}} = -\vec{G} - \overrightarrow{\sum F}/\text{Mass} \approx -\vec{G}. \tag{1}$$

The gyroscope measures, in radians per second (rad/s), the rotation speed around the axes OX, OY and OZ being positive, when the rotation around the axis is made in the anticlockwise direction seen by an observer located in a positive location of the axis.

For its part, the magnetic field sensor measures, in microTeslas (μT), the magnetic field in the device environment. The sensor provides the vector components resulting from the vector sum of the earth's magnetic field and the magnetic field generated by the objects in the environment (motors, wiring, etc.) (Eq. 2):

Table 1 Main characteristics of the sensors used, description of the recorded physical parameter, data generation interval of each sensor (sampling period) and variables extracted from each parameter

Sensor	Measured parameter	Sampling period	Variable
Acceleration	Linear acceleration vector	50 ms	AceX
Range $=\pm 19.6133$ m/s^2			AceY
Sensitivity $= 0{,}009570$ m/s^2			AceZ
Gyroscope	Rotation speed around the three axes	50 ms	GirX
Range $=\pm 8.7266$ rad/s			GirY
Sensitivity $= 0.0002661$ rad/s			GirZ
Magnetic field	Vector of the magnetic field	50 ms	MagX
Range $=\pm 4900$ µT			MagY
Sensitivity $= 0.6$ µT			MagZ
GPS	Bearing	1 s	GpsR
	Speed		GpsV

$$\overrightarrow{\text{Mag}} = \overrightarrow{\text{Mag}_{\text{Earth}}} + \overrightarrow{\text{Mag}_{\text{Environ}}}. \qquad (2)$$

Finally, the GPS (global positioning system) sensor is capable of recording information about its three-dimensional geographic location in the form of latitude, longitude and height based on a satellite network. It also registers a temporary mark of the measurement and generates additional processed information such as course and instantaneous speed, average speed and precision. In this work, it has only been considered necessary to store the bearing (GpsR) and the instantaneous speed (GpsV) every second during the vessel activity.

2.2 The mobile application

A mobile application (APP) has been developed to record the vessel's movement during its fishing activity. The integrated development environment (IDE) Android Studio has been used to implement the APP on the Android platform. The user interface allows to start the recording and to select the haul phase. Once the recording has begun, the interface shows the data recorded by the sensors numerically and graphically. In addition, it incorporates a configuration menu with data about the ship, the mobile device and the fishing gear. During the recording, the data provided by the sensors, linked to the haul phase, ship's information, fishing gear and mobile device, are temporarily stored in a local database (SQLite).

The communication with the mobile device is done through a REST API ("Application Programming Interface for the Representational State Transfer") service. When the mobile device has a connection to the server, it sends the information temporarily stored in its database and the server stores it in a MySql database (Fig. 1). The server allows access to information through the REST API service or through the WEB service using the RStudio application (Fig. 2).
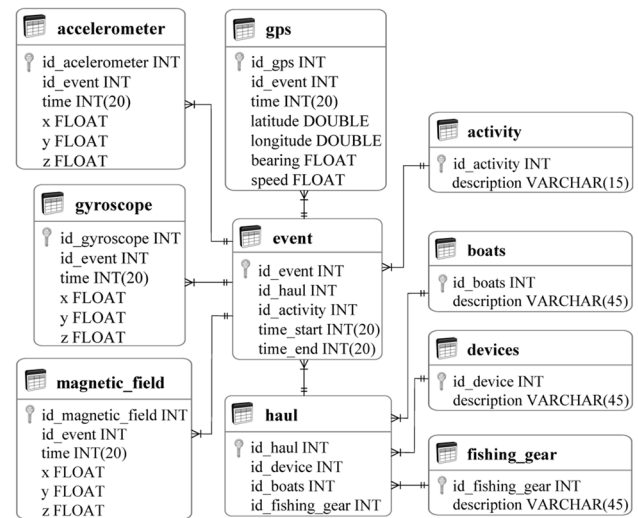
Fig. 1 Structure of the relational database that stores the data recorded by the sensors of the mobile device. The accelerometer table contains the acceleration applied to the mobile device (vector $\{x,y,z\}$) and the instant (time) in which it has been recorded. The *gyroscope* table contains the rotation speed around the three axes (vector $\{x,y,z\}$) of the mobile device and the instant (time) in which it was recorded. The magnetic field table contains the magnetic field around the mobile device (vector $\{x,y,z\}$) and the instant (time) in which it has been recorded. The GPS table contains the position (latitude and longitude), bearing and speed of the vessel and the instant (time) in which it has been recorded. The activity table contains the description of the four different activities of the vessel: shipping, setting, towing and hauling. The boats table contains the description of the fishing vessels involved in the study. The devices table contains the description of the mobile device used. The fishing gear table contains the description of the fishing gear used. Table haul contains the registered haul number. The event table allows to relate each data stored in the accelerometer, gyroscope, magnetic field and GPS tables to the activity, haul number, fishing vessel, fishing gear and mobile recording device

Fig. 2 Mobile application user interface and information flow

Fig. 3 Spatial distribution of the 22 hauls recorded along the coast of Castellón, Tarragona, Barcelona and Gerona (Spain)

2.3 Case study

The initial execution of the prototype was carried out on the fishing and oceanographic research vessel Miguel Oliver [20] of the General Fisheries Secretariat of Spain, with one length of 70.00 m, beam of 14.40 m, draft of 6.50 m, gross register tonnage of 2.495 GT, maximum speed of 14 knots and equipped with a small draft trawl GOC-73 with Morguere doors. The data survey was performed in 2016 during the MEDITS-ES [21] survey conducted annually by the Spanish Institute of Oceanography (IEO) to evaluate demersal resources along the continental shelf and slope of the Spanish Mediterranean coast from May 27 to June 5, 2016 in the coasts of Castellón, Tarragona, Barcelona and Gerona (Fig. 3). During the campaign, the ship movement was recorded along 22 trawl hauls, at 3 knots speed, distributed in 20 sets of 30 min in depths from 50 to 200 m, two sets of 60 min in depth greater than 200 m and an average capture of 60 kg. In each set, it was recorded at least 15 min before the haul (shipping), the complete periods of setting, towing and hauling, and 15 min shipping after the haul.

To record as accurately as possible, the roll and pitch of the ship, the device was placed on the bridge about 18 m above the ship's centre of gravity, crossing point of the pitch and roll axes [22]. Special interest was placed on the orientation of the device as the sensors offer vector information referenced to the coordinate axes linked to the screen, placing the axis OY parallel to the longitudinal axis (stern-bow), the axis OX parallel to the transverse axis (port-starboard) and the OZ axis parallel to the vertical axis (Fig. 4).
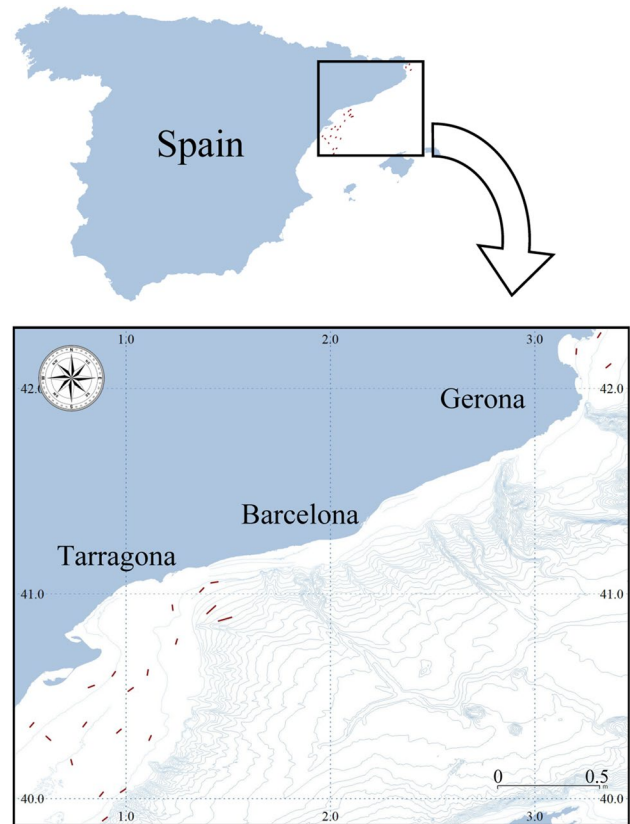
2.4 Data analysis

From the information recorded by the sensors and stored in the database, the data table was made by estimating the mean (X_m) and the standard deviation (X_s) in 1-min periods,
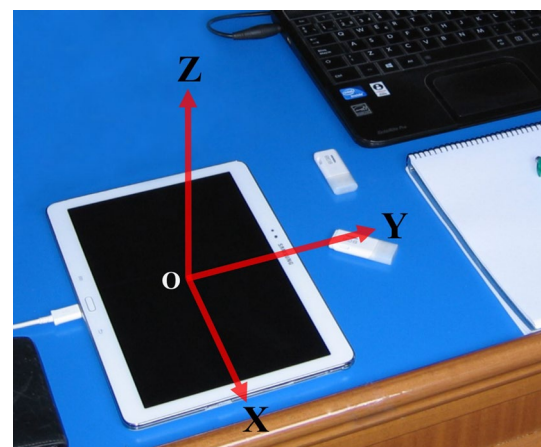

Fig. 4 Location of the mobile device on the bridge of the oceanographic vessel

thus obtaining 22 independent quantitative continuous variables. Likewise, a nominal qualitative independent variable named *Activity* composed of four categories (shipping, shooting the gear, towing and hauling the gear) was created. The data table was formed by 22 variables with 2453 data divided into 1140 shipping, 262 during the gear shooting, 786 towing and 265 during the gear hauling.

The data analysis focused on identifying the changes in the static and dynamic vessel behaviour during the haul, in identifying the sensors capable of measuring those changes, in locating the quantitative variables to be used to recognise said changes, and in analysing the variables behaviour in the different haul phases. Each quantitative variable was analysed using the kurtosis and Fisher's asymmetry coefficients, the Shapiro–Wilk test to contrast the normality, the box and density plots by haul phase to observe differences, their median and interquartile range by haul phase to detect differences in its central value and dispersion, its coefficient of variation of the absolute deviation of the median (MAD) to discard unrepresentative variables and the non-parametric test of Kruskall–Wallis to confirm the differences between the median values by haul phase. Likewise, quantitative variables were standardised to compare them with each other. The statistical package RStudio was used to perform all the analyses.

3 Results

This section shows the obtained results from the data table variables related to the trim of the ship and the four degrees of freedom of movement affected by the start of the haul and susceptible to be registered by the available sensors: pitch, roll, yaw and surge.

3.1 Trim

The presence of an active fishing gear causes changes in the ship's trim, modifying the angle formed by its longitudinal axis with the horizontal plane. The accelerometer is the sensor capable of recording information about the ship's orientation. The variables $AceX_m$, $AceY_m$ and $AceZ_m$ represent the mean gravity vector components and provide information on the ship's orientation with respect to the horizontal plane. The projection of the gravity vector on the centreline plane, represented by the variables $AceY_m$ and $AceZ_m$, makes it possible to evaluate the longitudinal angle of the ship trim.

The variable $AceY_m$ showed a value of kurtosis ($\beta = 4.27$) and a p value of the Shapiro–Wilk test ($p < 0.001$) that suggested that the data did not adjust to a normal distribution. The Kruskall–Wallis test ($p < 0.001$) showed the existence of significant differences in its central value and dispersion between activities. During the set, their median increased

by 42% and their interquartile range was reduced by 58% (Table 2, Fig. 5). On the other hand, $AceZ_m$ presented a clearly non-normal behaviour ($\gamma = 2.74$, $\beta = 21.58$, Shapiro–Wilk test $p < 0.001$) and its density by activity showed the absence of significant differences in its central value and dispersion (Kruskall–Wallis test $p = 0.749$) (Table 2; Fig. 6).

3.2 Roll and pitch

Pitch and roll movements are originated by external forces such as waves, wind, etc. In both cases, the vessel describes the rotational movements around its transverse and longitudinal axis, causing the alternate rise and fall of the bow and stern or of the port and starboard bands. The presence of an active fishing gear causes the displacement of the vessel mass centre (crossing point of the transverse and longitudinal axes) and modifies the pitch and roll movements. The accelerometer, gyroscope and magnetic field sensors detect these changes.

The accelerometer showed higher values of the coefficient of variation in $AceX_s$, $AceY_s$ and $AceZ_s$. Like $AceZ_m$, the deviations associated with the three axes of the accelerator present behaviours with non-normal distributions. In $AceX_s$ and $AceZ_s$ (Kruskall–Wallis $AceX_s$ test $p < 0.001$; Kruskall–Wallis $AceY_s$ test $p = 0.168$; Kruskall–Wallis $AceZ_s$ test $p < 0.001$) (Table 2; Figs. 7, 8).

The gyroscope measures the speed of rotation around the X, Y and Z axes. The sensors orientation on the ship indicates that the variable GirX records the pitch movement and the variable GirY the roll movement. In particular, the $GirX_s$ and $GirY_s$ variables of the data table must change in the variability or dispersion of the pitch and roll rotation speed. The $GirX_s$ and $GirY_s$ components showed a non-normal behaviour and at least one of them ($GirY_s$) pointed out the possibility of differentiating the navigation from the towing, shooting and hauling activities (Kruskall–Wallis $GirY_s$ test $p < 0.001$) (Table 2; Fig. 9), however, $GirX_s$ did not present any difference between the four activities (Kruskall–Wallis test $GirX_s$ $p = 0.198$).

For its part, the magnetic field sensor measures in micro-Teslas (μT), the magnetic field in the environment of the device. Therefore, all the components of the vector that register the sensor (MagX, MagY and MagZ) and particularly the variables $MagX_s$, $MagY_s$ and $MagZ_s$ are likely to show the pitch and roll changes in the different haul phases. The Shapiro–Wilk tests indicate that none of the three variables presented a normal distribution. However, a comparison of medians showed differences between the central value and the dispersion between shipping and towing activities in the variables $MagX_s$, $MagY_s$ (Kruskall–Wallis $MagX_s$ test $p < 0.001$) (Table 2, Figs. 10 and 11) and to a lesser extent in the variable $MagZ_s$.

Table 2 Statistical summary of the 22 variables under study

	$AceX_m$	$AceY_m$	$AceZ_m$	$GirX_m$	$GirY_m$	$GirZ_m$	$MagX_m$	$MagY_m$	$MagZ_m$	$GpsR_m$	$GpsV_m$
N	2453	2453	2453	2453	2453	2453	2453	2453	2453	2453	2453
Mean	− 1.0302	0.0412	9.8941	− 0.0001	0.0003	0.0000	− 29.2530	60.0077	− 15.2240	140.8623	4.9530
Median	− 1.0420	0.0440	9.8940	0.0000	0.0000	0.0000	− 33.5270	65.4720	− 16.0900	114.2770	3.4580
SD	0.0692	0.0162	0.0121	0.0005	0.0005	0.0027	11.6915	15.1598	3.0334	103.6299	2.9446
CV_Pearson	0.0672	0.3932	0.0012	4.7825	1.4049	62.8112	0.3997	0.2526	0.1992	0.7357	0.5945
MAD	0.0667	0.0163	0.0074	0.0000	0.0000	0.0000	12.7593	17.3242	2.8807	116.2092	1.5879
CV_MAD	0.0640	0.3707	0.0007	–	–	–	0.3806	0.2646	0.1790	1.0169	0.4592
Min	− 1.4660	− 0.0420	9.8300	− 0.0040	− 0.0010	− 0.0270	− 50.0170	34.9760	− 20.8220	2.6130	0.0580
Max	− 0.6170	0.0870	9.9930	0.0020	0.0020	0.0240	− 6.8830	81.2610	− 7.5720	356.0770	10.3920
Range	0.8490	0.1290	0.1630	0.0060	0.0030	0.0510	43.1340	46.2850	13.2500	353.4640	10.3340
Skewness	0.3780	− 0.8532	2.7354	− 1.7448	0.7071	0.0192	0.4113	− 0.3437	0.6319	0.5513	0.6494
Kurtosis	4.2573	4.2739	21.5804	14.2086	1.6542	36.7724	1.6587	1.5022	2.4319	2.0463	1.9805
Kr.Wallis p val	0.1177	0.0000	0.7489	0.0158	0.9760	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	$AceX_s$	$AceY_s$	$AceZ_s$	$GirX_s$	$GirY_s$	$GirZ_s$	$MagX_s$	$MagY_s$	$MagZ_s$	$GpsR_s$	$GpsV_s$
N	2453	2453	2453	2453	2453	2453	2453	2453	2453	2453	2453
Mean	0.0872	0.0680	0.0645	0.0039	0.0032	0.0019	0.4904	0.5637	0.4417	7.6063	0.2339
Median	0.0680	0.0540	0.0490	0.0030	0.0020	0.0020	0.3290	0.4030	0.3940	1.0430	0.0690
SD	0.0568	0.0395	0.0450	0.0019	0.0020	0.0013	0.7251	0.7283	0.3198	22.0062	0.4935
CV_Pearson	0.6516	0.5815	0.6973	0.4818	0.6135	0.6631	1.4785	1.2921	0.7240	2.8931	2.1097
MAD	0.0356	0.0297	0.0282	0.0015	0.0015	0.0015	0.0593	0.0771	0.1097	0.9563	0.0474
CV_MAD	0.5233	0.5491	0.5749	0.4942	0.7413	0.7413	0.1803	0.1913	0.2785	0.9169	0.6876
Min	0.0180	0.0090	0.0150	0.0020	0.0010	0.0010	0.2330	0.2820	0.2540	0.1040	0.0000
Max	0.4170	0.2610	0.3350	0.0130	0.0150	0.0140	16.2410	16.8470	11.6380	179.0620	4.3130
Range	0.3990	0.2520	0.3200	0.0110	0.0140	0.0130	16.0080	16.5650	11.3840	178.9580	4.3130
Skewness	1.8170	1.2624	1.9776	1.0152	1.9266	3.5891	9.6756	9.3140	20.9884	4.8946	4.5567
Kurtosis	6.9348	4.2777	7.8477	3.5099	7.5787	23.9303	140.6254	139.7816	653.1549	29.7713	28.2055
Kr.Wallis p val	0.0000	0.1677	0.0000	0.1977	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

The variable with subscript m represents the average of the data generated by the sensor in periods of 1 min. The variable with subscript s represents the standard deviation of the data generated by the sensor in periods of 1 min. For each variable, the table includes the number of samples (N), the mean, the median, the standard deviation (SD), the Pearson's Coefficient of Variation (CV_Pearson), the median absolute deviation (MAD), the coefficient of variation of the median absolute deviation (CV_MAD), the minimum value (min), the maximum value (max), the range of values (range), the skewness, the kurtosis and the p val of the Kruskal–Wallis normality test (Kr.Wallis p val) double

3.3 Yaw

The yaw movement is a rotation oscillatory movement around the ship OZ axis that causes small changes in the course of the vessel. The sensors capable of recording the yaw are the gyroscope, the magnetic field sensor and the GPS.

The Z component of the gyroscope measures the speed of rotation around the OZ axis and the variable $GirZ_s$ records the magnitude of the course changes during the fishing activity. The $GirZ_s$ data had a clearly asymmetric distribution ($\gamma = 3.59$) that caused the Shapiro–Wilk test to be highly significant ($p < 0.001$). Likewise, it showed a high variability around its median, which was reflected in the level of significance of the Kruskall–Wallis test ($p < 0.001$) (Table 2, Fig. 12).
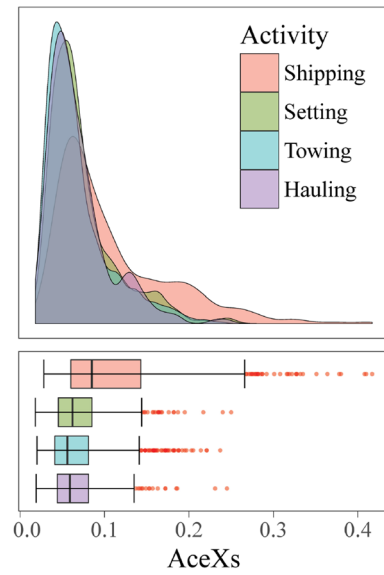
In turn, the magnetic field sensor X and Y components are the most sensitive ones to the ship yaw along with the variables described in the previous section $MagX_s$ and $MagY_s$ that reflect its magnitude.

The GPS sensor provided also information on vessel the course and speed. $GpsR_s$ is the more sensitive to yaw variable. This variable showed a high variability of 92% around its median (Table 2). The density diagram (Fig. 13) showed significant differences in its central value and dispersion among activities (Kruskall–Wallis test $p < 0.001$). Its median decreased by 16% in the shooting phase, increased by 53% in the towing phase and by 160% in the hauling phase.
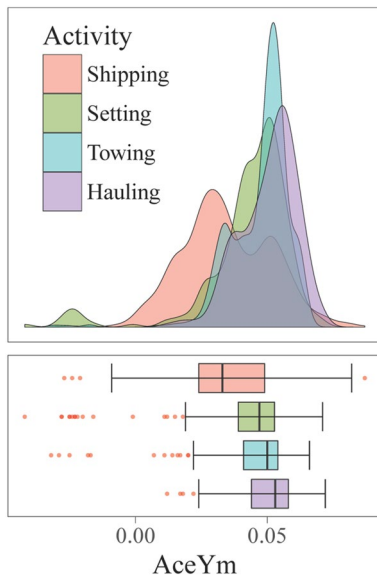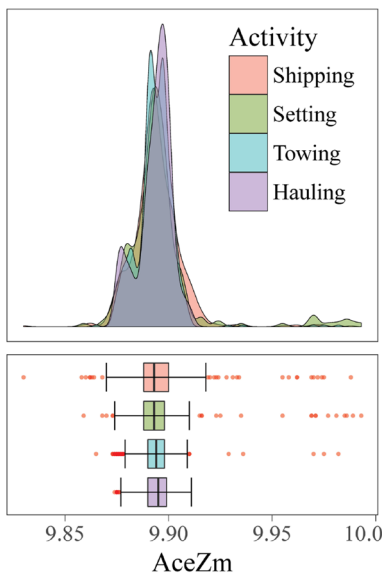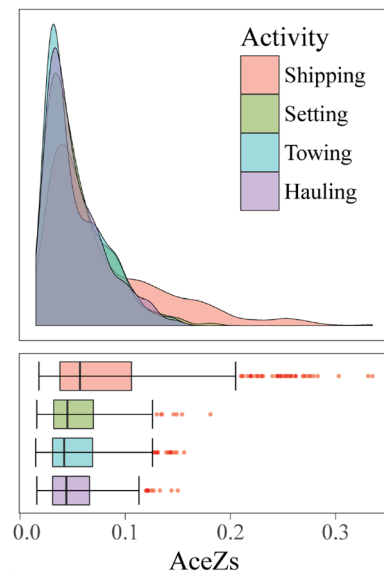
Fig. 5 Tukey boxplots and density by activity of the variable $AceY_m$

Fig. 6 Tukey boxplots and density by activity of the variable $AceZ_m$

Fig. 7 Tukey boxplots and density by activity of the variable $AceX_s$
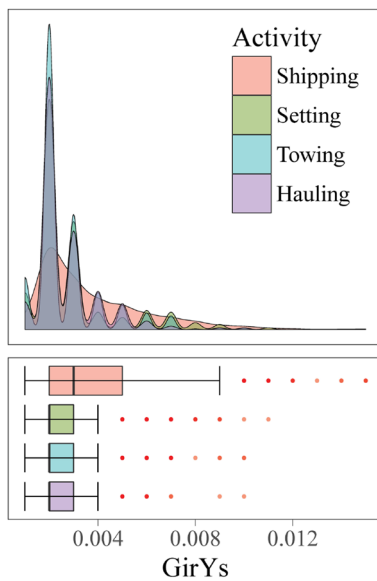

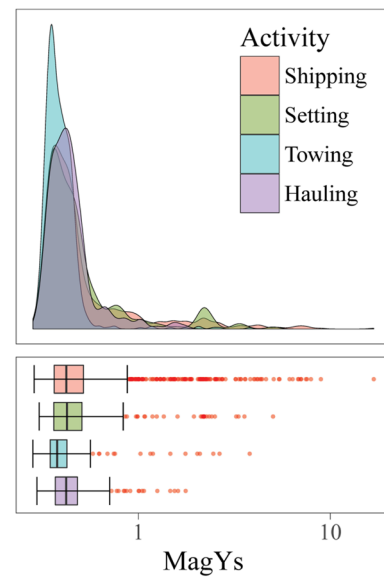**Fig. 8** Tukey boxplots and density by activity of the variable $AceZ_s$

3.4 Surge

The study of the surge speed during the different fishing activity phases focuses on analysing the average speed and speed stability against external forces. In this way, the sensor that measures the surge speed is the GPS and $GpsV_m$ is the variable to show the average speed evolution. This variable clearly discriminates between the navigation and the rest of the activity phases and this was reflected in significant differences of its central value and dispersion (Table 2; Fig. 14).

On the other hand, the surge stability is characterised by the variable $GpsV_s$ (speed variability). The analysis of this variable indicated that both, the asymmetry and kurtosis coefficients, were out of range ($\gamma = 4.56$, $\beta = 28.21$). A Shapiro–Wilk test ($p < 0.001$) confirmed that the data did not show a normal distribution. The variable showed a high variability of 69% around its median (Table 2). The activity density (Fig. 15) and the Kruskall–Wallis test ($p < 0.001$) showed significant differences in their central value and dispersion among the activities. Its median decreased by 24% in the towing phase and increased by 125% in the shooting phase and by 62% in

Fig. 9 Tukey boxplots and density by activity of the variable $GirY_s$

Fig. 11 Tukey boxplots and density by activity of the variable $MagY_s$. Logarithmic scale was used on the OX axis
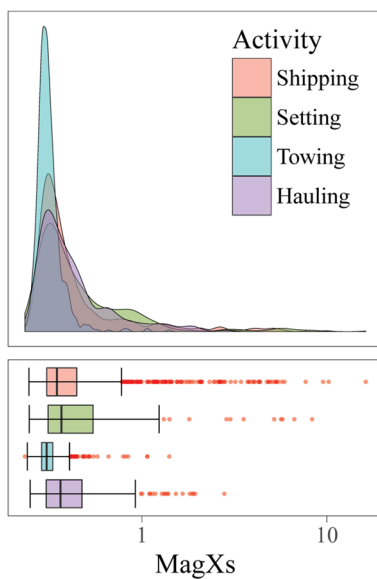
Fig. 10 Tukey boxplots and density by activity of the variable $MagX_s$. Logarithmic scale was used on the OX axis
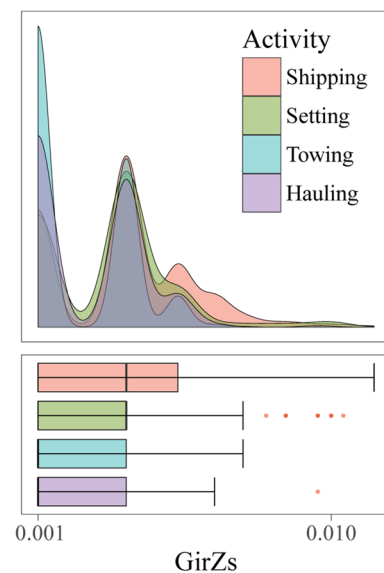

Fig. 12 Tukey boxplots and density by activity of the variable $GirZ_s$. Logarithmic scale was used on the OX axis

the hauling phase. Its interquartile range decreased by 33% in the towing phase and increased by 267% in the shooting phase and by 93% in the hauling phase.

4 Discussion

Significant progress in data collection involves the acquisition of this information for all vessels, including those that are not required to have a VMS system installed, for which it

is essential to develop effective, low-cost and easy-to-install systems. Faced with the high cost of current VMS devices, the solution we propose for each fishing vessel is a mobile device running a specific APP and a modem that allows its connection to a satellite network. In the coastal fleet, the cost can be reduced if the communication is done through the mobile phone network or using the WiFi network, when arriving at the port. The estimation of the cost has been calculated using a medium-range mobile phone and an Iridium
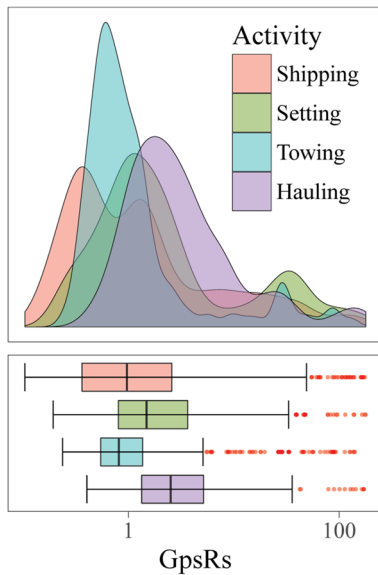
Fig. 13 Tukey boxplots and density by activity of the variable $GpsR_s$. Logarithmic scale was used on the OX axis

Fig. 15 Tukey boxplots and density by activity of the variable $GpsV_s$. Logarithmic scale was used on the OX axis

Fig. 14 Tukey boxplots and density by activity of the variable $GpsV_m$

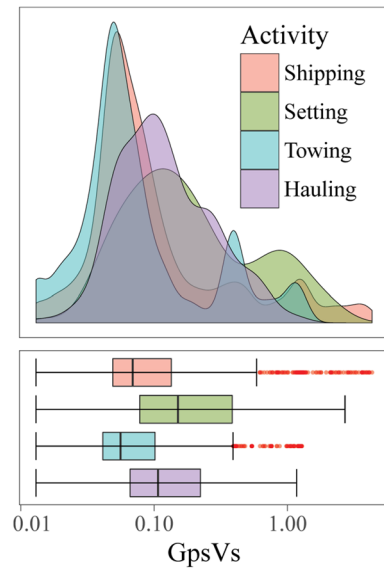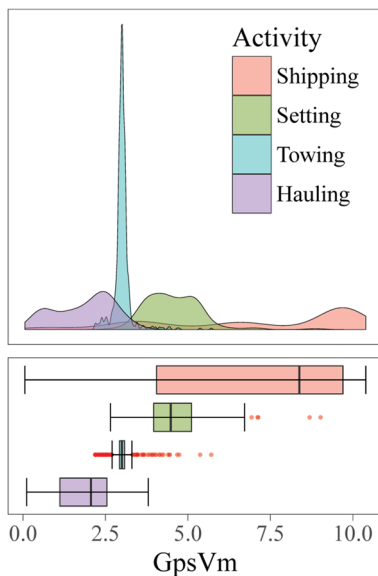modem with worldwide coverage, obtaining an estimated cost of 700 € for local equipment and 400 €/year of maintenance and data transmission. These numbers contrast with the current costs of the VMS for deep-sea fishing of 3500 € for local equipment and 800 €/year of maintenance, or the costs for coastal fishing such as the SLSEPA system of 1400 € for local equipment and 1200 €/year of maintenance [23].

To evaluate if the sensors of mobile devices have the necessary sensitivity to record changes in the movement of the vessel when the trawl is active, we found the most

unfavourable test environment, where these changes were minimal. The effect of the trawl fishing gear on the vessel depends on the size of the vessel, the size of the fishing gear and the volume of catch dragged [19]. On the one hand, the larger the vessel, the smaller the movement due to wind and waves. On the other hand, the smaller the fishing gear and the volume of catch with respect to the size of the vessel, the smaller its effect on the vessel and the differences in the dynamic behaviour of the vessel between each phase of the fishing set will be smaller. The oceanographic vessel Miguel Oliver during the Medits campaign meets these requirements, it is a large vessel equipped with a small sampling trawl, executing during the campaign sets of short duration and with a small volume of catch. A commercial trawler usually offers more favourable conditions for sensors to record differences in their dynamic behaviour during the set. Therefore, we can state that the results presented in this paper can be applied to most of the trawl fishing fleet. These results are analysed below.

4.1 Trim

During the set, the fishing gear is towed by two cables located on both sides of the stern of the vessel, each one exerting a pulling force on the vessel that causes changes in its the static and dynamic behaviour [19]. This force appears when the gear is released, increases during the trawling phase and is maximal when the trawling speed increases during the hauling phase. The vertical component of these forces causes a progressive ship centre of mass displacement aft and consequently an increase in the ship seat.

4.2 Pitch and roll

When a ship is in equilibrium position, the centre of buoyancy (point through which the force of buoyancy supporting the vessel acts vertically upwards) and the centre of gravity (point through which all of the weight of the vessel acts vertically downwards) are in the same vertical axis. When an external force inclines the ship, the submerged part of the hull is modified; the centre of the hull moves and a righting pair that opposes this inclination is generated [22, 24].

The presence of a fishing gear increases the righting torque and reduces the vessel pitch and roll [19]. In both movements, there are variations in speed and drag that cause an additional centre of masses displacement increasing its distance from the centre of the hull.

Our results show that the accelerometer, gyroscope and magnetic field sensors were able to detect a decrease in the pitch and roll level during the haul. In the case of the accelerometer, the variables $AceX_s$ and $AceZ_s$ reflected a significant reduction in their median and interquartile range. On the other hand, the $GirY_s$ gyroscope variable reflected a reduction of its median and interquartile range. Finally, the $MagX_s$ and $MagY_s$ magnetic field sensor variables recorded a reduction in its median and interquartile range during the towing phase.

4.3 Yaw

The yaw movement is a rotation oscillatory movement around the ship's OZ axis that causes small changes in the vessel's course. A starboard turn accelerates the port drag line and slows the starboard line, causing an imbalance in the drag forces that opposes the ship's turn. A turn to port provokes the same effect. As a result, during the haul, the yaw movement is reduced and the ship's course remains more stable [19].

The gyroscope, the magnetic field sensor and the GPS have been able to detect the reduction of the yaw angle during haul. In the case of the gyroscope, the $GirZ_s$ variable has shown a reduction in the towing phase in its median and interquartile range. As mentioned in the previous section, the $MagX_s$ and $MagY_s$ variables of the magnetic field sensor have also shown greater stability on the course during the haul. Finally, the results obtained with the GPS indicate that the variable shows small reductions of its median and a significant reduction of its interquartile range between the shipping and towing phases. During the setting and hauling phases, there is an increase of this movement due to the fact that in these phases, there are usually course changes to orient the vessel to the haul that will begin or to direct the vessel to the next fishing zone.

4.4 Surge

The study of the ship speed during the different fishing activity phases focuses on analysing the average speed reached voluntarily by acting on the propulsion system and observing the stability of the desired speed against external forces (waves, gusts of wind, etc.).

The variable $GpsV_m$ reflected, in each haul phase, velocity values similar to those found in other studies for trawl fishing [15, 17, 25], with medians of 8.4 knots shipping, 4.5 knots setting, 3 knots towing and 2.1 knots hauling. Furthermore, the expected variability of the speed is reflected in its interquartile ranges of 5.7 knots shipping, 1.2 knots setting, 0.16 knots towing and 1.4 knots hauling.

During the haul, small variations in speed cause changes in the drag force that opposes to the speed change. The trawl net shock absorbing effect allows the vessel to maintain a more stable forward speed [19]. The analysis of the variable $GpsV_s$ data shows that the speed value and speed variation are lower in the towing phase than in the shipping phase. During the set and haul phases, the speed varies voluntarily, which justifies an increase in the $GpsV_s$ median and dispersion.

5 Conclusions

Processing data into information and information into knowledge to optimise the decision-making process is one of the great challenges facing the managing bodies that plan the exploitation of fishery resources. In this context, the VMS system has become a relevant source of data, boosting research on the fishing activity control and planning as can be seen in Gerritsen [5] and Jennings [6]. In spite of this, the scarce information offered by the VMS system limits its future.

Analysing the results of this article, the proposal to use in the future, a mobile-like device to overcome the VMS systems seems more viable. As a first step, we have verified that the sensors of a mobile device are capable of detecting significant differences in the movement of the vessel, as stated and mathematically modelled by Sun [19]. The experience was realised under very unfavourable conditions, in a large tonnage oceanographic vessel, with a small fishing gear, a small catch volume, and in weather conditions that did not cause large movements of the vessel. In spite of this, the sensors of the mobile device recorded a progressive increase in the ship's seat, a reduction in the angle of balance, pitch and yaw, and an increase in the stability of surge and speed values according to the haul phase.

In short, the results of this work indicate that the mobile device's sensors have been able to detect, with a very high

accuracy level, the foreseeable changes in the movement of the vessel during the trawling manoeuvre.

6 Software availability

Name of software: FAMIS. Developers: Galotto-Tebar, M.M. Hardware requirements: Mobile device with Android 5.1 or higher. Programming languages: Java. Availability: The programme will be available up on request to Ms. Maria del Mar Galotto-Tebar.

References

1. FAO Fishery Resources Division (1999) Indicators for sustainable development of marine capture fisheries. FAO Technical Guidelines for Responsible Fisheries. No 8. Rome, p 68. http://www.fao.org/3/a-x3307e.pdf
2. European Commission (2007) Communication from the commission to the council and the European Parliament on improving fishing capacity and effort indicators under the common fisheries policy. Brussels, 5.2.2007 COM(2007) 39 final. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52007DC0039&from=LV
3. de Souza EN, Boerder K, Matwin S, Worm B (2016) Improving fishing pattern detection from satellite AIS using data mining and machine learning. PLoS One 11:e0158248. https://doi.org/10.1371/journal.pone.0158248
4. 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. Official Journal of the European Union, L343: 1–50. Strasbourg. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009R1224&rid=1
5. Gerritsen HD, Lordan C, Minto C, Kraak SBM (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. https://doi.org/10.1016/J.FISHRES.2012.06.019
6. Jennings S, Lee J (2012) Defining fishing grounds with vessel monitoring system data. ICES J Mar Sci 69:51–63. https://doi.org/10.1093/icesjms/fsr173
7. Shepperson J, Murray LG (2016) Use of a choice-based survey approach to characterise fishing behaviour in a scallop fishery. Environ Modell Softw 86:116–130. https://doi.org/10.1016/j.envsoft.2016.09.013
8. Palmer MC (2008) Calculation of distance travelled by fishing vessels using GPS positional data: a theoretical evaluation of the sources of error. Fish Res 89:57–64. https://doi.org/10.1016/J.FISHRES.2007.09.001
9. Deng R, Dichmont C, Milton D, Haywood M, Vance D, Hall N, Die D (2005) Can vessel monitoring system data also be used to study trawling intensity and population depletion? The example of Australia's northern prawn fishery. Can J Fish Aquat Sci 62:611–622. https://doi.org/10.1139/f04-219
10. Mills CM, Townsend SE, Jennings S, Eastwood PD, Houghton CA (2007) Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. ICES J Mar Sci 64:248–255. https://doi.org/10.1093/icesjms/fsl026
11. Bastardie F, Nielsen JR, 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. https://doi.org/10.1016/j.fishres.2010.06.016
12. 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. https://doi.org/10.1093/icesjms/fsq137
13. Hintzen NT, Bastardie F, Beare D, Piet GJ, 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. https://doi.org/10.1016/J.FISHRES.2011.11.007
14. Cojan M, Burgos C (2015) Análisis de la información proporcionada por los sistemas de localización vía satélite de la flota que explota la chirla (*Chamelea gallina*) en el Golfo de Cádiz sistemas de localización vía satélite de la flota que explota la chirla (*Chamelea gallina*). In: Teledetección: Humedales y Espacios. XVI Congreso de La Asociación Española de Teledetección, pp 550–553. http://www.aet.org.es/congresos/xvi/XVI_Congreso_AET_actas.pdf
15. Lee J, South AB, Jennings S (2010) Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. ICES J Mar Sci 67:1260–1271. https://doi.org/10.1093/icesjms/fsq010
16. Szostek CL, Murray LG, Bell E, Kaiser MJ (2017) Filling the gap: using fishers' knowledge to map the extent and intensity of fishing activity. Mar Environ Res 129:329–346. https://doi.org/10.1016/J.MARENVRES.2017.06.012
17. Marzuki MI, Garello R, Fablet R, Kerbaol V, Gaspar P (2015) Fishing gear recognition from VMS data to identify illegal fishing activities in Indonesia. In: MTS/IEEE OCEANS 2015. Institute of Electrical and Electronics Engineers Inc, Genoa, pp 1–5. https://doi.org/10.1109/OCEANS-Genova.2015.7271551
18. Politis PJ, DeAlteris JT, Brown RW, Morrison AT (2012) Effects of sea-state on the physical performance of a survey bottom trawl. Fish Res 123–124:26–36. https://doi.org/10.1016/J.FISHRES.2011.11.017
19. Sun X, Yin Y, Jin Y, Zhang X, Zhang X (2011) The modelling of single-boat, mid-water trawl systems for fishing simulation. Fish Res 109:7–15. https://doi.org/10.1016/J.FISHRES.2010.12.027
20. Buque oceanográfico Miguel Oliver (2019) Ministerio de agricultura, pesca y alimentación. https://www.mapa.gob.es/es/pesca/temas/proteccion-recursos-pesqueros/buques-secretaria-general/investigacion-pesquera-oceanografica/miguel-oliver/default.aspx. Accessed on 24 Jul 2019 **(Online)**

21. Bertrand JA, Gil de Sola L, Papaconstantinou C, Relini G, Souplet A (2002) The general specifications of the MEDITS surveys. Sci Mar 66:9–17. https://doi.org/10.3989/scimar.2002.66s29

22. Ibrahim RA, Grace IM (2010) Modelling of ship roll dynamics and its coupling with heave and pitch. Math Probl Eng 2010:1–32. https://doi.org/10.1155/2010/934714

23. Agricultura y Pesca C (2010) Location and track system for Andalusian Fishing Vessels (SLSEPA) (Technical report). Junta de Andalucía. http://www.faocopemed.org/pdf/reg_net_dbases/locationandtracksystemforandalusianfishingvessels(slsepa).pdf

24. Romano A, Marasco A (2014) Fluid dynamics and ship motion. Birkhäuser, New York, pp 429–461. https://doi.org/10.1007/978-1-4939-1604-7_13

25. Burgos C, Gil J, del Olmo LA (2013) The Spanish blackspot seabream (*Pagellus bogaraveo*) fishery in the Strait of Gibraltar: spatial distribution and fishing effort derived from a small-scale GPRS/GSM based fisheries vessel monitoring system. Aquat Living Resour 26:399–407. https://doi.org/10.1051/alr/2013068

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Expert Systems With Applications

Is the vessel fishing? Discrimination of fishing activity with low-cost intelligent mobile devices through traditional and heuristic approaches

Manuscript Number:	ESWA-D-20-06080R2
Article Type:	Full length article
Keywords:	Vessel monitoring system; mobile device; Sensor; trawl fishing; ship's behavior; Machine learning
Corresponding Author:	Maria del Mar Galotto-Tébar, PhD student Universidad Miguel Hernandez de Elche Elche, Alicante SPAIN
First Author:	Maria del Mar Galotto-Tébar, PhD student
Order of Authors:	Maria del Mar Galotto-Tébar, PhD student
	Alejandro Pomares Padilla, Ph.D
	Ivone Alejandra Czerwinski, Ph.D
	Juan Carlos Gutiérrez Estrada, Ph.D
Abstract:	Knowing the activity of fishing vessels accurately and in real time means a leap in quality in the management of fishing activity. This paper presents the development of a new comprehensive fisheries monitoring system (FAMIS) that can complement and overcome the limitations of current fishing vessel monitoring systems (VMS). FAMIS is developed on the basis of a low-cost mobile device with GPS sensors, accelerometer, gyroscope and magnetic field and integrates different statistical methods (discriminant functions) and heuristics (artificial neural networks and vectorial support machines) as techniques to classify the information recorded by the sensors of a mobile device during fishing activity. The results obtained with FAMIS indicate that, in general, heuristics have a high degree of discrimination of each of the phases of fishing operation and that, in particular, multilayer perceptrons (MLPs) are capable of correctly identifying 96.3% of towing phases using only GPS and gyro sensors.

Departamento de Ingeniería de Computadores
UNIVERSIDAD MIGUEL HERNÁNDEZ
Campus de Elche, Avenida del ferrocarril s/n, 03202 Alicante, SPAIN

Ms. Maria del Mar Galotto Tébar
Phone and fax: 34 966 65 83 96, E-mail: m.galotto@umh.es

March 26, 2021

Manuscript title: Is the vessel fishing? Discrimination of fishing activity with low-cost intelligent mobile devices through traditional and heuristic approaches.

Authors: Maria del Mar GALOTTO-TÉBAR, Alejandro POMARES-PADILLA, Ivone Alejandra CZERWINSKI, Juan Carlos GUTIERREZ-ESTRADA.

Recommend Editor
Dr. Binshan Lin
Editor-in-Chief, Expert Systems with Applications

Dear Dr. Lin:

Please find enclosed one copy of the revised manuscript "Is the vessel fishing? Discrimination of fishing activity with low-cost intelligent mobile devices through traditional and heuristic approaches " written by Maria del Mar GALOTTO-TÉBAR, Alejandro POMARES-PADILLA, Ivone Alejandra CZERWINSKI y Juan Carlos GUTIERREZ-ESTRADA.

I also attach a copy of the document "Detailed Response to Reviewers" with the response to the editors' and reviewers' comments.

Sincerely,

Ms. Maria del Mar Galotto-Tébar

Highlights:

- European fishing ships are controlled by a vessel monitoring system (VMS).
- FAMI system exceeds in reliability and accuracy the current VMS
- The mobile' sensors can be used by learning machines to predict fishing events.
- Multilayer perceptrons identified correctly 96.26% of the fishing events.

Table 1. Sensors used: main characteristics, data generation interval of each sensor (sampling period) and variables extracted from each sensor)

Sensor	*Range*	*Sensitivity*	Sampling period	Variable
Acceleration	$\pm 19.6133\ m/s^2$	*0,009570 m/s²*	50 ms	AceX AceY AceZ
Gyroscope	$\pm 8.7266\ rad/s$	*0.0002661 rad/s*	50 ms	GirX GirY GirZ
Magnetometer	$\pm 4900\ \mu T$	*0.6 µT*	50 ms	MagX MagY MagZ
GPS	-	-	1 s	GpsR GpsV

Table 2. *Accuracy*, *precision*, *recall* and *F1 score* (expressed as percentages) in the final test of the LDA, MLP, PNN and SVM learning machines. This analysis included the samples recorded during all categories of activity with all their attributes.

Activity	*Accuracy*	*Precision*	*Recall*	*F1 Score*
LDA				
Steaming	73.63	47.66	55.73	51.38
Setting	77.82	61.07	31.11	41.22
Towing	80.30	62.76	52.14	56.96
Hauling	78.76	55.05	82.05	65.89
MLP				
Steaming	72.15	46.96	70.17	55.93
Setting	80.00	73.00	32.43	43.66
Towing	87.24	71.46	81.97	76.26
Hauling	87.60	78.66	69.42	73.51
PNN				
Steaming	76.88	54.26	47.86	50.86
Setting	71.88	37.10	17.95	24.19
Towing	68.08	40.24	57.09	47.21
Hauling	71.45	44.16	53.68	48.46
SVM				
Steaming	83.80	99.05	35.56	52.33
Setting	86.45	76.69	65.81	70.84
Towing	68.59	44.20	97.61	60.84
Hauling	84.83	84.23	48.38	61.45

Table 3. *F1 score* (expressed as a percentage) in the final test of the LDA, MLP, PNN and SVM learning machines. The samples are included with all their attributes. *1-vs-1* includes samples from two categories. *1-vs-Rest* includes all the samples, maintaining the category on the left and grouping the other categories in the *Rest* category.

Type of partition	LDA		MLP		PNN		SVM	
1-vs-1								
Steaming-vs-Setting	59.37	35.58	69.48	44.82	34.46	68.11	57.65	73.77
Steaming-vs-Towing	74.34	77.12	84.94	85.04	51.52	75.39	63.56	78.87
Steaming-vs-Hauling	79.77	84.01	85.12	86.42	58.25	76.27	66.59	79.97
Setting-vs-Towing	90.66	91.05	92.49	91.84	25.89	70.14	87.24	88.58
Setting-vs-Hauling	95.21	95.22	94.14	94.47	81.59	83.44	93.81	93.17
Towing-vs-Hauling	81.23	80.28	66.37	75.58	75.88	54.79	84.83	79.27
1-vs-Rest								
Steaming-vs-Rest	69.43	63.15	75.85	67.77	48.51	73.05	61.26	71.79
Setting-vs-Rest	46.09	65.88	53.16	72.90	73.12	60.32	62.08	69.30
Towing-vs-Rest	60.54	61.67	88.72	88.54	69.60	22.46	71.03	34.95
Hauling-vs-Rest	80.09	70.55	81.84	79.71	74.73	64.62	77.48	61.32

Table 4. *F1 score* (expressed as a percentage) for the towing category in the final test of the LDA, MLP, PNN and SVM learning machines with all the attributes and after reducing the number of categories. The results without a reduction and with 3 and 2 categories. The symbol ">" written between two categories indicates that the category on the left is merged with that on the right, the label of the samples of the category on the left being changed to that of the one on the right.

Activity	LDA	MLP	PNN	SVM
Steaming-Setting-Towing-Hauling				
Unreduced	56.96	76.26	47.21	60.84
Steaming-Towing-Hauling				
Setting > Steaming	57.14	73.77	55.70	64.81
Setting > Towing	55.16	68.13	58.70	67.06
Steaming-Setting-Towing				
Hauling > Steaming	41.64	71.15	55.06	18.88
Hauling > Towing	77.75	84.48	64.95	71.16
Steaming-Towing				
Setting > Steaming / Hauling > Steaming	55.47	77.51	70.72	71.05
Setting > Steaming / Hauling > Towing	85.38	87.34	76.18	79.42
Setting > Towing / Hauling > Steaming	44.98	73.97	69.41	53.75
Setting > Towing / Hauling > Towing	76.42	77.81	76.02	79.35

Table 5. *Accuracy, precision, recall* and *F1 score* (expressed as percentages) for the towing category in the final tests of the LDA, MLP, PNN and SVM learning machines, with all the categories and after reducing attributes grouped by sensors. The results are shown before reduction (Ace-Gry-Mag-GPS) and after reduction using the set of sensors that provided the best *F1 score* for the *Towing* phase. The attributes were grouped by sensors as follows: Ace = (AceXm, AceXs, AceYm, AceYs, AceZm, AceZs), Gir = (GirXm, GirXs, GirYm, GirYs, GirZm, GirZs), Mag = (MagXm, MagXs, MagYm, MagYs, MagZm, MagZs) and Gps = (GpsRm, GpsRs, GpsVm, GpsVs).

Sensors	Accuracy	Precision	Recall	F1 Score
LDA				
Ace-Gir-Mag-Gps	80.30	62.76	52.14	56.96
Ace-Gir-Gps	82.35	60.29	86.15	70.94
MLP				
Ace-Gir-Mag-Gps	87.24	71.46	81.97	76.26
Gir-Gps	88.62	74.39	83.37	78.57
PNN				
Ace-Gir-Mag-Gps	68.08	40.24	57.09	47.21
Gps	66.67	42.55	95.21	58.82
SVM				
Ace-Gir-Mag-Gps	68.59	44.20	97.61	60.84
-	-	-	-	-

Table 6. *Accuracy*, *precision*, *recall*, *F1 score* and *Kappa index* for the towing category in the final test of the LDA, MPL, PNN and SVM learning machines. The results are shown before reduction (all activities/all sensors) and after the reduction in categories and attributes grouped by sensors that provide the best *F1 scores* in the *towing* phase.

Activities/Sensors	Accuracy	Precision	Recall	F1 Score	Kappa
LDA					
All activities / All sensors	80.30	62.76	52.14	56.96	0.60
Steaming-Towing / Ace-Gir-Gps	84.02	76.78	97.52	85.92	0.73
MLP					
All activities / All sensors	87.24	71.46	81.97	76.26	0.70
Steaming-Towing / Gir-Gps	89.69	85.11	96.26	90.33	0.83
PNN					
All activities / All sensors	68.08	40.24	57.09	47.21	0.34
Steaming-Towing / Gps	74.19	66.45	97.69	79.10	0.54
SVM					
All activities / All sensors	68.59	44.20	97.61	60.84	0.56
Steaming-Towing / All sensors	70.04	62.66	99.23	76.81	0.52

Figure 1

FIGURE 1

Figure 2
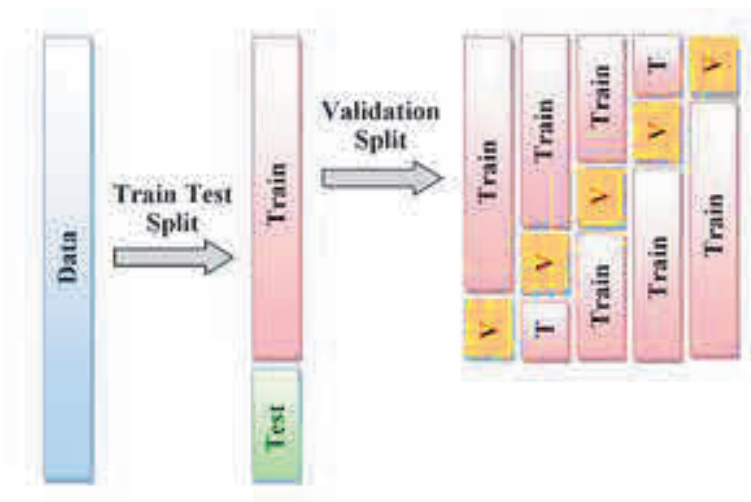
FIGURE 2

Figure 3

FIGURE 3

Figure 4

x_1

x_2

x_n

o_1

o_j

Input layer   Hidden layers   Output layer
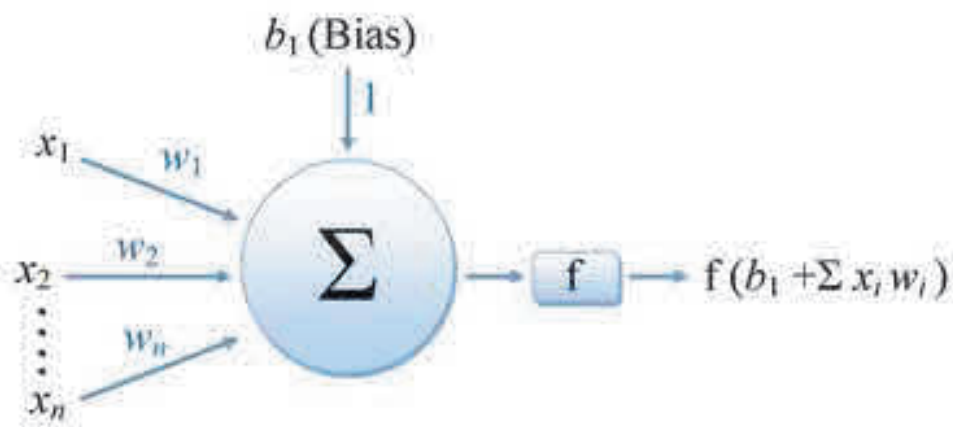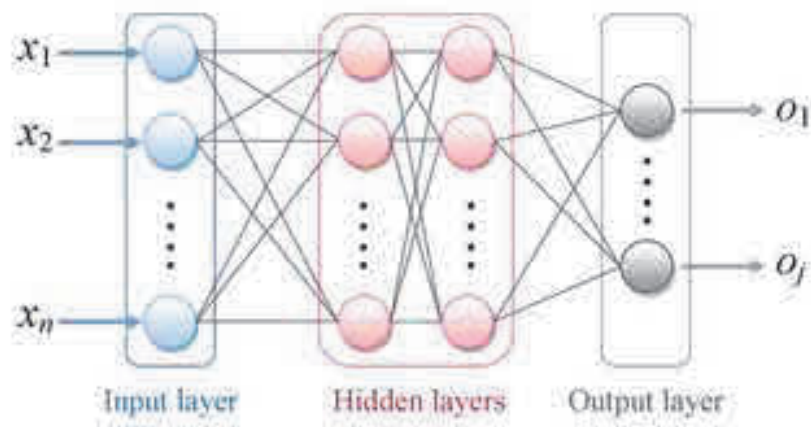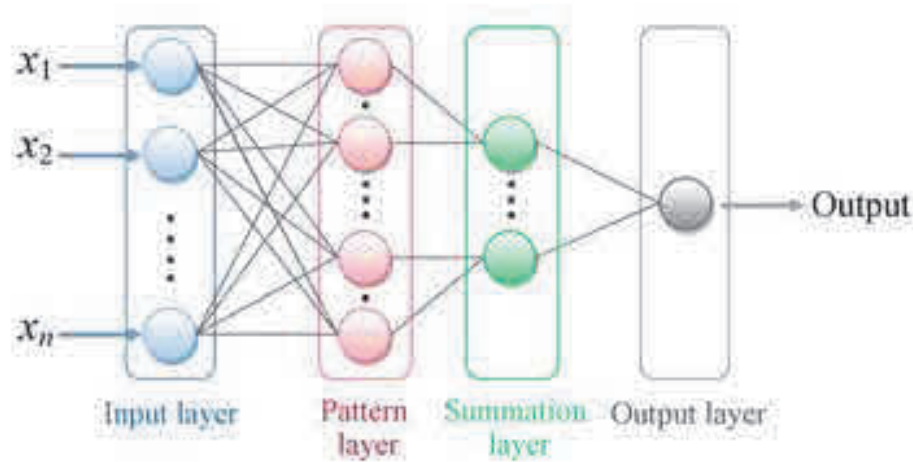
FIGURE 4

Figure 5

FIGURE 5

Figure 6
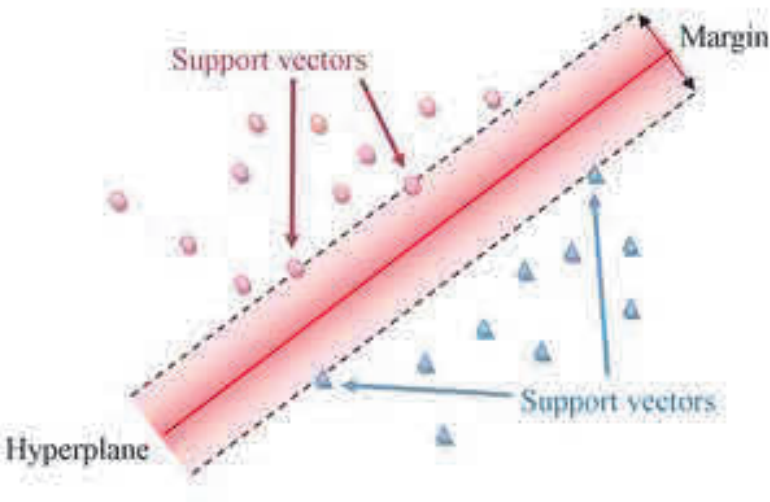
FIGURE 6

Figure Captions

Figure 1. FAMIS user interface installed in a smart phone and information flow.

Figure 2. K-fold cross-validation process. The samples from the table of data (Data) are split into training data (Train - T) to be used in the cross-validation process and test data (Test) to be used in the final model assessment. The 5-fold cross-validation process calculates the mean of the results from repeating the model training and validation process five times. Each process uses different data, ensuring that all the data have been used once in the validation phase.

Figure 3. Artificial neuron. It sums the inputs (x_i) multiplied by the weight associated with each one (w_i), adds the bias (b) and applies the activation function (f).

Figure 4. Architecture of the multi-layer perceptron neural network. Input layer with n neurons that receive input attributes (x_n). Hidden layers composed of a variable number of layers and neurons. Output layer that provides the response of the neural network (o_j). The lines represent the connections that transmit the data between layers.

Figure 5. Architecture of the probabilistic neural network. Input layer with n neurons that receive input attributes (x_n). Pattern layer with one neuron per sample. Summation layer with one neuron per class. Output layer with one neuron providing the output of the neural network.

Figure 6. Graphical representation of the hyperplane used by the support vector machine to discriminate samples of different classes. Support vectors show which samples are closest to the hyperplane and define the edge of the margin.

Figure 7. Graphical representation of the prediction of 4 hauls of the LDA, MLP, PNN and SVM models. The orange area represents the real hauls, the blue dots the *steaming* prediction and the red dots the *towing* prediction.

Figure 7
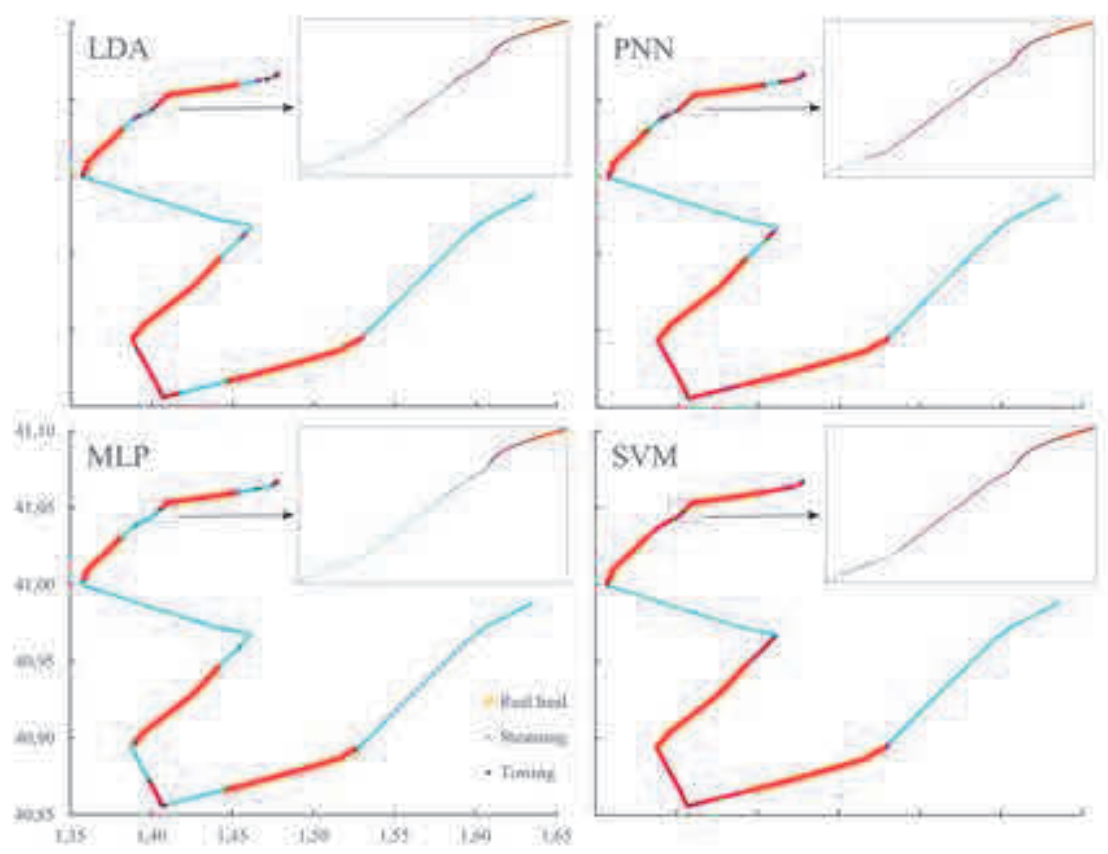
Figure 7

Maria del Mar Galotto Tébar

https://orcid.org/0000-0001-6045-4563

No public information available.
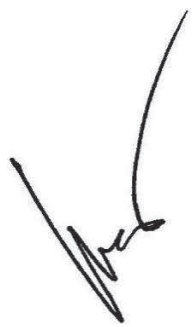
Record last modified Mar 25, 2021 8:58:28 PM

Credit Author Statement

Galotto-Tebar, M.M.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization; **Pomares-Padilla, A**.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Review & Editing, Visualization; **Czerwinski, I.A.**: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing, Visualization, Supervision; **Gutiérrez-Estrada, J.C**.: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing, Visualization, Supervision

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

1 **Is the vessel fishing? Discrimination of fishing activity with low-cost**

2 **intelligent mobile devices through traditional and heuristic approaches**

3

4 Galotto-Tebar, M.M.[1]; Pomares-Padilla, A.[1]; Czerwinski, I.A.[2]; Gutiérrez-Estrada, J.C[3]

5

6 [1]Dpto. de Ingeniería de Computadores, Universidad Miguel Hernández. Campus de
7 Elche, Avenida del ferrocarril s/n, 03202 Elche, Alicante, Spain (m.galotto@umh.es;
8 pomares@umh.es)

9

10 [2]Instituto Español de Oceanografía, Centro Oceanográfico de Cádiz, Puerto Pesquero,
11 Muelle de Levante, s/n, 11006 Cádiz, Spain (ivone.czerwinski@ieo.es)

12

13 [3]Dpto. de Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Campus de
14 El Carmen, Universidad de Huelva, 21007 Huelva, Spain (juanc@uhu.es)

15
16

17 **Abstract**

18

19 Knowing the activity of fishing vessels accurately and in real time means a leap in quality

20 in the management of fishing activity. This paper presents the development of a new

21 fishing activity monitoring integral system (FAMIS) that can complement and overcome

22 the limitations of current fishing vessel monitoring systems (VMS). FAMIS is developed

23 on the basis of a low-cost mobile device with GPS sensors, accelerometer, gyroscope and

24 magnetic field and integrates different statistical methods (discriminant functions) and

25 heuristics (artificial neural networks and vectorial support machines) as techniques to

26 classify the information recorded by the sensors of a mobile device during fishing activity.

27 The results obtained with FAMIS indicate that, in general, heuristics have a high degree

28 of discrimination of each of the phases of fishing operation and that, in particular,

29 multilayer perceptrons (MLPs) are capable of correctly identifying 96.3% of towing

30 phases using only GPS and gyro sensors.

31

32 *Keywords*

33

34 Vessel monitoring system; mobile device; sensor; trawl fishing; ship's behavior; machine

35 learning

36

37 **1. Introduction**

38

For the exploitation of living aquatic resources to be sustainable, there is a need to limit the impact of fishing fleets on marine ecosystems. To assess the impact of a fishing vessel during its activity, we use the concept of fishing effort defined as fishing capacity multiplied by the duration of fishing activity (EC, 2007). Fishing capacity can be quantified from the vessel's technical characteristics and the fishing gear used. On the other hand, assessment of the duration of fishing activity requires knowledge of the time during which the fishing capacity of a vessel is effectively operated, a parameter which is highly variable and specific to each type of fisheries and fishing operation. Therefore, to reliably identify the duration of activity of a vessel, we need to have information, beyond just the time spent at sea, that allow us to track the activity of a fishing vessel.

In 1995, the European Commission undertook a pilot project to assess the functionality and costs of various satellite systems that might make it possible to continuously track the position and activity of fishing vessels. The result of this pilot was the implementation, in 2000, of a Vessel Monitoring System (VMS) with the goal of improving the management and monitoring of fishing activity. Specifically, all fishing vessels ≥ 24 m long were required to use the VMS, this requirement being extended to vessels ≥ 15 m in 2003 and those ≥ 12 m in 2012 (EC, 2009). Data provided by the VMS (the vessel's geographical position, course and speed) are transmitted every hour to the Spanish Fisheries Monitoring Centre where all this information is stored and shared internationally in real time with parties in whose waters European vessels operate.

Based on the information provided by the VMS, several authors have proposed actions and protocols to strengthen the monitoring and planning of fishing activity (Gerritsen and Lordan, 2011; Gerritsen et al., 2012; Jennings and Lee, 2012), a large proportion of published studies having focused on estimating fishing effort (Rijnsdorp et al, 1998; Deng et al., 2005; Murawski et al., 2005; Salthaug and Johannessen, 2006; Walter et al., 2007; Fock, 2008; Hintzen and Brunel, 2008). Nonetheless, despite the large amount of information provided by the VMS and the indisputable advantages of its analysis for monitoring the fishing fleet, the long sampling period, the limited information provided in each sample and the exclusion of the coastal fleet limit the reliability and accuracy of the estimations that can be obtained from this information (Russo et al, 2016).

72 To overcome these limitations, there is a need for new devices and/or monitoring
73 procedures. For example, the Location and Track System for Andalusian Fishing Vessels
74 (SLSEPA), also known as the "green box" system (Junta de Andalucia, 2004) developed
75 by the regional government of Andalusia (Spain) uses the GPRS mobile network to
76 transmit data on location, course and speed of the artisanal fisheries to a control centre at
77 3-min intervals. The higher sampling rate improves the accuracy of the estimation of
78 fishing effort and spatial distribution of the fleet and the low data transmission costs make
79 the system feasible and operative for vessels <12 m long of the artesanal fleets (Cojan
80 and Burgos, 2015). Despite SLSEPA representing a significant improvement, however,
81 this type of system still relies on assigning the type of activity based on the recorded speed
82 alone (Lee et al., 2010; Burgos et al., 2013).

83

84 We intend to develop a new fisheries monitoring system that will provide better quality
85 information and at a low cost to facilitate their deployment across the entire fishing fleet.
86 Considering that fishing gear in operation has an effect on the dynamic behaviour of a
87 vessel (Sun et al., 2011; Russo et al, 2011) and that the manoeuvres of a vessel may be
88 related to its fishing activity, a device equipped with positioning and movement sensors
89 could potentially provide the information necessary to accurately determine when a vessel
90 is fishing.

91

92 Previous research has found that the sensors that are usually used in smart mobile devices
93 such as mobile phones and tablets (GPS, accelerometer, gyroscope and compass) are
94 capable of accurately detecting likely changes in the dynamic behaviour of a vessel during
95 each trawl phase (Galotto-Tebar et al., 2020). On the other hand, the processing capacity
96 of these devices enables us to include machine learning algorithms (artificial intelligence)
97 that are able to learn about the dynamic behaviour of a vessel during the different phases
98 of operation and identify the activity of a fishing vessel in real time, at any time.

99

100 Along these lines, some authors have succeeded in using various learning machines to
101 classify a situation or recognise movements of objects or people based on data provided
102 by movement sensors from smart mobile devices (Rodríguez-Martín et al., 2013; Tian et
103 al., 2019; Cust et al., 2019). In this paper we have selected four supervised machine
104 learning classification techniques: linear discriminant analysis (LDA) used to classify
105 linearly separable samples, multilayer perceptron (MLP) and support vector machine

106 (SVM) used to classify non-linearly separable samples and Bayesian classifier or
107 probabilistic neural network (PNN). The selected classifiers are commonly used in
108 various fields such as livestock (Rodero et al., 2012), fisheries (Bertrand et al., 2008;
109 Czerwinski et al., 2007; Gutiérrez-Estrada et al., 2000, 2007, 2008, 2010; Queirolo et al.,
110 2012; Robotham et al., 2010), hydraulic management (Pulido-Calvo and Portela, 2007),
111 economics (Pérez-Ramírez and Fernández-Castaño, 2007), quality control (Gutierrez and
112 Vázquez, 2013), etc.
113
114 Given this, the aims of this study were first to assess the viability of a range of statistical
115 and heuristic methods as tools to classify the data retrieved by sensors on a mobile device
116 during fishing activity, and secondly, to test their ability to identify which trawl phase a
117 fishing vessel is in at a given time.
118
119 **2. Material and Methods**
120
121 The models developed in this paper works with the Fishing Activity Monitoring Integral
122 System (FAMIS) (Galotto-Tebar et al., 2020). FAMIS is a mobile application (APP)
123 developed with Android Studio that records the vessel's movement during its fishing
124 activity. During the recording, the data provide by the sensors of devices like smart
125 phones and tablets are temporarily stored in a local database (SQLite) linked the fishing
126 phase, ship's informations and fishing gear. When the mobile device connects to a server,
127 it sends the information to a MySql database. The server allows access to information
128 through the REST API service which allows to analyse the data using RStudio application
129 (Figure 1).
130
131 *2.1. Mobile device and sensors*
132
133 In this study, we used a Samsung SM-P600 tablet with movement sensors that include as
134 standard: 6-axis inertial measurement unit (Bosch Sensortec BMI055) composed of a
135 digital triaxial 12-bit acceleration sensor and a digital triaxial 16-bit gyroscope, a 16-bit
136 3-axis magnetometer (Asahi Kasei Microdevices AK8963C) and a GPS receptor. Using
137 these sensors and sampling periods of 50 ms and 1 s, data were collected on 11 different
138 variables (Table 1).
139

140 With the acceleration sensor, we measured the acceleration $(\overrightarrow{Ace})$ of the mobile device
141 while it is moving, this movement being produced when there are external forces on the
142 device, the value recorded being the force acting on the mass of the sensor. Once the
143 mobile device is attached to the vessel's structure, we can say that the movement of the
144 device reflects that of the vessel. One of the forces always involved in the movement and
145 hence in the acceleration recorded by the sensor is gravity (G) (Equation 1). In this
146 equation, the mass is the sum of the mass of the device and that of the vessel; this term
147 can be ignored and we can consider that the accelerometer will record the components of
148 the gravity vector $(\vec{G})$ at each point.

149

150 $$\overrightarrow{Ace} = -\vec{G} - \overrightarrow{\sum F} / Mass \approx -\vec{G} \tag{1}$$

151

152 With the gyroscope sensor, we measure the speed of rotation around the three main
153 coordinate axes: OX, OY and OZ in radians per second (rad/s). If an observer placed on
154 the positive side of an axis detects that the rotation around the axis is anticlockwise, the
155 speed of rotation will be positive and when the rotation is clockwise, the speed of rotation
156 will be negative.

157

158 With the magnetometer, we measure the Earth's magnetic field and the magnetic field
159 around the device, in microTesla (μT). This sensor obtains the vector sum of the Earth's
160 magnetic field and the magnetic fields generated by objects around the device such as
161 engines, cables, etc., and provides the vector components resulting from the sum along
162 OX, OY and OZ axes (Equation 2):

163

164 $$\overrightarrow{Mag} = \overrightarrow{Mag_{Earth}} + \overrightarrow{Mag_{Environ}} \tag{2}$$

165

166 Using a network of satellites, the global positioning system sensor measures the 3-D
167 geographical position of the device, providing values for latitude, longitude and altitude.
168 The sensor also records the timepoint of the measurement and is able to process these
169 data and generate new information indicating the course, instantaneous and mean
170 velocity, and accuracy. For this study, we have only considered it relevant to record the
171 course (GpsR) and instantaneous speed (GpsV) every second during the activity of the
172 vessel.

173

2.2. Data collection

176   Data were collected from the fishing and oceanographic research vessel Miguel Oliver of
177   the Spanish General Secretariat for Fisheries, between 28 May and 1 June 1, 2016 in the
178   waters off Castellón, Tarragona, Barcelona and Girona, during the Mediterranean
179   International Trawl Survey (MEDITS)-Spain bottom trawl survey. This survey is
180   undertaken annually by the Spanish National Institute of Oceanography (IEO), to assess
181   demersal fish stock along the continental shelf and slope in the Spanish Mediterranean
182   Sea (Bertrand et al., 2002).

184   Seeking to record the position of a vessel as accurately as possible, the mobile device was
185   placed in the navigation bridge at around 18 m above the centre of gravity of the boat at
186   the point where the pitch and roll axes intersect (Ibrahim and Grace, 2010), aligning the
187   device with the OY axis parallel to the longitudinal axis (running from stern to prow), the
188   OX axis parallel to the transverse axis (running from port to starboard) and the OZ axis
189   parallel to the vertical axis. The vector information from the sensors is recorded with
190   respect to these coordinate axes. In relation to this, we previously developed a mobile
191   application (FAMIS) to allow users to activate sensors, select a vessel's phase of hauling
192   at each point in time and store the information recorded in a database (Galotto et al.,
193   2020).

195   We recorded the vessel's movements during 22 trawl hauls using a GOC 73 sampling
196   gear with Morgère trawl doors at a speed of 3 knots, of which 20 were 30-minute hauls
197   at depths of 50 to 200 m and 2 were 60-minute hauls at depths greater than 200 m. We
198   divided the vessel's fishing activity into four phases: steaming, setting, towing and
199   hauling, and recorded the start and end time of each phase of operation.

2.3. Automated learning techniques

203   Automated machine learning refers to a set of techniques that allow computer systems
204   themselves to create algorithms that are able to analyse data and acquire the knowledge
205   necessary to carry out tasks such as predicting, classifying, ranking and decision making,
206   without the need for defining initial rules to facilitate these tasks. For this study, we

selected four supervised automated learning classification techniques with very different strategies: (i) linear discriminant analysis (LDA), a statistical technique that classifies samples that are linearly separable and estimates the probability of being in each group; (ii) support vector machines (SVMs) that classify nonlinearly separable samples using a geometric approach; (iii) multilayer perceptrons (MLPs) that are universal proxies of nonlinear functions; and (iv) Bayesian classifiers or probabilistic neural networks (PNNs).

2.4. Data processing

The first step in data processing was filtering, to eliminate zeros and outliers in the data recorded by the sensors. Next, we created 22 independent continuous quantitative variables (attributes) with a mean (Xm) and standard deviation (Xs) of the data in 10-s periods, allowing sufficient time for the boat to complete half a pitch/roll cycle (Barras and Derrett, 2012): For the Accelerometer, the mean and standard deviation were recorded for the three axes, resulting in variables *AceXm, AceXs, AceYm, AceYs, AceZm,* and *AceZs*. Similar variables were recorded for the gyroscope (*GirXm, GirXs, GirYm, GirYs, GirZm, GirZs*) and the Magnometer (*MagXm, MagXs, MagYm, MagYs, MagZm, MagZs*). For the GPS the mean and standard deviation for the course and instanteneous speed were recorded: *GpsRm, GpsRs, GpsVm* and *GpsVs*.

We then created an independent nominal qualitative variable (label) called Activity to identify the launch phase of each sample from the official MEDITS-2016 survey data. The labelling of this variable follows a window-type criterion similar to that used by O'Farrell with VMS samples (O'Farrell et al, 2017), assigning the most repeated value during the 10 seconds prior to the sample.

The table of labelled data was created by merging the 22 attributes and the label, removing the records with at least one null, normalising attributes and balancing the samples by activity. As a result, we obtained a total of 6184 samples (1546 samples per activity).

To evaluate the classification performance of each model, we divided the table of data into two blocks: a training data set consisting of 3844 samples (62.16%) from the first 3 days of the survey, covering a total of 15 hauls, and a test data set consisting of 2340

7

241   samples (37.83%) from the last 2 days of the survey, during which 7 hauls were
242   completed. This procedure allows models to be trained and tested under different weather
243   conditions.
244

245   *2.5. Assessment metrics*

246
247   The classification performance of each model is related to the number of hits and misses.
248   In this study, we used five specific metrics to assess the predictive power of the classifiers:

249
250   *Accuracy*: Percentage of positive and negative predictions that are correct

251

252 $$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100 \tag{3}$$

253
254   *Precision*: Percentage of positive predictions that are correct

255

256 $$Precision = \frac{TP}{(TP+FP)} \times 100 \tag{4}$$

257
258   *Recall*: Percentage of true positives retrieved

259

260 $$Recall = \frac{TP}{(TP+FN)} \times 100 \tag{5}$$

261
262   *F1 Score*: Harmonic mean of precision and recall

263

264 $$F1\ Score = \frac{2*Precision*Recall}{(Precision+Recall)} \tag{6}$$

265
266   where *TP* stands for *True Positive*, the number of fishing sets correctly predicted to be a
267   member of a class, *TN* for *True Negative*, the number of fishing sets correctly predicted
268   to not be a member of a class, *FP* for *False Positive,* the number of fishing set incorrectly
269   predicted to be a member of a class, and *FN* for *False negative*, the number of fishing sets
270   incorrectly predicted to not be a member of a class.

271

272 *Kappa index*: Measure of the degree of agreement between the classifier's prediction and
273 the true classification (Cohen, 1960)

274

275
$$Kappa = \frac{p_o - p_e}{1 - p_e} \tag{7}$$

276

277 where p_o is the probability of success of the classifier and p_e is the probability of success
278 of a random classifier.

279

280 *2.6. Model hyperparameter tuning*

281
282 Automated learning models are parameterised in order that their behaviour can be
283 adjusted to fit a given problem. These models can have multiple parameters and finding
284 the optimal combination of values sometimes requires in-depth analysis that is beyond
285 the scope of this paper. In this study, we selected the values for each parameter using a
286 grid search procedure for hyperparameters, establishing a discrete range of values for
287 each parameter and creating a grid with all the possible combinations of hyperparameters
288 to methodically assess all the resulting models. The best model was selected based on
289 using the F1 score to compare the model's performance.

290

291 To avoid choosing hyperparameters of the model that best fit the test data set, the
292 validation process uses different data to train and validate each configuration, saving the
293 test data set for the final assessment of the model. The validation of each configuration
294 of the model was carried out using k-fold cross-validation resampling (Burman, 1989)
295 (Figure 2). To avoid overly small data sets, we used K=5. Having selected the parameters
296 that provided the best results in the validation process, the model was trained with all the
297 training data and then the classification performance is assessed with the test data.

298

299 *2.7. Simplification of the model*

300
301 The samples recorded during the vessel's fishing activity reflect the influence of trawling
302 on the dynamic behaviour of the vessel. Galotto et al. (2020) demonstrated that the
303 sensors of a mobile device were capable of identifying four distinct periods labelled:
304 steaming, setting, towing and hauling. The task of the models used in this study is to
305 classify the samples into one of these four categories. We should recall that the main

9

306 objective of this study is to identify the start and end of trawling to obtain real-time
307 information on fishing, and hence, the efficacy in recognising the samples taken during
308 the towing phase is key to achieving this objective. The remaining phases could
309 potentially be merged, thereby reducing the number of phases, if, by doing so, we improve
310 the performance of the model when classifying the towing category. In relation to this,
311 we validated the models with 4, 3, and 2 categories, always keeping the towing category.

312

313 The inclusion of too many attributes of samples may lead to learning machines
314 performing the classification task less well. Given that each attribute is linked to a sensor,
315 and in some cases, to the orientation of the mobile device with respect to the vessel, we
316 grouped the attributes by sensor, and analysed the response of the four learning machines
317 to all of the potential reductions in the group of attributes.

318

319 *2.8. Machine learning*

320
321 *2.8.1. Linear discriminant analysis (LDA)*

322
323 Linear discriminant analysis (LDA) is a supervised learning algorithm used for data
324 classification and dimension reduction. Based on a dependent qualitative variable and a
325 set of independent quantitative variables, discriminant analysis allows samples to be
326 classified into one of the groups established by the dependent variable.

327

328 This type of analysis provides classification procedures for new observations with an
329 unknown origin into one of the groups analysed, by providing discriminant scores from
330 which we can estimate the probability of being in each group. For this, the algorithm uses
331 new variables known as discriminant variables that are able to characterise and
332 differentiate between groups, described using discriminant functions which are linear
333 combinations of the original variables. More detailed descriptions of the method can be
334 found in Kim et al. (2007), Berstein et al. (2019) and Li et al. (2020).

335

336 *2.8.2. Artificial neural networks: multilayer perceptrons (MLPs) and probabilistic neural*
337 *networks (PNNs)*

338
339 Artificial neural networks are mathematical models inspired by the neural architecture of
340 the human brain (Rumelhart et al, 1986). To create an artificial neural system, we use

341 artificial neurons functionally organised in layers. The information flows through the

342 layers of the neural network via one-way connections that simulate the synapses between

343 neurons. Each connection has an associated weight (w_i) equivalent to the synapse

344 strength, and each input (x_i) through this connection is multiplied by this weight. The

345 receptor neurone calculates the weighted sum of all the inputs, adds a numerical value

346 known as bias (b) and applies an activation function to the result (*f*) to determine the

347 output (Figure 3).

348

349 In particular, multilayer perceptrons (MLPs) are one-way neural networks formed by an

350 input layer with as many neurons as there are input attributes, one or several hidden layers

351 with a variable number of neurons and output layer with the number of neurons required

352 to show the output (Figure 4). The data flow in one direction from the input layer towards

353 the output layer. The outputs from each neuron of the input layer and hidden layers are

354 connected to the input neurons of the following layer. MPLs are able to analyse complex

355 data sets and perform nonlinear classification into two or more groups, and given this,

356 have been widely used in a range of technical applications (Lek and Guegan, 1999;

357 Gutiérrez-Estrada et al., 2000; Dedecker et al., 2005; Goethals et al., 2007; Pulido-Calvo

358 and Portela, 2007; Gutiérrez-Estrada et al., 2008).

359

360 On the other hand, probabilistic neural networks (PNNs) are Bayes-Parzen classifiers

361 composed of four layers: an input layer with the same number of neurons as input

362 attributes; a first hidden layer, which is a pattern layer with the same number of neurons

363 as training or calibration samples; a second hidden layer, which is a summation layer with

364 the same number of neurons as classes; and an output layer, with a neuron that provides

365 the result of the classification (Figure 5) (Specht and Specht, 1990; Hajmeer and Basheer,

366 2002; Rodero et al., 2012).

367

368 When we supply an input to the PNN, the pattern layer assesses the distances from the

369 input vector to the training vectors using a vector of distances, the elements of which

370 indicate how close the input is to each training input. The summation layer (the second

371 layer) sums the patterns of each class generating a vector of probabilities which is then

372 used by a competitive transfer function at the output layer to select the most probable

373 class (Pérez-Ramírez and Fernándo-Castaño, 2007).

374

11

375 *2.8.3. Support Vector Machine*

376

377 The support vector machine (SVM) are statistical classifiers proposed by Vapnik (1995)

378 that belong to a family of linear classifiers. SVMs seek to identify hyperplanes that divide

379 the input feature space. At the algorithmic level, SVM learning is modelled as a quadratic

380 optimisation problem with linear constraints, the size of the problem depending on the

381 dimension of the feature space (Figure 6). While markedly less popular than the

382 aforementioned models, they have been used in a range of applications and shown a

383 discriminatory power similar to that of MLPs (Robotham et al., 2010; Rodero et al.,

384 2012).

385

386 *2.9. Calibration and validation procedures*

387

388 The construction of learning machine models starts with the processes of validation (k-

389 fold cross-validation) and parameter section allowing adaption of their behaviour to the

390 problem to be solved. In the case of the LDA modelling, given the use of balanced

391 samples, the initial probability of class membership is the same for all the classes. In MLP

392 modelling, the following were selected: a hidden layer with 8 neurons, a maximum of 30

393 iterations of learning, Randomize_Weights(0,1) as the initialization function,

394 Std_Backpropagation(0.2,0) as the learning function, Topological_Order(0) as the update

395 function, and Act_Logistic as the activation function of all hidden units and all output

396 units. As the initialisation function assigns random initial weights, producing small

397 differences between the models, we calculated the mean of 10 training and testing runs

398 of the MLP. In the PNN modelling, a value of 1.4 was set as the smoothing parameter for

399 the pattern-layer activation function. Lastly, in the SVM modelling, we chose a linear

400 kernel and set the constraint violation cost to 1.

401

402 For data processing, simulating the learning machines and obtaining the results, we used

403 the R programming language (https://www.r-project.org) with specific libraries to work

404 with learning machines such as Modern Applied Statistics with S (Ripley et al., 2020), R

405 Neural Networks using the Stuttgart Neural Network Simulator

406 (https://github.com/cbergmeir/RSNNS), probabilistic neural networks PNN (Chasset,

407 2016) and Misc Functions of the Department of Statistics Probability Theory Group

408 (E1071; Meyer, 2019).

12

409

3. Results

411

Table 2 shows the results concerning the performance in the final test of the four learning machines analysed in this study.

414

In general, MLPs and SVMs provided better results than LDA or PNNs. The best harmonic mean between precision and recall (F1 scores) was obtained with the SVM model, with a mean among the categories of 68.21% compared to the 65.41% observed in the MLP model. Further, the SVM model was better than the MLP model in precision, with a mean of 76.04% compared to 67.52%. In contrast, the MLP model was better in terms of accuracy and sensitivity, with mean percentages of 81.75% and 63.50% compared to 80.92% and 61.84% in the SVM model. On the other hand, the MLP model provided more balanced results than the SVM one, with regards to classifying the samples recorded during the trawling phase, with an accuracy of 87.24%, a precision of 71.46% and an F1 score value of 76.26%. The SVM model only performed better than the MLP model in terms of sensitivity, being able to identify 97.61% of the trawling activity samples, at the expense of a very high number of false positives. We should highlight the difficulties of the LDA, MLP and PNN models in identifying the setting category, these only being able to recognise 31.11%, 32.43% and 17.95%, respectively, of the samples recorded in this phase.

430

Table 3 presents the ability of learning machines to distinguish samples from two fishing phases. The analysis from using samples from two trawl phases (1 versus 1) is reported in the top half of the table and from using all the samples, three phases being merged in a new class called "the rest" (1 versus the rest) in the bottom half. This table provides F1 scores as a percentage in the final test of the four models. The results obtained reveal the weaknesses of the LDA, MLP and PNN models in different trawl phases. The LDA and MLP models struggle to differentiate samples from the steaming and setting phases, being able to identify less than 33% of the samples recorded during setting. On the other hand, the PNN model did not differentiate well between steaming and setting, but in this case, the poor results correspond to the steaming phase, this model identifying only 22.56% of the samples obtained during steaming. Similarly, the PNN provided poor results for

442 differentiating between setting and towing, being able to identify only 14.87% of the

443 samples from the setting phase.

444

445 Next, we explored the possibility of improving performance in the classification process

446 by reducing the number of categories, testing all four models, maintaining the towing and

447 steaming activities and removing the setting or hauling categories or both. Table 4 shows

448 that all the models performed better in classifying towing (F1 score) when setting and

449 hauling were eliminated and assigning the samples taken in these phases to steaming and

450 towing, respectively.

451

452 We also explored the response of the four models to a reduction in the number of input

453 attributes. For this purpose, we grouped the attributes by sensor and tested all the possible

454 combinations. Table 5 shows the F1 score for towing with each of the four models with

455 all the attributes and the best response to a potential reduction in attributes grouped by

456 sensor. We can observe the substantial improvement in the LDA model after removing

457 the variables based on data from the magnetometer while the PNN model provided better

458 results using only the GPS-based attributes. Similarly, the MLP model improved 3%

459 using only the gyroscope- and GPS-based attributes, while the performance of the SVM

460 model did not improve by reducing attributes.

461

462 Table 6 shows the results in the final test of the four models considering the reduction of

463 attributes and categories. The MLP model performs markedly better than the others, with

464 a hit rate of 89.69% in their positive and negative predictions concerning the towing

465 phase, a positive predictive rate of 85.11%, a towing phase sample identification rate of

466 96.26%, a harmonic mean precision and recall of 90.33%, and a Kappa index with a high

467 degree of agreement of 0.83 (p<0.0001). Figure 6 shows graphically a part of the final

468 test of the models.

469

470 **4. Discussion**

471

472 The information provided by the VMS (position, course and speed) to identify fishing

473 grounds and the activity of a fishing fleet can be improved using additional sensors as is

474 suggested in this study. In trawl fishing, precision is key to assessing haul duration and

475 track, the spatial and time distribution of the effort, and subsequently the catch per unit

476 effort in each fishing area. Deng (2005) stated that the errors in the prediction of trawl
477 fishing increase with the time interval between VMS records and that the loss in precision
478 is relatively high with sampling intervals longer than 30 minutes. On the other hand, Mills
479 et al. (2007) indicated that a vessel's speed alone is not a suitable criterion for identifying
480 trawling activity and confirmed that the sampling frequency of the VMS is too low to
481 allow trawl tracks to be characterised, most trawls being represented by just one record
482 and some not being detected at all, the haul occurring between two records. In addition,
483 for Peruvian anchoveta, Bertrand et al. (2008) reported that speed provided by VMS data
484 leads to the number of fishing events being overestimated by nearly 182%. These findings
485 warrant the development and deployment of new low-cost monitoring systems,
486 complementary to VMS, that are flexible and easily upgradeable and that allow accurate
487 estimation of the activities of fishing vessels.

488

489 Since the deployment of the VMS in 2000, many researchers have studied how to process
490 and complement the data provided by this system to improve the management and
491 monitoring of fishing activity. For example, Szostek et al. (2017) included information
492 extracted from interviews of fishermen about their experience fishing king scallops in the
493 English Channel, while Bastardie et al. (2010) combined the data provided by the VMS
494 with information extracted from logbooks through a linking process considering the
495 degree of mismatch. In these ways, these authors were able to obtain disaggregated
496 fishing effort data at a fine geographical scale. Overall, they indicate that the procedures
497 they used significantly improve the delimitation of the catchment area as well as the
498 assessment of fishing effort in time and space, but to achieve this, there was a need for
499 offline processing of external data and its subsequent integration with data generated by
500 the VMS. In relation to this, the data provided by FAMIS (Galotto et al., 2020), the system
501 we propose, is at the same level as that from VMS, and hence, it is fully complementary
502 and therefore easy to integrate.

503

504 The use of the FAMIS in fishing vessels would significantly improve on the use of the
505 VMS alone, as it adds movement sensors on the vessel and processing capacity to allow
506 the use of artificial intelligence algorithms to identify the activity of a vessel in situ.
507 Galotto et al. (2020) showed that standard low-cost devices such as mobile phones and
508 tablets do record different data in each trawl phase, and building on that, in this study, we

509 demonstrate that certain models, such as MLPs or SVMs, yield classifiers that are good
510 tools for automated classification.

511

512 The four types of models analysed in this study use very different classification strategies,
513 and in line with that, their classification performance varied. The results obtained with
514 samples recorded in the four trawl phases and attributes associated with the four sensors
515 indicated that SVMs and MLPs provide better results than LDA and PNNs. These results
516 are similar to those obtained by other authors. Robotham et al. (2010) found that MLPs
517 and SVMs performed significantly better than PNNs or LDA, when attempting to
518 distinguish between four pelagic fish species caught off the northern coast of Chile based
519 on echograms of fish shoals. Further, Rodero et al. (2012) compared the same four types
520 of models for distinguishing between four Andalusian cattle breeds based on
521 morphometric variables and found that MLPs and SVMs had better classification
522 performance. This is attributable to the fact that both MLPs and SVMs associate highly
523 nonlinear discriminant functions with the input patterns. This is a result of the large
524 number of variables in each sample, the physical magnitudes that they represent and the
525 complexity of the identification task, in our case, the identification of movement patterns
526 of a vessel exposed to forces from the waves, wind, engine, rudder and fishing gear.

527

528 The high nonlinearity of the input data is clearly reflected in the inability of LDAs to find
529 a discriminant function that distinguishes the samples from different classes. In relation
530 to this, the PNN-based model, which takes a radial approach, identifying new samples
531 based on similarities with training samples, provided the poorest results with a mean F1
532 score of just 44%. In contrast, the MLP as a universal classifier, with a single hidden layer
533 and a small number of neurons, provided an F1 score of 65.41%, being able to identify
534 nearly 82% of samples in the trawl phase. Finally, the most efficient configuration of the
535 SVM model with a linear kernel finds spaces with new dimensions where it is possible to
536 separate different classes with hyperplanes, providing an acceptable mean F1 score of
537 68.21% and excellent recall of the trawl phase of over 97%.

538

539 These results significantly improve on the classification performance found by previous
540 authors based on VMS data. Gerritsen and Lordan (2011) reported that the use of vessel
541 speed as a classification criterion for fishing activity provided a low rate of false negatives
542 but a very high rate of false positives in the classification (32%), only 68% of cases being

16

543 correctly classified. These authors highlighted that the potential reasons for these results

544 are related to the fact that a boat may travel at a speed similar to that when trawling when

545 sail at a slow speed while waiting for the right tide or because of poor weather conditions,

546 among other factors.

547

548 In relation to this, the four trawl phases identify periods in time when the fishing vessel

549 is involved in different activities. In each phase, the dynamic behaviour of the vessel is

550 influenced by various elements involved such as speed, course, waves, fishing gear, catch,

551 etc. These elements have a different impact at each phase. For example, the steaming

552 phase ends when the boat is manoeuvring and slowing down until it positions itself where

553 it is going to start trawling. After that, the setting phase starts by shooting the net; then

554 the doors are set and the towing cable paid out. During the towing phase, the drag on the

555 vessel grows as the growing volume of the catch is added to the resistance of the fishing

556 gear itself. Next, in the hauling phase, the towing cable is winched in, the doors are raised,

557 and finally, the fishing gear and the catch are brought on board. At the end of the haul,

558 the vessel starts a new steaming phase, manoeuvring and increasing its speed to move to

559 the next fishing spot. Given all this, the effect of the different manoeuvres within each

560 phase on the behaviour of the vessel cannot be detected by the speed data provided by the

561 VMS unless this information is complemented with movement sensor data such as those

562 recorded through FAMIS.

563

564 On the one hand, homogenous behaviour of the vessel within a phase and marked changes

565 between phases should facilitate the task of classifying the samples and identifying the

566 start and end of the trawl phases. The start of setting is the most problematic time point.

567 This is because until the doors are in the water, the resistance of the nets is low and the

568 behaviour of the vessel is similar to that in the steaming phase. In contrast, the towing

569 phase starts when the fishing gear, moving at the towing speed, reaches the correct depth,

570 the cable stops being paid out and the speed of the vessel is reduced to the towing speed.

571 This change of speed enables the models to identify the first samples of the towing phase

572 and therefore the start of the real phase of fishing. Similarly, the hauling phase is easy to

573 identify, given that it starts with a reduction in the vessel's speed to make it easier to bring

574 the fishing gear on board and finishes with the fishing gear inactive and an increase in the

575 vessel´s speed.

576

17

577 The results of the models analysed are significantly better when we simplify the
578 description of the activity of the fishing vessel to two phases, a common approach for
579 discriminating the activity in a fishing boat when onboard observers log whether the boat
580 is "fishing" or "not fishing" as a function of whether the fishing gear is deployed in the
581 water (Chang and Yuan, 2014).
582
583 By merging the phases steaming+setting and towing+hauling, we obtained F1 scores of
584 over 76% in all the models. These proposed mergers avoid the difficulty of identifying
585 the samples at the start of the setting phase and considers that the vessel is fishing from
586 when it starts towing until the fishing gear is taken out of the water. This partially
587 consistent with the results of Joo et al. (2011) who used MLPs to estimate fishing events
588 in Peruvian anchoveta fisheries, based on VMS parameters and validated in-situ by
589 onboard observers. In that study, observers validated two segments of activity in the speed
590 data series provided by VMS, identified as acceleration between the previous and current
591 pace (which would coincide in part with the merging of steaming+setting) and the
592 acceleration between the current and the following pace (approximately towing+hauling).
593
594 In relation to the sensor systems used in this study, we should indicate that all four sensors
595 provided relevant information about changes in the vessel´s behaviour in the four trawl
596 phases (Galotto et al., 2020). The reduction in computational costs associated with
597 removing variables had a relatively small impact on the behaviour of the models. The
598 sensitivity analysis indicated that the GPS sensor is undoubtedly the device that provides
599 the most important information, given that using this device alone, the models assessed
600 achieved a mean F1 score of over 61%. However, the use of information from the other
601 sensors improved the behaviour of some models. For example, the MLP model was able
602 to correctly identify 96.26% of the samples recorded by the GPS+gyroscope in the towing
603 phase. Specifically, for a standard 60-minute haul, the model identified the act of fishing
604 during 57 minutes and 45 seconds, in the worst-case scenario with the false negatives
605 concentrated at the start or end of trawling. The prediction error of 2 minutes and 45
606 seconds translates to the identification of a shorter-duration hauling operation and an
607 advance or delay in the start or end of the hauling. On the other hand, we should bear in
608 mind that characteristics of the boat where we performed our experiment (Miguel Oliver)
609 mean that it had a higher vessel size to fishing gear ratio than that in fishing boats. This
610 suggests that the dynamics in a fishing boat may be even more affected by the fishing

611 gear, and hence, the difference between the trawl phases will be easier to identify using
612 this system.
613
614 **Conclusions**
615
616 The results of this work indicate that the use of the proposed device (FAMIS) on the basis
617 of a mobile device with GPS, accelerometer, gyroscope and magnetic field sensors
618 significantly improves the accuracy of current VMS systems.
619
620 The 4 models analysed have shown different capacities to classify fishing activity. The
621 results obtained with samples recorded in the 4 phases of the set (steaming, setting, towing
622 and hauling) and attributes associated with the 4 sensors (Gps, accelerometer, gyroscope
623 and magnetic field) indicate that the SVM and MLP models offer better results than the
624 LDA and PNN models.
625
626 On the other hand, the reduction to 2 phases (steaming and towing) considering that the
627 towing phase starts when the fishing gear reaches the correct depth and trawling speed,
628 and ends when the fishing gear is on the deck of the vessel, facilitates the task of
629 classification and substantially improves the results of the 4 models. Furthermore, the
630 reduction of the sensors involved in each model also improves the quality of their
631 predictions.
632
633 In summary, the results of this work indicate that the proposed system is capable of
634 classifying the information recorded by the sensors of a mobile device during fishing
635 activity and identifying the phase of the set in which the vessel is at any given time.
636

637 **References**
638 Alfaro-Cortés, E. (2006). Combinación de clasificadores mediante el método boosting.
639   Una aplicación a la predicción del fracaso empresarial en españa. PhD Thesis,
640   University of Castilla la Mancha (Spain).

641 Barrass, C. B., & Derrett. D.R. (2012). Ship Stability for Masters and Mates. Elsevier
642   Ltd.

643 Bastardie, F., Nielsen, J.R., Ulrich, C., Egekvist, J., & Degel, H. (2010). Detailed mapping
644   of fishing effort and landings by coupling fishing logbooks with satellite-recorded
645   vessel geo-location. *Fisheries Research, 106(1)*, 41-53.

646 Bernstein, R., Osadchy, M., Keren, D., & Schuster, A. (2019). LDA classifier monitoring
647   in distributed streaming systems. *Journal of Parallel and Distributed Computing, 123,*
648   156-167.

649 Bertrand, J. A., Gil de Sola, L., Papaconstantinou, C., Relini, G., & Souplet, A. (2002).
650   The general specifications of the MEDITS surveys. *Scientia Marina, 66(2),* 9-17.

651 Bertrand, S., Díaz, E., & Lengaigne, M. (2008). Patterns in the spatial distribution of
652   Peruvian anchovy (*Engraulis ringens*) revealed by spatially explicit fishing data.
653   *Progress in Oceanography, 79(2-4),* 379-389.

654 Burman, P. (1989). A comparative study of ordinary cross-validation, v-fold cross-
655   validation and the repeated learning-testing methods. *Biometrika, 76(3),* 503-514.

656 Chang, S.K., & Yuan, T.L. (2014). Deriving high-resolution spatiotemporal fishing effort
657   of large-scale longline fishery from vessel monitoring system (VMS) data and
658   validated by observer data. *Canadian Journal of Fisheries and Aquatic Sciences, 71,*
659   1363–1370.

660 Chasset, P.O. (2016). Package pnn. https://cran.r-project.org/web/packages/pnn/pnn.pdf.

661 Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and*
662   *Psychological Measurement, 20*, 1.

663 Cojan, M., & Burgos, C. (2015). Análisis de la información proporcionada por los
664   sistemas de localización vía satélite de la flota que explota la chirla (Chamelea gallina)
665   en el Golfo de Cádiz. In *Teledetección: Humedales y Espacios Protegidos. XVI*
666   *Congreso   de   la   Asociación   Española   de   Teledetección,   550-553*
667   http://ocs.ebd.csic.es/index.php/AET/2015/schedConf/presentations.

668 Council of the European Union. (2009). Establishing a Community control system for
669   ensuring compliance with the rules of the common fisheries policy, amending
670   Regulations. *Official Journal of the European Union, L343,* 1-50.

671 Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning
672   for sport-specific movement recognition: a systematic review of model development
673   and performance. *Journal of Sports Sciences, 37(5),* 568-600.

674 Czerwinski, I.A., Gutiérrez-Estrada, J.C., & Hernando-Casal, J.A. (2007). Short-term
675   forecasting of halibut CPUE: Linear and non-linear univariate approaches. *Fisheries*
676   *Research, 86(2-3),* 120-128.

677 Dabiri, S., Marković, N., Heaslip, K., & Reddy, C.K. (2020). A deep convolutional neural
678   network based approach for vehicle classification using large-scale GPS trajectory
679   data. *Transportation Research Part C: Emerging Technologies, 116,* 102644-102644.

680 Deng, R., Dichmont, C., Milton, D., Haywood, M., Vance, D., Hall, N., & Die, D. (2005).
681   Can vessel monitoring system data also be used to study trawling intensity and
682   population depletion? The example of Australia's northern prawn fishery. *Canadian*
683   *Journal of Fisheries and Aquatic Sciences, 62(3),* 611-622.

684 European Commission. (2007). On improving fishing capacity and effort indicators under
685   the common fisheries policy. http://eur-lex.europa.eu/legal-
686   content/EN/TXT/PDF/?uri=CELEX:52007DC0039&from=EN.

687 Galotto-Tebar, M.M., Pomares-Padilla, A., Czerwinski, I.A., & Gutiérrez-Estrada, J.C.
688   (2020). Using mobile device's sensors to identify fishing activity. *Journal of Marine*
689   *Science and Technology, 25,* 978-989.

690 Gerritsen, H., & Lordan, C. (2011). Integrating vessel monitoring systems (VMS) data
691   with daily catch data from logbooks to explore the spatial distribution of catch and
692   effort at high resolution. *ICES Journal of Marine Science, 68(1),* 245-252.

693 Gerritsen, H.D., Lordan, C., Minto, C., & Kraak, S.B.M. (2012). Spatial patterns in the
694   retained catch composition of Irish demersal otter trawlers: High-resolution fisheries
695   data as a management tool. *Fisheries Research, 129-130,* 127-136.

696 Gloaguen, P., Woillez, M., Mahévas, S., Vermard, Y., & Rivot, E. (2016). Is speed
697   through water a better proxy for fishing activities than speed over ground? *Aquatic*
698   *Living Resources, 29 (210),* 1-8.

699 Global Fishing Watch (2012). https://globalfishingwatch.org.

700 Gutierrez, P.T., & Vázquez, J.A., (2013). Aplicación de la red neuronal probabilística
701   para la clasificación de productos conforme a sus especificaciones. *Innovation in*
702   *engineering, technology and education for competitiveness and prosperity, 1-10.*

703 Gutiérrez-Estrada, J.C., Pulido-Calvo, I., & Prenda, J. (2000). Gonadosomatic index
704   estimates of an introduced pumpkinseed (*Lepomis gibbosus*) population in a

705   Mediterranean stream, using computational neural networks. *Aquatic Sciences, 62(4),*
706   350-363.

707   Gutiérrez-Estrada, J.C., Silva, C., Yáñez, E., Rodríguez, N., & Pulido-Calvo, I. (2007).
708   Monthly catch forecasting of anchovy (*Engraulis ringens*) in the north area of Chile:
709   Non-linear univariate approach. *Fisheries Research, 86(2-3),* 188-200.

710   Gutiérrez-Estrada, J. C., Vasconcelos, R., & Costa, M. J. (2008). Estimating fish
711   community diversity from environmental features in the Tagus estuary (Portugal):
712   Multiple Linear Regression and Artificial Neural Network approaches. *Journal of*
713   *Applied Ichthyology, 24(2)*, 150-162.

714   Gutiérrez-Estrada, J.C., & Bilton, D. T. (2010). A heuristic approach to predicting water
715   beetle diversity in temporary and fluctuating waters. *Ecological Modelling, 221(11)*,
716   1451-1462.

717   Hajmeer, M., & Basheer, I. (2002). A probabilistic neural network approach for modeling
718   and classification of bacterial growth/no-growth data. *Journal of Microbiological*
719   *Methods, 51(2),* 217-226.

720   Heyn, H.M., Blanke, M., & Skjetne, R. (2019). Ice Condition Assessment Using Onboard
721   Accelerometers and Statistical Change Detection. *IEEE Journal of Oceanic*
722   *Engineering,* 1-17.

723   Hintzen, N.T., Bastardie, F., Beare, D., Piet, G.J., Ulrich, C., Deporte, N., Egekvist, J., &
724   Degel, H. (2012). VMStools: Open-source software for the processing, analysis and
725   visualisation of fisheries logbook and VMS data. *Fisheries Research, 115-116,* 31-43.

726   Ibrahim, R. A., & Grace, I.M. (2010). Modeling of Ship Roll Dynamics and Its Coupling
727   with Heave and Pitch. *Mathematical Problems in Engineering, 2010,* 1-32.

728   Ichimura, S., & Zhao, Q. (2019). Route-Based Ship Classification. *IEEE 10th*
729   *International Conference on Awareness Science and Technology (iCAST),* 1-6.

730   Jennings, S., & Lee, J. (2012). Defining fishing grounds with vessel monitoring system
731   data. *ICES Journal of Marine Science, 69(1),* 51-63.

732   Joo, R., Salcedo, O., Gutierrez, M., Fablet, R., & Bertrand, S. (2015). Defining fishing
733   spatial strategies from VMS data: Insights from the world's largest monospecific
734   fishery. *Fisheries Research, 164,* 223-230.

735   Junta de Andalucía (2004). Localización y seguimiento de embarcaciones pesqueras
736   (SLSEPA).http://www.juntadeandalucia.es/organismos/agriculturapescaydesarrolloru
737   ral/areas/pesca-acuicultura/slsepa.html.

738   Kim, H., Drake, B.L., & Park, H. (2007). Multiclass classifiers based on dimension
739   reduction with generalized LDA. *Pattern Recognition, 40(11),* 2939-2945.

740   Lee, J., South, A.B., & Jennings, S. (2010). Developing reliable, repeatable, and
741   accessible methods to provide high-resolution estimates of fishing-effort distributions
742   from vessel monitoring system (VMS) data. *ICES Journal of Marine Science, 67(6),*
743   1260-1271.

744   Li, C.N., Shao, Y.H., Yin, W., & Liu, M.Z. (2020). Robust and Sparse Linear
745   Discriminant Analysis via an Alternating Direction Method of Multipliers. *IEEE*
746   *Transactions on Neural Networks and Learning Systems, 31(3),* 915-926.

747   Marzuki, M.I., Garello, R., Fablet, R., Kerbaol, V., & Gaspar, P. (2015). Fishing gear
748   recognition from VMS data to identify illegal fishing activities in Indonesia. In
749   *MTS/IEEE OCEANS 2015 - Genova,* 1-5.

750   Masters, T (1995). Advanced algorithms for neural networks : a C++ sourcebook. Wiley.

751   Meyer, D. (2019). Package e1071. https://cran.r-
752   project.org/web/packages/e1071/e1071.pdf.

753   Mills, C.M., Townsend, S.E., Jennings, S., Eastwood, P.D., & Houghton, C.A. (2007).
754   Estimating high resolution trawl fishing effort from satellite-based vessel monitoring
755   system data. *ICES Journal of Marine Science, 64(2),* 248-255.

756   Neural Networks in R using the Stuttgart Neural Network Simulator.
757   https://github.com/cbergmeir/RSNNS.

758   O'Farrell, S., Sanchirico, J. N., Chollett, I., Cockrell, M., Murawski, S. A., Watson, J. T.,
759   Haynie, A., Strelcheck, A., & Perruso, L. (2017). Improving detection of short-
760   duration fishing behaviour in vessel tracks by feature engineering of training data.
761   *ICES Journal of Marine Science, 74(5),* 1428-1436.

762   Palmer, M.C. (2008). Calculation of distance traveled by fishing vessels using GPS
763   positional data: A theoretical evaluation of the sources of error. *Fisheries Research,*
764   *89(1),* 57-64.

765   Pérez-Ramírez, F.O., & Fernández-Castaño, H. (2007). Las redes neuronales y la
766   evaluación del riesgo de crédito. *Revista Ingenierías Universidad de Medellín, 6(10),*
767   77-91.

768   Politis, P.J., DeAlteris, J.T., Brown, R.W., & Morrison, A.T. (2012). Effects of sea-state
769   on the physical performance of a survey bottom trawl. *Fisheries Research, 123-124,*
770   26-36.

771   Pulido-Calvo, I., & Portela, M.M. (2007). Application of neural approaches to one-step
772   daily flow forecasting in Portuguese watersheds. *Journal of Hydrology, 332(1-2),* 1-
773   15.

774   Queirolo, D., Hurtado, C.F., Gaete, E., Soriguer, M.C., Erzini, K., & Gutiérrez-Estrada,
775   J.C. (2012). Effects of environmental conditions and fishing operations on the
776   performance of a bottom trawl. *ICES Journal of Marine Science, 69(2),* 293-302.

777   Ripley, B., Venables, B., Bates, D.M, Hornik, K., Gebhardt, A., & Firth, D. (2020).
778   Package mass. http://www.stats.ox.ac.uk/pub/MASS4/.

779   Robotham, H., Bosch, P., Gutiérrez-Estrada, J. C., Castillo, J., & Pulido-Calvo, I. (2010).
780   Acoustic identification of small pelagic fish species in Chile using support vector
781   machines and neural networks. *Fisheries Research, 102(1-2)*, 115-122.

782   Rodero, E., González, A., Luque, M., Herrera, M., & Gutiérrez-Estrada, J.C. (2012).
783   Classification of Spanish autochthonous bovine breeds. Morphometric study using
784   classical and heuristic techniques. *Livestock Science, 143 (2-3),* 226-232.

785   Rodriguez-Martin, D., Samà, A., Perez-Lopez, C., Català, A., Cabestany, J., &
786   Rodriguez-Molinero, A. (2013). SVM-based posture identification with a single waist-
787   located triaxial accelerometer. *Expert Systems with Applications, 40(18),* 7203-7211.

788   Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning representations by
789   backpropagation errors. *Nature, 323,* 533-536.

790   Russo, T., Parisi, A., Prorgi, M., Boccoli, F., Cignini, I., Tordoni, M., & Cataudella, S.
791   (2011). When behaviour reveals activity: Assigning fishing effort to métiers based on
792   VMS data using artificial neural networks. *Fisheries Research, 111*, 53-64.

793   Russo, T., D'Andrea, L., Parisi, A., Martinelli, M., Belardinelli, A., Boccoli, F., Cignini,
794   I., Tordoni, M. & Cataudella, S. (2016). Assessing the fishing footprint using data

795   integrated from different tracking devices: Issues and opportunities. *Ecological*
796   *Indicators, 69*, 818-827.

797   Shepperson, J., Murray, L.G., Mackinson, S., Bell, E., & Kaiser, M.J. (2016). Use of a
798   choice-based survey approach to characterise fishing behaviour in a scallop fishery.
799   *Environmental Modelling & Software, 86,* 116-130.

800   Simeone, O. (2018). A Very Brief Introduction to Machine Learning With Applications
801   to Communication Systems. *IEEE Transactions on Cognitive Communications and*
802   *Networking, 4(4),* 648-664.

803   Specht D. F. (1990). Probabilistic Neural Networks. *Neural Networks, 3,* 109-118.

804   Sun, X., Yin, Y., Jin, Y., Zhang, X., & Zhang, X. (2011). The modeling of single-boat,
805   mid-water trawl systems for fishing simulation. *Fisheries Research, 109(1),* 7-15.

806   Szostek, C.L., Murray, L.G., Bell, E., & Kaiser, M. J. (2017). Filling the gap: Using
807   fishers' knowledge to map the extent and intensity of fishing activity. *Marine*
808   *Environmental Research, 129,* 329-346.

809   The R Project for Statistical Computing. https://www.r-project.org.

810   Tian, Y., Wang, X., Chen, W., Liu, Z., & Li, L. (2019). Adaptive multiple classifiers
811   fusion for inertial sensor based human activity recognition. *Cluster Computing, 22(4),*
812   8141-8154.

813   Vapnik, V.N. (1995). The Nature of Statistical Learning Theory. *Springer New York*.