FROM VMS DATA TO TUNA DISTRIBUTION MAPS AND INDICES OF ABUNDANCE.

Nicolas Bez[1], Emily Walker[1,2], Daniel Gaertner[1], Jacques Rivoirard[3] and Philippe Gaspar[2].

Abstract

Vessel Monitoring Systems (VMS) data were to analyse the different fishing activities of the French fleet of purse seiners at a small scale (i.e. stop, track and cruise), generally embedded into the definition of a nominal fishing effort. A state-space model (run in a Bayesian framework) was applied to speeds and turning angles from Vessel Monitoring Systems (VMS) data to identify the different "states" of the fishing behaviour of a purse seiner over a fishing trip. The activities "fishing", "tracking", and "cruising" were estimated on the vessels trajectories and aggregated at the fleet level. Summary statistics of the various components of the fishing effort were calculated and highlighted the variability in time and space of the fishing effort from French tuna purse-seiners in the Indian Ocean.
Distribution maps of the presence of tuna were derived and aggregated by stratum to build VMS-based indices of abundance. A comparison with traditional CPUE is performed for the study period (2006-2008)

Keywords

Catch, fishing effort, CPUE, index of abundance, Trajectory, VMS data.

Introduction

The lack of direct estimate of index of tuna abundance is usually bypassed by the scientific community by the development and the use of Catch Per Unit of Effort (CPUE). However this can be appropriate if and only if one knows the proper fishing effort by which dividing catch for the CPUE to be (reasonably) proportional to the abundance. In the case of the tuna purse seiner fishery in areas where the proportion of associated catch is important, this latter concern is particularly thorny. As a matter of fact, when dealing with FAD-associated catch, the definition of an effective fishing effort is still missing. In particular, the scientific community does not know what part of the vessel activity prior to the catch must be considered and how. In addition to this, the spatio-temporal attraction capacity of FADs would be required to also to define a relevant FAD-CPUE. In the end, given the above mentioned difficulties, the assessments are tuned by long-liners CPUEs.

[1] UMR 212 EME, IRD, Centre de Recherche Halieutique Méditerranéenne et Tropicale, Avenue Jean Monnet, BP 171, 34203 Sète, France. Nicolas.Bez@ird.fr.
[2] CLS, Parc technologique du Canal, rue Hermès, 31520 Ramonville St-Agne, France.
[3] Ecole des Mines de Paris, 35 rue Saint Honoré, 77305 Fontainebleau, France

With the objective to build an alternative index of abundance not based on catch and effort data but based on an other source of information, we analysed the VMS data provided by the French tropical purse-seiners. The individual tracks of purse seiners were analysed to identify which part of days spent at sea corresponded to real fishing activity, searching in area of aggregation (tracking) or searching while cruising (cruising). The objective of this study was to aggregate the information provided by the interpreted trajectories at the fleet level.

Material

Since 2000, the European Commission legislated that all European fishing vessel longer than 24 meters should be equipped with a Vessel Monitoring System (VMS). The Global Positioning System (GPS) positions of the vessels were registered every hour and transmitted on shore by satellite (Argos or Inmarsat). Being GPS positions, the data were accurate (error smaller than few tens of meters). Speeds (in knots) and turning angles (in radians) between consecutive daylight positions were readily calculated from VMS data.

The data used for this study were collected from the French purse-seiners based in the Seychelles islands, targeting tropical tuna species: yellowfin tuna (*Thunnus albacares),* skipjack *(Katsuwonus pelamis),* and bigeye tuna *(Thunnus obesus*) in the Western Indian Ocean (14 vessels in 2006, 18 vessels in 2007, with a total catch around 100 000 tons per year).

The material used in this study included 141 489 GPS positions recorded daily from the 1st of January 2006 to the 31st December 2008. They corresponded to 131 fishing trips performed by the 18 French purse seiners that operated in the Indian Ocean during that period of time. Hourly speeds (Fig. 1a) were distributed over two clear modes corresponding respectively to stillness and full-speed (around 12 knots). Meanwhile, turning angles (Fig. 1b) had a mode around 0 radian (straight line).

The VMS database was complemented by observers' data that helped founding the structure of the model on empirical characteristics and that allowed validating the model outputs. The scientific observers' program conducted in the Indian Ocean is being undertaken in the framework of the European Data Collection Framework. This framework targets that 10% of the trips have an observer on board. Observers on board recorded the position of the vessel every hour, or each time a change in speed or in turning angle (cap) occurred. The beginning and the ending time of each fishing operation were also reported as well as the fishing mode (i.e., non-associated school or free-swimming school and FAD school). Eleven trips (corresponding to 301 days at sea and 265 sets) were available for the analysis.

Method

Expert knowledge on the fishing tactics of purse seiners led us to considered three different states (Table 1). Each state corresponded to particular speed and angle distributions whose combination amounts to the overall histograms observed for the all fleet. First, we expected a purse seiner to move quickly through abundance-poor areas. These "*cruising"* phases were associated to large speeds and to turning angles being predominantly around 0 radian (Fig. 3).

It is worth mentioning that visual check of tuna school detection is permanent on board fishing vessel during daylight so that cruising phases have to be considered as effectively contributing to fishing effort; fruitless but effective. On the opposite, within areas where tuna schools are abundant, skippers try to track schools. In these "*tracking*" phase, apparent hourly speeds are expected to be smaller on average and turning angles should be widely distributed over the full circle. Finally, purse seiners can remain "*still*" for a while either to effectively fish, to observe the behavior of school prior to its fishing, or to repair engine break down. This latter is considered seldom as skipper would tend as much as possible to postpone reparation at night. We thus considered that each of these three states contributed to the fishing effort with however unknown respective weight.

The model, for which a thorough description can be found in Walker and Bez (2010), consisted in a state-space model (Buckland et al. 2004, Royer et al. 2005, Patterson et al., 2008). Based on empirical evidences provided by observers' data, states were assumed to follow an order one Markovian process. States were inferred in a Bayesian framework (Gelman et al., 2004) knowing both vessel speeds and turning angles. Estimates corresponded to the state having the maximum a posteriori probability to occur.

The subset of VMS data corresponding to observers' data, was extracted with the objective to tune some of the model parameters and, more importantly, to validate the model outputs before its application to the entire data set. 97% of the fishing sets declared by observers were detected by the model. However, some sequences (e.g. three fishing sets of two hours each separated by only a quarter of an hour) were viewed as only one long sequence of six fishing steps.

We further attributed one of the two possible main activities compatible with the absence of movement, namely "*fishing*" and "*stopping*" to steps estimated in the still state (Table 1). The "*fishing*" activity was attributed to steps where fishing was the dominant estimated activity, whereas "*stopping*" corresponded to a waiting time near a free-swimming school or near a FAD. Stops are required either to maintain electronic equipments located on FADs or to evaluate the effective presence of fishable schools.

The model outputs were pooled along various key dimensions for their analyses namely, time (month, trimesters or years) , space (grid of regular small cells of 0.2° x 0.2° or strata used by the relevant Regional Fisheries Organization, the Indian Ocean Tuna Commission), vessel ID. All results are given in hours spent in the components of the fishing effort. They were also looked at in terms of the covered distances. They were highly similar and are thus not represented.

The stops at sea, i.e. still phases that were estimated not to be fishing sets, represented 3% of the (daylight) time spent at sea by the studied fleet. They were thus removed from the rest of the analysis so that remained only three types of activities: fishing, tracking and cruising.

Time spent in each category of activity by vessel and by trimester were analysed via time series of boxplots. To account for seasonal changes in the total time spent at sea, the time spent by each vessel and trimester in each of the three activities were expressed in percentage of the total and represented in triangles of proportions like compositional data (Aitchison, 1982) and regression were done with the R-package "compositions".

As far as spatial characteristics are concerned, we used monthly distribution maps of the model outputs aggregated (summed) at the fleet level and at the scale of 0.2° x 0.2° squares. We choose to use monthly time windows as a compromise to get enough data in order to get a relevant spatial coverage and to have as much as possible small time windows in order to get as much as possible synoptic views of the fishing effort spatial distributions. The size of the square was chosen in reference to the lateral detection of tuna schools by visual inspection (binocular). Given this later value (~15 n.m) , a 0.2° x 0.2° square was considered as exhaustively sampled as far as a vessel visited it. Being variance like parameters, all the statistics presented further are dependent on the size of the square chosen for the analysis.

The characteristics of each component of the fishing effort (fishing, tracking, and cruising) and of each pair of such components were summarised by one statistical and one geostatistical criteria respectively (Table 2). The heterogeneity of the (statistical) distribution of the time spent in each cell and in each effort component was quantified by a Gini's like index, also called spreading area (Woillez et al., 2007). This index quantifies the uniformity of the statistical distribution of the grid cell values. It equals 1 when all grid cells have the same value (pure uniformity) and 0 when only one grid cell is filled in with a non 0 data. The analysis of each component was complemented by the computation of an omnidirectional variogram (Matheron, 1971; Rivoirard et al., 2000; Petitgas, 2001) for each monthly spatial distribution. The sets of 36 standardised variograms were represented by boxplots allowing the analysis of the dominant spatial features that emerged.

The relationship between each component of the fishing effort (pair by pair) was summarised by the use of local index of collocation (Bez and Rivoirard, 2000). Like the coefficient correlation, but based on sums rather than on averages, this index is sensitive to the fact that large (versus low) values of each component occur in the same grid cell. It equals 0 when there is a strict alternation between the two variables being positive when the other one is null and reciprocally. It equals 1 when the two variables are identical or proportional. The statistical analysis of the correlation was also completed by a geostatistical analysis of all omnidirectional cross-variograms of gridded values.

Results

3.1. Time series of time spent per activities

The time spent fishing, cruising or tracking tuna, schools pooled by trimester from 2006 to 2007, fluctuated around 800 hours with strong variations upon vessels and trimesters (Figure 4a). All over, the percentage of this overall time spent at sea devoted to fishing tuna, tracking tuna schools or cruising looking at schools were estimated around 25%, 50% and 25% respectively (Figure 4b, c, and d). A pattern emerged within each year showing that the percentage tended to decrease during the completion of the year, while cruising increase and while the tracking phase did not show any temporal trend. All in all, a significant negative regression was found (Fig. 5; p-value = 0.03; slope = -0.53) between fishing and cruising indicating that the more fishermen spent time setting the net and fishing, the less they devoted energy cruising in the area. This was not the case between cruising and tracking which seemed to be independent.

3.2. Analysis of monthly spatial distributions of activities

Monthly distribution maps (the results for august 2007 were represented as an example; figure 6) allowed a fine scale analysis of fishing effort. Over the all studied period, the extrema observed for the time spent in a 0.2° x 0.2° cell and by month fishing, tracking or cruising were respectively 90 hours, 71 hours and 10 hours. The coefficients of variation of the full set of time spent per cell and per month were respectively 4.2, 1.7 and 1.1.
The spreading areas (Fig. 7) showed similar levels for fishing and tracking activities with however a significantly larger average value for tracking (Wilcoxon test; p-value ≈ 2.10^5); no clear temporal pattern emerged. Cruising activity depicted large spreading areas and was largely less patchy than fishing and tracking activities. A significant slight decrease over time appeared indicating that the cruising activity tended to be slightly more patchy at the end than at the beginning of the studied period (2006-2008).

At the grid cell scale chosen for this study, i.e. 0.2° x 0.2° cells, the co-occurrence of fishing and tracking was large without any temporal trend (Fig. 8). It was medium between tracking and cruising and small between fishing and cruising.

Spatial structures of each of the three activities (Fig. 9a) were strongly consistent over time (small expansion of the boxplot vertically speaking). This stability of the spatial structures of the vessels activities was observed despite a large month-to-month geographical shifts of the fleet in the Indian Ocean (results not shown). The same spatial features were thus repeatedly observed for each of the three fishing activities considered whatever the location of the fleet in the western half of the Indian Ocean. The spatial ranges happened to be about 0.6° for the fishing activity (36 n.m.) and 1° (60 n.m.) for the two others. The roughness (local heterogeneity) of the spatial distributions was quantified by the nugget effect of the empirical variogram. It was small for tracking (smooth maps on average) and large for fishing and cruising (rough maps on average; Fig. 6).

The time spent (per cell and per month by the entire fleet) fishing or tracking tuna schools got consistent cross-variogram (Fig. 9b) depicting positive spatial co-variations, i.e. a positive increase of one variable being associated, on average, to a positive increase of the other, up to a certain distance estimated equal to 0.6° (36 n.m.) in the present case, where their variations were no longer correlated. The features were less clear and consistent for the two other pairs of activities, their cross-variograms showing strong month-to-month variations. However, on average, they were negative indicating opposition in their spatial co-variation (one increasing when the other decreases).

The co-variations, positive between fishing and tracking, negative between the other two pairs, are consistent with the co-occurence of the states. However it is interesting to notice that the nugget components of the simple variograms disappear in the cross-variograms: the very local variabilities represented by the nugget components are not correlated to each other. This means that the co-variations between the times spent in each activity are due to their spatial variabilities but without the very local variabilities.

The various summary statistics used in this study are recapped Table 3.

Elements of Discussion

Defining a relevant fishing effort is a dilemma in the general case. This problems is particularly salient in tuna fisheries i) because stocks assessment models used routinely use

CPUEs as an index of abundance - a definition of an effective fishing effort is thus required, ii) because there are no field truth to help and validate the definition of any index of abundance - any proposal is legitimate and can not be evaluated properly, and iii) because purse seiners do use two different modes of fishing (free swimming schools versus logged schools under FADs) to which two different effective efforts should be associated.

Some behaviors of tuna purse-seiners were already studied, as “pursuing of the school” or “changing of fishing zone” (Gaertner et al., 1999), but none considered the three main activities “fishing”, “tracking” and “cruising” which can be estimated from vessels trajectories.

There were many applications of VMS data in different fishery studies (Bertrand et al., 2005, Witt et al., 2007, Mills et al., 2007) but none focussed on the determination of the different movement states of a fishing trip. The model on VMS data has been validated on the activities “fishing” versus “no fishing”, on the subset of VMS corresponding to observers’ data (Walker and Bez, 2010). The misdetection rate which was obtained on the steps was about 10%, and the rate of under-detection on the sets was about 3%. This reasonably low level of misdetection makes possible the use of VMS data to infer information on the vessel activity and on fishing effort.

Compared to logbooks in which only one point per day is recorded, this opens at least the possibility to locate the time spent at sea. But after trajectories were properly interpreted, this also opens the possibility to analyze the time spent not to fish, i.e. searching tuna schools either in a tracking or a cruising mode, and its spatial allocation. While computing spreading areas it was also possible to quantify the minimum area required to recover half of the total time spent in a given activity. We found that half of the time spent fishing was on average concentrated in 15% of the surface area occupied by this activity. It was double (31%) for cruising.

As far as CPUE are concerned a proposal is done on the possible used of VMS to derive an index of abundance. This has the advantage to integrate all the information related to the the enormous time spent in between sets. As a matter of fact, logbooks only provide information on sets.
At the present stage a strong drawback is that this does not differentiate yet between species. An index of abundance is provided irrespective of the species. Improvement towards a specific index of abundance could be considered via the use of the type of set performed and/or through the use of the TTT algorithm.

Bibliography

Aitchison 1982. The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 44, No. 2. (1982), 139-177.

Ariz J, Delgado A, Fonteneau A, Gonzalez Costas F, Pallares P. 1999. Logs and tunas in the Eastern Tropical Atlantic. A review of present knowledge and uncertainties. In: Scott MD, Bayliff WH, Lennert-Cody CE, Schaefer KM (eds) Proceedings of the International Workshop on Fishing for Tunas Associated with Floating Objects, La Jolla, CA, February 11–13, 1992. Inter-Am Trop Tuna Comm Spec Rep 11:21–65

Bertrand, S., Burgos, J., Gerlotto, F., Atiquipa, J., 2005. Lévy trajectories of Peruvian purse–seiners as an indicator of the spatial distribution of anchovy (Engraulis ringens). ICES Journal of Marine Science 62, 447–482.

Bez N. and J. Rivoirard. 2000. Indices of collocation between populations. In : Checkley, D.M., J.R. Hunter, L. Motos, and C.D. van der Lingen (eds). Report of a workshop on the use of Continuous Underway Fish Egg Sampler (CUFES) for mapping spawning habitat of pelagic fish. GLOBEC Report 14, 1-65.

Buckland S.T., Newman K.B., Thomas L., Koesters N.B. 2004. State-space models for the dynamics of wild animal populations. Ecol. Model. 171, 157–175.

Fonteneau, A., Gaertner, D., and Nordstrom, V. 1999. An overview of problems in the catch per unit of effort and abundance relationship for the tropical purse seine fisheries. Coll. Vol. Sci. Pap. ICCAT 49 (3): 258-278.

Gaertner, D., Pagavino, M., Marcano, J. 1999. Influence of fisher's behaviour on the catchability of surface tuna schools in the Venezuelan purse-seiner fishery in the Caribbean Sea. Can. J. Fish. Aquat. Sci. 56 :394-406.

Gelman A., Carlin J.B., Stern H.S., Rubin D.B. 2004. Bayesian Data Analysis. Chapman & Hall/CRC, Second Edition, 698 pp.

Hallier JP, Parajua JI 1999 Review of tuna fisheries on floating objects in the Indian Ocean. In: Scott MD, Bayliff WH, Lennert-Cody CE, Schaefer KM (eds) Proc Int Workshop on Fishing for Tunas Associated with Floating Objects, La Jolla, CA, February 11–13, 1992. Inter-Am Trop Tuna Comm Spec Rep 11:195–221

Jonsen I.D., Flemming J.M., Myers R.A. 2005. Robust state-space modelling of animal movement data. Ecology 86, 2874–2880.

Matheron G., 1971. The Theory of Regionalized Variables and their Applications. Les cahiers du Centre de Morphologie Mathématique, Fascicule 5. Centre de Géostatistique, Fontainebleau, Paris, 212 pp.

Mills C.M., Townsend S.E., Jennings S., Eastwood, P., Houghton, C.A. 2007. Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. ICES Journal of Marine Science 64, 248–255.

Morales J.M., Haydon D.T., Frair J., Holsinger K.E., Fryxell J.M. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology 85, 2436–2445.
Moreno G., Dagorn L., Sancho G., Itano D., 2007, Fish behaviour from fishers' knowledge: the case study of tropical tuna around drifting fish aggregating devices (DFADs). Can. J. Fish. Aquat. Sci. 64, 1517–1528.

Patterson T.A., Thomas L., Wilcox C., Ovaskainen O., Matthiopoulos J. 2008. State-space models of individual animal movement. Trends in ecology & evolution, 23, 2, 87-94.

Petitgas P., 2001. Geostatistics in fisheries survey design and stock assessment: models, variances and applications. Fish and Fisheries, 2, 231-249

Polacheck, T. 1988. Analyses of the relationship between the distribution of searching effort, tuna catches, and dolphin sightings within individual purse seine cruises. Fish Bull US. 86, 2, 351-366

Soto M., Pallarés P., Delgado de Molina A., Gaertner D. 2008. Standardized CPUE for juveniles yellowfin, skipjack and bigeye tuna from the european purse Seine fleet in the Atlantic ocean from 1991 to 2006. ICCAT Sci. Doc. SCRS/2008/116.

Rivoirard J., J. Simmonds, K. Foote, P. Fernandes and N. Bez, 2000. Geostatistics for estimating fish abundance. Blackwell Science Ltd , Oxford, 206 p.

Royer F., Fromentin J.M., Gaspar P. 2005. A state-space model to derive bluefin tuna movement and habitat from archival tags. Oikos 109, 473–484.

Walker E. and Bez N. 2010. A pioneer validation of a state-space model of vessel trajectories (VMS) with observers' data. Ecological Modelling 221, 17, p. 2008-2017.

Woillez M., J.Ch., Poulard, J. Rivoirard, P., Petitgas and N. Bez, 2007. Indices for capturing spatial patterns and their evolution in time with an application on European hake (Merluccius merluccius) in the Bay of Biscaye. ICES Journal of Marine Science, 64, 537-550.

Witt M.J., Godley, B.J. 2007. A Step Towards Seascape Scale Conservation: Using Vessel Monitoring Systems (VMS) to Map Fishing Activity. PLoS ONE 2, 10.

TABLES

Table 1: Definition of the dominant states and activities per step. Discrimination between effective fishing and various elements contributing to prospecting.

			Dominant state per step				
			1	2	3		
			still	meander	straight		
Dominant activity per step	1	fishing	x			Fishing.	Recorded by observers. Coverage: 10% of the fleet
	2	stop	x			Non fishing. Components of fishing effort.	Not recorded by observers.
	3	tracking		x			
	4	cruising			x		

Table 2. Statistics used to summarize the characteristics of each component of the fishing effort (Fishing, Tracking, Cruising) and of each pair of such components.

	Component by component	Pairs of components
Statistics	Spreading area	Local Index of Collocation (LIC)
Geostatistics	Variogram	Cross-variogram

Table 3. Levels (small versus LARGE) of the summary statistics for the three elementary components of the fishing effort: time spent (hours), surface of occupancy (nm²), spreading area (no unit) and spatial range (degree).

	Fishing	Tracking	Cruising
Time spent	small	LARGE	small
Surface of occupancy	small	LARGE	LARGE
Spreading Area	small	small	LARGE
Nugget	LARGE	small	LARGE
Spatial range	Small	LARGE	LARGE

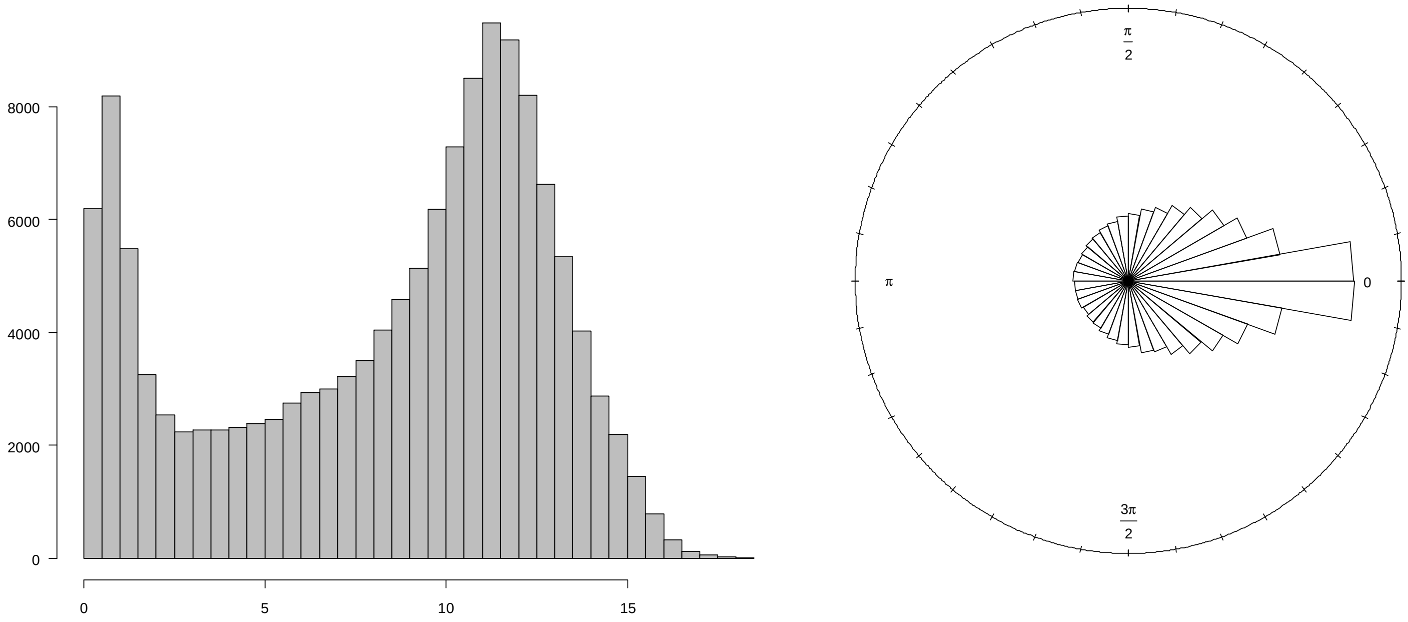

Figure 1: Histogram of a) hourly speeds (in knots) and b) turning angles (in radian) derived from VMS data (GPS positions). All fleet, 2006-2008.

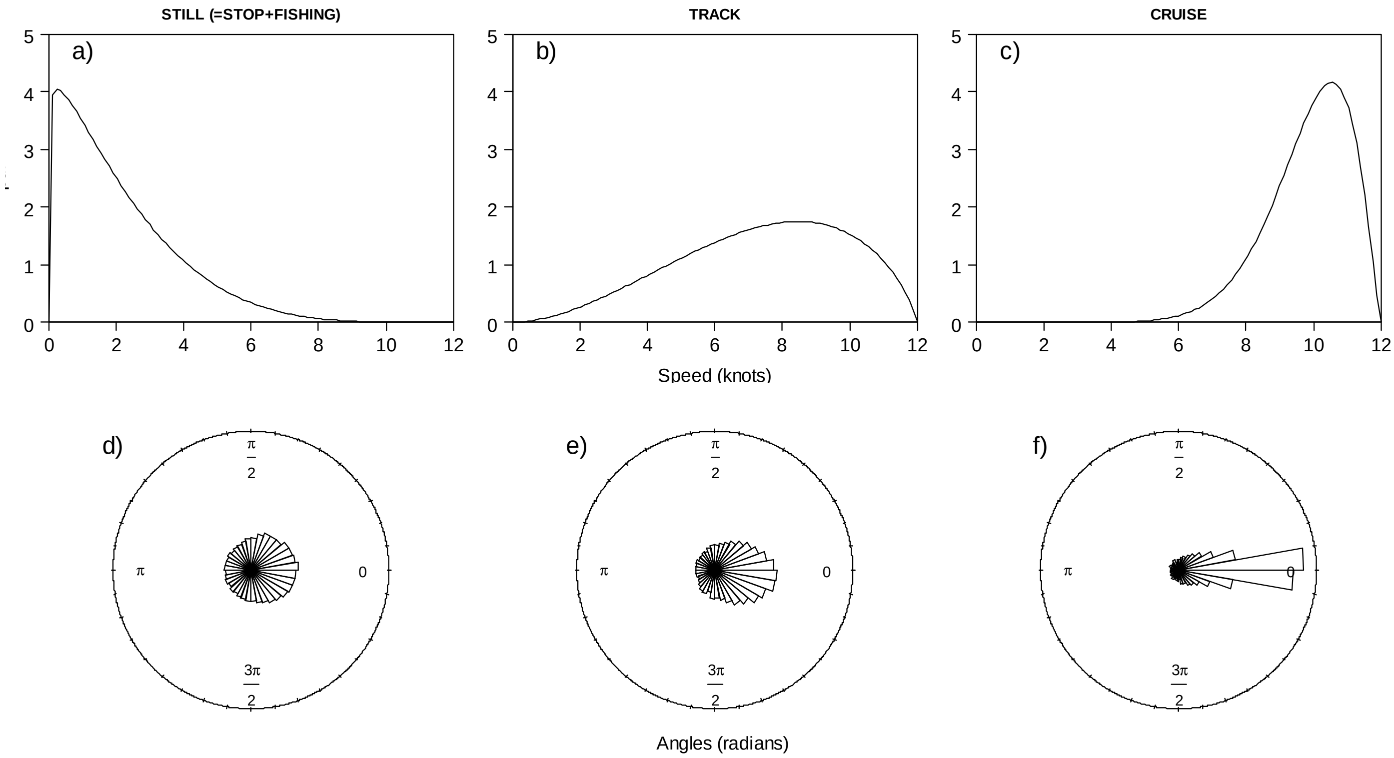

Figure 2 Distributions of speeds (a, b, c) and turning angles (c, d, e) for the three estimated states (fishing (left), tracking (middle), cruising (right)) estimated from the Bayesian model.

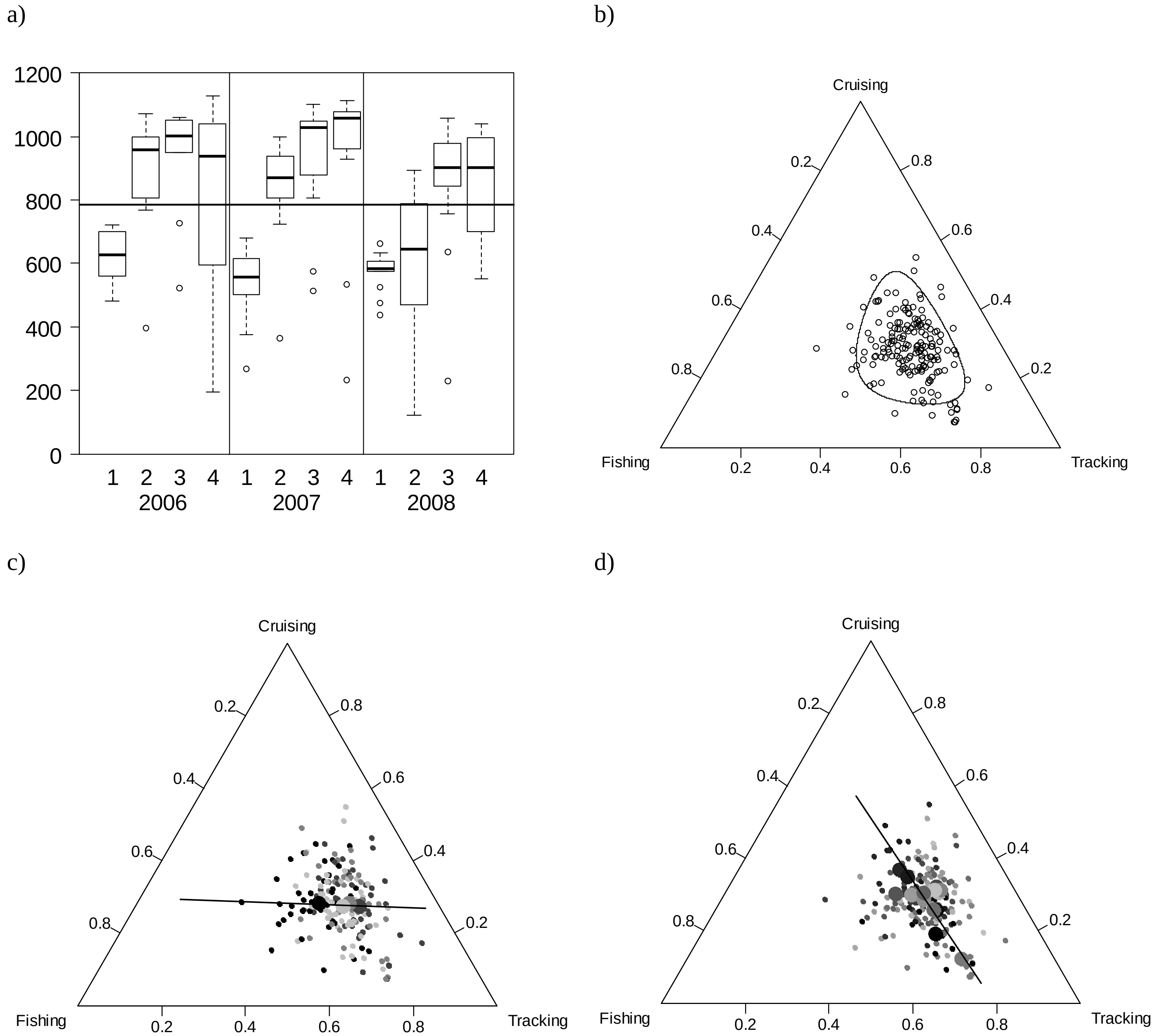

Figure 3: Cumulated time spent in the various components of the fishing effort by vessel and by trimester. (a) Total time spent at sea (hours of daylight); (b) composition of the activity, (c) trimester effect (d) vessel effect.

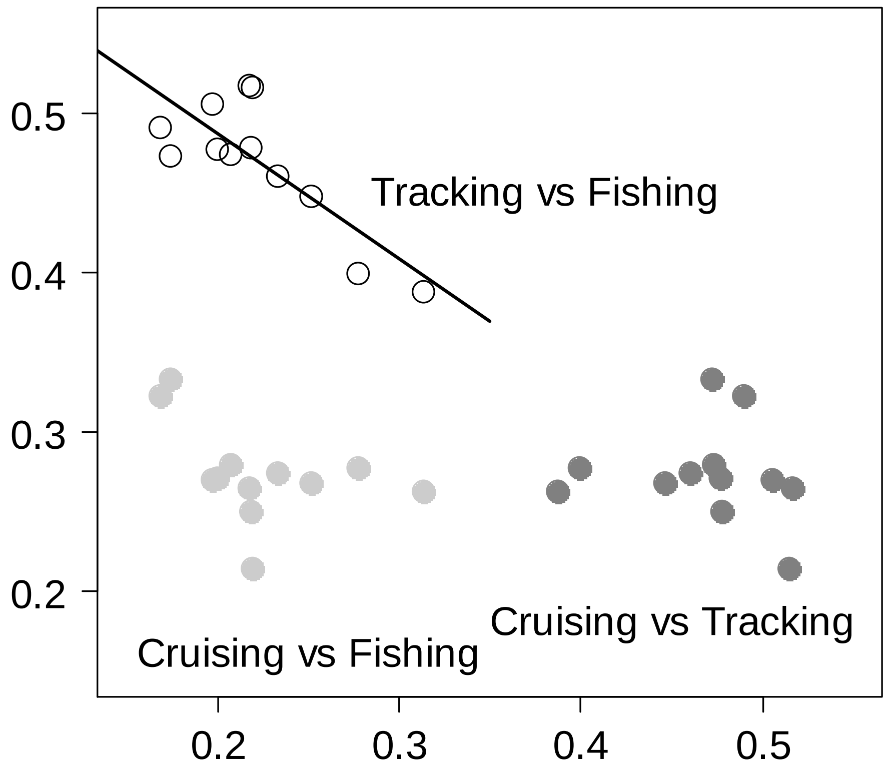

Figure 4 Correlations between the percentages of time spent in the different component of the fishing effort (Fishing, Tracking, Cruising). Data values correspond to the averages of time spent by all the vessels of the fleet during each trimester of the period 2006-2008 (N=12 values). Only the significant regression is represented (p-value=0.02, slope=-0.78)

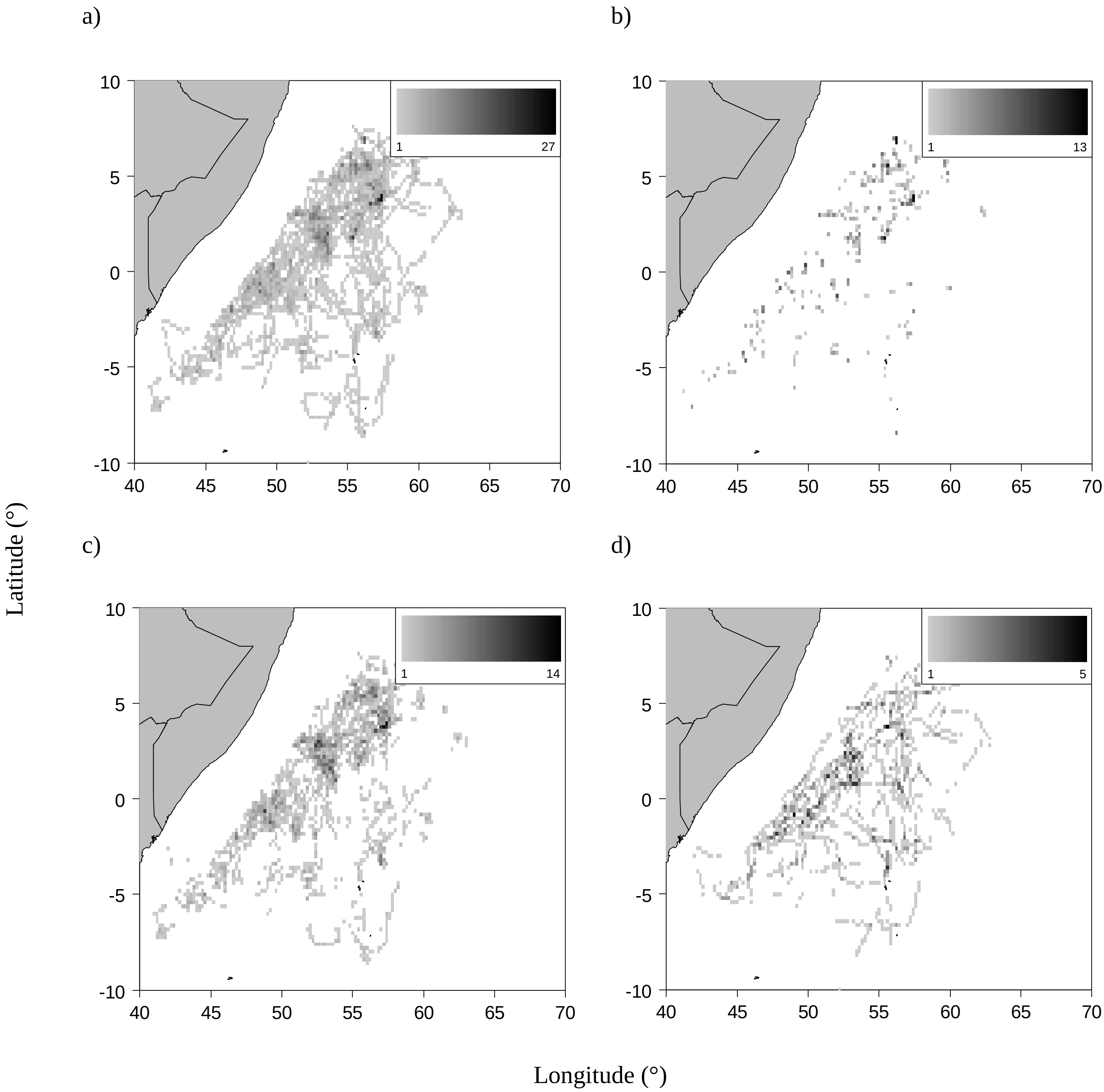

Figure 5 Spatial repartition of the fishing effort (hours) at the scale of 0.2 ° by 0.2 ° grid cells. August 2007. a) the time spent by the fleet irrespective of its activity representing the sampling effort distribution, b) fishing operations (sets), c) tracking phases and d) cruising phases.

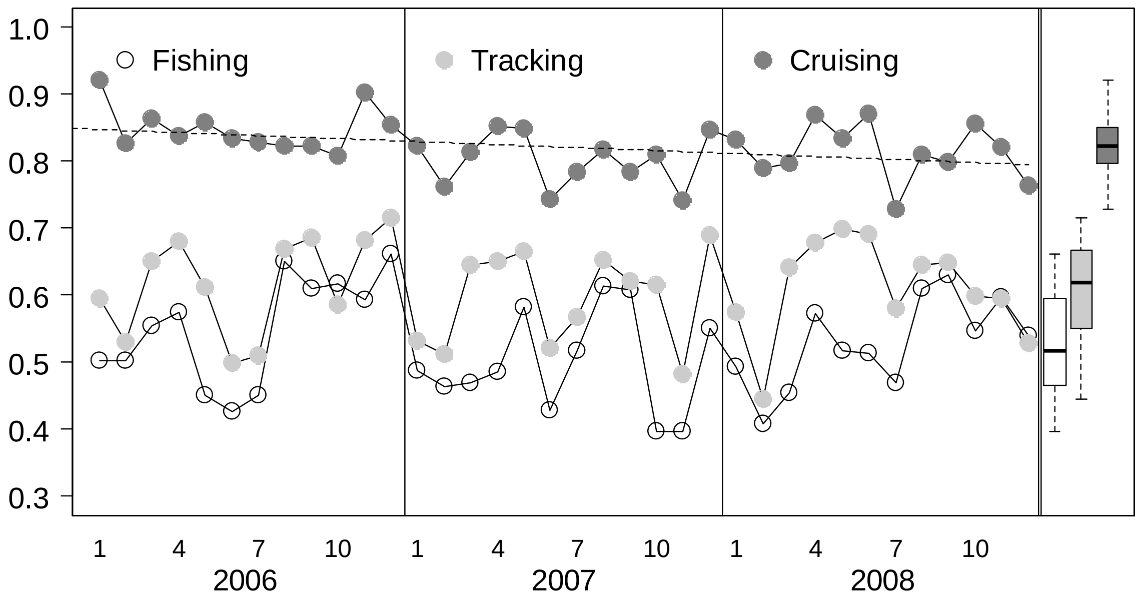

Figure 6 Summary statistics: spreading area (SA) of the various components of the fishing effort (Fishing, Tracking, and Cruising). The y_axis represents the spreading area (1 when all grid cells have the same value - pure homogeneity, and 0 when all is concentrated in one grid cell). The x-axis represents months for the period 2006-2008. Boxplots of each data set are represented on the right end panel.

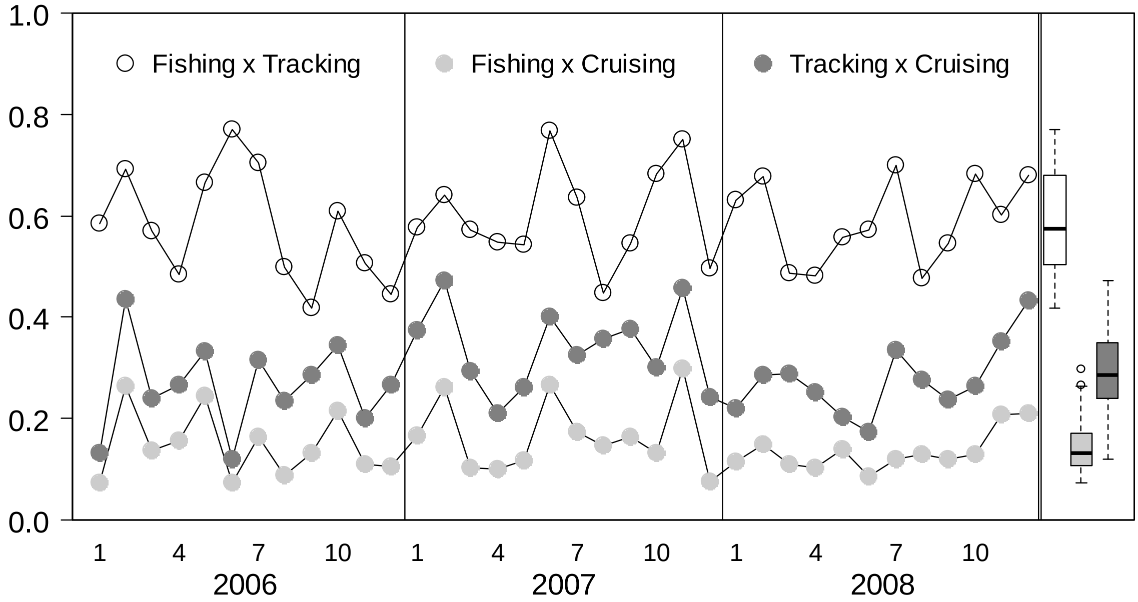

Figure 7 Summary statistics: local index of collocation (LIC) of the various components of the fishing effort (Fishing, Tracking, Cruising). Boxplots are built on the LIC values obtained for each the 36 available monthly maps.

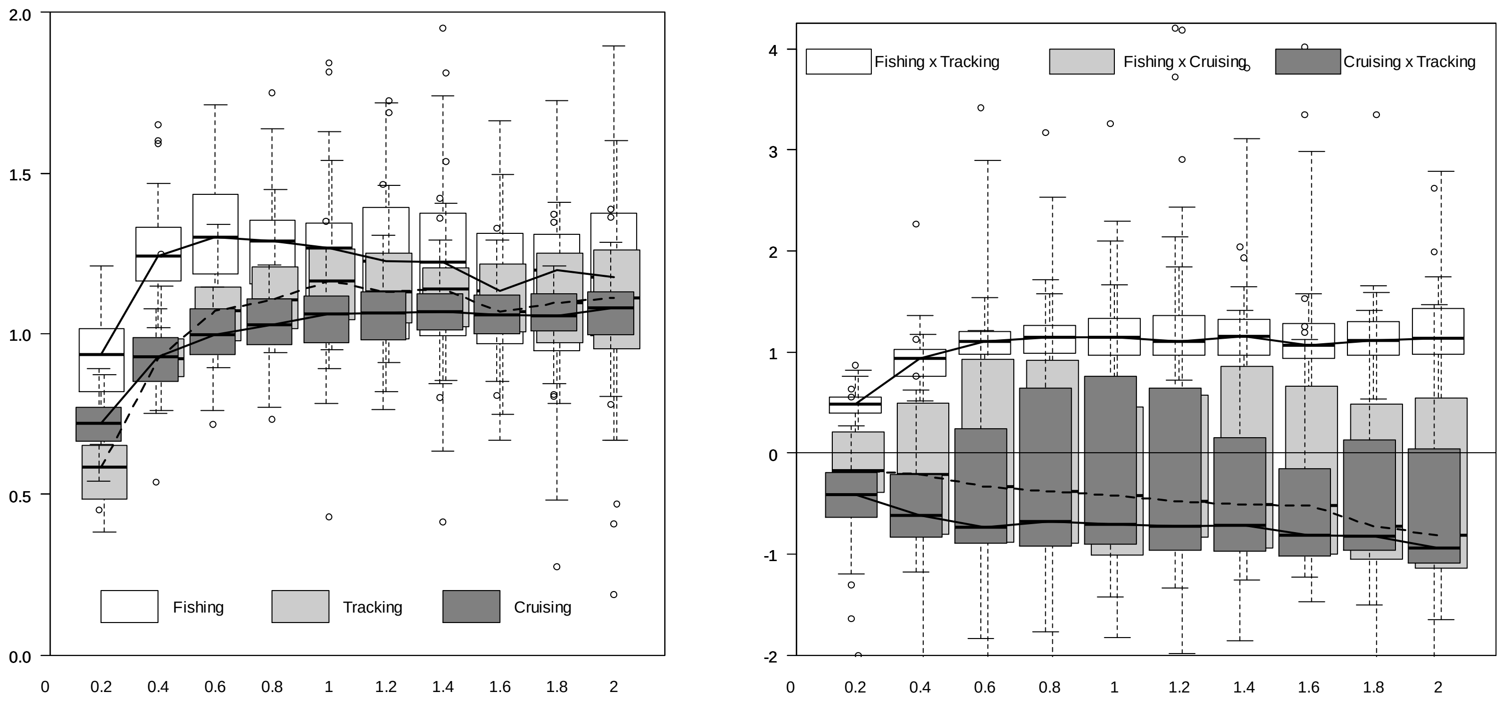

Figure 8 Simple (left panel) and cross (right panel) variograms (standardised by the monthly variances and covariances) of the various components of the fishing effort (Fishing, Tracking, Cruising). Boxplots are based on monthly omnidirectional variograms. They are slightly shifted horizontally to make them visible. The x-axis represents the distance in degrees.