Estudios de Economía Aplicada

VOL. 37-3 2019 PP. 60-80

CLUSTERING BASQUE FISHING COMMUNITIES[1]

IKERNE DEL VALLE ERKIAGA

Departamento de Economía Aplicada V, UNIVERSIDAD DEL PÁIS VASCO, ESPAÑA
E-mail: ikerne.delvalle@ehu.eus

KEPA ASTORKIZA IKAZURIAGA

Departamento de Economía Aplicada, UNIVERSIDAD DEL PÁIS VASCO, ESPAÑA
E-mail: ebpasikk@ehu.eus

Recibido: 10 de Septiembre de 2019
Aceptado: 21 de Septiembre de 2019

ABSTRACT

The main objective of this paper is to identify the taxonomy of Basque local fishing communities (FC) using a set of either, hierarchical (i.e. Ward, average and complete linkage), non-hierarchical (i.e. k-means and k-medoids) and mixed hierarchical-kmeans clustering algorithms; and two alternative fishing related variates at fishing community level, $\{\mathbb{X}\}$ and $\{\mathbb{Y}\}$. The former $\{\mathbb{X}\}$ includes a set of input, output and fleets' structure variables (i.e. the value of landings (PQ), the number of vessels (NB), the estimated capital value (K), the number of fishermen (L), the local level percentage of the small scale artisanal vessels (NBA); while the latter, $\{\mathbb{Y}\}$, exclusively incorporates economic performance productivity ratios (i.e. PQ/NB, PQ/K, PQ/L). Since the variables in both the variates are highly correlated, we are applying a two-step principal component clustering approach in order to find potential groups of homogeneous FC described on a multivariate profile. Our results support 4 FC typologies. The classification is robust to alternative methods and algorithms.

Keywords: local fishing community; artisanal fleet; industrial fleet; two step principal component-clustering; Basque fishing.

RESUMEN

El principal objetivo de este trabajo es identificar la taxonomía de las comunidades pesqueras locales del País Vasco (FC) a partir de un conjunto de algoritmos de clúster tanto jerárquicos (i.e. Ward, enlace promedio y completo) como no jerárquicos (k-medias y k-medois) o mixtos (jerárquico-kmedias) y dos conjuntos alternativos de variables a nivel local, $\{\mathbb{X}\}$ and $\{\mathbb{Y}\}$. El primero $\{\mathbb{X}\}$ incluye un conjunto de variables input, output y de estructura de la flota (i.e. valor de las capturas (PQ), número de barcos (NB), valor estimado del capital (K), número de pescadores (L), porcentaje de embarcaciones artesanales (NBA)); mientras que el segundo $\{\mathbb{Y}\}$, incorpora exclusivamente ratios de productividad económica (i.e. PQ/NB, PQ/K, PQ/L). En tanto que las variables en los dos conjuntos de datos están altamente correlacionadas se opta por una aproximación componentes principales-clúster en dos etapas, para así analizar la existencia de potenciales grupos homogéneos de FC descritas en un perfil multivariante. Nuestros resultados avalan la existencia de 4 tipologías de FC. La clasificación resultante es robusta al método y algoritmo empleado.

Palabras clave: comunidades pesqueras locales; flota artesanal; flota industrial; análisis de componentes principales-clúster en dos etapas.

JEL: A14; C40; C43; Q22

[1] This study has received financial support from the Spanish Ministry of Economics and Competitiveness (Project Ref: RTI2018-099225-B-I00).

1. 1. INTRODUCTION

Spain is one of the main European fishing powers. It agglutinates 19% of the professional fishing of the EU, and, with 1.14% of the catches, occupies the 19th position in the world fishing ranking. Spain has the most important fleet in terms of capacity (21.2%) and is the third in number of vessels (11%). In addition, it generates 31% of EU fisheries employment (SOFIA Report, 2018; MAPA, 2018). However, the fishing sector is not homogeneously distributed along the Spanish coast.

Despite its strong industrial background, the Basque Country, together with Galicia and Andalusia, is one of the principal Spanish fishing powers. With 27% of the capacity, the Basque Country leads tonnage records, after Galicia (41%), and ahead of Andalusia (10%), and the Canary Islands (7%). However, Basque Country barely brings together 3% of professional fishing vessels, ranking led by Galicia (49%), Andalusia (16%), Canary Islands (8%) and Catalonia (8%) (European Fleet Register, 2018). This asymmetry highlights the marked industrial background of the Basque fishing fleet compared to other Spanish fishing autonomous communities. In addition, the analysis of the age pyramid of the Spanish fishing fleet corroborates that, with an average age of 18, the Basque is the youngest fleet in the Spanish State, followed by the Cantabrian (20 years) and Asturian (22 years) fleets, ranking that, on the other end, close the Balearic Islands, Catalonia and the Canary Islands (with average ages that respectively reach 35, 38 and 43 years). Moreover, the Basque Country (17%), after Galicia (35%), is the second Spanish autonomous community according to the value of the landed fish, closely followed up by Andalusia (15%) and Catalonia (10%). Certainly, two Basque fishing ports, Ondarroa and Pasaia, occupy the 9th and 16th positions of the 1.621 official fish landing points collected in EUMOFA, only surpassed by the Galician ports of Vigo (2), A Coruña (6), Burela (7) and Cillero (15).

Although some indicators support that the Basque Country may be catalogued, after Galicia, as the second Spanish autonomous community in importance of its fishing sector, even so, fishing does not even represent 0.5% of its GDP. Even in Galicia, the outstanding leading fishing region in Spain and Europe, no matter the indicator we are using, fishing barely reaches 2% of its GDP. Moreover, although in any country of the world, fishing rarely reaches two digits of the GDP, nevertheless, at local level, fishing is an essential economic activity that determines the evolution, sustainability and survival of fishing communities (FC). In the European common fisheries policy framework, the overall strategy of the fleet's adjustment to the real situation of the fish stocks has been mainly oriented to the aggregate country and regional level, paying a secondary attention to the potential local level external effects caused by the policy itself. Nowadays, after subsequent reforms of the European fisheries policy, we cannot ignore the substantial changes in the local picture of fishing communities.

The local dimension of the global crisis in the fishing sector, and the urgency of facing it in depth, has been a breeding ground for the development of a new orientation in the academic socio-economic research in fisheries, in which, in the framework of the *three bottom line* approach, fishing communities became a central axis of the analysis, covering both theoretical fundamentals as well as empirical cases studies both, at European and international levels (see among others, Agrawal and Gibson, 1999; Bene, 2003, 2009; Ferse, 2010; Granek, 2008; Halpern et al. 2013; Himes, 2003; Nilsson et al., 2016; et al. 2015; Sachs 2012; Symes et al.; Tuler et al. 2013; Visbeck et al., 2014; Rickels et al, 2019). The need to support and strengthen the adaptation of fishing communities to face the vulnerability that the fleets' adjustment policies has placed on them, highlights the need to understand and meet the triple challenge of sustainability in the economic, social and ecological domain (Halpern et al. 2013; Anderson et al. 2015). Additionally, the monitoring of the quantitative and qualitative measurement of these challenges requires new types of data and procedures that, going together with the traditional ones, are essential to build knowledge-based robust policies.

This inherent reality in the fishing sector itself demands a change in the focus from the aggregate or the regional towards the local level. There are plenty of studies analysing the evolution of fishing in the different Spanish regions (Amigo, et al., 2009; Asche and Guillen 2012; Castilla and García del Hoyo, 2006; del Valle et al. 2003; García - Enríquez et al. 2014; Garza and Amigo, 2013; Garza, et al., 2017; Jiménez-Toribio and García del Hoyo 2006; Lostado, 1997; Lostado, 2000; Pascual, 1991; Pascual 2003; Pita and Villasante, 2019); Rodríguez and Villasante, 2012; Villasante et al., 2011), but, to the best of our knowledge, there is no paper addressing the fisheries socioeconomic issues from a local perspective. We intend to cover this gap analysing contemporaneous Basque fishing communities, which in the first place implies to define, identify and classify the FCs based on a set of primary numerical indicators (i.e. output (catches), input (vessels), labour (fishers) as non-numerical (i.e. presence or absence of fish first sale, institutions linked to the sector, etc.); and indicators derived from the crossing of the primary variables indicated above (i.e % of fishers to the local population, age structure of the fleet, predominance of the artisanal or industrial fleet, etc.). This local approach will allow us to analyse some of the keys that condition the vulnerability of FCs in a general framework characterized by the decline and leaving behind of the fishing activity, the lack of generational relief (i.e grey fishing) and the development of tourism as a complementary and/or substitute activity for professional fishing. From the preceding analysis we will derive the ranking or relative position that each FC occupies in the set of FCs and through classification techniques we will delve into the heterogeneity of Basque FCs and in different potential groupings that allow us to identify the taxonomy of Basque FCs.

The remainder of the paper is organised as follows. After this introduction, Section 2 gives an overall, as well as a local community level overview of the fishing in the Basque Country, including the severe adjustment that the Basque fishing sector that has taken place in the last 30 years. In section 3 we first focus on the data generation process and the discussion of the variates to be incorporated in the clustering process, and afterwards, based on a two-step principal component clustering approach, the taxonomy of the Basque fishing communities is discussed. Section 4 summarises the major points made in the paper and concludes.

2. BASQUE FISHING SECTOR IN A NUTSHELL

At present, the Basque fishing fleet is composed of 200 fishing boats (NB), a capacity of 62,945 gross tonnages (GT), and 1,876 fishers (L), distributed in three subsectors (i.e. *inshore*, *offshore*, and *distant offshore* (NASDAP, 2018) (see Table 1 and Figure 1)), and 15 fishing communities (FC). Overall, with an average age of 18 years, 50% of the vessels in the 15-year-old and 75% in the 22-year-old, it can be said that it the Basque is a relatively young fleet and, accordingly, it may have a certain future projection (Figure 2 (a)). The fleet distribution by length (LEN) (Figure 2 ((b)) shows the dominance of the fleet's segment below 40 meters (86%), while the set classified as artisanal according to the legal definition accepted by the European Commission (LEN<12, excluding trawlers) represents around 25% of the total Basque vessels. On the right side of the distribution, 28 vessels exceed 40 meters.

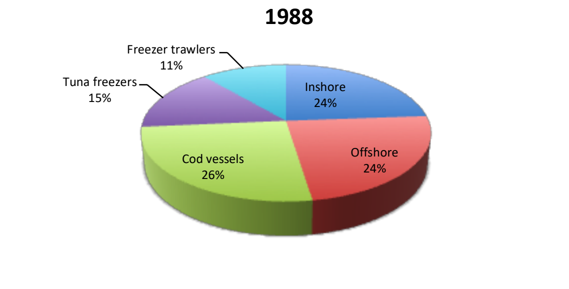
Table 1. Basque fishing fleet (1988-2018)

	1988			2018			Adjustment (%)		
Fishing sub-sector	NB	GT	L	NB	GT	L	NB	GT	L
Inshore fleet	528	26740	4146	156	10017	1131	-70.4%	-62.5%	-72.7%
Offshore fleet	116	26896	1740	17	3876	210	-85.3%	-85.5%	-87.9%
Cod fishing fleet	24	13572	578	2	1650	52	-91.6%	-87.8%	-91%
Freezer trawler fleet	45	31362	933	0	0	0	-100%	-100%	-100%
Tuna freezer fleet	32	21182	707	25	47402	483	-21.8%	223.7%	31.6%

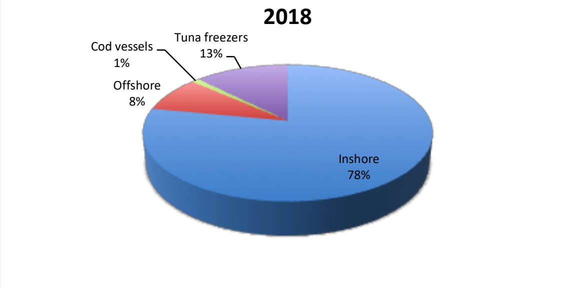
Source: NASDAP

Currently, the Basque inshore fleet comprises 156 fishing boats, a capacity of 10,017 gross tonnages (GT), and 1,131 fishers (L). The average inshore vessel has 64 GT, a power of 310 kilowatts (KW), a length of 20 metres (LEN), a crew of 7 (L), and operates in the Cantabrian and Northwest fishing grounds. The inshore fleet encompasses three main fishing modalities: gillnets, longline-trolling and live bait-seine. According to the number of vessels (NB) by fishing modality, the gillnet and longline-trolling (typical modalities of the so-called artisanal fleet), represent 45% and 17% respectively, while 38% of the fleet is dedicated to live bait, gears typically used by the so-called Basque inshore purse seine fleet. As for the labour employed, the live bait-seine represents 71% of the total crewmembers, followed by gillnets (19%) and the longline-trolling occupies 10%. The main target species of the inshore fleet are white tuna (36%) and anchovy (23%), timely complemented by the mackerel (15%) and sardine (7%), species mostly caught by the purse seine fleet. Hake (5%) and others (11%) (demersal species such as monkfish, turbot, sea bass, etc.) constitute the leading target species of the artisanal fleet segment included in the inshore fleet (Figure 3 (a)).

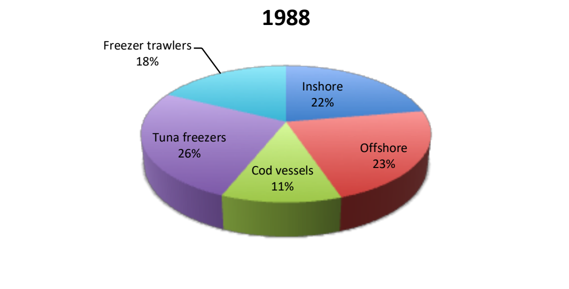
Figure 1. The structure of Basque fishing fleet (1988-2018)



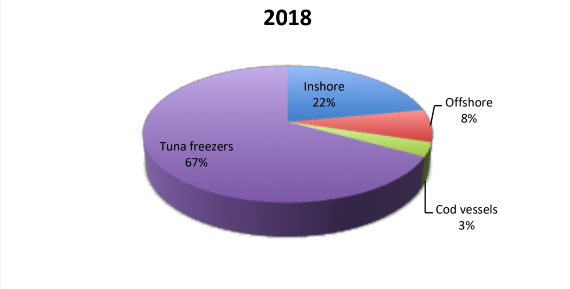
a) Number of vessels (NB) by subsector



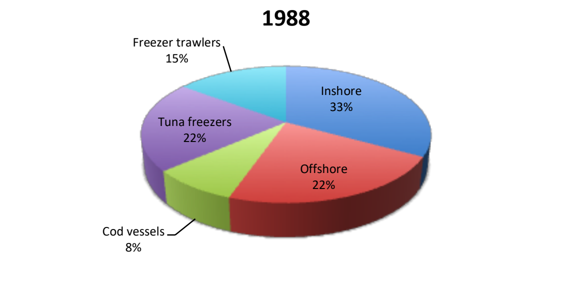
a) Number of vessels (NB) by subsector



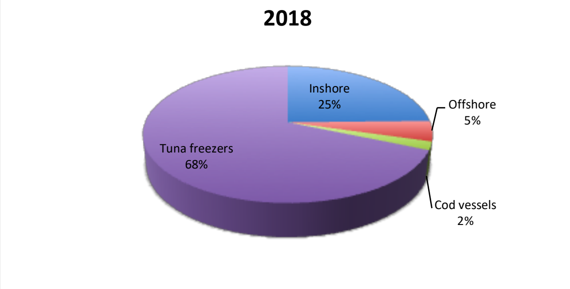
b) Fishing capacity (GT) by subsector



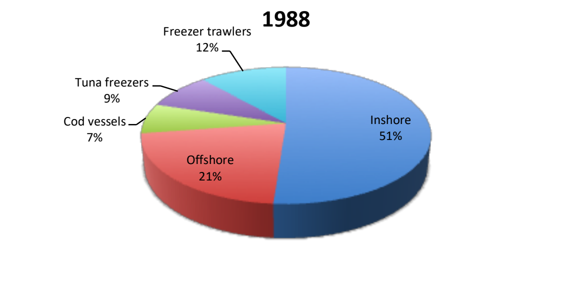
b) Fishing capacity (GT) by subsector



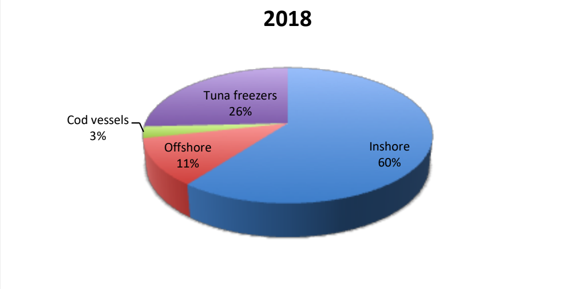
c) Fishing power (KW) by subsector



Fishing power (KW) by subsector



d) Number of fishers (L) by subsector



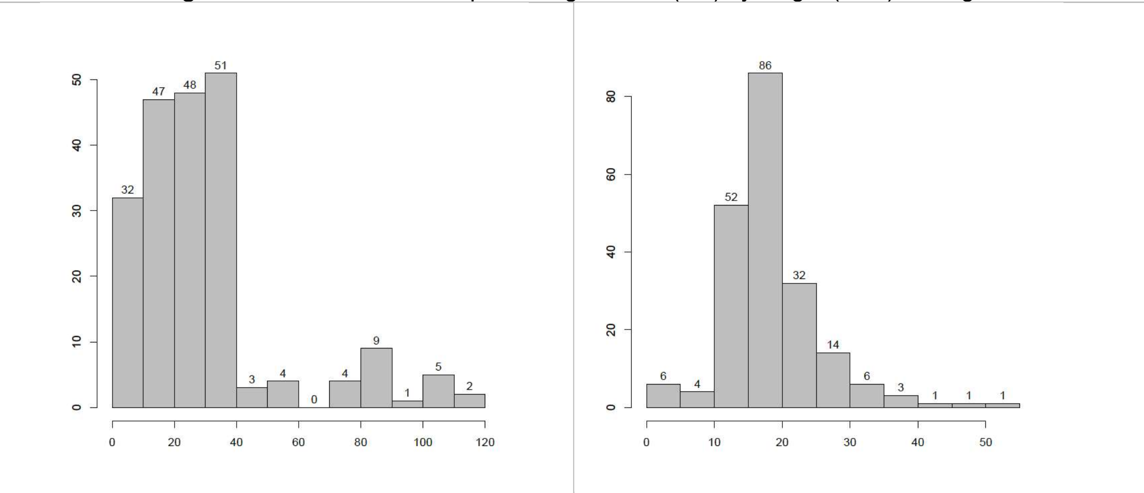
d) Number of fishers by subsector

Source: NASDAP

The offshore fleet is made up by 17 trawlers, a capacity of 3,876 GT and 210 L. The average boat has 228 GT, 552 KW, 35 LEN, 12 L, and operates in the Grand Sol, the fishing ground located in the Irish and British waters of the North Atlantic. The Basque offshore fleet operates with two alternative modalities, pair-trawling and baka, modalities that cover 59% and 41% of vessels and 58% and 41% of the fishers. The target species of the offshore fleet are hake (46%), monkfish (22%) and sea bass (7%) (Figure 3 (b)).

Finally, the distant offshore fleet brings together the so-called cod fishing fleet and the tuna freezer fleet. Nowadays, the former only has two fishing units, 1,650 GT and 52 L, while the tuna freezer fleet comprises 25 NB, 47,402 GT and 483 L. The cod fishing fleet operates in the northeast Arctic (particularly in Svalbard, Division IIB of the ICES area), while the tuna fleet captures in tropical areas of the Indian (68%), Atlantic (26%) and Pacific (6 %) Oceans. The average cod fishing vessel reaches 825 GT, 1,901 KW, 52 LEN, 26 L, and an average tuna freezer 1,896 GT, 5,392 KW, 86 LEN and 35 L. The tune freezer fleet operates with fish aggregation device (FAD) and purse seine or encircling gears, while the cod fishing vessels mainly uses gillnets. As its own denomination indicates, the cod and tuna fleets respectively concentrate their activity on cod and haddock and different varieties of tuna.

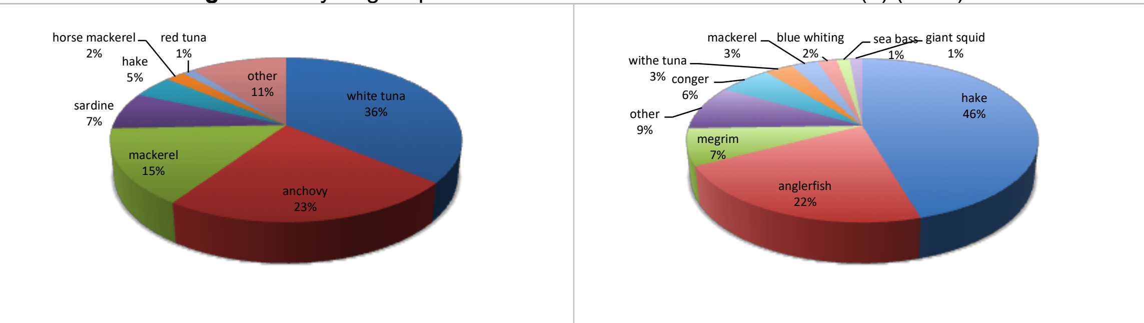
Figure 2. Distribution of Basque fishing vessels (NB) by length (LEN) and age



a) Lenght (LEN) distribution | b) Age distribution

Source: European FLEET REGISTER

Figure 3: Key target species for the inshore and offshore fleets (€) (2018)



a) Species for the inshore fleet (€) | b) Species for the offshore fleet (€)

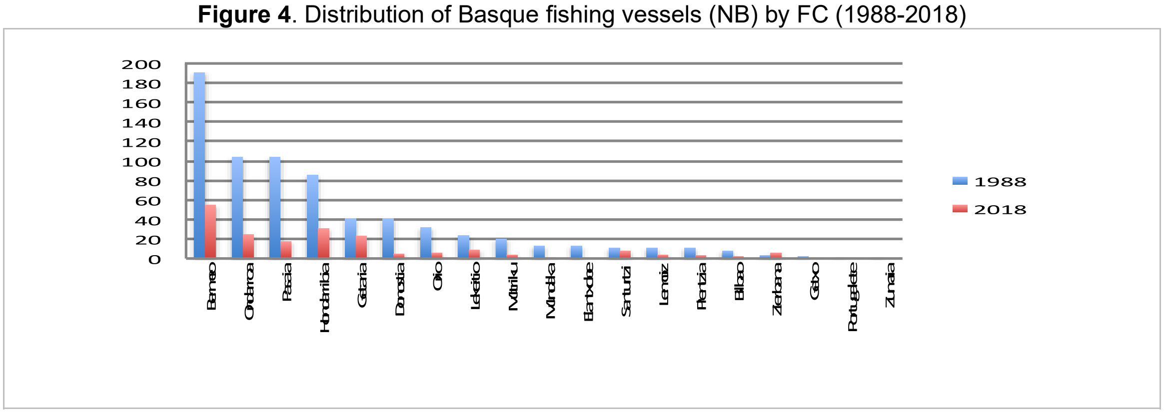
Source: NASDAP | Source: OPPAO

Since Basque inshore and offshore fishing subsector develop their activity in EU waters, to a large extent, their evolution has been subject to the framework established by the Common Fisheries Policy (CFP), including the policies for adjusting fishing fleets to fishing resources, multiannual guidance plans (MAGP) and related capacity limitation programs, incentives for scrapping, limits to the construction of new vessels, and aids to promote innovation and modernization of the fleets. For its part, the evolution of the distant offshore fleets has been linked both, to the global crisis of cod stocks, and the consequent difficulty in obtaining licenses through fisheries agreements with third countries, in whose waters Basque vessels were traditionally fishing from immemorial time.

In the last 30 years, the Basque fleet has undergone a radical adjustment, from 745 units in 1988, to only 200 in 2018. Although the decrease has affected each and every one of the subsectors (inshore, offshore, cod fishing vessels, freezer trawlers and tuna freezers), however, the degree of affectation has been significantly different. The tuna freezer fleet exhibits the most stable behaviour, which, despite the decrease in NB and L has almost doubled its fishing capacity. This allows understanding the dynamism and the technological changes that have been adopted in this subsector. In addition, the decrease in NB and L is much lower than that observed in the other Basque fishing subsectors. At the opposite end, the disappearance of the freezer trawler fleet took place in 2002. In the rest of the segments, there is a gradation in the downward adjustment, led by the cod fishing fleet, with a decrease of around 90%, both in the number of units (NB), capacity (GT) and personnel on board (L); which is closely followed by the offshore fleet, with a reduction of approximately 85%, no matter the indicator we are using. Although the inshore fleet has shown greater resilience, in 2018 it has remained at approximately 75% of what it was 30 years ago.

In addition to the notable adjustment of the Basque fishing fleet, there has been a deep sectorial reconversion that has led to the readjustment in the relative importance of each of the segments, as well as a territorial reweighting of the fishing subsectors. Certainly, the disappearance and /or marginalization of some of the Basque fishing sub-sectors has led to changes in the relative shares of the rest. In this sense, while the freezer trawlers fleet went from 11% of the total Basque vessels to its disappearance, the cod fishing fleet gave 25%, and the offshore fleet went from 24% to only the third (8%); on the other side of the balance, the inshore fleet increases its weight from 24% to 78% of the total NB. The translation of this notorious change in the structure of the Basque fleet to the workforce sphere, leads us to a noticeable increase in the relative weight of the personnel employed in the tuna fleet (from 9% to 25%) and in the inshore subsector (from 51% to 60%), while the relative weight of the offshore and distant offshore subsectors goes from 40% in 1988 to only 14% today (Figure 1).

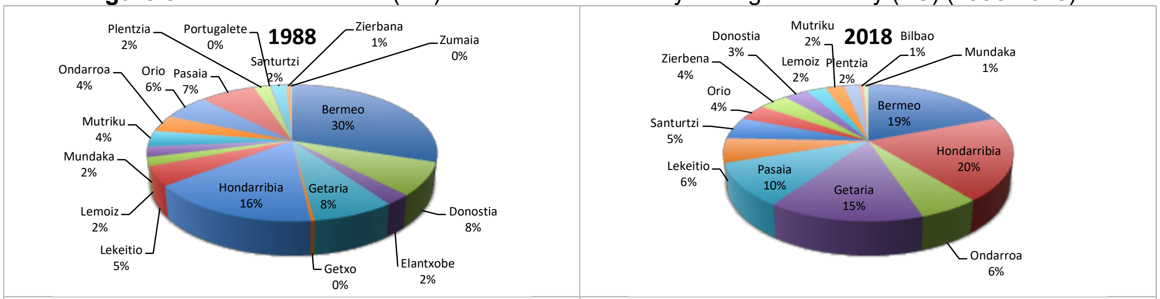
Figure 4. Distribution of Basque fishing vessels (NB) by FC (1988-2018)



Currently, the Basque Country has 15 fishing communities (FC). A FC is understood to be a local entity that accounts operative professional fishing vessels in its local census, with official catches and/or official first sale points. The 15 FC (i.e. Armintza (AR), Bermeo (BE), Bilbao (Erandio) (BI), Donostia (DO), Getaria (GE), Hondarribia (HO), Lekeitio (LE), Mundaka (MUN), Mutriku (MU), Ondarroa (ON), Orio (OR), Pasaia (PA), Plentzia (PL) and Santurtzi (SA) and Zierbena (ZI)) are rather heterogeneous attending both, their current weight to the total Basque fishing and, their dynamic evolution. In the last 30 years, 21% of Basque FCs have disappeared, specifically, Elantxobe, Getxo, Portugalete and Zumaia (Figure 4). In addition, some of the Basque FCs lost their local first sale fish market, although they still maintain fishing vessels in their respective fishing census. This is, for example, the case of Bilbao (Erandio), Mundaka, Orio and Donostia.

The Basque fishing sector has a high spatial concentration, both from the point of view of the fleet and the first sale of fish. Currently, the 5 top FCs according to NB (i.e. Bermeo, Hondarribia, Ondarroa, Getaria and Pasaia) agglutinate more than 75% of the fishing vessels. This concentration is absolute in the offshore and distant offshore fleets. Currently, the offshore fleet is entirely located in Ondarroa (in 1988 it was distributed between the ports of Ondarroa (72%), Pasaia (26%) and Bilbao (3%)). For its part, the cod fishing fleet (with its two current units) and freezer trawlers (currently missing) have been exclusively registered in Pasaia, and the tuna freezer fleet in Bermeo. It is the inshore fleet the one with the most homogeneous presence in all of the 15 Basque FCs, especially Hondarribia (20%), Bermeo (19%), Getaria (15%) and Pasaia (10%). There is also a substantial change in the local structure of the Basque fleet, and a marked relative weight gain of the inshore fleet in FC such as Getaria, Hondarribia and Pasaia, offset by a substantial reduction in the inshore fleet in Bermeo (Figure 5). The commercialization of fish offers (as it will be developed in the next section) a even more significant concentration, with two fish first sale markets (Pasaia and Ondarroa), clearly leading the first sale Basque fish market, comprising around 76% of the fish landed (€) in the Basque Country.

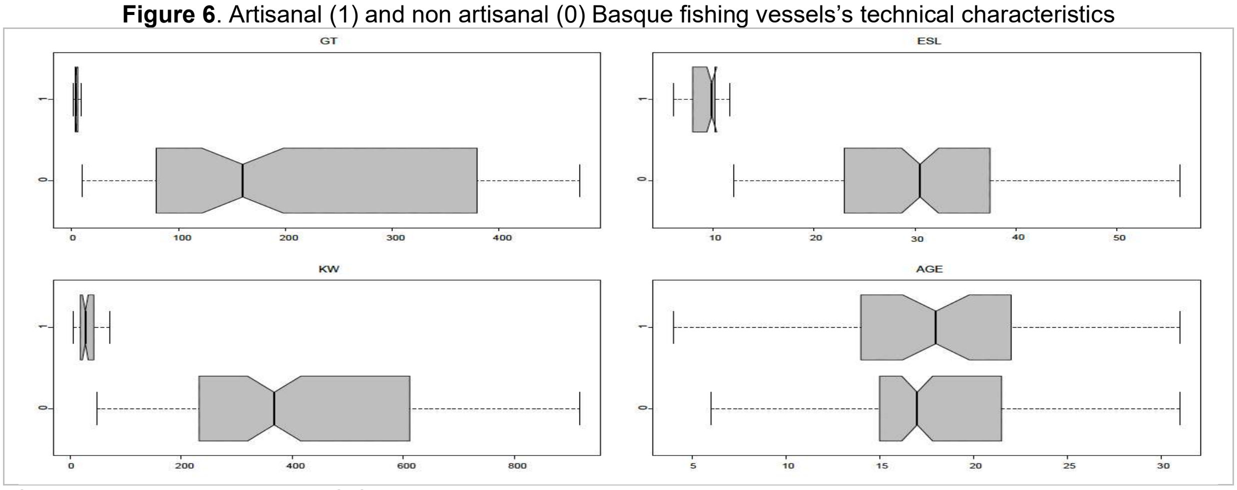
Figure 5. Vessel distribution (NB) for the inshore fleet by fishing community (FC) (1998-2018)



a) NB by FC	b) NB by FC

Source: NASDAP

Figure 6. Artisanal (1) and non artisanal (0) Basque fishing vessels's technical characteristics

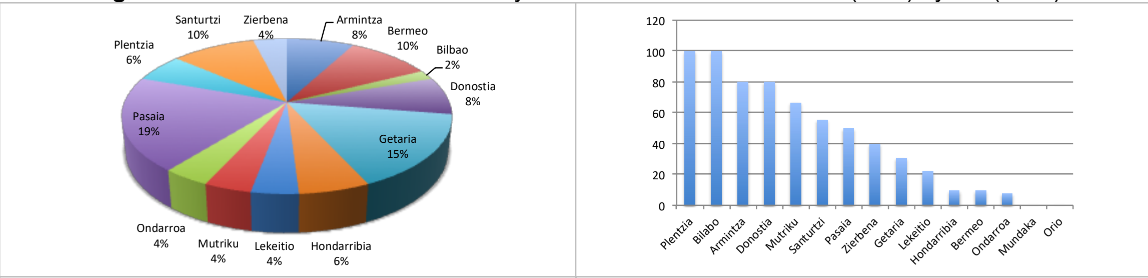


Source: European FLEET REGISTER

The operational definition of an artisanal coastal fishing vessel within the EU framework (fishing vessels with a total length of less than 12 meters and not using trawling gear) allows us to infer the characterization of each FC as artisanal FC or industrial FC by means of the percentage of vessels shorter than 12 meters with respect to the total vessels in the FC (NBA). The artisanal segment of the Basque fleet brings together 25% of the total vessels and is made up by an average vessel of 4.5 GT and 9 LEN. Although the differences in capacity, power and length between artisanal and non-artisanal vessels are obviously significant, there are no meaningful differences in the age of artisanal and non-artisanal vessels (Figure 6).

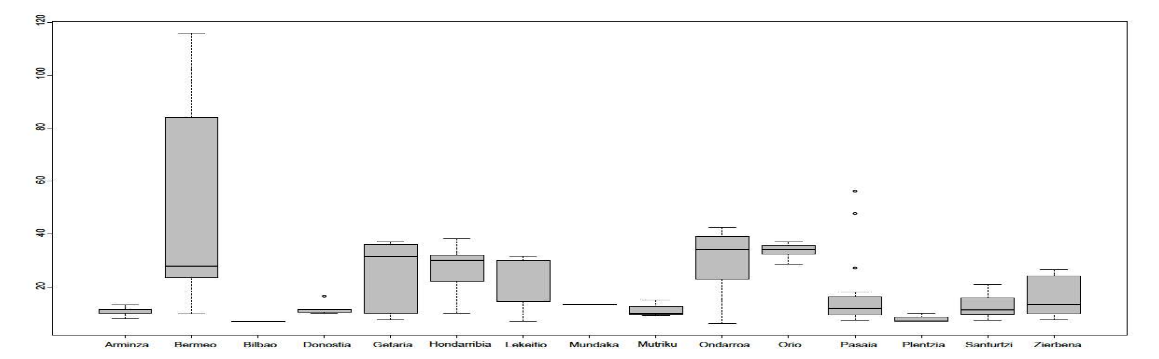
Artisanal vessels are present in virtually all Basque FC (except Orio and Mundaka) (Figure 7 (a)). In locations such as Plentzia, Arminza, Bilbao (Erandio) or Donostia artisanal vessels represent a percentage above 80%, not arriving 25% in Getaria, Lekeitio, Hondarribia, Bermeo Orio or Mundaka. Finally, in Mutriku, Santurtzi, Pasaia and Zierbana, artisanal boats ranges between 40-60% (Figure 6 (b)). From the percentage that represents artisanal fishing with respect to the total (NBA), we infer that Plentzia, Arminza, Bilbao or Donostia can be classified as artisanal FCs; while Getaria, Lekeitio, Hondarribia, Bermeo, Orio and Mundaka can be typified as industrial FCs. Finally, Mutriku, Santurtzi and Pasaia present a mixed artisanal-industrial structure. The average length of the fleet by FC (Figure 8(a)) allows complementing and, at the same time, corroborating the division between artisanal, industrial and mixed FC derived from the percentage of artisanal vessels over the total (NBA). The distribution of fleet ages by FC (Figure 8 (b)) does not offer markedly significant differences in the age of the vessels. However, it can be seen that the average age of vessels in artisanal FCs is slightly above.

Figure 7. Number of artisanal vessels by FC and % of artisanal to total (NBA) by FC (2018)

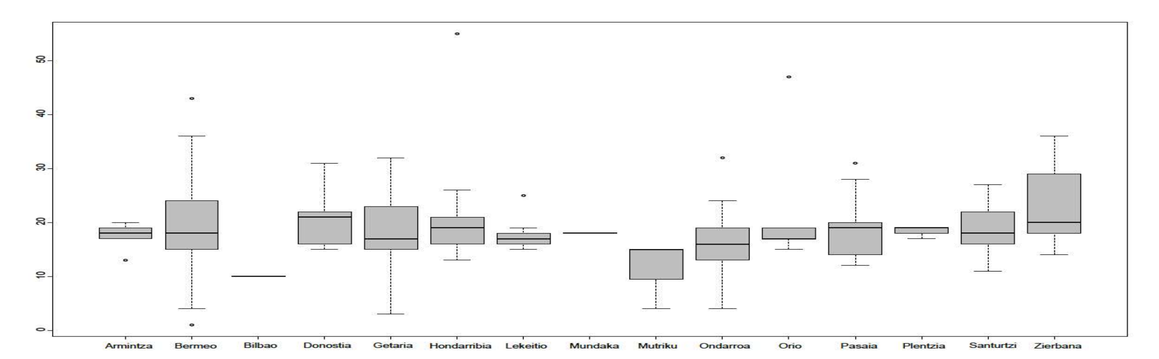


a) Number of artisanal vessels (NB) by FC | b) % of artisanal to total (pNBA) by FC

Figure 8: Length (LEN) and AGE distributions by FC (2018)



a) Length (LEN) distribution by FC



b) AGE by FC

Source: European FLEET REGISTER

3. MATERIALS AND METHODS

3.1. Variables in the clustering process

Cluster analysis is a helpful method for quantifying the structural characteristics of a set of objects (in our case fishing communities (FC)). As such, it has strong mathematical properties, but not statistical foundations. Hence, the usual requirements of normality linearity, and homoscedasticity that are so essential in other quantitative techniques really have little bearing on cluster analysis. Attention should be paid, however, on other key issues such as representativeness of the sample, the presence and treatment of outliers, and the potential correlation in the cluster variate (see Millingan (1996) for a complete conceptual framework).

Clustering results entirely depend on the set of variables included in the analysis. Accordingly, variates in applied clustering should be selected and weighted carefully, or to put in another words, only those variables that are believed to help to discriminate the data should be included. Since our clustering process intends to categorise Basque FC, just fisheries related variables will be incorporated in the analysis. Due to the small size of the population we are working with (N=15), two separate variates will be considered, $\{\mathbb{X}\}$ and $\{\mathbb{Y}\}$. Specifically, the value of landings (PQ), the number of vessels (NB), the estimated capital value of the vessels (K), the number of estimated fishermen (L) and the percentage of artisanal vessels in the fleet (pNBA) make up the set of variables in the variate

{𝕏}; while the economic performance ratios, PQ/NB, PQ/K, and PQ/REM comprise the second variate {𝕐}.

Table 2 shows the descriptive statistics of the set of variables included in {𝕏} and {𝕐}. Different sources have been used and made compatible in order to get our local level database. Specifically, the value of the catches (PQ) comes from EUMOFA (data accessed 12/2/2019). The source for the number of vessels (NB), their technical characteristics, the fishing subsector each vessel belongs to and the vessels' AGE used to estimate the capital stock for each Basque FC (K) is FLEET REGISTRER (data accessed 26/2/2019). We are taking advantage of the hedonic cost function related to GT, AGE and fishing subsector estimated by del Valle and Astorkiza (2013) for the Basque fishing sector to derive fleet's capital (K) at FC level. To that end, following Kirkley and Squires (1988) and del Valle and Astorkiza (2013), vessel characteristics for all the vessels in the Basque 2018 census have been inserted into the estimated hedonic function to obtain estimations of capital stock per vessel, and afterwards, the vessel level capital estimates were summed over all the vessels in each fishing community to obtain FC level measures of the stock of capital inherent in fisheries[2].

The number of fishermen at community level (L) has been inferred from the workforces inscribed in the special regime of the sea (REM) in the Spanish Social Security Institute. The potential overestimation of the fishermen in the ports of special interest for the Spanish State (such as Santurtzi (also known as Port of Bilbao) and Pasaia) have been property corrected to match the dimension of their real fishing fleets using the average number of fishermen by fishing sub-fleets and fishing gears estimated from the data of the Department of Economic Development and Infrastructures of the Basque Govern[3].

Table 2: Descriptive statistics of variables in variates {𝕏} and {𝕐}

	{𝕏}					{𝕐}		
	PQ	NB	K	L	NBA	PQ/NB	PQ/K	PQ/L
Min	0	1	1093	5	0	0	0	0
1Q	0	4	72854	7.5	9.7	0	0	0
Median	24046	6	321772	59	40	7368	0.49	4008
Mean	11011102	13.5	2143450	168.8	43.5	472447	5.16	37879
3Q	9280962	23	2788599	255.5	73.3	261506	1.6	26945
Max	74220667	52	11547196	933	100	3711033	60.12	296883
sd	22256292	14.6	3392520	249.3	35.6	1039219	15.36	80169

To obtain the percentage of the artisanal vessels to the total FC fleet (NBA) we have used the length distribution of vessels lengths derived form data of FLEET REGISTER (data accessed 26/2/2019). NBA is a key indicator of the fleet's structure of each Basque FC. Following EU Commission (Directorate Fisheries) we are categorizing as artisanal (or small-scale coastal vessels) the fishing vessel of less than 12 meters in length (excluding trawlers and draggers). Based on the estimated small-scale or artisanal coefficient for each FC, Plentzia Lemoiz or Donostia (NBA > 65%) may be typified as artisanal fishing community; while, at the other end of the spectrum, Bermeo, Getaria, Hondarribia, Lekeitio or Orio (NBA < 30) are clearly dominated by the non-artisanal fleet

[2] Following the hedonic approach, vessels (capital input in fishing) are considered to be composite production inputs made up of a heterogeneous set of attributes, whose cost of acquisition (C_{it}) or market sale price at time t (P_{it}) (notice that $C_{it} = P_{it}$) is determined by a set of characteristics such as gross tonnage (GT), AGE, VINTAGE, the SECTOR the vessel belongs to, as well as the potential total price effects (including both the supply and demand side ones) that the European Fisheries Structural Policy, via multi annual guidance plans ($MAGP_t$) , may have exercised in the second hand market. Therefore, when acquiring a vessel, we can consider its sale price (P_{it}) to be the sum of the price paid for each one of its attributes [(GT, AGE, VINTAGE, SECTOR), TIME = $MAGP_t$)], so that an implicit price, or hedonic price exists for each one of the attributes defining the vessel.

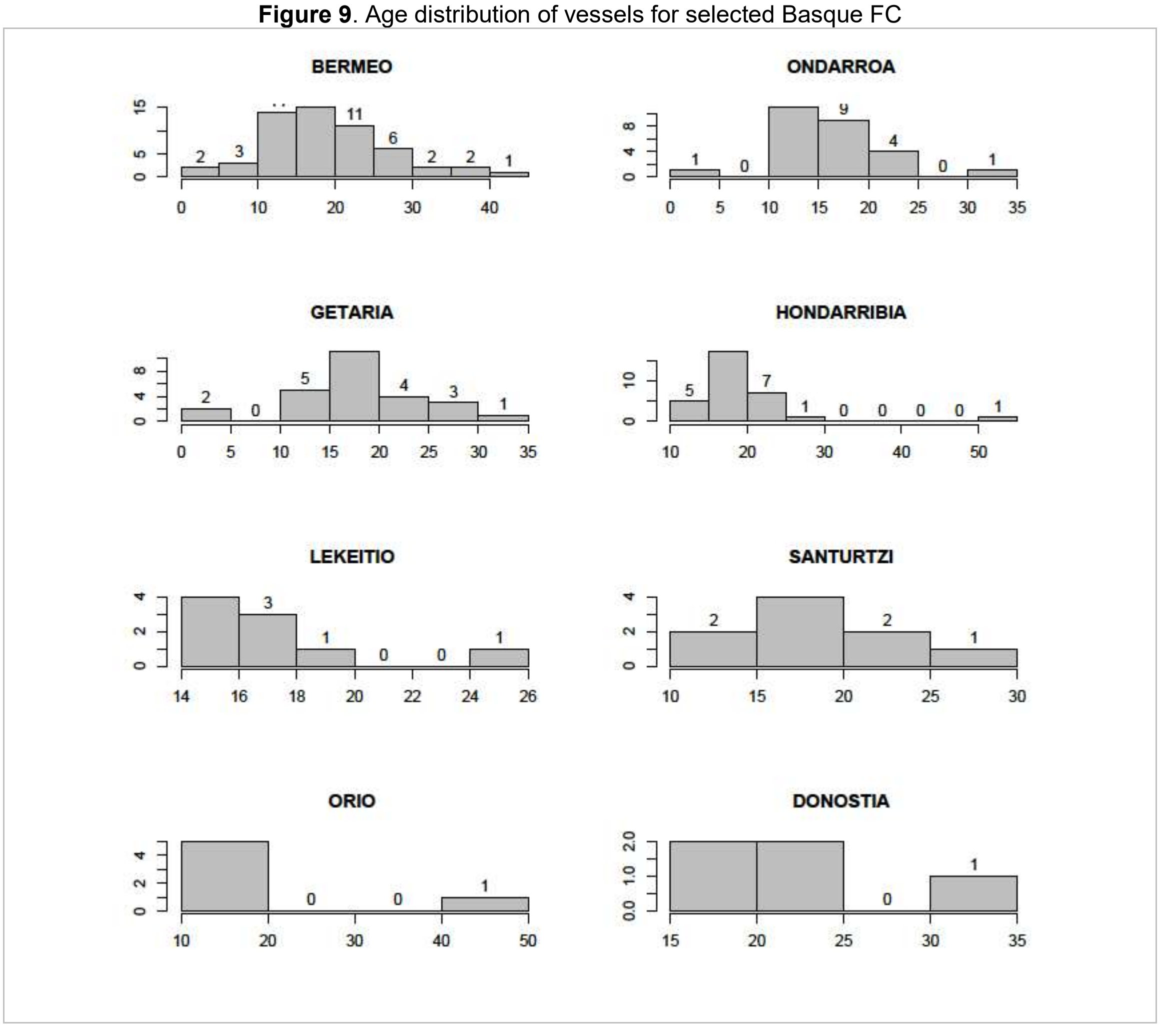
[3] In addition to the fishermen specifically enrolled in the fishing vessels, the special regime of the sea (REM) also includes other workers linked to the maritime sector such as merchant shipping, dockworkers, on board auxiliary health and cooking service, employees in Cofradías and other fishing associations, port practitioners, fishing observers or security members. In order to exclude the workers properly not related to the fishing sector from the ones registered in the REM, a maximum amount of 20, 15, 10 and 5 fishers will be considered depending whether the local level average fleet's length is ≥20, [15-20), [10-15) or ≤5 metres. Hence, when the data coming from the social security records disproportionately outnumbers the local fleet dimension this correction will be used.

segment. Pasaia (NBA=50%) and Zierbena (NBA=40%) exhibit a fairly balanced mixture of artisanal and non-artisanal vessels.

Even if some variables such as the number of new (AGE<5) and old (AGE>30) vessels were a priori conceived relevant and differentiating to be included because they may well show the survival potential of FCs; however, they have been finally excluded. This was due to the fact that the age distribution of the fleet by FC is rather homogeneous (see Figure 9), and accordingly a slight presence of a testimonial aged vessel significantly biased the clustering results. Notice that, although not directly, the AGE of the vessels has been already taken into account in the measure of capital (K) at FC level.

Although nowadays fishing is residual and not even testimonial in communities such as Bilbao (Erandio), nevertheless, we are incorporating all the contemporaneous 15 Basque fishing communities in the clustering process, hereafter, Armintza (AR), Bermeo (BE), Bilbao (Erandio) (BI), Donostia (DO), Getaria (GE), Hondarribia (HO), Lekeitio (LE), Mundaka (MUN), Mutriku (MU), Ondarroa (ON), Orio (OR), Pasaia (PA), Plentzia (PL) and Santurtzi (PA) and Zierbana (ZI). Obviously, this limited population size (N=15) should be taken into account when addressing the number of maximum potential clusters (k). Of course, the outliers should not be eliminated in this clustering process, because, in fact, they may represent outstanding and relevant leading FC that deserves detailed and specific attention.

Figure 9. Age distribution of vessels for selected Basque FC



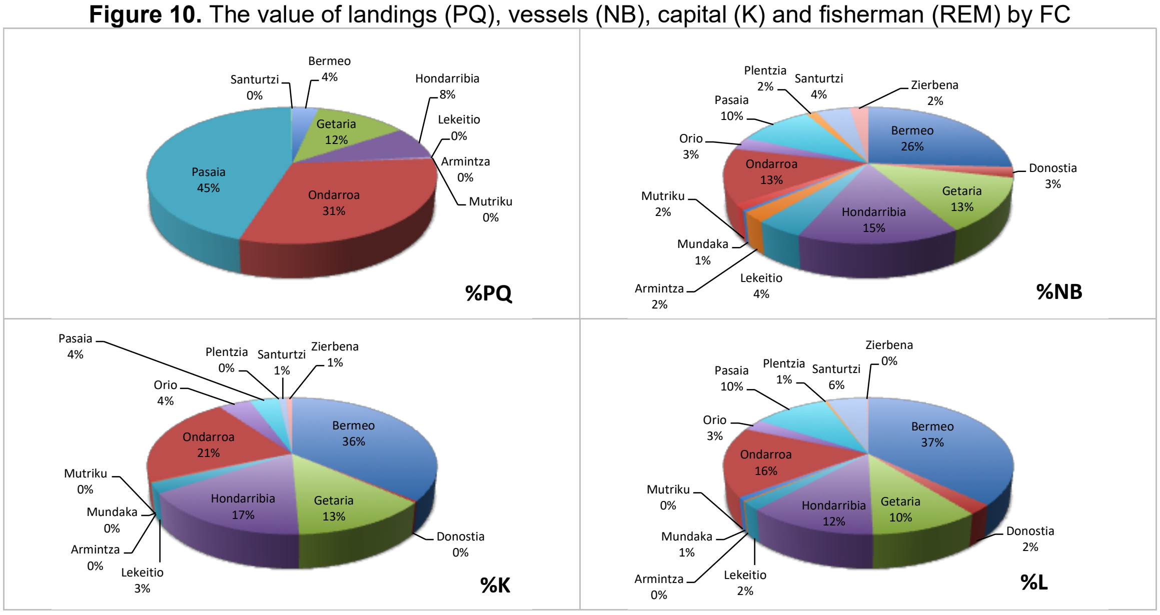
It is straightforward to see that the relevance of fishing regarding the four dimension's vector (PQ, NB, K, L) is not uniformly distributed among the 15 FCs (Figure 10), and that the core of the Basque FC is in fact comprised by Bermeo (BE), Getaria (GE), Hondarribia (HO), Ondarroa (ON) and Pasaia (PA). Moreover, the Basque fishing sector shows a strong concentration no matter the indicator we are using. The C5 concentration indices of PQ, catches (Q), NB, K and L are respectively 99.5%, 99%, 77%, 91% and 86% (Table 3). This concentration is particularly high in the commercialisation of fish, with two FC, Pasaia (PA) and Ondarroa (ON), representing the 76% and 64% of the total value and fish quantity of landed in the Basque Country.

Table 3. Concentration indices and related FC ranking (FCr) for selected variables (%)

	PQ	FCr	Q	FCr	NB	FCr	K	FCr	L	FCr
C_1	45	PA	40	ON	26	BE	36	BE	37	BE
C_2	76	ON	64	PA	41	HO	57	ON	53	ON
C_3	88	GE	81	GE	54	GE	74	HO	66	HO
C_4	96	HO	94	HO	67	ON	87	GE	76	GE
C_5	99.5	BE	99	BE	77	PA	91	OR	86	PA
Herfindahl	0.32	-	0.25	-	0.13	-	0.22	-	0.20	-
Rosenbluth	0.34	-	0.28		0.14	-	0.23	-	0.19	-

The set of variables in variates {𝕏} and {𝕐} are highly correlated (Figure 11), which may bias the partitions resulting from clustering. Notice that multicollinearity is really a weighting process not apparent to the observer, but affecting the analysis nonetheless. Accordingly, we are factoring the variables using principal component analysis (PCA) prior to clustering, and using the resulting factor scores as cluster variables. Recall that principal components are uncorrelated. Before applying PCA variables in {𝕏} and {𝕐} are typified by subtracting their respective mean and dividing by their standard deviation[4].

Figure 10. The value of landings (PQ), vessels (NB), capital (K) and fisherman (REM) by FC



3.2. Methods

3.2.1. Assessing clustering tendency

Although the descriptive statistics of {𝕏} suggest a sound group polarisation of the Basque FCs according to catches (PQ), capital value (K) and the number of fishers (L), we are checking whether the selected {𝕏} and {𝕐} exhibit an underlying clustering structure by means of Hopkins test[5] (Hopkins and Skellam, 1954; Lawson and Jurs, 1990) and a battery of modality tests[6] including

[4] Cluster analysis is quite sensitive to different scales or magnitudes among the variables and variables with larger standard deviation have in general more impact on the final similarity value. Accordingly, when clustering, variables that are not in the same scale should be standardised to avoid instances where variable's influence on the cluster solution is greater that it should be.

[5] The Hopkins statistic tests the spatial randomness of the data by measuring the probability that a given data set is generated by a uniform data distribution. The test compares the distances between the data points and the nearest neighbours from a sample of pseudo points and their nearest neighbours. If the data are not distributed in clusters, then both sets of distances should be similar on average. Clusterability can be inferred by comparing to a threshold calculated based on the distribution of Hopkins statistic.

[6] Multimodality tests initially assume that data is generated from a unimodal distribution (the null) and accordingly the p-value is the probability of observing the given input or a more extremely multimodal input under the null. If only a single

Hartigan (1985), Cheng and Hall (1998) Fisher and Marron (2001) and Hall and York (2001) (Table 4). The R package *modetest* (Ameijeiras et al. 2018) has been used to obtain modality tests. The value of Hopkins statistic is not far from 1, so we can conclude that our dataset is significantly clusterable. Cheng-Hall and Fisher-Marron modality and multimodality tests also give support of strong clustering structure. Moreover, the multimodality test of Fisher and Marron suggest a multimodal structure with at least 4 modes. However, based on Hartigan and Hall and York tests there is no evidence against the dataset is uniformly distributed. Despite this ambiguity, taking into account the small population size of our data set we will accept that our data exhibits a clusterable pattern.

Figure 11: Correlation among the selected variables

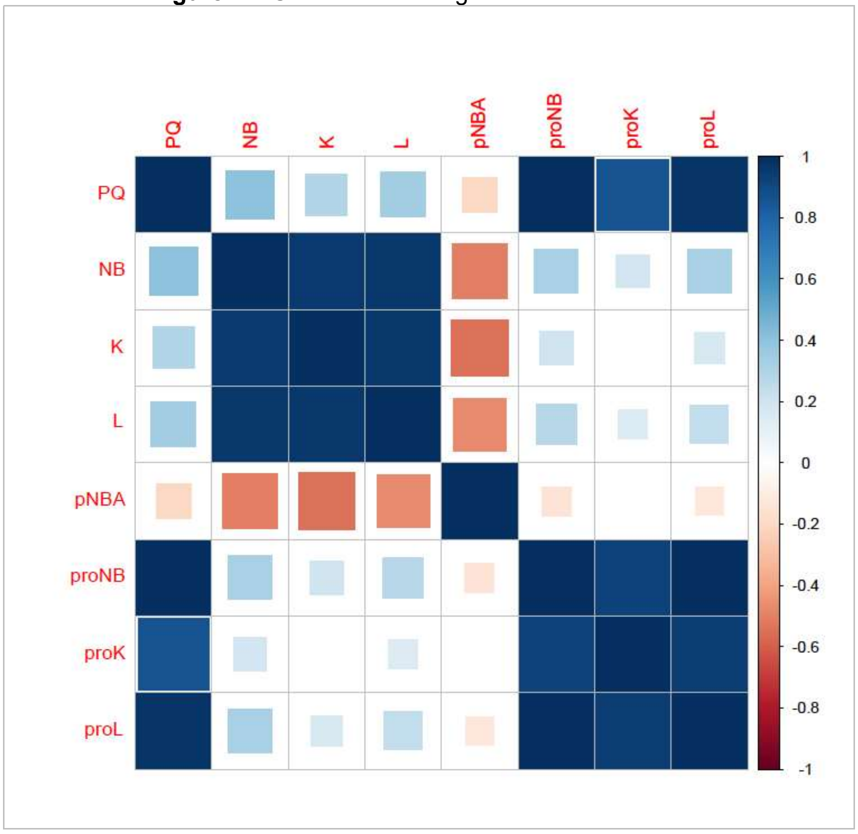


Table 4. Testing for clusterability

	$\{\mathbb{X}\}$		$\{\mathbb{Y}\}$	
	Statistics	p-value	Statistics	p-value
Hopkins	0.2017	-	0.1461	-
Hartigan dip test for unimodality’	0.0460	0.258	0.0577	0.0075***
Cheng and Hall excess of mass test	0.09216	0.042*	0.1579	0.0000***
Hall and York critical bandwidth test	0.5197	0.138	0.9125	0.046**
Fisher and Marron test’’	0.6128	0.000***	1.649	0.0000***
Fisher and Marron test’’’	0.3365	0.006***	0.9215	0.0000***
Fisher and Marron test’’’’	0.25	0.006***	0.5671	0.0000***

‘ Alternative hypothesis: non-unimodal, i.e., at least bimodal simulated p-value based on 2000 replicates.
‘ Null hypothesis: unimodality. Alternative hypothesis: at least 2 modes. B=100 bootstrap replicas.
‘’ Null hypothesis: 2 modes. Alternative hypothesis: at least 3 modes B=100 bootstrap replicas.
‘’’ Null hypothesis: 3 modes. Alternative hypothesis: at least 4 modes B=100 bootstrap replicas.

3.2.2. Principal Components Analysis (PCA)

The general goal of PCA is to find the linear combinations with large variance formed by the original variate, which may be useful when the variables within the data set are highly correlated. Accordingly, PCA is often used before clustering to reduce the original variables into a smaller and

mode is present, then the p-value should be large, indicating that the underlying data is deemed unclusterable. By contrast, small p-values make us the question the original assumption of unimodality and instead conclude that multiple modes (and multiple clusters) are present.

more parsimonious set of *new* variables (or principal components (PC)) explaining most of the variance in the original variate (see for example Anderson, 1984; Raychaudhuri et al., 2000, Brusco el al, 2017). Although, in fact, all the PC will be required to reproduce the total system variability of the data, a common practice is to replace the initial variables with a limited number of PC. This means that just certain number of PC will conform the effective inputs to perform the cluster analysis (Jolliffe et al., 1980; Johnson and Wichern, 1988). As a common rule of thumb originally suggested by Kaiser (1958), we will retain eigenvalues[7] >1 and limit the number of components to that number that accounts for al least 85% of the total variance explained.

Table 5 includes eigenvalues, percentages and cumulative percentages of projected variances of the variates $\{\mathbb{X}\}$ and $\{\mathbb{Y}\}$. The first two factors (PC1 and PC2) account for 85% of the total variance of $\{\mathbb{X}\}$ and 99% of $\{\mathbb{Y}\}$, which means that 85% and almost 100% of the information contained in each of the variates is retained by the two first principal components. Thus, we may consider that the variance corresponding to the remaining axes may be mainly random noise (Lebart et al. 1984). Accordingly, we will retain PC1 and PC2 $\{\mathbb{X}\}$ and PC1 $\{\mathbb{Y}\}$ to proceed with the cluster analysis.

Table 5. PCA analysis

Eigenvalues and percentages of the projected variances

	$\{\mathbb{X}\}$					$\{\mathbb{Y}\}$		
	PC1	PC2	PC3	PC4	PC5	PC1	PC2	PC3
Eigenvalues	3.4277	0.8483	0.6571	0.0361	0.0308	2.9093	0.085	0.004
Standard deviation	1.8514	0.9210	0.8106	0.1900	0.1755	1.7057	0.29306	0.0692
Proportion of variance	0.6855	0.1697	0.1314	0.0072	0.0062	0.9698	0.02863	0.0016
Cumulative proportion	0.6855	0.8552	0.9866	0.9938	1	0.9698	0.9984	1

Coordinates and variable-factor correlations

	$\{\mathbb{X}\}$					$\{\mathbb{Y}\}$			
	PC1	PC2	PC3	PC4	PC5		PC1	PC2	PC3
PQ	0.47	-0.88	0.12	0.01	-0.01	proNB	0.98	-0.15	-0.04
NB	0.97	0.03	-0.17	-0.16	0.01	proK	0.97	0.23	-0.01
K	0.97	0.16	-0.15	0.06	-0.13	proL	0.99	-0.07	0.05
NBA	-0.65	-0.21	-0.73	0.00	-0.01	-	-	-	-
L	0.96	0.09	-0.23	0.09	0.12	-	-	-	-

3.2.3. Clustering

Starting from the PCA performed above, we carry out a classification analysis respectively using the scores of the first two (PC1 and PC2) and first principal components (PC1) for $\{\mathbb{X}\}$ and $\{\mathbb{Y}\}$ for (PC1).

In the hierarchical procedures, the clustering algorithm starts out by putting each observation into its own separate cluster. It then examines all the distances between all the observations and pairs together the two closest ones to form a new cluster. Accordingly, the result at the earlier stage is always nested within the results at a larger stage, creating a similarity tree or dendogram. The most popular agglomerative algorithms are Ward's, complete and average linkage methods[8]. In contrast to hierarchical methods, non-hierarchical procedures such as k-means do not involve a treelike construction process. Instead they assign objects into clusters once the number of clusters is specified. Thus, the first task is to identify the cluster seeds (or starting points) for each cluster and afterwards, based on similarity, assign each observation to one of the cluster seeds. Since the final result of k-

[7] Eigen values (also known as singular values) are derived for each dimension and measure the variability retained by each principal component. It is large for the first PC and small for the subsequent PCs.

[8] In the complete linkage method cluster similarity is based on maximum distance between observations in each cluster, while in the average linkage procedure similarity of any two clusters is the average similarity of all individuals in one cluster with all individuals in another. Accordingly, this second algorithm depends less on outliers and tend to generate clusters with approximately equal within-group variance (Hair et. al. 2014). Ward's method differs from the two previous methods in that the similarity between two clusters is not a single measure of similarity, but rather, the sum of squares within the clusters summed over all variables. The selection of which two clusters to combine is based on which combinations of cluster maximises the within-cluster sum of squares across the complete set of disjoint or separate clusters. Since the sum of squares is directly related to the number of observations involved, Ward's method tends to combine clusters with a small number of observations and tends to produce clusters with approximately the same number of observations. Moreover, the use of a sum of squares measure makes this method easily distorted by outliers (Milligan, 1980; Hair et. al. 2014).

means clustering is sensitive to the starting points, we are specifying 50 different random starting assignments and then select the best results corresponding to the one with the lowest within cluster variation. Instead of the means, k-medois uses medois[9] as cluster centers, which implies that it is less sensitive to noise and outliers. The most common k-medois clustering method is PAM algorithm (Kaufman and Rousseeuw, 1990). With pam, the sums of the distances between objects within a cluster are constantly recalculated as observations move around, which will hopefully provide a more reliable solution. See, among others, Hair et al, 2014; Kassambara, 2017; Brusco et al. 2017 for further detailed surveys about clustering algorithms.

Choosing the optimal number of clusters that best describes our FC is fundamental to derive the taxonomy of Basque FC. A substantial number of empirical clustering studies use an informal rule of thumb from Lehmann (1979) [N/60)>k<N/30], where N is the number of objects to be clustered. However, our limited population size (N=15) implies that this empirical practise is clearly inappropriate, and that, in practice, we should tentatively consider a maximum of no more than 4 clusters. In order to check the proper number of clusters constrained to the effective dimension of our data, that is, k=2, k=3 or k=4, we are using a number of standard internal cluster validation procedures, such as the popular elbow[10] and silhouette methods (Rousseeuw 1987; Kaufman and Ronssenw, 1990)[11], as well as a set of additional indices including CH (Calisnki and Harabasz, 1974)[12], D (Dunn, 1974; Halkidi et al., 2002)[13], average Pearson gamma (Halkidi, 2001)[14], entropy (Meila, 2007)[15] and WB ratio[16]. Tables 6 and 7 show the optimal decisions based on each of this indices. At this stage we are using the R package *fpc* (Henning, C. (2019).

In any case, regardless of the result, our limited population size makes reasonably manageable to check in parallel the resulting k=2, k=3, and k=4 cluster membership before concluding about a taxonomy of the Basque FC. Cluster membership related to each of the partitioning hierarchical and non-hierarchical methods and number of clusters has been reported in Tables 8 and 9. The certainly outstanding result is that the partitions of FC are quite robust to the clustering algorithm we are applying. This is particularly truth for the four clusters solution (with 100% matching) no matter the variate we are clustering, $\{\mathbb{X}\}$ or $\{\mathbb{Y}\}$.

Concerning to the optimal number of clusters, we see contradicting results between the different methods and variates. While the k=4 clusters solution dominates for $\{\mathbb{X}\}$, the two-cluster solution seems to be more appropriate when we are using the performance ratio indicators included in $\{\mathbb{Y}\}$. However, the partition resulted with the two-cluster solution, although highly informative, does not help to conclude about a clear taxonomy for the Basque FC. Accordingly, four clusters were ultimately determined by balancing the performance of the cluster statistics and the informative capacity of the resulting partitions. However, for completeness 2 clusters and 3 clusters related taxonomies will be also discussed (see Table 10 and 11).

---

9 The medoid refers to an object within a cluster for which the average distance between it and all the other members of the cluster are minimal. In corresponds to the most centrally located point of the cluster.

10 The basic idea behind partitioning methods is to define clusters such that the total intra-cluster variation or the total within-cluster sum of squares (WSS) is minimised. The elbow method looks at the WWS as a function of the potential number of clusters (k). Then, one should choose a number of clusters so that adding another cluster doesn't improve much better the total WSS.

11 The silhouette method measures how well each object lies within its cluster. Average silhouette method computes the average silhouette of observations for different values of k. The higher the average silhouette width, the better the clustering quality. Then, the optimum number of clusters k is the one that maximise the average silhouette over a range of possible values for k.

12 CH index or Pseudo F is the ratio of between cluster to within cluster variance, CH= (BSS/(k-1)/(WSS/(N-K), where k is the number of clusters, N the number of observations, and BSS the between cluster sum of squares. WSS measures how close the points in a cluster are to each other, while BSS measures how far apart the clusters are from each other. Accordingly, a good clustering has a small WSS and a large BSS. Thus, large values of Pseudo F indicate close-knit and separated clusters. Then, CH should be maximized at the optimal k.

13 Dunn's index (Dunn, 1974) is defined as the minimum of the ratio of the dissimilarity measure between two clusters (i.e. the minimum inter-cluster distance) to the diameter of the cluster (maximum cluster size), where the minimum is taken over all the clusters in the data set. Based on the definition, large values of the index indicate the presence of compact and well separate clusters. Then, a higher Dunn index implies better clustering.

14 Pearsongamma.best <- cs_metrics$cluster.number[which.max(cs_metrics$pearsongamma)]

15 Entropy.best <- cs_metrics$cluster.number[which.min(cs_metrics$entropy)]

16 A smaller within-between cluster ratio indicates a better fitting.

Table 6. Internal cluster validation measures for {𝕏}

	k=2	k=3	k=3	k=4
	km=pam=hc=hkm	kmvs=hc=hkm	pam	km=pam=hc=hkm
between ss	3.7217	3.7363*	3.6993	3.6824
within ss	21.7579	10.1553	13.1375	5.0186*
silhouette	0.6138*	0.6047	0.5916	0.5957
CH	22.7677	29.3690	21.3403	40.0707*
dunn	0.3034	0.5150	0.4435	0.9910*
dunn2	1.3323	1.5558	1.4390	1.9758*
entropy	0.6365*	0.8033	0.8033	0.9533
Pearson gamma	0.7510	0.8190	0.7974	0.8340*
wb ratio	0.2897	0.2290	0.2419	0.2008*

Note: *optimal cluster choices

Table 7. Internal cluster validation measures for {𝕐}

	k=2	k=3	k=3	k=4
	km=pam=hc=hkm	km=hkm	pam=hk	km=pam=hc=hkm
between ss	6.1615*	3.3944	4.3153	2.7848
within ss	6.6227*	1.2005	1.5957	0.2055
silhouette	0.8323*	0.7454	0.7243	0.7190
CH	69.3120	203.5742	151.6706	744.3940*
dunn	1.7469*	0.4551	1.0344	0.8381
dunn2	10.3460*	1.3930	7.5816	1.5585
entropy	0.2449*	0.6277	0.4851	0.8572
Pearson gamma	0.9231*	0.7557	0.8547	0.6604
wb.ratio	0.0966	0.0504	0.0711	0.0256*

Note: *optimal cluster choices

Table 8. Cluster membership {𝕏} by cluster algorithm used

	k=2	k=3	k=4
kmeans	{BE GE HO ON PA} {R}	{BE} {GE HO ON PA} {R}	{BE} {PA}{GE HO ON} {R}
PAM	{BE GE HO ON PA} {R}	{PA} {BE GE HO ON} {R}	{BE} {PA}{GE HO ON} {R}
Ward.2	{BE GE HO ON PA} {R}	{BE} {GE HO ON PA} {R}	{BE} {PA}{GE HO ON} {R}
Average	{BE}{R}	{BE} {GE HO ON PA} {R}	{BE} {PA}{GE HO ON} {R}
Complete	{BE}{R}	{BE} {GE HO ON PA} {R}	{BE} {PA}{GE HO ON} {R}
hkmeans	{BE GE HO ON PA} {R}	{BE} {GE HO ON PA} {R}	{BE} {PA}{GE HO ON} {R}

Table 9. Cluster membership {𝕐} by cluster algorithm used

	k=2	k=3	K=4
kmeans	{PA} {R}	{PA} {GE ON} {R}	{ON} {PA}{GE ON} {R}
PAM	{PA} {R}	{PA} {ON}{R}	{ON} {PA}{GE ON} {R}
Ward.2	{PA} {R}	{PA} {ON}{R}	{ON} {PA}{GE ON} {R}
Average	{PA}{R}	{PA} {ON}{R}	{ON} {PA}{GE ON} {R}
Complete	{PA}{R}	{PA} {ON}{R}	{ON} {PA}{GE ON} {R}
hkmeans	{PA} {R}	{PA} {GE ON} {R}	{ON} {PA}{GE ON} {R}

Focusing on the FC's output (PQ), inputs (NB, K, L) and fleet structure (NBA) indicators included in variate {𝕏}, Basque FC may be divided in four clusters. Bermeo (cluster 1) and Pasaia (cluster 2) have been isolated alone, no matter the algorithm we are using in the clustering process, and constitute two sound single differentiated groups. Hondarrabia, Getaria and Ondarroa make up (cluster 3), and the rest 10 Basque FC have been grouped in (cluster 4). Bermeo, with just 10% of artisanal vessels, exhibits a clear industrial fishing pattern dominated by the tuna freezer fleet, and concentrates the majority of the Basque fishing input potential, including fishing units, fishermen and capital value. However, although not in the tail of the highest value of landings, the value of landings in Bermeo is significantly below the average of the FC in clusters 2 and 3. Compared to Bermeo, Pasaia shows the opposite behaviour. With a mixture of artisanal and non-artisanal (NBA=50%) and a moderate input size significantly below the figures of Bermeo and cluster 3 members, Pasaia may be catalogued as a leading landing port that attracts vessels of other FC (including French vessels). Cluster 3 is conformed by a set of industrial FCs (the cluster average percentage of the artisanal vessels is 16%) with a more balanced input-output distribution compared to the opposite asymmetry of cluster 1 and 2. The dendogram related to the hierarchical kmeans algorithm (Figure 12) shows that Getaria and Hondarribia have been grouped before, which is an indication of the major similarity between this two fishing communities. Finally, cluster 4 is comprised by a mixture of 10 heterogeneous FC, where fishing (although in different degrees) is nowadays almost testimonial compared to the core FCs in clusters 1, 2 and 3. Again, the dendogram is helpful to identify internal specific patterns within cluster

4. Before joining the second right side subgroup in cluster 4, the left side group in cluster 4 (Bilbao, Plentzia, Lemoniz, Donostia, Mutriku) in mainly conformed by artisanal FC that, due to reduced fleets, lost their first sale fish local markets. The communities belonging to this subgroup of cluster 4 are prone to leave the professional fishing activity in the near future. The remaining FC in cluster 4 share a marked pattern of non-artisanal fleet structure and (apart from Orio) they still maintain their first sale market. Compared to the 4 cluster solution, in the 3 cluster solution Pasaia joins Getaria, Hondarribia and Ondarroa; while the 2 cluster solution groups the 5 top Basque FC in the same cluster. The original cluster 4 in the 4 cluster partition remains unchanged.

Clustering process on economic performance indicators in variate $\{\mathbb{Y}\}$ shows an interesting complement to the taxonomy of the Basque FCs that resulted from variate $\{\mathbb{X}\}$. In the four cluster solution Pasaia (cluster 1) and Ondarroa (cluster 2) have been isolated alone, and constitute two sound single differentiated clusters. Hondarrabia and Getaria make up (cluster 3), and Bermeo joins the tails-end of basque FC. The outstanding average productivity indicators of Pasaia and Ondarroa reinforce their role as key commercialisation ports, the former with a mixed artisanal-industrial fleet structure and the later with a marked industrial one. In an intermediate situation, the cluster made up by Getaria and Hondarribia gives around the 60% of the ratios of Ondarroa, but show a deep gap with respect to the ratios of the FC in cluster 4. Special attention deserves the case of Bermeo. As mentioned above, the fishing potential of Bermeo is highly determined by the tuna freezers. Since this fleet operates in distant waters, their landings may be miss-recorded in the European registers.

Table 10. FC taxonomy $\{\mathbb{X}\}$: average variate values by cluster

k=4

cluster	Cluster membership	PQ	NB	K	NBA	L
1	{BE}	5815517	52	11547196	10	933
2	{PA}	74220667	20	1234434	50	250
3	{GE HO ON}	28094076	28	5447978	16	328
4	{ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	84811	5	302619	54	36

k=3

cluster	Cluster membership	PQ	NB	K	NBA	L
1	{BE}	5815517	52	11547196	10	933
2	{GE HO ON PA}	39625724	26	4394592	25	309
3	{ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	84811	5	302619	54	36

k=2

cluster	Cluster membership	PQ	NB	K	NBA	L
1	{BE GE HO ON PA}	32863682	31	5825113	22	434
2	{ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	84811	5	302619	54	36

Table 11. FC taxonomy $\{\mathbb{Y}\}$: average variate values by cluster

k=4

cluster	Cluster membership	proNB	proK	proL
1	{PA}	3711033.33	60.13	296882.67
2	{ON}	1977297.61	7.53	124478.78
3	{GE HO}	592627.26	3.59	59048.33
4	{BE, ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	19374.61	0.23	2611.54

k=3

cluster	Cluster membership	proNB	proK	proL
1	{ON, GE}	1375688.90	6.16	100795.11
2	{PA}	3711033.30	60.13	296882.67
3	{BE, HO, ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	52024.60	0.41	5809.34

k=2

cluster	Cluster membership	proNB	proK	proL
1	{BE, HO, ZI, BI, DO, LEK, LEM, MUN, MUT, OR, PL, SA}	241119.5	1.232634	19378.74
2	{PA}	3711033.3	60.125262	296882.67

The 3 and 2 cluster solutions show increasing members in the tail-end. Specifically in the 3 cluster solution Hondarribia abandons the group of Getaria and Ondarroa, and finally in the 2 cluster solution all the FC but Pasaia conform a unique cluster (Figure 13).

Figure 12. Dendogram {𝕏} (k=4)

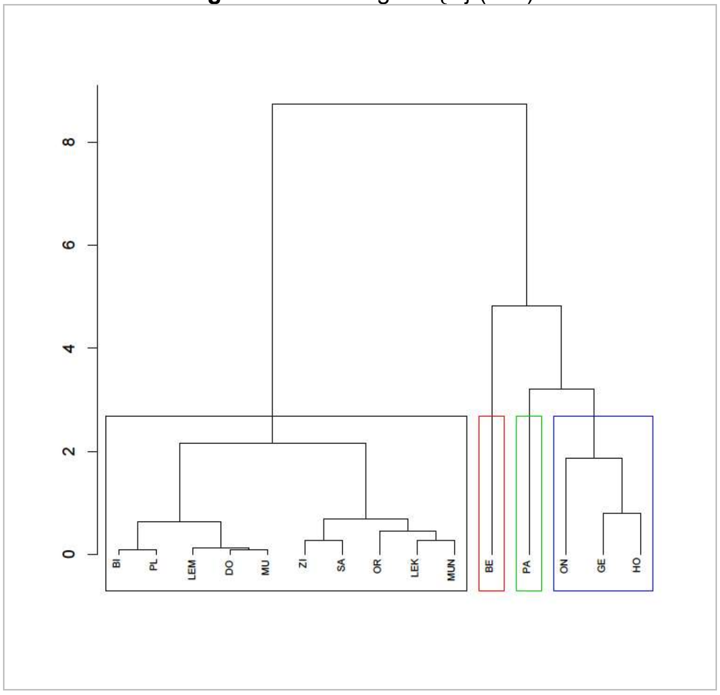
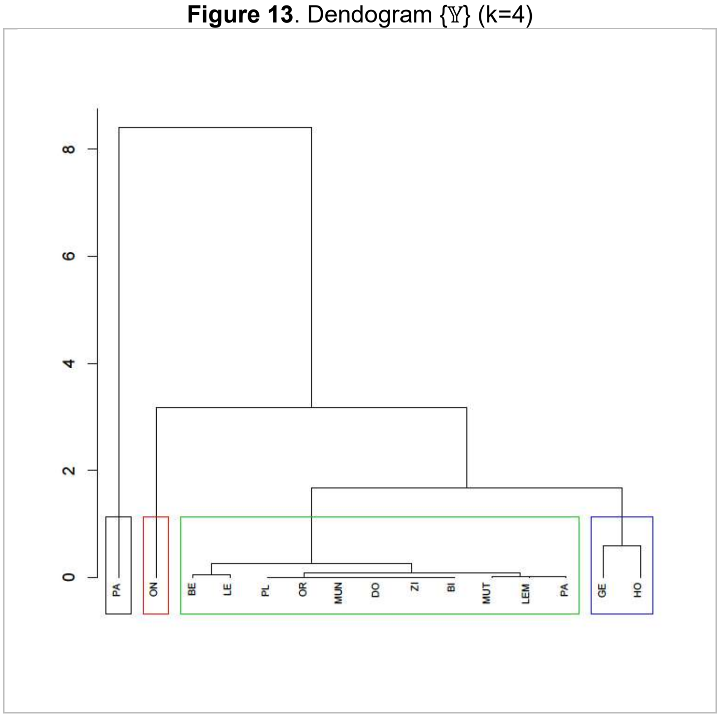


Figure 13. Dendogram {𝕐} (k=4)



4. CONCLUDING REMARKS AND DISCUSSION

Although the participation of the fishing sector in the GDP of the Basque Country does not even reach 0.5%, however, in relative terms, it is still one of the Spanish Autonomous Communities that comparatively has a relevant specific weight in all coastal communities of the Spanish state. Thus, the capacity of its fishing fleet (27%) is the second largest after Galicia (41%), partly as a result of the noticeable industrial pattern of the own Basque fishing fleet. Besides, the Basque is the youngest fishing fleet in average, which, to some extent, shows its potential future projection. Moreover, with the 17% of the value of total fish landed in Spain, the Basque Country is also the second largest after Galicia (35%). These figures indicate that, in comparative terms, Basque fishing is a dynamic socio-economic sector.

The Basque fishing sector exhibits a sound spatial concentration in 15 local fishing communities, which account for professional fishing vessels in the official census of professional fishing, although only 9 of them still maintain the official local first sale market. Herfindahl and Rosenbluth indices for the value of catches, number of vessels, estimated capital value and fishers are around (0.2-0.35), while the related C5 concentration indices for the 5 leading Basque fishing communities respectively reach 99.5%, 77%, 91% and 86%. Concentration is particularly high in the commercialisation of fish, with two local communities, namely, Pasaia and Ondarroa, representing 76% of the total value of the fish landed in the Basque Country.

In the last 30 years, like the whole Spanish fishing sector, the Basque has undergone a profound reconversion, so that its size has been drastically reduced to more than 75%, no matter the indicators taken into consideration. As a result of this radical adjustment, the macro-level figures that have been presented cover a reality that, at the local level, is much less flourishing, raising the issue of the potential for the resilience of the Basque fishing communities. For this reason, in the framework of the three-pillar foundation for ecological, economical and social sustainability, we have changed our attention from the usual regional or fishery level to the local level. Based on a two-step principal component clustering approach the taxonomy of the Basque fishing communities has been identified. The resulting classification is robust to the alternative methods and algorithms we have used, including hierarchical agglomerative (i.e. Ward, average and complete linkage), non-hierarchical (k-means and k-medoids) and the mixed hierarchical-kmeans.

Our results support 4 fishing community typologies in the Basque Country. Group 1 is made up only of Bermeo, community with unique characteristics that differentiate it from all the other Basque fishing communities. The participation of the artisanal fleet in the whole of its fleet is around 10%, and, at the same time, it is the base port of the tuna freezers operating in distant waters. Bermeo concentrates around 80% of the Basque fishing capacity, and therefore may be catalogued as a sound industrial fishing town. However, according to the value of its landings, Bermeo is clearly below the average of the rest of the leading fishing communities and even second level ones, because, obviously, tuna freezers do not make their landings in the local fish market. Another single fishing community, namely Pasaia, comprises group 2. Pasaia is characterized by a balanced distribution of artisanal and non-artisanal vessels, a moderate concentration of fishing inputs (around 2.5% of total capacity) but with a clear leading position in the value of landings (45%), which indicates the role of Pasaia as a pole of attraction for the catches of the fleets of other fishing communities, including the close French ones. Group 3 is comprised by a subset of fishing communities, namely Getaria, Hodarrabia and Ondarroa, with a clear and sound domain of non-artisanal vessels (84%) over artisanal ones (16%), but with a more balanced distribution between fishing inputs and outputs than Bermeo (outstanding in fishing inputs) and Pasaia (outstanding in fishing outputs). The closest fishing communities in this third group are Getaria and Hondarrabia (with around 4% of the capacity and 10% of the landings). Ondarroa with around 8% of the capacity is also the base port of the offshore fleet operating in the Grand Sol and the second key commercialization point for the fish in the Basque Country, with 31% of the value of the total fish landed. Finally, group 4 is a more heterogeneous set of 10 secondary and (in some specific cases) marginal fishing communities, in which two subgroups with differentiable patterns can be identified. On the one hand, Bilbao (Erandio), Plentzia, Armintza, Donostia and Mutriku (with less than 0.07% of the capacity share dominated by Donostia) have lost their first sale market, and accordingly show the weakest position within the group. The rest of communities (Lekeitio, Orio, Santurtzi and Zierbena) have non-artisanal and mixed fleet representing 0.6%, 1.3%, 0.17% and 0.20% of the total capacity and (with the exception of Orio), they still maintain their first sale local fish market, although with very modest figures below 1% of the total value of landings.

Summarising, communities in groups 1,2,3, namely Bermeo, Pasaia, Ondarroa, Getaria and Hondarribia, constitute the core of the strongest Basque fishing communities, with a high-medium potential to survive in their role of fishing communities in the near future. Some of them have adapted themselves to the changing circumstances, such as the disappearance and tighten of certain fleet segments, but still maintaining a powerful and leading first sale local market, while others have opted for changing the local fleet structure favouring distant water fishing and/or fight trying to strengthen their status quo. Although some communities belonging to group 4 show a resilient and a rather stable dynamics, specifically the ones having lost their local fish market are prone to give in to their fishing past and abandon professional fishing activity.

However, fishing is not only the act of catching fish, but also a way of understanding life, interacting with people and living with the natural environment. Fishing encompasses all aspects of culture and characterizes society as a whole. It brings together a set of knowledge, skills, and techniques that pass from generation to generation; It also includes the creation of a set of infrastructures, instruments, devices, rigging, arts, fishing favours the development of specific human organizations, collective and operational rules of action, as well as the promotion of other strongly linked activities, such as tourism, folklore or the gastronomy itself. All these elements are part of the Fishing Cultural Heritage (PCP). PCP includes identity markers common to all fishing communities, as well as specific, unique and exclusive markers of each community (Jiménez and García, 2018). Specific attention should be paid to preserve this heritage before is too late, not only in the local fishing communities prone to disappear (group 4), but also in the leading ones. This might also favour cultural tourism to complement the current socio-economic activity and reinforce local sustainability and resilience.

REFERENCES

AGRAWAL, A.; and GIBSON C.C. (1999): Enchantment and disenchantment: the role of community in natural resource conservation. *World Development*; 27: 629–649.

AMEIJEIRAS-ALONSO, J.; CRUJEIRAS, R.M. and RODRÍGUEZ-CASAL, A. (2018): *R package "multimode". Mode Testing and Exploring*. CRAN.R-project.org.

AMIGO, L.; GARZA, Mª D.; and VARELA M. (2009): Transmisión de precios en la cadena comercial de la pesca costera gallega. El caso del pulpo, *Revista Española de Estudios Agrosociales y Pesqueros*, Nº 222: 1-24.

ANDERSON, J.L.; ANDERSON, C.M, CHU J.; MEREDITH J.; ASCHE F.; SYLVIA G.; SMITH M.D.; ANGGRAENI D.; ARTHUR R.; GUTTORMSEN A.; MCCLUNEY J.K.; WARD T.; AKPALU W.; EGGERT H.; FLORES J.; FREEMAN M.A.; HOLLAND D.S.; KNAPP G.; KOBAYASHI M.; LARKIN S.; MACLAUCHLIN K.; SCHNIER K.; SOBOIL M.; TVETERAS S.; UCHIDA H.; and VALDERRAMA, D. (2015): The Fishery Performance Indicators: A Management Tool for Triple Bottom Line Outcomes. *PLoS ONE*; 10;5: e0122809

ANDERSON, T. (1984): *An Introduction to Multivariate Statistical Analysis*. 2nd Edition. New York Wiley & Sons.

ASCHE, F.; and GUILLEN J. (2012): The importance of fishing method, gear and origin: The Spanish hake market. *Marine Policy*; Vol. 36; 2, 365-369.

BENE, C. (2003): When Fishery Rhymes with Poverty: A First Step Beyond the Old Paradigm on Poverty in Small-Scale Fisheries. *World Development*, Vol. 31, No. 6, pp. 949–975.

BENE, C. (2009): Are fishers poor or vulnerable? Assessing economic vulnerability in small-scale fishing communities. *The Journal of Development Studies*; Vol. 45; 6, 911-933.

BHATNAGAR ET AL (2017): *R Package 'eclust'. Environment Based Clustering for Interpretable Predictive Models in High Dimensional Data*. CRAN.R-project.org.

BRUSCO, M.J.; SINGH, R.; DRADIF, J.D.; and STEINLEY, D. (2017): Cluster Analysis in Empirical OM Research: Survey and Recommendations. *International Journal of Operations & Production Management,* vol 37 3, 300-320.

CALISNKI, T.; and HARABASZ, J. (1974): A Dendrite Method for Cluster Analysis. *Communications in Statistics-Theory and Methods*, 3 (1), 1-27.

CASTILLA, D.; and GARCIA DEL HOYO J.J. (2006): Medición de la capacidad de pesca de la flota voraz del Estrecho de Gibraltar Enfoques paramétricos y no paramétricos. *Revista de Estudios Agrosociales y Pesqueros*. Nº 210; 115-154.

CHENG, M.Y.; and HALL, P. (1998): Calibrating the Excess Mass and Dip Tests of Modality. *Journal of the Royal Statistical Society*: Series B (Statistical Methodology) 60 (3) 579–589.

DEL VALLE, I.; ASTORKIZA, I.; and ASTORKIZA, K. (2003): Fishing Effort Validation and Substitution Possibilities Among Components The Case Study of the VIII Division European Anchovy Fishery. *Applied Economics*; 35; 63-67

DEL VALLE, I.; and ASTORKIZA, K. (2013): Assessing Changes in Capital and Investment as a Result of Fishing Capacity Limitation Programs. *Environmental & Resource Economics*, 52. 223-260.

DUNN, J. (1974): Well Separated Clusters and Optimal Fuzzy Partitions. *Journal of Cybernetics*, 4(1), 95-104.

EUROPEAN COMMISSION. EUMOFA. European Market Observatory for Fisheries and Aquaculture Products (https://www.eumofa.eu/).

EUROPEAN COMMISSION. FLEET REGISTER. http://ec.europa.eu/fisheries/fleet.

FERSE, S. (2010): Allies, not aliens: increasing the role of local communities in marine protected area implementation. *Environmental Conservation*; Vol 37; 1, 23-34.

FISHER, N.I.; and MARRON, J.S. (2001): Mode Testing Via the Excess Mass Estimate. *Biometrika*, 88, 419–517.

GARCÍA-ENRÍQUEZ, J.; HUALDE, J.; ARTECHE, J.; and MURILLAS, A. (2014): Spatial integration in the Spanish mackerel market, *Journal of Agricultural Economics*, Nº. 65; 1, 234-256.

GARZA, D.; and AMIGO, L. (2013): Transmisión vertical de precios en el mercado de productos pesqueros: el caso de la merluza europea en puertos de Galicia, *Revista Galega de Economía*, Vol. 22, Núm. 2: Pág. 79-98.

GARZA, D.; SURÍS, J.; and VARELA, M.M. (2017): Evaluando los Efectos Económicos de las Actividades Pesquera y Acuícola en la Economía de Galicia. *Working Paper*: 17/02. September 2017. UniversidadeVigo.

GOBIERNO DE ESPAÑA. Ministerio de Trabajo, Migraciones y Seguridad Social. (http://www.seg-social.es/wps/portal/wss/internet/EstadisticasPresupuestosEstudios/

GOBIERNO VASCO (NASDAP). Departamento de Desarrollo Económico e Infraestructuras Agricultura, Pesca y Política Alimentaria (https://www.euskadi.eus/gobierno-vasco/agricultura-pesca-industria-alimentaria/inicio/).

GRANEK, E. F. (2008): Engaging Recreational Fishers in Management and Conservation: Global Case Studies. *Conservation Biology*; Vol. 22; 5, 1125-1134.

HAIR, J.F.; BLACK, W.C.; BABIN, B.; and ANDERSON, R.E., (2014): *Multivariate Data Analysis*. Seventh Edition. Pearson Education Limited.

HALL, P.; and YORK, M. (2001): On the calibration of Silverman's test for multimodality. *Statistica Sinica*, 11, 515–536.

HALPERN, B.S.; KLEIN, C.J.; BROWN, C.J.; BEGER, M.; GRANTHAM, H.S.; MANGUBHAI, S.; RUCKELSHAUS, M.; TULLOCH, V.J.; WATTS, M.; WHITE, C.; and POSSINGHAM, H.P. (2013): Achieving the Triple Bottom Line in the Face of Inherent Trade-Offs among Social Equity, Economic Return, and Conservation. *PNAS*; vol.110;15, 6229–6234.

HAIKIDI, M.; BATISTAKIS, I.; and VAZIRGIANNIS, M. (2001): On Clustering Validation Techniques. *Journal of Inteligent Information Systems*, 17, 107-145.

HARTIGAN, J.A. (1985). The Dip Test of Unimodality, *Ann. Statist*. 13, 70–84.

HARTIGAN, J. A.; and HARTIGAN, P. M. (1985): The Dip Test of Unimodality. *Journal of the American Statistical Association*, 86, 738–746.

HENNING, C. (2019): *The R package 'fpc'. Flexible Procedures for Clustering*. CRAN.R-project.org.

HIMES, A. H. (2003): Small-Scale Sicilian Fisheries: Opinions of Artisanal Fishers and Sociocultural Effects in Two MPA Case Studies. *Coastal Management*, 31:4, 389-408.

HOPKINS, B.; and SKELLAM, J.G. (1954): A New Method for Determining the Type of Distribution of Plant Individuals. *Annals of Botany*, 18 (70) 213–227.

JIMÉNEZ DE MADARIAGA, C.; and GARCÍA DEL HOYO, J.J. (2018): Enhancing of the cultural fishing heritage and the development of tourism: A case study in Isla Cristina (Spain); *Ocean and Coastal Management* 168 (2019) 1–11.

JIMÉNEZ-TORIBIO R; and GARCÍA DEL HOYO J. J. (2006): Evidence of market price leadership in the Spanish red seabream value chain: implications for fisheries management, *Fisheries Research*, 81; 51-59.

JOLLIFFE, I. T. (1986): *Principal Component Analysis.* New York: Springer-Verlag.

JOHNSON, R.; and WICHWEN, D. (1988): *Applied Multivariate Statistical Analysis*. Prentice Hall.

KAISER, H.F. (1958): The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psycometrika,* 23 187-200.

KASSAMBARA, A. (2017): *Practical Guide to Cluster Analysis in R. Unsupervised Machine Learning*. Publised by STHDA.

KAUFMAN, L.; and ROUSSEEUW, P.J. (1990): *Finding Groups in Data: An Introduction to Cluster Analysis*. 2005 John Wiley & Sons, Inc.

KIRKLEY, J.; and SQUIRES, D. (1988): A Limited Information Approach for Determining Capital and Investment in a Fishery. *Fish Bull*, 86(2):339–349.

LAWSON, R.G.; and JURS, P.C. (1990): New Idex for Clustering Tendency and its Application to Chemical Problems. *Journal of Chemical Information and Computer Sciences*, 30 36–41.

LEBART, L.; MORINEAU, A.; and WARWICK, K. (1984): *Multivariate Descriptive Statistic Analysis*. New York, John Wiley and Sons.

LEHMANN, D.R. (1979): *Market Research and Analysis*. Irwin, Homewood, IL.

LOSTADO R. (1997): Recursos y pesquerías en el Mediterráneo español. *Papeles de Economía Española.* Nº 71;133-145.
LOSTADO R. (2000): Reflexiones sobre actividad pesquera y desarrollo sostenible, el caso valenciano. *Noticias de la economía pública social y cooperativa*. Nº. 32, págs. 39-45.
MAECHLER, M. (2019): *R package "cluster". Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al*. CRAN.R-project.org.
MEILA, M. (2007): Comparing Clustering. An Information based Distance. *Journal of Multivariate Analysis*, 98, 873-895.
MILLIGAN, G.W. (1996): Clustering Validation: Results and implications for applied analyses. In Phipps Arabie, Lawrence Hubert & G. DeSoete (Eds.), *Clustering and classification*, River Edge, NJ: World Scientific, p. 341-375.
NILSSON, D.; BAXTER G.; JAMES R.A.; BUTLER J.; and MCALPINE C. (2016): How do community-based conservation programs in developing countries change human behaviour? A realist synthesis. *Biological Conservation*; 200, 93–103.
PASCUAL, J.J. (1991): *Entre el mar y la tierra: Los pescadores artesanales canarios*. Santa Cruz de Tenerife: Ministerio de Cultura-Interinsular Canaria.
PASCUAL, J.J. (2003): Del "mar es de todos" al mar reservado: turistas, poblaciones de pescadores y reservas marinas en Canarias. Pasos: *Revista de Turismo y Patrimonio Cultural*, Vol. 1, Nº. 1; 65-78.
PITA, P. and VILLASANTE, S. (2019): The building of a management system for marine recreational fisheries in Galicia (NW Spain); *Ocean & Coastal Management* 169; 191-200.
R CORE TEAM, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (ISBN 3-900051-07-0, URL http://www.R-project.org/).
RAYCHAUDHURI, S.; STUART, J.M; and ALTMAN, R.B. (2000): Principal Components Analysis to Summarize Microarray Experiments: Application to Sporulation Time Series. In *Pacific Symposiumon Biocomputing*, vol.5.
RICKELS, W.; WEIGAND, C.; GRASSE, P.; SCHMIDT, J.; and VOSS, R. (2019): does the European Union achieve comprehensive blue growth? Progress of EU coastal states in the Baltic and North Sea, and the Atlantic Ocean against sustainable development goal 14; *Marine Policy*; article in press.
RODRI'GUEZ, D.; and VILLASANTE, S. (2012): The Spanish fishing fleet and the economic value of Southern stock of European hake fishery (Merluccius merluccius); *Ocean & Coastal Management*; Volume: 70 Issues: 59—67.
ROUSSEEUW, P. (1987): SILHOUETTES: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.
SACHS, J. (2012): From Millennium Development Goals to Sustainable Development Goals. *The Lancet*; 379, Issue 9832; 2206-2211.
SILVERMAN, B. W. (1981): Using Kernel Density Estimates to Investigate Multimodality. *Journal of the Royal Statistical Society*. Series B, 43, 97–99.
SYMES D.; PHILLIPSON J.; and SALMI P. (2015): Europe Coastal Fisheries Instability and the Impacts of Fisheries Policy. *Sociologia Ruralis*, Vol 55; 3.
TULER, S; AGYEMAN, J.; PINTO DA SILVA, P.; ROTH LORUSSO, K.; KAY, R. (2013): Assessing Vulnerabilities Integrating Information about Driving Forces that Affect Risks and Resilience in Fishing Communities. *Ocean & Coastal Management* 71;131-140.
VILLASANTE, C.S.; GARCÍA-NEGRO, Mª.C., GONZÁLEZ-LAXE, F. (2011): Overfishing and the Common Fisheries Policy: (un)successful resultants from TAC regulation?. *Fish and Fisheries*, Vol. 12; 1; 34-50.
VISBECK, M.; KRONFELD-GOHARANI, U.; NEUMANN, B.; RICKELS, W.; SCHMIDT, J.; VAN DOORN, E.; MATZ-LÜCK, N.; OTT, K. and QUAAS, M.F. (2014): Securing blue wealth: The need for a special sustainable development goal for the ocean and coasts. *Marine Policy* 48; 184–191.